Journal of Pathology Informatics 14 (2023) 100330

Contents lists available at ScienceDirect

Journal of Pathology Informatics

journal homepage: www.elsevier.com/locate/jpi

Short Communication

Automated HL7v2 LRI informatics framework for streamlining genomics-EHR data integration

Robert H. Dolin [a,1], Rohan Gupta [b,a], Kimberly Newsom [c], Bret S.E. Heale [d], Shailesh Gothi [c], Petr Starostik [c], Srikar Chamala [c,e,f,*]

[a] Elimu Informatics, Inc, Richmond, CA, USA
[b] Department of Computer Science, Shri Mata Vaishno Devi University, Katra, Jammu and Kashmir, India
[c] Department of Pathology, Immunology and Laboratory Medicine, University of Florida, Gainesville, FL, USA
[d] Humanized Health Consulting LLC, Salt Lake City, UT, USA
[e] Department of Pathology and Laboratory Medicine, Children's Hospital Los Angeles, CA, USA
[f] Department of Pathology, University of Southern California, CA, USA

ARTICLE INFO

Keywords:
Clinical genomics
Electronic health record
EHR
Laboratory informatics
Genomics module
HL7

ABSTRACT

While VCF formatted files are the lingua franca of next-generation sequencing, most EHRs do not provide native VCF support. As a result, labs often must send non-structured PDF reports to the EHR. On the other hand, while FHIR adoption is growing, most EHRs support HL7 interoperability standards, particularly those based on the HL7 Version 2 (HL7v2) standard. The HL7 Version 2 genomics component of the HL7 Laboratory Results Interface (HL7v2 LRI) standard specifies a formalism for the structured communication of genomic data from lab to EHR. We previously described an open-source tool (vcf2fhir) that converts VCF files into HL7 FHIR format. In this report, we describe how the utility has been extended to output HL7v2 LRI data that contains both variants and variant annotations (e.g., predicted phenotypes and therapeutic implications). Using this HL7v2 converter, we implemented an automated pipeline for moving structured genomic data from the clinical laboratory to EHR. We developed an open source hl7v2GenomicsExtractor that converts genomic interpretation report files into a series of HL7v2 observations conformant to HL7v2 LRI. We further enhanced the converter to produce output conformant to Epic's genomic import specification and to support alternative input formats. An automated pipeline for pushing standards-based structured genomic data directly into the EHR was successfully implemented, where genetic variant data and the clinical annotations are now both available to be viewed in the EHR through Epic's genomics module. Issues encountered in the development and deployment of the HL7v2 converter primarily revolved around data variability issues, primarily lack of a standardized representation of data elements within various genomic interpretation report files. The technical implementation of a HL7v2 message transformation to feed genomic variant and clinical annotation data into an EHR has been successful. In addition to genetic variant data, the implementation described here releases the valuable asset of clinically relevant genomic annotations provided by labs from static PDFs to calculable, structured data in EHR systems.

Introduction

Next-generation sequencing (NGS) is a versatile tool that has been rapidly adopted in the clinical laboratory for pharmacogenomics, inherited disorders, tumor mutation testing, and more. One of the driving forces behind this is the massive amount of data that can be provided with a single test. NGS testing offers detailed information for a variety of genetic variations and can provide insight to therapeutic targets. However, a major barrier for implementing NGS in a clinical setting is the return of NGS data into the electronic health record (EHR) in a way that can be useful to a clinician. Information provided from labs is not yet standardized and may have variances in variant classification, annotation and reporting, making clinical NGS reports complicated and difficult to read. Additionally, data, even when structured within the lab, is not easily integrated into the EHR and is usually scanned in as a PDF that lacks the utility of structured data. PDF formatting of NGS data makes this data difficult to manage and limits

* Corresponding author at: Department of Pathology and Laboratory Medicine, Children's Hospital Los Angeles, 4650 Sunset Blvd, Los Angeles, CA 90027, USA.
E-mail addresses: bdolin@elimu.io (R.H. Dolin), kja3@ufl.edu (K. Newsom), bheale@humanizedhealthconsulting.com (B.S.E. Heale), starostik@pathology.ufl.edu (P. Starostik), schamala@chla.usc.edu (S. Chamala).
[1] Robert H. Dolin and Rohan Gupta contributed equally to the manuscript and are declared as co-first authors.

http://dx.doi.org/10.1016/j.jpi.2023.100330
Received 29 March 2023; Received in revised form 8 July 2023; Accepted 9 August 2023
Available online 15 August 2023

the overall usefulness of this data. Structured data can be easily manipulated making this a highly sought after goal for clinical NGS data, particularly where that structured data adheres to or is convertible to a standard format that can be consumed by EHRs.

In the work reported here, we sought to develop and implement an automated pipeline for moving standards-based structured genomic data (both variants and annotations) directly from the testing lab to the EHR.

Technical background

Reporting clinical NGS data requires sophisticated tools and specially trained personnel for the evaluation of and interpretation of genetic data.[1] The Association for Molecular Pathology, American Society of Clinical Oncology, College of American Pathologists, and American College of Medical Genetics have set standards and guidelines for the interpretation and reporting of sequence variants.[2,3] These guidelines provide recommendations for information that should be included for each variant. This information, including gene symbol, variant location, variant type, HGVS nomenclature for cDNA sequence changes, and predicted protein sequence alterations allow for clarity of the results. The standard format for the representation of detected variants is the variant call format (VCF).[4] However, for this data to be clinically useful, it also requires variant annotation and interpretation. A critical step for reporting clinically useful data is the annotation of VCF data using genomic coordinates to cDNA and protein changes using clinically accepted transcripts which ensures that the correct variants are reported according to HGVS nomenclature. These variants can then be categorized based on clinical impact. The inclusion of the fully annotated variant data into the EHR allows for providers to have a full understanding of the variants being reported and reduces the possibility of variant misidentification which could negatively impact patient care.

Returning sequencing data to the EHR in a structured format requires the data to be in a format that is compatible with the EHR and for the EHR to be designed to handle this relatively new type of data. Historically, for many assays, the data returned back to the EHR would consist of only a few values. For NGS, there can be thousands of lines of data with different information for each variant and there can also be different types of variants. We previously described an open source tool that converts VCF files into HL7 FHIR format,[5] an increasingly popular format for EHR interoperability. However, many EHRs have limited support today for HL7 FHIR, whereas most EHRs support earlier HL7 interoperability formats such as those based on the HL7 Version 2 standard.

There have been several standards-based approaches to genomics-EHR integration including approaches to integrate using HL7 FHIR genomics, SMART-on-FHIR, CDS Hooks, and the HL7 Version 2 genomics component of the HL7 Laboratory Results Interface (HL7v2 LRI based on HL7 Version 2.5.1).[6–10] As an EHR vendor, EPIC supports a specification based on the HL7v2 LRI Clinical Genomics component.[11–13] The HL7v2 LRI Clinical Genomics component builds on the HL7v2 syntax, with genomics specific LOINC codes, and specific extensions/constraints on the base HL7v2 specification. The HL7v2 LRI Clinical genomics specification is a genomics-specific component that requires specific handling. As a standard approach, HL7v2 lab order results include a header section (MSG), order information (OBR segments), and report of observed data (OBX segments). It is the OBX segments where much of the HL7v2 LRI Clinical Genomics specification focuses. EPIC has created a further interface specification based on the HL7v2 LRI Clinical Genomics specification that enables populating its Genomic Module, which makes some genetic variation and annotation data available in the clinical workflow.

Overall structure of the OBXs in an HL7 LRI report is shown in Fig. 1. Section 1 of a genomics report will contain observations that apply to the

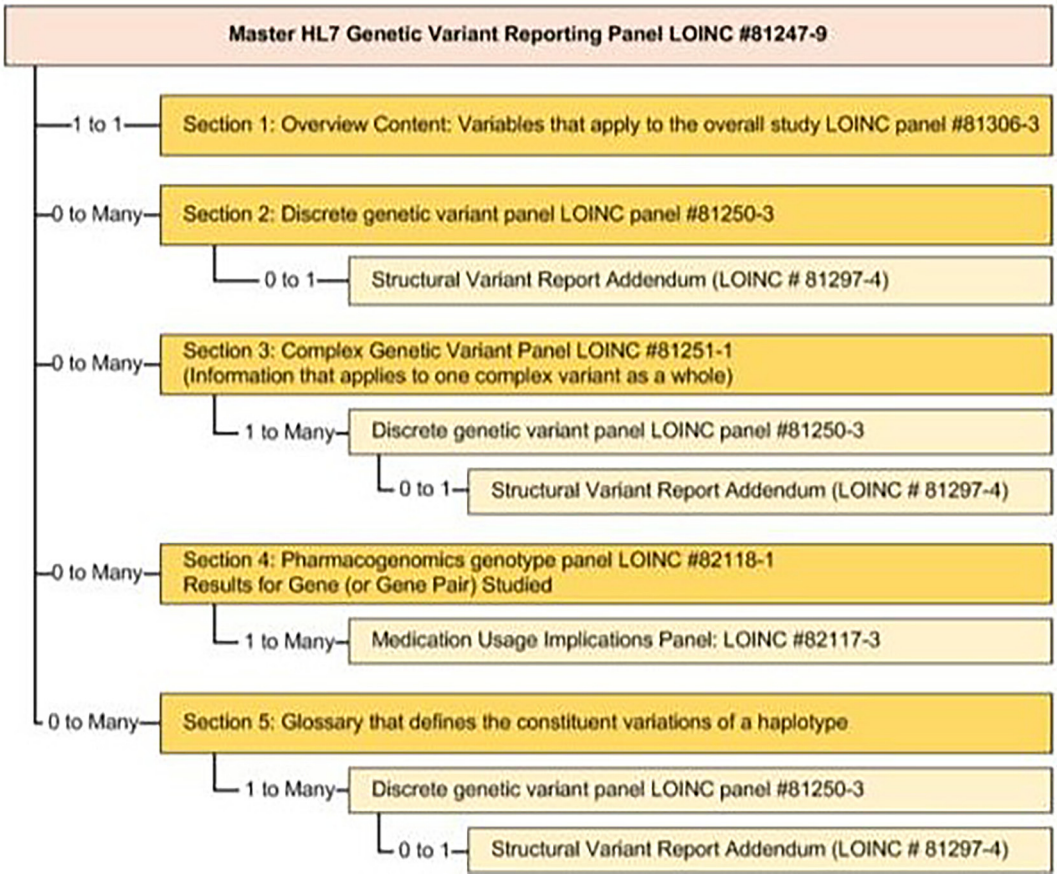

Fig. 1. Graphical structure of the HL7 LRI message. The Master HL7 Genetic Variant Reporting Panel is comprised of observations related to the overall study, to discrete variants, to complex variants, and to pharmacogenomic findings and implications (Used with permission from HL7 International).

```
OBX|23|ST|83005-9^Variant Category^LN|2b|Simple|
OBX|24|ST|47998-0^Variant Display Name^LN|2b|NC_000017.11:7673766:C:T|
OBX|25|ST|48018-6^Gene Studied^LN|2b|HGNC:11998^TP53^HGNC|
OBX|26|ST|48004-6^DNA Change c.HGVS^LN|2b|NM_000546.6:c.853G>A|
OBX|27|ST|48005-3^Amino Acid Change p.HGVS^LN|2b|NP_000537.3:p.Glu285Lys|
OBX|28|CWE|48006-1^Amino Acid Change [Type]^LN|2b|LA6698-0^Missense^LN|
OBX|29|ST|48013-7^Genomic reference sequence^LN|2b|NC_000017.11|
OBX|30|ST|69547-8^Genomic ref allele^LN|2b|C|
OBX|31|NR|81254-5^Genomic allele start-end^LN|2b|7673766|
OBX|32|ST|69551-0^Genomic alt allele^LN|2b|T|
OBX|33|ST|48002-0^Genomic Source Class^LN|2a|LA6684-0^Somatic^LN|
OBX|34|ST|53037-8^Variant Significance^LN|2b|LA6668-3^Pathogenic^LN|
OBX|35|ST|69548-6^Genetic Variant Assessment^LN|2b|Present|
OBX|36|CWE|81259-4^Phenotype^LN|2a|699346009^Hereditary cancer predisposition^SCT|
OBX|37|NM|81258-6^Allelic Frequency^LN|2a|0.15|
OBX|38|NM|82121-5^Allelic Read Depth^LN|2a|320|
```

Fig. 2. HL7v2 LRI message. The observations describe a somatic missense mutation where base 'C' has changed to base 'T' in the TP53 gene at chromosome position 7673766.

overall study—such as the reason for the study, genetic conditions being assessed, genes or regions studied by the test, and overall interpretation. Section 2 reports discrete (simple or structural) variants. Variants can be represented in VCF-like fields, as HGVS expressions, or with public identifiers (e.g., a variant's ClinVar ID). Details of the variant, such as the gene location, the predicted amino acid change, germline vs. somatic, detection method, clinical significance, associated phenotype, and allelic state, can be included. An example of an HL7v2 LRI representation of a variant a somatic missense mutation where base 'C' has changed to base 'T' in the TP53 gene at chromosome position 7673766) is shown in Fig. 2. Section 3 reports complex variants, where many discrete (simple or structural) variants taken together have one effect or phenotype. Potential observations that can be included represent phenotype, clinical significance, allelic state, and the individual variants that comprise the complex variants. Section 4 reports pharmacogenomic (PGx) observations. A PGx panel of observations contains information about the variants or gene haplotypes, along with phenotype information (e.g., 'intermediate metabolizer of clopidogrel') and therapeutic recommendations. Section 5 provides a way for a laboratory to define the constituent variants of a haplotype.

Approach

For this project, we developed and implemented an automated pipeline for converting structured genomic data (both variants and annotations) directly from the testing lab into HL7v2 LRI format (Fig. 3 and Fig. 4) that further conformed to Epic's HL7v2 LRI import specification.

We began by modifying our previously published open-source vcf2fhir conversion tool,[14] into a similar open-source conversion utility named hl7v2GenomicsExtractor.[15] Conceptually, the hl7v2GenomicsExtractor utility takes a variant annotation file as input and outputs a set of HL7v2 OBX observations, as shown in Fig. 3. Software converts simple variants (SNV, MNV, and Indel), autosomes, sex chromosomes, and mitochondrial DNA. These OBX observations can be then incorporated into a complete HL7v2 lab message. It should be noted that OBX observations are an important part of a complete HL7 message, but that there are other necessary components (e.g., message header, patient information) in a complete message. As such, the hl7v2GenomicsExtractor utility can be thought of as a component of a larger workflow, such as is shown in Fig. 4, where an HL7v2 engine calls hl7v2GenomicsExtractor specifically to construct

Fig. 3. hl7v2GenomicsExtractor conversion. Conceptually, the hl7v2GenomicsExtractor utility takes a VCF file as input and outputs a set of HL7v2 OBX observations.

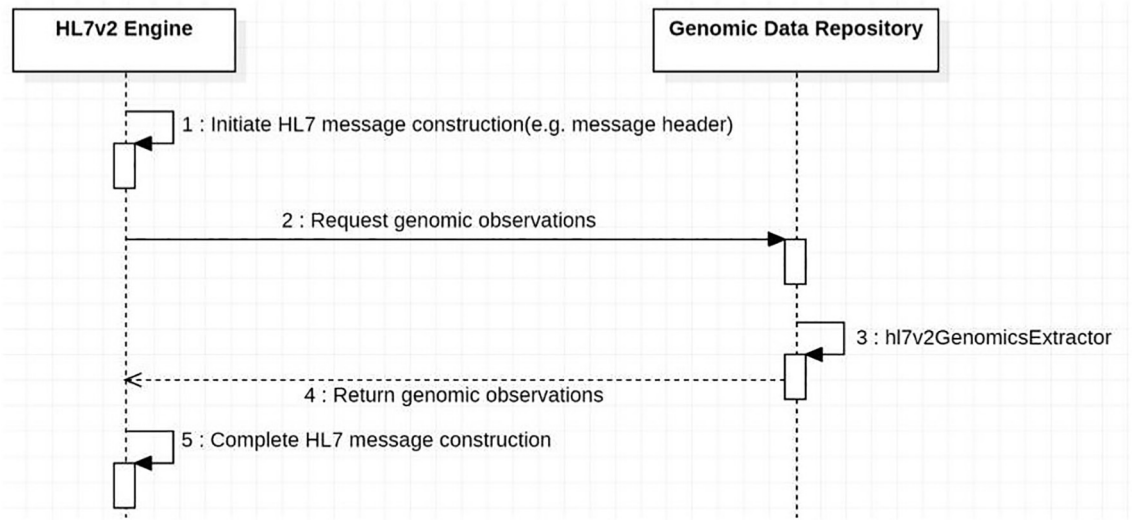

Fig. 4. HL7v2 LRI utilities as components of a larger workflow. The hl7v2GenomicsExtractor utility can be thought of as a component of a larger workflow, where an HL7v2 engine calls the extractor specifically to construct component observations within the overall HL7v2 message.

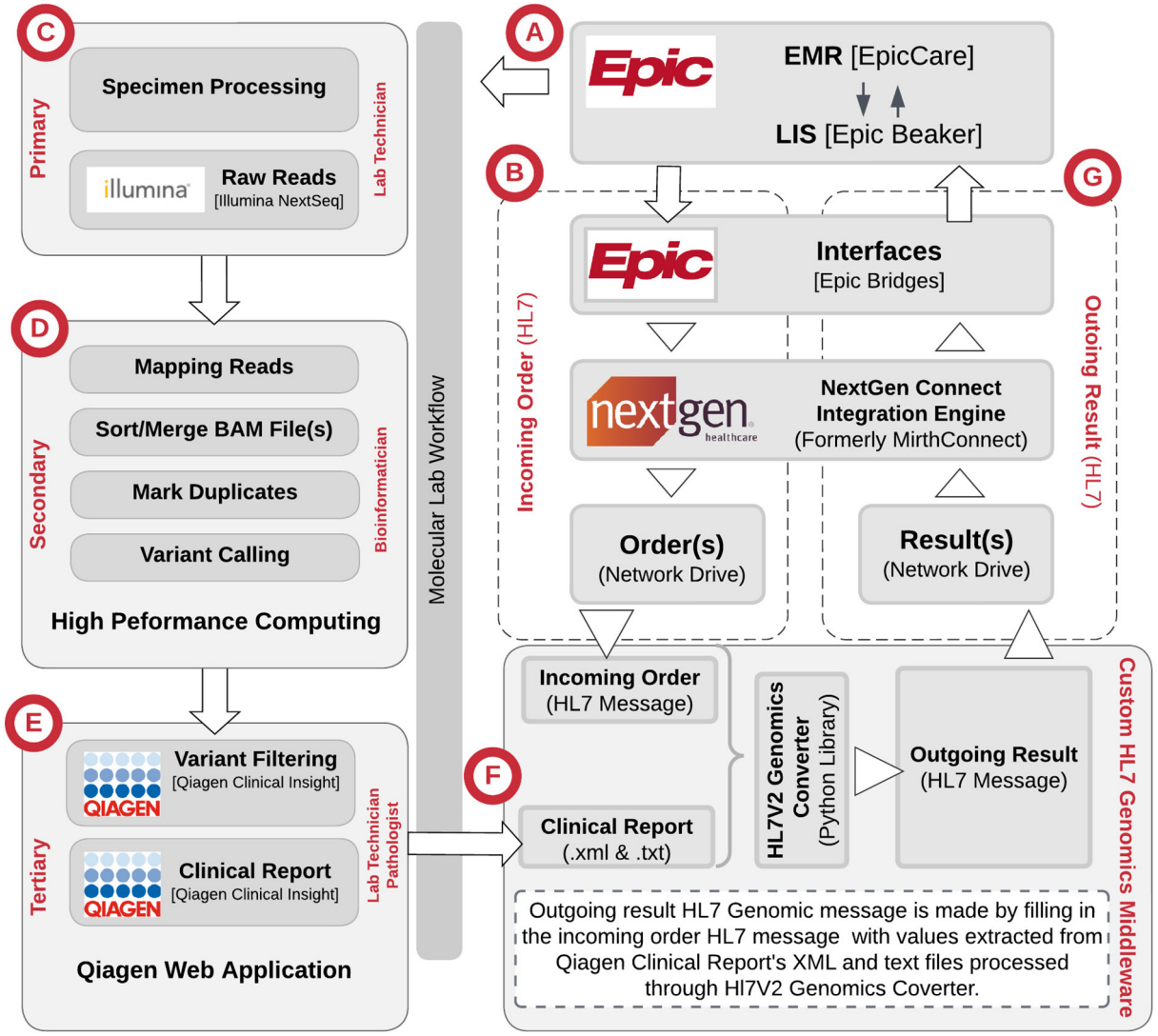

Fig. 5. The complete sequence of steps, from ordering a genomic lab test to generating a genomics lab report, and its subsequent integration into the electronic health record (EHR) using genomics converter and middleware.

component OBXs within the overall HL7v2 message. The specific pipeline used at our location to add message header and additional site-specific OBXs can be found in our hl7v2GenomicsLabMiddleware public github repository.[16]

Subsequent to building hl7v2GenomicsLabMiddleware, we extended the conversion capabilities so that the converter can accept as input not only a VCF file, but also the converter can accept as input XML files generated by Qiagen Clinical Insight. We modularized the conversion code to simplify future support for other source annotated VCF files. All source code, including SnpEff[17] and Qiagen extensions, is available here.[15] We have previously described a custom middleware solution for moving genomic data into the EHR.[18–21] For the work described in this manuscript, the middleware solution was modified, with the revised solution shown in Fig. 5 and summary of tools described in Table 1. The previously employed middleware solution centered around storing genomic data as lab–result–value pairs, emulating the strategy utilized in complete blood count tests. In that model, we denoted each gene as a distinct result component, with its corresponding genomic variant serving as its value. However, in the current paper, we've represented genomic data using the HL7V2 LRI tailored specifically for genomic data. This differs from our previous method, which relied on generic laboratory HL7 results.

Genomics lab middleware

Generation of variant file, annotation and XML export: In Step A, as soon as genomic testing is ordered and a specimen is collected, it appears on the pathology laboratory worklist in Epic Beaker LIS. We configured our electronic interface software EPIC Bridges and NextGen Connect for simultaneously sending an order HL7 message (Step B) to a network folder. Once the specimen is received in the laboratory, the DNA is processed (Step C). The resulting raw DNA sequencing data are run through our clinical grade custom bioinformatics pipeline on high-performance computing infrastructure (Step D). This will output genomic variants (in VCF format) which are automatically uploaded to Qiagen Clinical Insight (QCI) Interpret. Laboratory technologist(s) and molecular pathologist(s) use QCI Interpret to shortlist the genomic variants based on their known evidence of clinical actionability. These clinically actionable genomic variants and descriptions of their clinical impact are output as XML and text files (Step E).

Parsing XML into LRI format: The XML files from Step E are passed to our middleware solution, 'hl7v2GenomicsLabMiddleware.py'. The first action of the middleware is to leverage our 'hl7v2GenomicsExtractor' python library to convert XML into result component OBXs in HL7 LRI format.

Combining OBXs to generate Full HL7 genomic test result message: Genomic lab middleware then combines HL7 LRI format OBXs with the incoming HL7 order message elements to create an outgoing HL7 result message (Step F). This outgoing HL7 message is placed on the network drive and is automatically picked up by NextGen Connect, pushed into Beaker (Step G), and finally delivers the HL7 message to the Epic EHR.

Genomics result message data are stored in LIS

The ability to display and store NGS data in a detailed structured format simplifies the resulting and reporting of data. The Epic Beaker LIS genomics module has 7 different categories of information for each variant. This includes basic information (variant name, frequency, and clinical significance), sequence change (HGVS information), variant coding system, region, allele information, sequence information, and associated phenotype information. Without an automated system, this information would be impossible to capture and only a fraction of the data supported here would be able to be reported.

Figs. 6–9 shows the detailed level of information that is accepted in the Beaker genomics module. The required data is shown with an exclamation point in a red circle and the suggested data is shown with an exclamation point in a yellow triangle. All other data is optional according to the user's needs.

Epic defines additional constraints on HL7v2LRI in order to receive the genomic information.[12] The OBX segment is used to transmit a single observation or observation fragment. The HL7 data type is the expected data type for the observation's value. To return all of this data dynamically, each result type is assigned an Observation ID (OBX-3). The Observation ID can either be a LOINC Code or a local code can be created and mapped when no LOINC is available or a local code set already exists. The observation Sub-ID field (OBX-4) is used to distinguish between multiple OBX segments with the same observation ID. Each variant is grouped by a Sub-ID (OBX-4) which ties variant data across multiple OBX segments together so each single variant should have the same sub-ID. The result for each item is stored in the Observation Value (OBX-5) and should match the HL7 Data Type. Fig. 10 shows the view in Beaker when the HL7 message is transmitted to the EHR. From here, the data can be reviewed in Beaker prior to sign-out and release to EPIC.

Beaker supports additional components that can be used to store data types other than variant information. This data can be posted to the chart in EPIC or it can be suppressed and stored only as a discrete data field that can be later retrieved if needed. Fig. 11 shows an example of the types of data that can be stored.

The genomic data is integrated into the EHR

When the data is reviewed and finalized within Epic Beaker LIS, it is directly reported into the clinician-facing Epic genomic module. The SnapShot dashboard Fig. 12 within Epic genomic module is where genomic test results are presented where variant data is organized by priority. Within this module, there is an option for visualizing expanded/detailed variant information.

Another key point for migrating to this resulting system is the storage of data into discrete data fields. Since the data is structured, it is also searchable using standard reporting workbench queries Fig. 13 where the report logic defined by the user can provide useful reports based on variants.

Table 1
Summary of the various tools used in this pipeline, each with their specific purpose, to facilitate the transfer of genomic data from the molecular pathology laboratory to the electronic health record system.

Tool name	Purpose/Description
EpicCare	A comprehensive, integrated electronic health records system developed by Epic Systems Corporation. This is where clinicians place genomic test orders and final results are reported back and stored in the Genomic Module of EpicCare.
EpicBeaker	A Laboratory Information System provided by Epic Systems Corporation, is the platform where molecular and genomic tests are administered and processed.
EpicBridges	Epic Bridges is the integration engine for Epic Systems software products. It is used to connect the various Epic modules with external systems and applications, facilitating the seamless exchange of data between them.
NextGen Connect	Formerly known as Mirth Connect, is an open-source, cross-platform interface engine used in the healthcare industry.
hl7v2GenomicsExtractor	Python library code for converting genomic variant annotations into result component OBXs in HL7 LRI format (version 0.0.4).
hl7v2GenomicsLabMiddleware	Python wrapper code for adding HL7 message header and footer to those observations created by hl7v2GenomicsExtractor.
Qiagen Clinical Insight	This is variant annotation and reporting tool for clinical genomic testing used by molecular pathology laboratories.

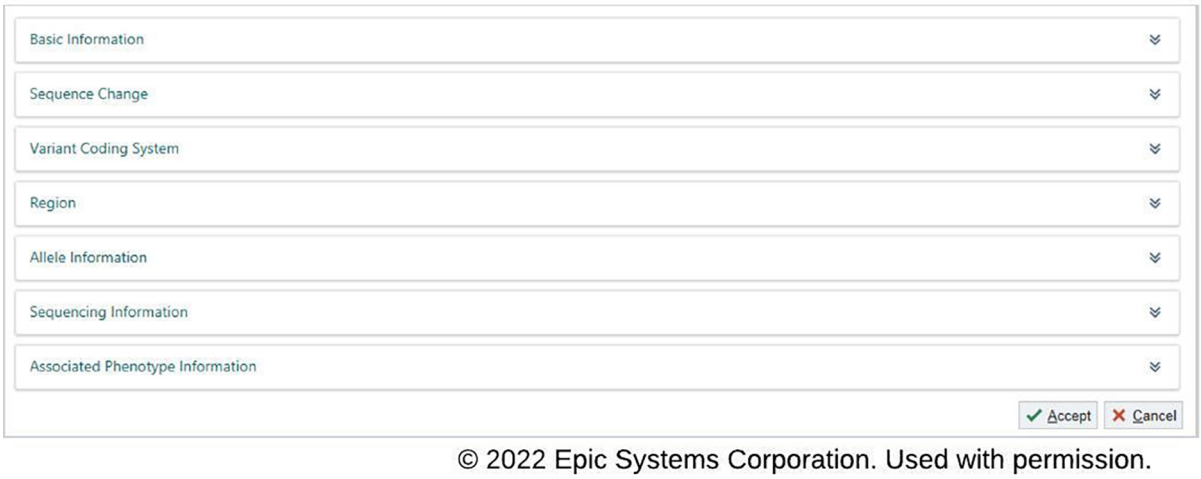

Fig. 6. Collapsible view of all genomic data sections within the Epic Beaker Laboratory LIS, where HL7V2 LRI data has been imported.

Discussion

Several issues were encountered in the development and deployment of the HL7v2 converter—2 having to do with data variability, and 1 having to do with practical integration considerations.

Inconsistencies in variant annotation representation pose challenges to the development of an automated HL7v2 conversion utility. While many popular automated annotation pipelines (e.g., SnpEff,[22] ANNOVAR,[23] and Ensembl's VEP[24]) have adopted a common annotation representation,[25] many annotation pipelines (e.g., Illumina Nirvana,[26] SnpSift[22]) define alternative annotation formalisms. hl7v2GenomicsExtractor currently converts annotations from SnpEff and SnpSift, and from Qiagen XML source files. The code is modularized to simplify future support for other source annotated VCF files. Ultimately, standards such as the FHIR Genomics reporting guide[10] and the GA4GH Variant Annotation specification[27] hold promise for remediating this issue.

Inconsistencies in the representation of variants themselves, not only across different variant formats such as HGVS and VCF, but even across VCF files, also pose a barrier to computational annotation, particularly where the intent is to automatically look up a patient's variant in a knowledge base to obtain annotation information. Studies show significant inconsistency in variant representation across tools and databases,[28] with different sources either not normalizing variants, or normalizing variants using different algorithms (e.g., VEP, 1000 Genomes,[29] and gnomAD[30]) use a 'left-aligned' variant normalization strategy[31]). In our implementation, annotation fidelity lies within QCI Interpret. A consistent normalization approach, such as with NCBI's SPDI normalization algorithms[32] will likely enable higher fidelity annotations.

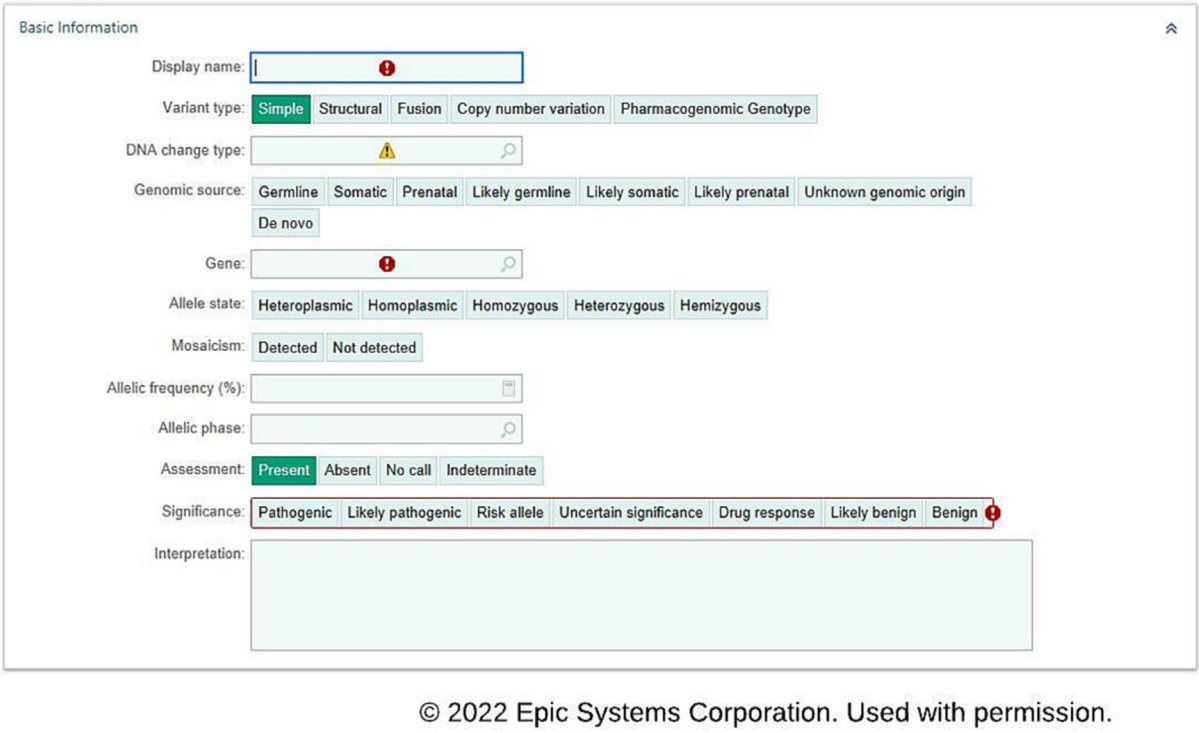

Fig. 7. Expanded view of the 'Basic Information' section of the genomic data in the Epic Beaker LIS, which is populated with imported HL7V2 LRI data transmitted from the genomics middleware.

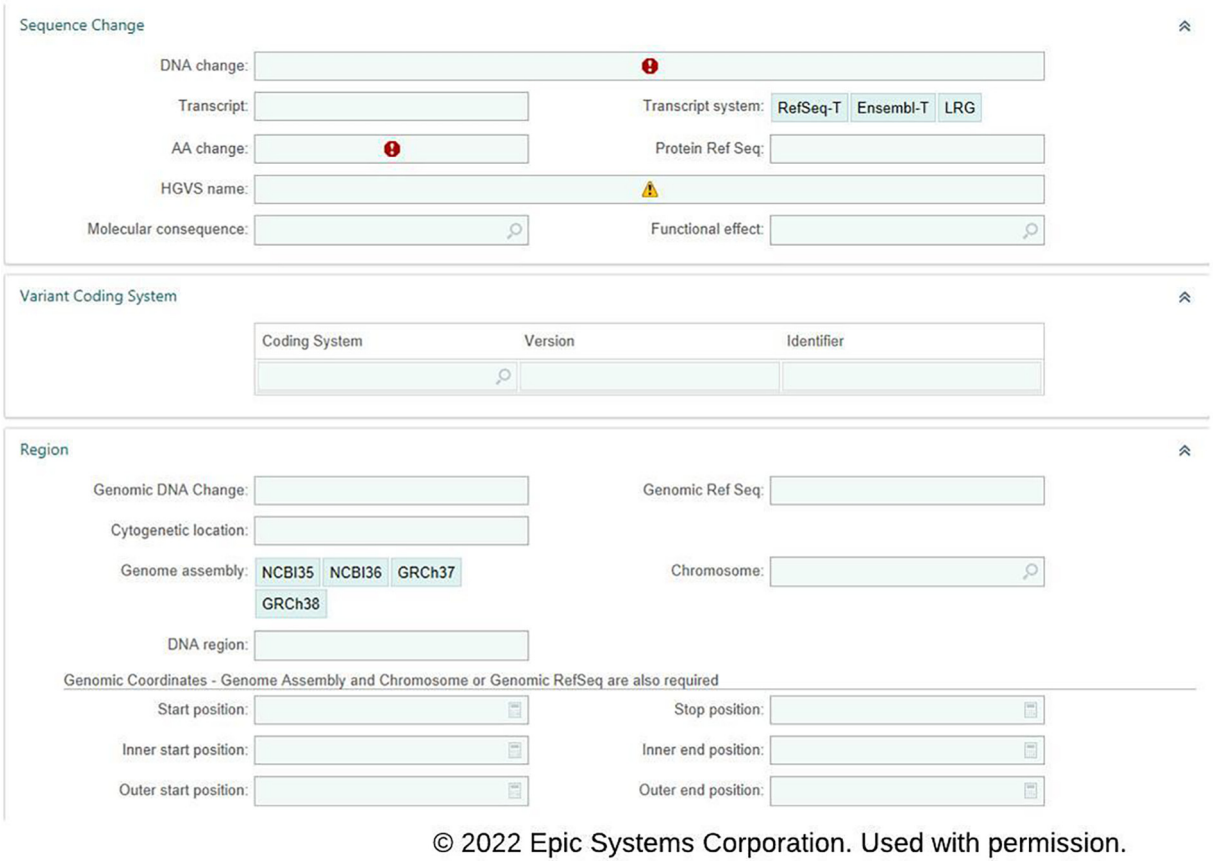

Fig. 8. Expanded view of the 'Sequence Change', 'Variant Coding System', and 'Region' sections of the genomic data in the Epic Beaker LIS, which is populated with imported HL7V2 LRI data transmitted from the genomics middleware.

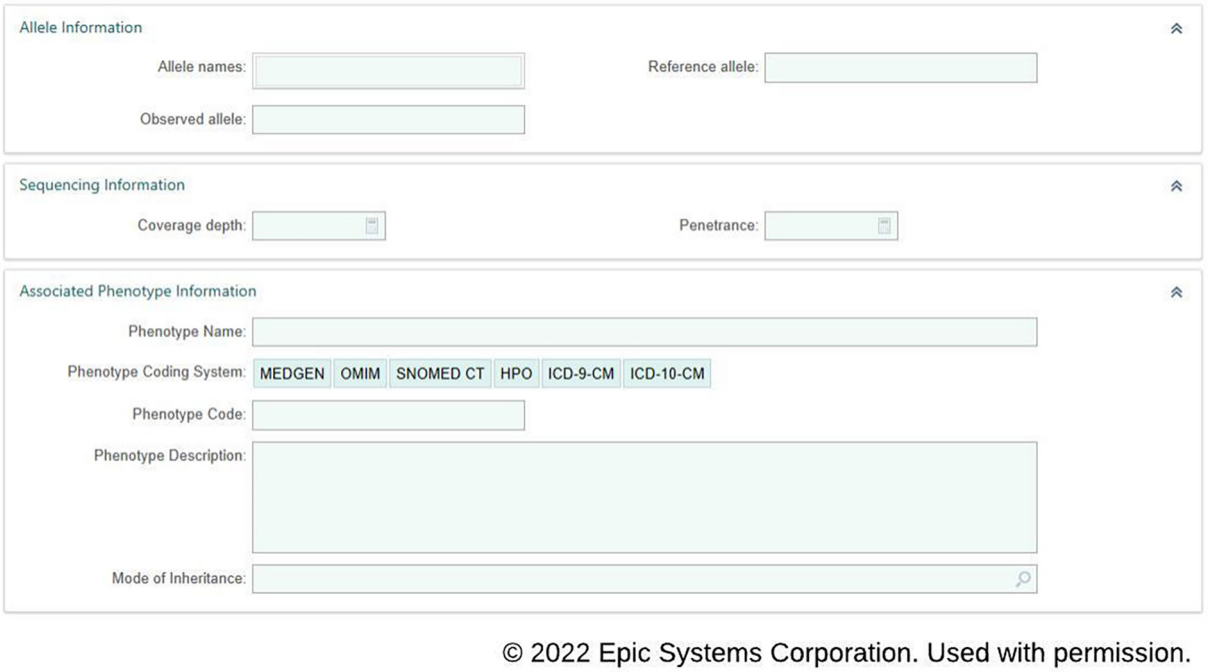

Fig. 9. Expanded view of the 'Allele Information', 'Sequence Information', and 'Associated Phenotype Information' sections of the genomic data in the Epic Beaker LIS, which is populated with imported HL7V2 LRI data transmitted from the genomics middleware.

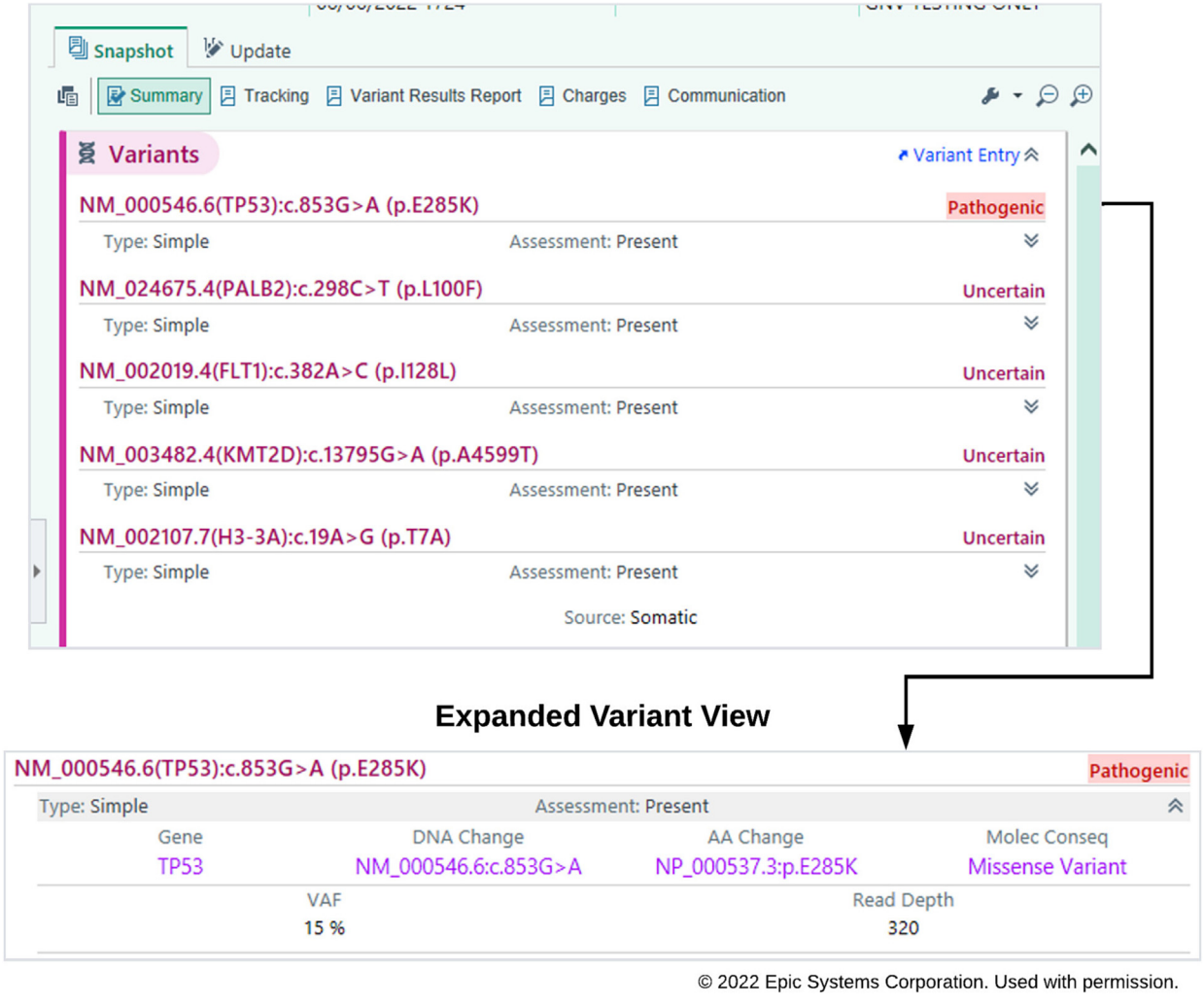

Fig. 10. Summary and detailed views of genomic variants within the Epic Beaker LIS (sample data from Epic test environment).

The approach taken here is likely to be generalizable to a wide range of EHRs and EHR integration solutions, given that the work is founded on a core semantic mapping between VCF on the one hand, and current and emerging EHR interoperability standards (particularly HL7 FHIR and HL7 v2) on the other hand. As we have shown, open source vcf2fhir[14] and the HL7 v2 extractor code[17] can be extended where necessary in order to accommodate different inputs and/or to customize the output based on an EHR's specific import specification.

The current use cases shown for this software include simple variants, as our NGS cancer panel is currently reporting simple variants. However, the capabilities can be readily expanded to other variant types supported by HL7 V2 genomic LRI. Core logic in vcf2fhir has

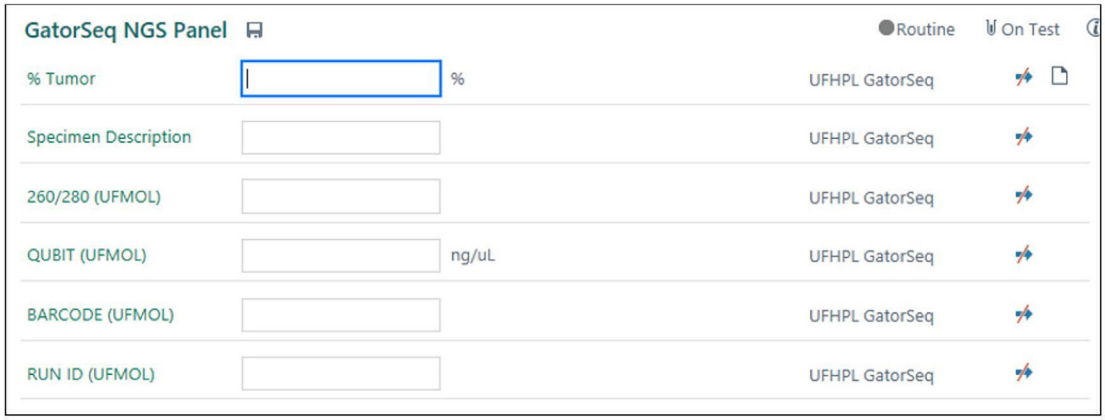

Fig. 11. Example of Epic Beaker LIS view for storing data types beyond variant information.

Fig. 12. Snapshot view of genomic variant data in the Epic EHR Genomic Module, received from the Epic Beaker LIS (sample data from Epic test environment).

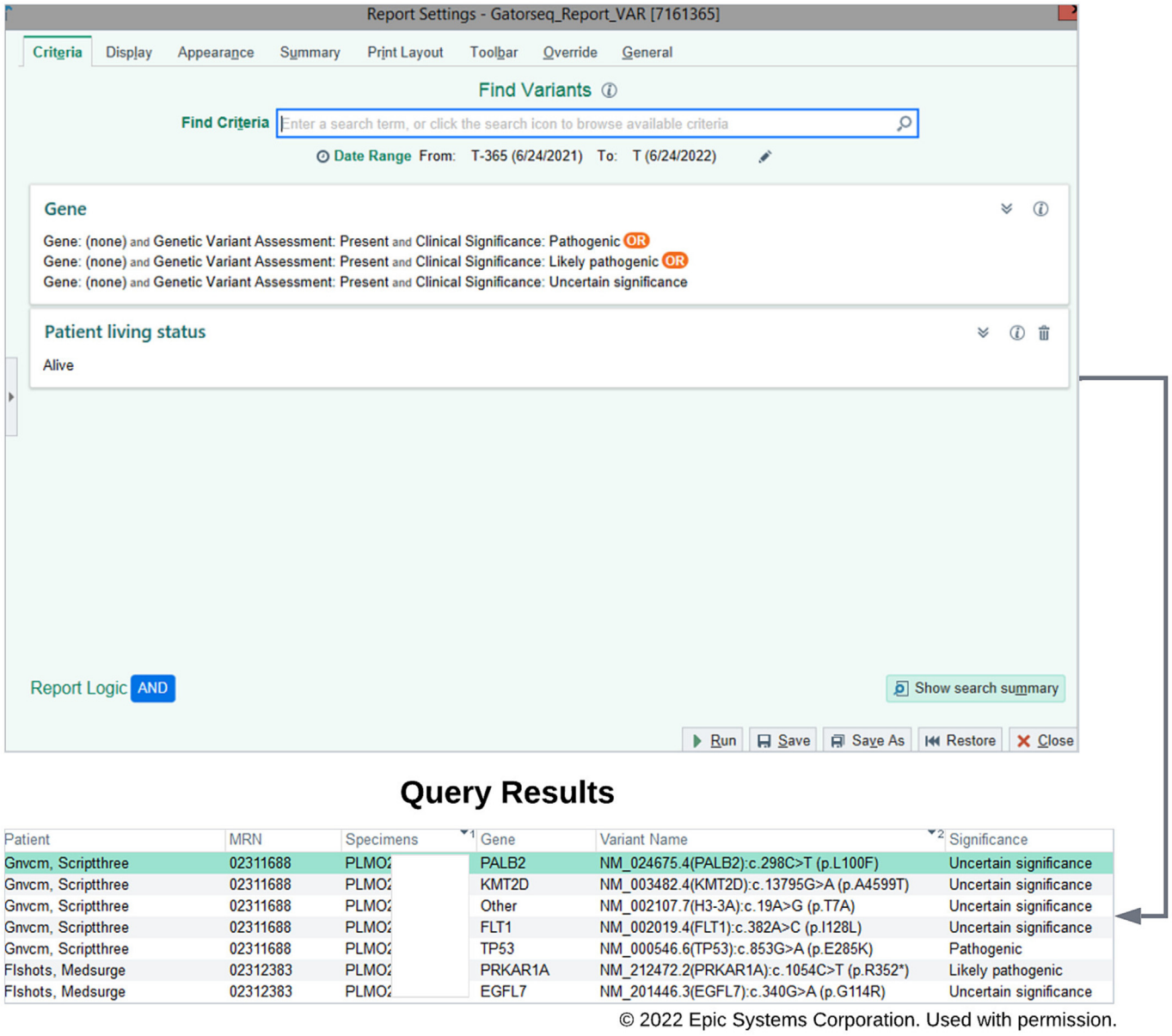

Fig. 13. Setting search criteria and generating results for genomic variant reports of patient cohorts using the reporting workbench in the Epic EHR system (sample query results from Epic EHR test environment).

been extended to address structural variant translation, and our implementation of FHIR Genomics Operations[34] also demonstrates FHIR-based translations of pharmacogenomic star alleles.

Conclusion

Our pilot confirmed the feasibility of an automated standards-based Genomics–EHR integration pipeline. As we have described, the technical implementation of a VCF (or other custom input formats such as Qiagen XML) to HL7v2 message transformation to feed genomic variant and clinical annotation data into an EHR has been successful. In addition to genetic variant data, the implementation described here releases the valuable asset of clinically relevant genomic annotations provided by labs from static PDFs to calculable, structured data in EHR systems.

While the software translates patient variants and statically associated annotations (e.g., diagnostic and therapeutic implications that are provided as part of a lab report), it also demonstrates the feasibility of automatically computing annotations at runtime. This can be particularly useful in order to present clinicians with up to date clinical implications. Such capabilities for dynamic computation of variant annotations have been further developed in HL7 FHIR, as part of the evolving HL7 FHIR Genomics Operations,[33,34] which represent a standard set of APIs for accessing genomic data stores. In particular, whereas FHIR 'genotype' operations return FHIR-encoded variants, FHIR 'phenotype' operations return statically or dynamically computed FHIR-encoded diagnostic and therapeutic implications of those variants.

Funding

The authors received no financial support for the research, authorship, and/or publication of this article.

Declaration of Conflicting Interests

The authors declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Acknowledgments

We gratefully acknowledge Epic System's permission in allowing us to reproduce screen captures, and Qiagen's permission in allowing us to include sample VCF and XML files in the github repository. The authors thank Ashley Chandler, Erica Hughes, Patricia Rizzo, Ravin Mehta, Riyaz Shaik, and Scott Nelson and other members from the UF Health enterprise IT services for their contributions to some of the informatics components detailed in this paper.

References

1. Walton NA, Johnson DK, Person TN, Chamala S. Genomic data in the electronic health record. Adv Mol Pathol 2019;2(1):21–33.
2. Li MM, Datto M, Duncavage EJ, et al. Standards and guidelines for the interpretation and reporting of sequence variants in cancer. J Mol Diagn JMD 2017;19(1):4-23. https://doi.org/10.1016/j.jmoldx.2016.10.002.
3. Richards S, Aziz N, Bale 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 Off J Am Coll Med Genet 2015;17(5):405–424. https://doi.org/10.1038/gim.2015.30.
4. VCFv4.3.pdf. Accessed June 15, 2022: https://samtools.github.io/hts-specs/VCFv4.3.pdf.
5. Dolin RH, Gothi SR, Boxwala A, et al. vcf2fhir: a utility to convert VCF files into HL7 FHIR format for genomics-EHR integration. BMC Bioinformatics 2021;22(1):104. https://doi.org/10.1186/s12859-021-04039-1.
6. 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–e123. https://doi.org/10.1055/s-0038-1676466.
7. 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. https://doi.org/10.1200/PO.19.00232. PO.19.00232.
8. Abul-Husn NS, Kenny EE. Personalized medicine and the power of electronic health records. Cell 2019;177(1):58–69. https://doi.org/10.1016/j.cell.2019.02.039.
9. HL7 Standards Product Brief - HL7 Version 2.5.1 Implementation Guide: Laboratory Results Interface, Release 1 STU Release 3 - US Realm | HL7 International. Accessed June 15, 2022: http://www.hl7.org/implement/standards/product_brief.cfm?product_id=279.
10. HL7.FHIR.UV.GENOMICS-REPORTING\Home Page - FHIR v4.0.1. Accessed June 15, 2022: http://build.fhir.org/ig/HL7/genomics-reporting/index.html.
11. Lau-Min KS, Asher SB, Chen J, et al. Real-world integration of genomic data into the electronic health record: the PennChart Genomics Initiative. Genet Med 2021;23(4):603–605. https://doi.org/10.1038/s41436-020-01056-y.
12. open.epic:: Ancillary Systems. Accessed June 15, 2022: https://open.epic.com/Ancillary/Results.
13. Walton N, Johnson D, Heale B, Person T, Williams M. Creating a home for genomic data in the electronic health record. AMIA Annu Symp Proc 2022;2021:1196–1197.
14. elimuinformatics/vcf2fhir. Published online May 27, 2022. Accessed June 15, 2022: https://github.com/elimuinformatics/vcf2fhir.
15. HL7V2-Genomics-Conversion-Utilities/hl7v2GenomicsExtractor. Published online March 8, 2023. Accessed March 15, 2023: https://github.com/HL7V2-Genomics-Conversion-Utilities/hl7v2GenomicsExtractor.
16. HL7V2-Genomics-Conversion-Utilities/hl7v2GenomicsLabMiddleware. Published online March 8, 2023. Accessed March 15, 2023: https://github.com/HL7V2-Genomics-Conversion-Utilities/hl7v2GenomicsLabMiddleware.
17. Cingolani P. Variant annotation and functional prediction: SnpEff. Methods Mol Biol Clifton NJ 2022;2493:289–314. https://doi.org/10.1007/978-1-0716-2293-3_19.
18. Chamala S, Majety S, Mishra SN, et al. Indispensability of clinical bioinformatics for effective implementation of genomic medicine in pathology laboratories. ACI Open 2020;04(2):e167–e172. https://doi.org/10.1055/s-0040-1721480.
19. Chamala S, Maness HTD, Brown L, Adams CB, Lamba JK, Cogle CR. Building a precision oncology workforce by multidisciplinary and case-based learning. BMC Med Educ 2021;21:75. https://doi.org/10.1186/s12909-021-02500-6.
20. Balasubramani B, Newsom KJ, Martinez KA, Starostik P, Clare-Salzler M, Chamala S. Pathology informatics and robotics strategies for improving efficiency of COVID-19 pooled testing. Acad Pathol 2021;8. https://doi.org/10.1177/23742895211020485.23742895211020484.
21. Chamala S, Flax S, Starostik P, et al. Optimizing COVID-19 testing capabilities and clinical management using pathology informatics. JAMIA Open 2020;3(4):523–529. https://doi.org/10.1093/jamiaopen/ooaa055.
22. SnpEff and SnpSift. Accessed June 15, 2022: http://pcingola.github.io/SnpEff/.
23. Wang K, Li M, Hakonarson H. ANNOVAR: functional annotation of genetic variants from high-throughput sequencing data. Nucleic Acids Res 2010;38(16), e164. https://doi.org/10.1093/nar/gkq603.
24. McLaren W, Gil L, Hunt SE, et al. The ensembl variant effect predictor. Genome Biol 2016;17(1):122. https://doi.org/10.1186/s13059-016-0974-4.
25. Cingolani P, Cunningham F, McLaren W, Wang K. Variant annotations in VCF format. 9.
26. Nirvana. Published online June 22, 2022. Accessed July 5, 2022: https://github.com/Illumina/Nirvana.
27. VA-spec. Published online May 9, 2022. Accessed July 5, 2022: https://github.com/ga4gh/va-spec.
28. Yen JL, Garcia S, Montana A, et al. A variant by any name: quantifying annotation discordance across tools and clinical databases. Genome Med 2017;9(1):7. https://doi.org/10.1186/s13073-016-0396-7.
29. Fairley S, Lowy-Gallego E, Perry E, Flicek P. The International Genome Sample Resource (IGSR) collection of open human genomic variation resources. Nucleic Acids Res 2020;48(D1):D941–D947. https://doi.org/10.1093/nar/gkz836.
30. Karczewski KJ, Francioli LC, Tiao G, et al. The mutational constraint spectrum quantified from variation in 141,456 humans. Nature 2020;581(7809):434–443. https://doi.org/10.1038/s41586-020-2308-7.
31. EMILY. Cool stuff the VEP can do: normalisation. Ensembl Blog. Published June 22, 2018. Accessed July 5, 2022: https://www.ensembl.info/2018/06/22/cool-stuff-the-vep-can-do-normalisation/.
32. Holmes JB, Moyer E, Phan L, Maglott D, Kattman B. SPDI: data model for variants and applications at NCBI. Bioinformatics 2020;36(6):1902–1907. https://doi.org/10.1093/bioinformatics/btz856.
33. HL7.FHIR.UV.GENOMICS-REPORTING\Operations - FHIR v4.0.1. Accessed September 1, 2022: http://build.fhir.org/ig/HL7/genomics-reporting/operations.html.
34. Dolin RH, Heale BS, Alterovitz G, et al. Introducing HL7 FHIR Genomics Operations: a developer-friendly approach to genomics-EHR integration. J Am Med Inform Assoc 2023;30(3):485–493.