Pilot Implementation of Clinical Genomic Data into the Native Electronic Health Record: Challenges of Scalability

Nephi A. Walton[1,2]   Darren K. Johnson[1]   Thomas N. Person[1]   Jonathon C. Reynolds[1]   Marc S. Williams[1]

[1] Genomic Medicine Institute, Geisinger, Danville, Pennsylvania, United States

[2] Intermountain Precision Genomics, Intermountain Healthcare, St. George, Utah, United States

Address for correspondence Nephi A. Walton, MD, MS, FACMG, Intermountain Precision Genomics, Intermountain Healthcare, 600 S Medical Center Dr, St. George, UT 84790, United States (e-mail: nephi.walton@imail.org).

Abstract

Background Recently, electronic health record (EHR) vendors have allowed for the storage of discrete genomic data within the EHR. With this new capability there remain many challenges related to adoption and implementation. Geisinger has genomic data available for approximately 145,000 patients. With the goal of integrating genetic data into the EHR for the entire sequenced population, we performed a pilot study on a subpopulation to assess the plausibility of large-scale deployment.

Objectives

1. Import genetic results into the EHR for seven pharmacogenes, Center for Disease Control tier one conditions, and hereditary hemochromatosis.
2. Build decision support and information resources for each condition.
3. Identify barriers to scaling the deployment.
4. Identify short-term and long-term solutions to overcome identified barriers.

Methods Discrete genomic variants were imported into the EHR. Maintenance guidelines and information resources were developed for each genetic condition and best practice alerts were built for pharmacogenomic variants. Weekly calls were held to address barriers to implementation. A list of challenges was maintained throughout the process. Solutions were proposed and categorized into short-term and long-term solutions.

Results Challenges were identified and solutions proposed for five discrete areas:

1. Transmitting data from the laboratory to the EHR.
2. Maintenance of information and decision support.
3. Disparate implementation needs of geneticists versus primary care.
4. Differing perceptions/uses of discrete genomic data versus a PDF report.
5. Classification and reclassification of variants.

Conclusion Scaling the use of genomic data in the EHR requires the engagement of hospitals, open standards communities, laboratories, EHR vendors, and genomic information resources.

Keywords
► electronic health records and systems
► patient records
► precision medicine
► process management tools
► HL7
► genetics
► translational research
► methodologies
► evaluation
► laboratory information systems
► clinical decision support

Background and Significance

The integration of genetic information into health care is considered a major priority with several national and international initiatives that have been launched to pursue this goal.[1] Despite this, only recently have electronic health record (EHR) vendors allowed for the storage and access of discrete genomic data within the EHR ecosystem. Even with this new capability, there remain many challenges related to adoption and implementation of such technologies. As part of the Electronic Medical Records and Genomics (eMERGE) network[2] and through our MyCode® Community Health Initiative,[3] Geisinger has exome sequencing data available for more than 145,000 patients. With the goal of integrating genetic data into the EHR for our entire sequenced population, we first performed a pilot study on 1,317 patients from the eMERGE patient cohort to assess the plausibility of scaling the deployment to more genes and a larger sequenced population.

Objectives

1. Import genetic results into the EHR for Center for Disease Control (CDC) tier one conditions, hereditary hemochromatosis, and 7 pharmacogenes for 1,317 patients.
2. Build clinical decision support (CDS) and both patient and provider facing information for the imported genetic results.
3. Identify barriers to scaling the deployment to more genes and the larger sequenced population of 145,000 patients.
4. Identify short-term and long-term solutions to overcome identified barriers.

Methods

We presented a plan to Information Technology Governance (ITG) to import discrete pathogenic and likely pathogenic genomic variants for the CDC tier one conditions (hereditary breast and ovarian cancer syndrome (*BRCA1*, *BRCA2*), Lynch syndrome (*MLH1*, *MSH2*, *MSH6*, *PMS2*, *EPCAM*), and familial hypercholesterolemia (*APOB*, *LDLR*, *PCSK9*), and variants from seven pharmacogenes (*CYP2C19*, *SLCO1B1*, *DYPD*, *TPMT*, *IFNL3*, *CYP2C9*, and *VKORC1*). Variants for hereditary hemochromatosis (*HFE*) were included to assess for challenges related to the implementation of autosomal recessive conditions. We constructed context-sensitive maintenance guidelines in the EHR for each of the disease conditions and built best practice alerts for medications impacted by the pharmacogenomic variants. Patient and provider facing information were developed for all genetic conditions to be accessible through both the physician chart and the patient portal. Through the implementation process, weekly calls were held with pertinent stakeholders to address challenges and barriers to implementation as they arose. A list of these challenges was maintained throughout the process. After implementation, we analyzed the challenges and potential solutions to scaling our implementation to more genes and our entire sequenced population.

Solutions were categorized into short-term and long-term solutions.

Results

We were able to successfully implement all conditions for 1,317 patients with minimal compromises. Support from organizational leadership was critical to securing technical resources to achieve implementation and prioritization in the queue of many EHR implementation projects. The proposal was passed through ITG after a thorough review of the need and potential impact on patient care and was based on the contingency that there would be dedicated resources from the Genomics Institute to maintain and support CDS and information resources pertaining to genomics.

Due to the inability of our laboratory information system (LIS) to store or transmit genomic data in a structured format, our research team wrote custom Python scripts to convert structured laboratory data into Health Level Seven (HL7) OBX segments for import into the EHR. The resulting messages were reviewed and passed through several methods of validation by the EHR technical team.

As there was no single external resource that provided adequate patient or provider facing information, this content was developed by internal domain experts and formatted in an Excel spreadsheet for use by the EHR technical team to deploy in the EHR. Provider facing information contained more detailed management information, whereas patient facing information was more succinct and provided contact information for providers and links to outside patient-centered resources. Health maintenance guidelines were developed by domain experts and deployed in the EHR by the technical team to alert providers to needed studies at appropriate intervals (i.e., breast magnetic resonance imaging, colonoscopy, etc.). A pharmacist worked directly with the EHR technical team to develop active CDS for the seven pharmacogenes. Where applicable, the pharmacogenomic CDS would present an option to order an alternate medication.

Our initial plan was to develop slight variations in interfaces and workflows for primary care providers (PCPs) compared with clinical geneticists (CG). After reviewing workflows and needs of both groups, we decided to abandon the CG interface and workflow for this deployment as the needs were very disparate and pursuit of both workflows would hamper our ability to complete the implementation in the given time frame and within the allotted budget.

The EHR vendor offered an automated process to classify variants through ClinVar that was determined to be insufficient. All variants were assessed and independently validated by both an external Clinical Laboratory Improvement Amendments laboratory and an internal variant scientist prior to import into the EHR.

Several challenges were encountered that are anticipated to affect our ability to scale the process. We encountered some vendor-specific informatics challenges, including the inability to store discrete variants for pharmacogenomics (the native functionality stores the star allele diplotypes) and having inadequate infrastructure for autosomal recessive

conditions where the patient is a compound heterozygote. Although these were vendor-specific issues, the vendor's data implementation was based on the HL7 V2.5 Clinical Genomics Report standard, which had deficiencies that contributed to these problems. A list of global challenges and solutions is outlined in ►Table 1.

Discussion

We were able to capture discrete genomic data from our primary genetic testing laboratory; however, neither the current FHIR molecular sequence resource nor the HL7 V2.5 Genomic report standard was adequate for reporting as neither encapsulated all the information in the genetic test report. Because of this our primary genetic testing laboratory used its own proprietary JSON structured format. The passage of data into the EHR became a significant bottleneck in our process with the EHR interface team responding that this process could delay the implementation by several months. Having a research team that was well versed in both informatics and genetics, we were able to rapidly build a translator for an existing interface that converted structured laboratory data into HL7 messages. This effort required a significant amount of coordination and trust between our research team and the EHR technical team and underwent thorough validation prior to implementation. Even with the ability to pass discrete genetic data, our efforts in this area were hampered by lack of standardized genetic phenotypes. We developed the ability to pass the genetic phenotypes included in the pilot study for a single laboratory. To scale this to include any result returned on exome, we would have to specifically define and build out any additional phenotypes that might be discovered on exome sequencing. We would then have to map these phenotypes to each sequencing laboratory used. We opted in the short term to only allow the passage of secondary findings from exome where we had genetic phenotypes built out and limit our process to two laboratories. The initial role of our LIS was unclear, since laboratory data currently enters the EHR from the LIS, and our LIS did not have the ability to store discrete genomic variants. We opted to keep the existing pathway for ordering and reporting genomic testing through the LIS intact and build an additional messaging pathway that bypassed the LIS and delivered data directly to the EHR.

Long-term maintenance of gene/variant-specific information and CDS is a major challenge in a rapidly changing field. Changes to patient/provider facing information and CDS require support from vendor-specific technical personnel. Engaging this technical team can be challenging, given their responsibility to maintain all aspects of the EHR for the entire health care system. These personnel have little to no experience in genomics, requiring communication between the clinical and technical teams to make any changes to the system. EHR vendors adhering to more open standards for genomics and CDS, including SMART on FHIR and CDS Hooks, would allow for the development of interfaces to external applications where maintenance can be performed by staff

with genetics training and less technical skill. This is particularly important while we are still largely in the discovery phase of genomics, where many of the applications of genomic data are still being developed and may be governed and deployed in substantially different ways among institutions. Without nontechnical interfaces, it was critical that we had strong support from leadership to be able to maintain priority in the technical team's queue of projects. Another important component to this is the need for maintained resources such as ClinGen and PharmGKB that are Infobutton accessible and can allow for rapid access to the most current information on a specific gene or variant. To our knowledge, no vendor has implemented Infobutton support for genomics, and outside of pharmacogenomics, no such Infobutton compliant resources exist.

Early in the process, the clinical genomics team was brought on board to review workflows. It was clear in the first stages of deployment that the needs of CGs compared with PCPs were markedly different. Due to the complexity of implementation of CG workflows and the relative impact on care versus population genomics, we opted to focus on population genomics workflows and delay implementation for CGs. It remained critical to continue to engage the clinical genetics team in the population genomics deployment.

Although discrete genetic data fields and PDF reports contain identical information, genetic data as it exists in a PDF seemed to have a perceived protective effect where the data remained more in the domain of specialists, whereas adding a variant discretely to the patient summary subjected it to more visibility and potential action by providers with less experience. This contributed to a fear of providing uncertain information or providing information without clear guidelines and decision support within the EHR to support the actions that should be taken on the variant. In addition, the structured information can trigger CDS, which should not be triggered in cases where the classification is not certain.

We opted to bypass automated processes to classify variants and manually reviewed each variant laboratory classification prior to allowing import to the EHR. This process does not scale well; however, we felt that challenges in variant classification warranted extra caution. Reclassification of variants based on clinician reinterpretation presented several challenges. Vendor-specific challenges included the inability to reclassify at the variant level, requiring the updating of each patient with the same variant separately. Although variant reclassification was allowed and tracked appropriately in the EHR, this presented the possibility of having structured genomic data that was discordant with the scanned PDF. This demonstrates the need for bidirectional communication between the clinician and the laboratory when variants are reclassified, particularly if both parties submit their classification to ClinVar. While the need was expressed to automatically reclassify all patients with a given variant once reclassification was warranted, there was no clear process to manage this at scale.

Table 1 Challenges and long- and short-term solutions

Challenge category	Challenge	Short-term solution	Long-term solution
Transmitting data from the laboratory to the EHR			
	Need for standardization of genetic phenotypes (i.e., disease, metabolizer status, gene-specific phenotype)	Create manual genetic phenotype mappings to each genetic testing laboratory for all genetic conditions imported	Engage the data standards and genomics communities to develop standards for genetic phenotypes
	Poorly defined role for the LIS in delivery of discrete genomic data	Bypass the LIS, sending discrete genomic information directly to the EHR	Engage LIS vendor to discuss future plans for genomic data
	Inadequate and discordant standards for genomics in the HL7 genomic report format and the Fast Healthcare Interoperability Resource (FHIR) molecular sequence resource	Develop a laboratory interface that maps JSON structured data to the HL7 genomic report format	Engage HL7 and genomics community to address deficiencies
Maintenance of clinical information and CDS in a rapidly changing field			
	Lack of publicly available resources for patient and provider facing information for genetic conditions	Develop and maintain internal resources for a limited number of conditions	Engage appropriate stakeholder groups to encourage development of standard resources for genomics
	Maintenance of CDS and clinical information requires technically competent individuals with training specific to the EHR	Create an external database that can be curated by content experts (geneticists, genetic counselors) with limited technical skills, and can then be used by the technical team to update the EHR	Work with EHR vendor to adopt an open API that will allow for third party specialized tools for maintenance of clinical information and CDS through nontechnical interfaces
	Lack of open APIs to access external information resources and third-party tools in the EHR vendor deployment	Continue ongoing discussions with vendor regarding implementation of OpenInfobuttons	Encourage EHR vendor to implement a more mature adoption of CDS hooks, FHIR, and OpenInfobuttons
Disparate implementation needs for clinical geneticists (CGs) compared with primary care providers (PCPs)			
	PCPs are interested in genomics for population health and variant-guided health maintenance while CGs are interested in variants for diagnostic indications	Prioritize implementation for primary care/population health, delaying CG implementation for later stages in development, given the heightened complexity	Work with EHR vendor and CGs to develop a workflow specific to CG's needs
	CGs need access to variants of uncertain significance (VUS), while PCPs want variants with pathogenic interpretations only, as interpretation is outside of their training		
	PCPs favor a limited set of disorders with well-defined clinical recommendations CGs need to support rare disease where more limited knowledge and resources are available		

(*Continued*)

Table 1 (*Continued*)

Challenge category	Challenge	Short-term solution	Long-term solution
Differing perceptions/uses of discrete genomic data versus PDF of the genetic report			
	Inability to hide variants from public/patient view prior to clinical validation	Clinical expert to manually review all variant classifications prior to import into the EHR	Work with EHR/LIS vendor to develop variant queue that allows for clinical expert review before release to chart, Alternately develop external application for variant review
	CDS will fire from any variant classified by the laboratory as pathogenic		
	Risk of misinterpretation or inappropriate action from attributing a VUS to a patient	Do not import any variants that are not considered pathogenic	Study context sensitive delivery methods of variants based on variant pathogenicity and their impact on care
Classification and reclassification of variants			
	No process in place from an operational perspective to handle the global reclassification of variants	Develop thorough manual process for variant reclassification	Work with EHR vendor to develop automated workflows for reclassification
	Potential for discrete variant classification in the EHR to be discordant with the scanned PDF	Communicate with the laboratory after reclassification to consider issuing an updated report with changes	Work with EHR vendor to create a mark or note overlaying the scanned report that the variant has been reclassified by a provider in the system and allow for access to reclassification history
	The inability of EHR vendor solution to change the classification of all patients with a given variant, allowing for patients with the same variant to have discordant classifications	Manually track and change variant interpretations as needed	Engage stakeholders in clinical genomics and laboratory genomics communities to develop standard methodologies for reclassification of variants by clinical providers

Abbreviations: CDS, clinical decision support; EHR, electronic health record; HL7, Health Level Seven; LIS, laboratory information system.

Conclusion

Geisinger has developed extensive processes for the management and return of clinically actionable genetic results that may be able to be optimized through EHR integration. Efficiently utilizing Geisinger's existing resources would require interfaces that allow nontechnical clinical experts to manage the information and CDS around genetic variants. Moving genetic data from the laboratory to the EHR and managing it in a scalable way requires adequate standards for storing and transmitting genomic data and standard APIs that allow ancillary systems, including those in laboratories, to interface with the EHR. Classification and reclassification of variants remain a challenge. Scaling the use of genomic data in the EHR is a complicated endeavor that requires the engagement of hospitals, open standards communities, laboratories, EHR vendors, and curators of genomic information resources.

Clinical Relevance Statement

Integration of genomic data into the EHR allows for improved health through targeted clinical management using genetic sequences. Achievement of this has been a recent target of the scientific community.

Protection of Human and Animal Subjects

All research activities reported in this publication were reviewed and approved by the Geisinger Institutional Review Board.

Conflict of Interest

None declared.

References

1 Stark Z, Dolman L, Manolio TA, , et al. Integrating genomics into healthcare: a global responsibility. Am J Hum Genet 2019;104 (01):13–20
2 McCarty CA, Chisholm RL, Chute CG, , et al; eMERGE Team. The eMERGE Network: a consortium of biorepositories linked to electronic medical records data for conducting genomic studies. BMC Med Genomics 2011;4(01):13
3 Carey DJ, Fetterolf SN, Davis FD, et al. The Geisinger MyCode community health initiative: an electronic health record -linked biobank for precision medicine research. Genetics in medicine 2016;18(09):906