CSER and eMERGE: current and potential state of the display of genetic information in the electronic health record

Brian H Shirts[1], Joseph S Salama,[2] Samuel J Aronson[3], Wendy K Chung[4],
Stacy W Gray[5,6], Lucia A Hindorff[7], Gail P Jarvik[2,8], Sharon E Plon[9], Elena M Stoffel[10],
Peter Z Tarczy-Hornoch[11], Eliezer M Van Allen[6,12], Karen E Weck[13,14], Christopher G Chute[15],
Robert R Freimuth[15], Robert W Grundmeier[16], Andrea L Hartzler[17], Rongling Li[7], Peggy L Peissig[18],
Josh F Peterson[19], Luke V Rasmussen[20], Justin B Starren[20], Marc S Williams[21], Casey L Overby[21,22]

ABSTRACT

Objective Clinicians' ability to use and interpret genetic information depends upon how those data are displayed in electronic health records (EHRs). There is a critical need to develop systems to effectively display genetic information in EHRs and augment clinical decision support (CDS).

Materials and Methods The National Institutes of Health (NIH)-sponsored Clinical Sequencing Exploratory Research and Electronic Medical Records & Genomics EHR Working Groups conducted a multiphase, iterative process involving working group discussions and 2 surveys in order to determine how genetic and genomic information are currently displayed in EHRs, envision optimal uses for different types of genetic or genomic information, and prioritize areas for EHR improvement.

Results There is substantial heterogeneity in how genetic information enters and is documented in EHR systems. Most institutions indicated that genetic information was displayed in multiple locations in their EHRs. Among surveyed institutions, genetic information enters the EHR through multiple laboratory sources and through clinician notes. For laboratory-based data, the source laboratory was the main determinant of the location of genetic information in the EHR. The highest priority recommendation was to address the need to implement CDS mechanisms and content for decision support for medically actionable genetic information.

Conclusion Heterogeneity of genetic information flow and importance of source laboratory, rather than clinical content, as a determinant of information representation are major barriers to using genetic information optimally in patient care. Greater effort to develop interoperable systems to receive and consistently display genetic and/or genomic information and alert clinicians to genomic-dependent improvements to clinical care is recommended.

Keywords: genetics, electronic health records, translational research, clinical decision support, survey

<div style="writing-mode: vertical">RESEARCH AND APPLICATIONS</div>

INTRODUCTION/BACKGROUND

With increasing numbers of clinically useful genetic discoveries and decreased sequencing costs, genetic and genomic information (for convenience referred to collectively as "genetic information" referring to all forms of genetic testing unless specifically referring exome and genome-scale testing) is increasingly used in many aspects of medical care. Concurrent with the increased application of genetics in medicine, electronic health record (EHR) systems have been implemented by most physician practices and are nearly universal in large health-care systems.[1] The way genetic information is displayed influences clinicians' ability to use that information appropriately.[2] The advent of next generation sequencing has dramatically increased our ability to generate genetic information. Storage and display of genomic-scale data present separate but related issues.[3–9] Although it is not feasible to store all genetic information in the EHR,[10] the EHR should display genetic information to clinicians in a rational, organized manner that is ideally linked with both appropriate clinical decision support (CDS)[11–14] and systems managing underlying patient genetic information.

The National Institutes of Health (NIH)-sponsored Clinical Sequencing Exploratory Research (CSER)[15] and Electronic Medical Records & Genomics (eMERGE)[16] consortia independently established EHR integration working groups to explore the storage and display of genomic information in the EHR. The CSER consortium includes nine NIH-funded programs to explore the use of genomic-scale data in clinical medicine, including generating clinical genomic sequencing data, interpreting and presenting genomic results to clinicians and patients. Among the goals of the CSER consortium EHR working group is facilitating cross-site collaboration to integrate genetic information and CDS into the EHR.[17] The eMERGE Network consists of 15 institutions and links 10 biorepositories and EHRs for the purpose of conducting genetic discovery studies. In recent years, eMERGE has focused on implementing pharmacogenetics and other genetic medicine initiatives using the EHR.[16] The eMERGE EHR integration working group focuses on the storage and display of genomic information. For the work described here the 2 consortia working groups joined to explore the current state and future potential for integration of genetic information in the EHR across these sites.

There are many categories of genetic information in the EHR that are likely to be used by different clinicians in different ways.[18] Because of this diversity, it appears that results could be represented in many parts of the EHR; however, to our knowledge, this has never been comprehensively evaluated. Informed by existing work[9,19–25] and our experiences,[17,26–29] we aimed to 1) determine how genetic

information is currently displayed in the EHR among a sample of institutions that return genetic information, 2) envision optimal uses for different types of genetic information, and 3) prioritize areas for EHR improvement.

METHODS

The collaboration between CSER and eMERGE was initiated in discussions in October 2013, and formalized in February 2014. All members of the CSER and eMERGE EHR working groups, or site representatives from each CSER and eMERGE site, were invited to participate in a multiphase, iterative process to determine the current handling of genetic information in the EHR and to outline future priorities for EHR improvements. In phase one, we sought to define types of genetic information that could be displayed by the EHR as they relate to clinical practice and assess the location and format genetic information in the EHR, independent of whether the genetic information were generated as part of the funded CSER or eMERGE project. In phase two, we sought to identify recommendations for improving the display of genetic information in the EHR, and prioritize recommendations for using the EHR more effectively to display genetic information. We administered 2 separate surveys to accomplish these goals, one during each phase.

In phase one, the members of the CSER and eMERGE working groups initiated a series of semi-structured phone discussions to define functionally distinct types of genetic information that could be displayed in the EHR. For each type of genetic information, we proposed definitions and provided a description for how the genetic information could be used in clinical care in further discussions. We then iteratively designed and developed a survey tool to determine how each of these types of genetic information was presented in the EHR and to identify the determinants of how genetic information is displayed. The phase one survey contained 31 items and covered domains such as the type of EHR and laboratory information management systems used, sources of genetic information, location of genetic information in the EHR, process for determining and changing information location, ancillary genetic databases, challenges of reporting genetic information, automatic alerts about genetic information, ability to mark genetic information as sensitive and add genetic information to problem lists, and methods for transmitting genetic information to family members (See supplemental material.) It was pilot tested with 2 participants, refined for clarity, and finalized by the CSER and eMERGE working groups. We then surveyed CSER and eMERGE investigators to evaluate the current state of how genetic information is displayed in multiple EHRs at 17 academic institutions and health systems across the country. Surveys were delivered to CSER and eMERGE investigators by email and responses recorded on excel spreadsheets by respondents. Respondents included individuals in CSER and eMERGE EHR working groups involved in developing the survey and other CSER and eMERGE investigators who worked with others at their institutions to gather institutional information. In order to ensure adequate site representation, we asked that at least one investigator from each site respond to the surveys. For this phase one survey when there were multiple EHR working group members from the same institution, we asked them to collaborate to determine institutional practices.

In phase two, the CSER and eMERGE working groups reviewed the data from the first survey and discussed revised definitions and descriptions according to consensus in a conference-call meeting. Then, using the information from phase one, we held 2 CSER and eMERGE working group teleconferences about the potential ways to improve the display of genetic information from the EHR. Working group members provided 29 suggestions; these were cluster sorted into 20 distinct recommendations by BHS. Cluster sorted recommendations were

then reviewed, modified, and approved by combined CSER and eMERGE working groups. We then conducted the phase two survey among individuals from CSER and eMERGE institutions with multiple respondents from the same institution allowed. We asked respondents to prioritize recommendations, to rank their top four recommendations, and rate the feasibility of implementing the recommendations (see supplemental material). As before, this survey was part of an informal consensus building process, and the survey instrument was not formally validated. Multiple working group members then analyzed the phase two survey results to identify key themes in recommendations for future EHR improvement. The initial ranking exercise failed to produce a clear consensus; therefore, an informal consensus building process of working group phone discussions was used to explore feasibility and prioritization of recommendations based on survey results.

RESULTS

Categories and use cases for genetic information stored in the EHR

Genetic information has many overlapping use-cases and applies to many clinical situations. Table 1 enumerates categories of genetic information that the CSER and eMERGE communities agreed constitute discrete nonexclusive applications of genetic information in clinical care. While some categories overlap and to some extent differ only in degree (e.g., risk actionable vs low risk not actionable, theoretically actionable), most categories correspond to distinct clinical or public health use cases. Categories were not designed to be exclusive, so genetic information may fit multiple categories (e.g., sensitive – diagnostic results; actionable – incidental findings, etc.) and some types of genetic information may fit into different categories depending on the clinical context.

Key characteristics of current practice: Phase one survey

Current State for Display of Genetic Information in the EHR Summary

Table 2 summarizes survey participant responses regarding how genetic information is reported in the EHR system at their institution. Overall, respondents indicated that genetic information comes from a range of sources. For 65% of respondents, the laboratory performing the genetic test was a major factor that determines where genetic information is documented in the EHR. Other factors that influence where genetic information is documented in the EHR were the hospital laboratory interfacing the results with the EHR (47%), ordering clinician's department (41%), purpose of information (35%), pathology department (29%), and source tissue (6%). When genetic information is reported in the EHR, the majority indicated that genetic information is displayed in multiple places (81%) with fewer than half of respondents (42%) indicating any effort by the institution to consolidate genetic information into one place. Forty-seven percent responded that there is an institutional process for changing where a piece of genetic information is displayed in the EHR. When asked about locations to store genetic information outside of the EHR, more than half of respondents (56%) indicated that genetic information, generally lab reports, was viewable from a clinician or patient portal, which is usually tethered to the EHR; however, only 9 institutions provided a response. Institutional genomic results databases and web accessible databases of the testing laboratory were identified as other locations to store genetic information by 19% and 31% of respondents, respectively. The majority (71%) of respondents indicated that their health IT infrastructure is evolving in a way that may change how genetic test results are reported either by implementing new EHRs or modifying existing EHRs and ancillary systems.

Most respondents indicated that all categories of genetic information defined in Table 1 are viewable in the EHR, with the exception of "uninterpreted variants," which were only displayed by 3 institutions,

Table 1: Genetic categories

Genetic category[a]	Definition	Proposed clinical use example
Disease defining/diagnostic	Genetic variation associated with high likelihood of disease manifestation (high penetrance), usually in a dominant or recessive manner. For symptomatic individuals these variants may explain the cause of their disease. Individuals may be asymptomatic at the time of testing depending on age-related disease onset. Family member testing may be warranted to identify other at risk family members.	A patient presents with hypertrophic cardiomyopathy. A genetic cardiac panel is ordered and a pathogenic cardiomyopathy variant is identified. The overall test result, laboratory report, structured variant data, and associated interpretations are placed in the EHR's genetic summary screen. An alert is sent to the ordering clinician indicating the result is available. The ordering geneticist reviews the report and determines that it is appropriate to add "Genetic Predisposition to Cardiomyopathy" to the patient's problem list. The clinician establishes a plan for monitoring the patient. In addition, the clinician generates material the patient can use to contact other family members so they can consider testing.
Risk actionable	Genetic information that is medically actionable and should trigger a change to medical care (including treatment, surveillance, or avoiding agents). Most variants in this category are associated with a significant risk of morbidity and/or mortality (e.g., cancer, sudden cardiac death) and are relevant to family members.	A 29-year-old woman with no history of cancer has exome sequencing, and a known deleterious truncating mutation is identified in *BRCA1*. After consultation with a cancer geneticist, hereditary breast ovarian cancer syndrome is placed on the problem list and appropriate surveillance, including prophylactic surgery, is discussed. A careful family history indicates that more distant female relatives on the paternal side had early breast cancer. Follow-up genetic testing identifies presence of the *BRCA1* mutation in her paternal aunt, but not in her 2 adult sisters.
Low risk not actionable, theoretically actionable	Genetic variants with modest effect sizes, low penetrance, and/or unclear functional implications. Some may be associated with intermediate endpoints (e.g., lipid levels) as opposed to discrete clinical endpoints; may also include simple nucleotide polymorphisms (SNPs) from genome-wide association studies as well as aggregated risk scores of such variants. These results maybe reported by some exome sequencing laboratories or genotyped as part of a gene panel (e.g., thrombophilia) or in support of research interests (e.g., cardiovascular risk scores).	A patient's exome is sequenced and variants in several candidate genes related to lipid metabolism and cardiometabolic disease are identified and reported in the EHR laboratory result section. These variants are aggregated into a risk score which suggest a 1.5-fold increase in risk of myocardial infarction. The physician is likely to follow American Heart Association (AHA) and other professional organization recommendations for inclusion of these results in clinical management.
Large chromosomal changes and cytogenetic test results	Copy number variants (deletions and duplications) and large structural rearrangements (inversions and translocations) that are detectable by standard karyotyping and/or chromosomal microarray. At present these are poorly characterized using exome sequencing approaches but are reported from whole genome analyses.	A male infant is born with multiple congenital anomalies and physical features suggestive of Down syndrome. A peripheral blood karyotype reveals trisomy 21 due to a Robertsonian translocation involving chromosomes 14 and 21 (46,XY rob(14;21)). The result is placed in the laboratory report section and the diagnosis of Down syndrome is placed on the Problem List. Parental testing is recommended. The infant's mother is found to carry the translocation although this result is likely to appear in a separate EHR of the mother (also the normal result of the father). Appropriate genetic counseling for recurrence risk is provided.
Pharmacogenomics	Genetic variants that affect an individual's ability to respond to drug therapy. Pharmacogenomic variants may alter therapeutic efficacy (e.g., whether or not a given drug will "work" for a patient) or safety (e.g., hypersensitivity reactions, changes to dose or drug selection).	A patient experiences an acute coronary syndrome and is prescribed clopidogrel to inhibit platelet aggregation. Clopidogrel is activated by the CYP2C19 enzyme which is encoded by the *CYP2C19* gene. Several common *CYP2C19* loss-of-function alleles are associated with higher risk of major adverse cardiovascular events in patients treated with clopidogrel. At the time the prescription is written a clinical decision support alert prompts the clinician to order *CYP2C19* testing to help predict therapeutic response. The patient is found to be a *CYP2C19* "poor metabolizer." An alternative drug therapy is recommended based on the test results.

(continued)

RESEARCH AND APPLICATIONS

Table 1: Continued		
Genetic category[a]	Definition	Proposed clinical use example
Carrier recessive	Genetic variants that are disease causing only if a patient carries 2 deleterious mutations in the same gene, but for which the patient only has one copy. These may be important for reproductive decision making but usually do not directly influence the patient's health.	An individual undergoes exome testing and is found to be heterozygous for one *CFTR* mutation. The laboratory result is noted in the EHR and flagged as carrier status results. There is no information about her partner's *CFTR* status in the patient's EHR.
Somatic/tumor genetics	Knowledge of somatic alterations in tumor specimens may have immediate therapeutic implications with regard to US Food and Drug Administration (FDA) approval of targeted agents. In addition, mutations in pathways maybe suggestive of therapeutic efficacy. However, many oncologists may not be aware of how these mutations fit along pathway-specific therapeutic modules. Ongoing research is underway to determine how somatic mutations may impact treatment outcome for cancer patients. Thus, EHR representation of these tumor-specific mutations is important for ongoing improvement in cancer treatment.	A patient has a panel of multiple genes sequenced in her colon tumor. Analysis reveals a somatic mutation in *KRAS* codon 12 and a novel *PIK3CA* mutation in a codon where other mutations are know to cause *PIK3CA* activation. *KRAS* mutations are associated with lack of response to therapies that target the Epidermal Growth Factor (EGF) receptor in colon cancer patients. Activating mutations in *PIK3CA* have been associated with responsiveness to aspirin therapy. These findings are reported in EHR in a way that facilitates clinician access to decision support on how these classes of mutations may alter cancer care.
Incidental	Clinically significant genetic findings that are unrelated to the medical condition for which the genetic sequencing test was ordered.	A cancer risk panel of 40 genes is ordered for a male patient with a personal history of colon cancer. The patient is found to have a *BRCA2* truncating mutation. The result is reported in the EHR and to the clinician as an incidental finding. There is a mechanism to alert the clinician about additional, unexpected follow up that may be necessary such as genetic counseling counseled about increased cancer risk and appropriate evaluation of family members.
Variants of uncertain significance (VUS)	Genetic variants that cannot be classified definitively as pathogenic or benign at this time. Many are missense sequence variants that alter a single amino acid or in noncoding portions of genes. Many VUS are previously undescribed novel variants. VUS are reported on a variety of genetic testing platforms. Over time, VUS may be reclassified as benign or pathogenic; however, laboratories differ in whether VUS results are amended on clinical reports.	A 43-year-old female patient with a personal and family history of breast cancer undergoes sequencing analysis of *BRCA1* and *BRCA2*. A missense VUS is reported in *BRCA1* and reported as a VUS. Therefore it is not recommended that testing for this variant be used to determine risk in relatives of this patient. Nine months later, a revised laboratory report reclassifies the variant as pathogenic based on additional evidence. The EHR is updated to now follow the recommendations found in Diagnostic and Actionable categories.
Uninterpreted variants	Genetic sequence variation in an individual that differs from the reference genome for which no analysis is performed to determine clinical significance. Large-scale genomic analysis will reveal many thousands of sequence variants, many of which are intergenic (that is not within known genes), or in genes with no known clinical relevance. Genomic testing laboratories have different policies about whether this information is included in laboratory reports or the EHR.	An individual with a progressively debilitating neuromuscular disorder of unknown diagnosis undergoes whole exome sequencing. A diagnostic finding in the *GCH1* gene (nonsense mutation) is identified and reported in the EHR as consistent with a diagnosis of dopa-responsive dystonia. The detailed whole exome sequencing (WES) report also indicates that 77 055 sequence variants (including 17 299 coding variants) were detected but the policy of the testing laboratory is not to include the detailed genomic information in the EHR. The laboratory retains this information for potential re-analysis for 7 years in accordance with state requirements.

(continued)

Table 1: Continued

Genetic category[a]	Definition	Proposed clinical use example
Newborn Screening	Newborn screening is a public health program designed to facilitate the prevention of developmental impairments, delayed physical growth, severe illness, and death through early detection and intervention of a select group of conditions	Newborn screening is performed on an infant to detect his risk for the conditions screened for in his state. Upon entering results of all of the screening tests into the newborn screening (NBS) state program laboratory information management system, the NBS laboratory professional is alerted that a confirmatory test needs to be run for one test [phenylketonuria (PKU)]. The reasoning for the recommendation, a "borderline abnormal" result, is displayed in the alert message. The NBS program sends the results to the EHR of the physician of record and the birthing hospital and initiates phone contact with the health professionals and family to communicate the need for a repeat test. The repeat test is obtained and the phenylalanine level is elevated confirming the diagnosis of PKU. The result is entered into the state newborn screening database and the result is sent to the EHR of the ordering physician for review. The state Newborn Screening program follows up by phone with the family and provider to insure that the infant has been referred to an accredited metabolic center. The report is placed in the genetic results section of the laboratory tab in the EHR. The diagnosis of PKU is placed on the problem list.
Sensitive genetic information	Genetic information that patients have decided they do not want their providers to know about, or that they do not want to know about themselves, such as Huntington's disease risk, Alzheimer's disease risk, prenatal testing, or cancer risk. Genetic results in this category may also belong in other categories	A patient has his exome sequenced, but chooses not to know his *APOE* results because of concerns about Alzheimer's risk information. These results are not reported to the requesting hospital or displayed in his medical record; when his physician searches for them in the EHR she finds a note in the advanced directives tab that the patient's preference is to not know the results for this genetic locus.

[a]Genetic categories were designed to illustrate clinical or public health use cases and are not exclusive. Genetic results may fit multiple categories, e.g. sensitive – diagnostic results; actionable – incidental findings, etc,and some categories are clinical context dependent.

presumably in conjunction with CSER or eMERGE clinical research. Table 3 illustrates where genetic information is displayed in the EHR. Overall, respondents indicated that genetic information is often displayed on a laboratory tab (>50% across all categories). Also, genetic information is often displayed within genetics clinic notes (>50% across all categories with the exception of "uninterpreted variants"). Respondents noted that clinic notes were the "catch-all" location for results that could not be entered into the EHR via automated mechanisms. One respondent stated, "A well defined place, such as a genetics tab in the labs section, [many genetic results] . . . would belong there, but this location does not exist in our EHR." Among our respondents, the problem list was a common location for displaying "disease defining/diagnostic" and "risk actionable" genetic information in the EHR (>50%). A genetics tab, pathology tab, pharmacogenomics/drug interaction tab, and outside medical records were listed less often (14–38%, 0–33%, 0–18%, and 0–38% across all categories, respectively).

Table 4 summarizes formats in which genetic information are viewed. Overall, respondents indicated that most types of genetic information were displayed as PDFs of lab reports (>50% across all categories). Paragraphs of text were also a common way to display genetic information (>50% across all categories, with the exception of "uninterpreted variants"). Multiple variants from a single test largely are not displayed (<18% across all categories of genetic information) except in PDF laboratory reports. Display of test results in a discrete field depended on the category of genetic information (ranging from 13% for "uninterpreted variants" to 65% for "pharmacogenomics"). The specific genetic information that was most likely presented in discrete fields the EHR were most often routine genetic tests offered by

affiliated laboratories of local institutions (e.g., results for factor V Leiden with descriptions such as "homozygous" or "heterozygous" or "negative" transmitted as discrete fields from within a laboratory information system).

Mechanisms to modify display and specific categories of genetic information

All respondents indicated that they have manual mechanisms to capture disease defining genetic information in the EHR via the problem list (none have automatic mechanisms to do this; see Table 5). Seventy-one percent of institutions have a general mechanism for high-risk medically actionable information to trigger an alert in the EHR even though they may have limited or no functioning alerts for genetic information. Fifty-three percent had a mechanism to redact or exclude from the EHR sensitive genetic information that a patient does not wish to share with clinicians, but many respondents commented that this process is onerous. Only 18% of institutions' EHR implementations supported genetic risk scores. No respondents indicated that they had mechanisms for automatically transmitting genetic information to other family members, even in the specific case of migrating fetal results from prenatal testing of a mother into a child's medical record.

Recommendations for EHR improvement: Phase two survey

The CSER-eMERGE EHR working group provided recommendations for improving the display of genetic information in the EHR, which were cluster sorted into 20 distinct recommendations. Working group members were asked to prioritize these recommendations, to rank their top 4 recommendations, and to rate the feasibility of implementation

RESEARCH AND APPLICATIONS

Table 2: Reporting genetic results in EHR

Survey questions		N	Respondents reporting "Yes" (%)
1)	What are the different sources that enter genetic information that into your EHR?		
	a) Local Hospital Laboratory	17	17 (100)
	b) Reference Laboratory	17	17 (94)
	c) Independent genetic testing laboratories	17	17 (94)
	d) Physicians notes	17	17 (82)
2)	What factors determine the location in your EHR where genetic information is displayed?		
	a) Source laboratory (laboratory performing genetic testing)	17	11 (65)
	b) Source tissue (laboratory where the biospecimen originated)	17	1 (6)
	c) Pathology department (department within which the clinical laboratory is housed)	17	5 (29)
	d) Hospital Laboratory (clinical laboratory interfaced with hospital EHR)	17	8 (47)
	e) Ordering provider's department (department to which the provider ordering the clinical sequencing belongs)	17	7 (41)
	f) Purpose of information (indication for which the sequencing was performed)	17	6 (35)
	g) Other	17	7 (41)
3)	Are there instances where genetic information is displayed in multiple places in your EHR?	16	13 (81)
4)	Has there been any effort to consolidate where genetic information in your EHR is stored or displayed?	12	5 (42)
5)	Is there an institutional process for changing where a piece of genetic information is displayed in the EHR?	17	8 (47)
6)	Do you have a system other than the EHR to store genetic information where clinicians can view genetic results?	16	8 (50)
	a) Patient portal	9	5 (56)
	b) Clinical laboratory website	16	5 (31)
	c) Genomic result website outside EHR	16	3 (19)
	d) Physical copy	8	1 (13)
7)	Do you see your hospital IT infrastructure evolving in a way that will change the answers to the above questions?	17	12 (71)

aIncludes: context of user task, timing of test, Oversight Committee, and format.

(supplemental material). The top recommendations for improving the display of genetic information that were judged to be feasible with current EHR systems are summarized below. Each of these recommendations was ranked as a high priority recommendation by at least 50% of respondents:

- develop effective CDS for genetic results that are medically actionable;
- develop a decision support knowledge base to recommend appropriate actions (e.g., treatment, confirmatory diagnostic testing, specialty consultation);
- develop a mechanism for medically actionable genetic information to trigger an alert to the treating clinician;
- develop mechanisms to trigger an alert about pharmacogenomic information related to drug reactions if a relevant drug is prescribed; and
- develop effective CDS for genetic results that are diagnostic/disease defining.
- Additional recommendations that were ranked within the top four priority recommendations by at least 25% of respondents are listed below:

- provide mechanisms for EHRs to access external CDS knowledge bases and rules engines;
- develop ancillary CDS engines that store genomic information and interface with EHR systems;
- improve how institutions engage with clinician end users to develop priorities for EHR functionality; and
- implement systems capable of receiving and storing structured genetic sequence data according to standardized sequence ontology.

CDS systems were viewed an essential element for the successful implementation of personalized medicine.[12] Among high priority recommendations that were judged to be feasible with current systems were several variations of recommendations about developing effective CDS for medically actionable genetic information, diagnostic or disease defining genetic information, and pharmacogenetic information if a relevant drug is prescribed.

A key feature for genetic reporting and EHR integration that emerged through this analysis is the need to link genetic information to disease-specific knowledge bases that place variants in the appropriate context within the EHR. This would include linking variants to

RESEARCH AND APPLICATIONS

Table 3: Location of genetic information

Survey Question	n	Tab category						
		Laboratory	Genetics	Pathology	Problem list	Pharmacogenetics/drug interaction	Clinic notes	Outside medical records
Where is genetic information found in EHR?[a]								
Disease defining/diagnostic	17	88% (64–99)	29% (10–56)	18% (4–43)	59% (33–82)	0% (0–20)	59% (33–82)	24% (7–50)
Risk actionable	17	88% (64–99)	29% (10–56)	18% (4–43)	59% (33–82)	6% (0–29)	59% (33–82)	24% (7–50)
Low risk not actionable …	13	92% (64–99)	23% (5–54)	15% (2–45)	38% (14–68)	0% (0–25)	62% (32–86)	23% (5–54)
Large chromosomal changes …	16	81% (54–96)	25% (7–52)	25% (7–52)	50% (25–75)	0% (0–21)	56% (30–80)	31% (11–59)
Pharmacogenomics	17	82% (57–96)	24% (7–50)	18% (4–43)	41% (18–67)	18% (4–43)	59% (33–82)	24% (7–50)
Carrier recessive	17	88% (64–99)	29% (10–56)	18% (4–43)	35% (14–62)	0% (0–20)	65% (38–86)	35% (14–62)
Somatic/tumor genetics	15	87% (60–98)	27% (8–55)	33% (11–62)	40% (16–68)	0% (0–22)	60% (32–84)	33% (11–62)
Incidental	12	83% (52–98)	25% (5–58)	17% (2–48)	42% (15–72)	0% (0–26)	67% (35–90)	25% (5–58)
Variants of uncertain significance	13	85% (55–98)	38% (14–68)	15% (2–45)	38% (14–68)	0% (0–25)	62% (32–86)	38% (14–68)
Uninterpreted variants	7	57% (18–90)	14% (1–58)	0% (0–41)	43% (10–82)	0% (0–41)	29% (4–71)	0% (0–41)
Newborn Screening	16	81% (54–96)	25% (7–52)	19% (4–46)	38% (15–65)	0% (0–21)	69% (41–89)	31% (11–59)
Sensitive Genetic Information	15	87% (60–98)	33% (11–62)	13% (2–40)	40% (16–68)	0% (0–22)	60% (32–84)	27% (8–55)

[a]Genetic categories were designed to illustrate clinical or public health use cases and are not exclusive. Genetic results may fit multiple categories, e.g., sensitive – diagnostic results; actionable – incidental findings, etc., and some categories are clinical context dependent. Display locations are also not exclusive. Genetic information may be displayed in multiple locations simultaneously. Percentages are given with 95% bionomial confidence intervals.

RESEARCH AND APPLICATIONS

Table 4: Format of Genetic Information

Survey Question		n	Defined tests w/ discrete results (%)	Multiple variants from 1 test (%)	Text blobs (%)	PDFs (%)
	How is the genetic information stored?					
	Disease defining/diagnostic	17	53	12	59	88
	Risk actionable	17	53	12	59	88
	Low risk not actionable . . .	12	33	17	58	83
	Large chromosomal changes . . .	16	38	13	75	81
	Pharmacogenomics	17	65	18	53	65
	Carrier recessive	17	41	12	65	76
	Somatic/tumor genetics	14	50	14	71	79
	Incidental	13	38	8	62	77
	Variants of uncertain significance	15	40	13	53	67
	Uninterpreted variants	8	13	0	25	38
	Newborn Screening	16	44	13	56	81
	Sensitive Genetic Information	13	31	15	62	69

annotation-oriented knowledge bases that can effectively describe the variant in terms of clinical use (see Table 1), expected function (activating, inactivating, unknown), classification (e.g., synonymous, nonsynonymous, point mutation, insertion/deletion), and origin (somatic, germline), among other data types. In addition, EHR linkage would also include interpretation-oriented knowledge bases such as the ClinVar database[30] and the developing Clinical Genome Resource (ClinGen), which has a goal of providing well-researched and curated data about actionability and level of evidence that links a given gene with disease.[31] These 2 types of data sources would require continual review, updating and input from expert panels defined by disease type, and could be utilized to inform CDS for a given variant-disease relationship enabling an EHR to trigger clinical actions by the clinician based on the relationship between a variant's annotated function and disease relevance.

There is limited capacity to share even simple CDS across different EHR vendor products and few examples of successfully sharing CDS between health systems using the same EHR vendor. Beyond vendor supplied CDS rules (e.g., drug and allergy interaction alerts), most health systems manually implement and maintain their own CDS rules. Sharing CDS for genetic information will present an even greater challenge due to the amount of the data and the rapid changes in our knowledge. In our study, 40% of respondents ranked "access to external CDS" among their top 4 priorities, and 45% of respondents felt implementation is feasible at present (response of 4 or 5 on Likert scale).

Not surprisingly, the recommendations judged to be most feasible leverage systems that currently exist in most if not all EHR systems— that is, alert/reminder systems. This capability has been reported in both homegrown and proprietary systems, and has been utilized by many of the eMERGE sites for implementing pharmacogenomic decision support. The only function that was not implemented by any of the respondents was automated population of the problem list with diagnostic genetic information. This function is technically feasible requiring ICD or SNOMED codes to trigger alerts, which in some systems creates a scalability issue. Also, because the problem list is used by clinicians to manage patients there is some sensitivity to automated manipulation of the list.[32] This may contribute to perception that this option is less feasible or desirable.

Another recommendation that emerged from this study is the consensus for heightened engagement between developers in the EHR ecosystem and end user clinicians with respect to displaying complex genetic information. These data are unlike more common laboratory testing that is quantitatively represented in most EHRs, and genetic information can impact multiple layers of the EHR and have many different use cases depending on the type of end user (e.g., genetic counselor versus medical oncologist); thus, a thorough and continual dialogue between institutions that are implementing EHRs and the diverse set of genetic information types and clinician end users was judged as important to ensure that the EHR is equipped with appropriate genetic functionality that suits the diverse needs of the group.

DISCUSSION

We have shown that genetic information spans many use cases and applies to many distinct clinical situations, which may have different requirements for optimal representation and display in the EHR. For example, solutions for pharmacogenetic variants may require different display and CDS parameters than those for pathogenic variants that increase risk for disease. This latter class may require CDS that provides information such as screening recommendations and physical examination, laboratory, radiological, and procedure screening compliance. In addition, non-pharmacogenetic variants would not require pharmacy alerts.

Our results demonstrate that there is substantial heterogeneity in how genetic information enters and is documented in EHR systems at institutions across the United States. One of the most notable findings of the study is that genetic information, with the exception of pharmacogenetic information, is more commonly displayed in the EHR as PDFs and in paragraphs of free text rather than as structured data elements. For example, approximately 80% of our respondents noted that diagnostic, newborn screening, and somatic (tumor) genetic

Table 5: Display of specific categories of genetic information			
Survey Question		N*	Respondents reporting "Yes" (%)
a)	Which of the listed categories of genetic information are currently displayed in your EHR application?		
	Disease defining/diagnostic	17	17 (100)
	Risk actionable	17	17 (100)
	Low risk not actionable . . .	17	12 (71)
	Large chromosomal changes . . .	16	16 (100)
	Pharmacogenomics	17	17 (100)
	Carrier recessive	17	17 (100)
	Somatic/tumor genetics	17	14 (82)
	Incidental	17	11 (65)
	Variants of uncertain significance	17	14 (82)
	Uninterpreted variants	17	4 (24)
	Newborn Screening	17	15 (94)
	Sensitive Genetic Information	17	14 (88)
b)	Is this information further annotated in the EHR by physicians or other health professionals?		
	Disease defining/diagnostic	17	17 (100)
	Risk actionable	17	15 (88)
	Low risk not actionable . . .	16	11 (69)
	Large chromosomal changes . . .	17	16 (94)
	Pharmacogenomics	17	15 (88)
	Carrier recessive	16	14 (88)
	Somatic/tumor genetics	17	13 (76)
	Incidental	17	11 (65)
	Variants of uncertain significance	17	11 (65)
	Uninterpreted variants	17	5 (31)
	Newborn Screening	17	14 (88)
	Sensitive Genetic Information	17	11 (73)
c)	If a genetic test result is disease defining, could it be automatically be added to the EHR problem list?	17	0 (0)
d)	If a genetic test result is disease defining, would the ordering clinician be responsible for determining whether it should be added and adding it to the problem list?	17	17 (100)
e)	Is there a mechanism for high risk, medically actionable information to trigger an alert in the EHR?	17	12 (71)
f)	Can results be marked sensitive in your EHR?	17	7 (41)
g)	Can results be marked at physician preference?	17	5 (29)
h)	Can sensitive genomic results be redacted from the EHR or excluded from release of medical information?	15	8 (53)
i)	Are genetic risk scores supported in EHR implementation?	17	3 (18)
j)	Do prenatal genetic test results on the fetus migrate into the child's medical record after birth?	14	3 (21)
k)	For information that is relevant to other family members, is there a method for transmitting results to other family members?	17	0 (0)

* Not all respondents answered all questions

RESEARCH AND APPLICATIONS

information is displayed in their EHRs as PDFs whereas only about half reported that those data are displayed in their EHRs as structured data. The same test result may be generated by different laboratories and appear in different locations and different formats—for example, a *CFTR* gene mutation result from a targeted gene test or identified in an exome report. How genetic information is stored and displayed has profound implications for how those data can be used.

One disadvantage to genetic information stored in an unstructured way is that results are often difficult to locate, making it more likely that clinicians will be unable to find and act upon important genetic information months or years after initial testing. In addition, structured data can be utilized in ways that text or scanned PDFs cannot. Structured data is often needed to trigger CDS, and can be more easily aggregated for research purposes and quality initiatives, which helps to maximize the value of genetic information for populations as well as individual patients. Several of our respondents prioritized the need for EHR systems that are capable of receiving and storing structured data while many of our respondents noted that more robust CDS will be needed for the implementation of genetic medicine. Given the current state of informatics, most IT systems cannot reliably extract data from free text or scanned PDFs and require structured data input. Going forward, it will be critically important ensure that genetic information is stored and represented in the EHR in a way that maximizes their value.

Within the context of pharmacogenomics, as of this writing, the Clinical Pharmacogenetics Implementation Consortium (CPIC) has identified over 170 gene-drug pairs that either have been or may be targeted for guidelines.[33] As our understanding of genetics grows it will become increasingly difficult for clinicians to track how pharmacogenetic information can improve safety (by identifying genetic predisposition to adverse events) and provide individualized drug selection and dosing decisions for each patient. Similar to CPIC, ClinVar is accepting practice guidelines for variants as they are developed by professional societies, established *CFTR* variants are a current example.[30] As a result, many institutions are implementing these recommendations as CDS rules. As noted previously, relevant test data must be accessible to the CDS algorithm for this check to occur, which speaks to the need to support structured transmission and storage of genetic information.[34]

In addition, the lack of genetic education among health care providers and clinicians' low confidence in their ability to use genetic information offers both a challenge and an opportunity.[35–37] CDS alerts must be brief and convey complex knowledge simply. For example, identifying a variant as causing warfarin sensitivity indicates a need for interactive guidance that is difficult to convey with a simple alert, and could inadvertently lead a clinician to not to prescribe warfarin if unaware that guidelines recommend modification to the starting dose. The opportunity is that in an increasingly complex medical environment the EHR can be an important tool for supporting clinicians in making optimal decisions using information for which they may have a limited understanding.

As genome-scale testing becomes more common, there is increasing recognition of the possibility of reporting of incidental findings that are viewed as medically actionable. For example, the American College of Medical Genetics and Genomics recommends the reporting of pathogenic variants in 56 gene–disease pairs for any clinical germline test of the exome or genome.[38] Thus, consistent with this recommendation the survey respondents prioritized the need to develop standardized ways to display this type of putatively medically actionable genetic information in the EHR so that clinicians caring for the patient are away of this finding over time independent of the patient's current medical problems or family history.

We have documented substantial diversity in current practice with regard to handling of clinical genetic and genomic information in EHRs among large institutions (see Tables 2–4 and Supplemental Table A). One limitation of our study is that it was limited to large, predominantly academic institutions interested in advancing genomic healthcare. There would undoubtedly be even greater diversity if our survey had included smaller community sites. We believe the information from this survey can be used to develop surveys to get quantitative information from a broader group of organizations in the future. This heterogeneity is one of the main barriers to optimal genetics implementation in the EHR. However, EHR vendors could support general, interoperable solutions. The EHR mandates of the Health Information Technology for Economic and Clinical Health (HITECH) Act have led to a rapid expansion of EHR implementations and required vendors to attend to a diverse array of EHR needs; unfortunately, genetics has not been recognized on this list.[39] Separately, site-by-site solutions could be designed and built for sharing. Supportive resources would need to be prioritized and benefits to the developing institution may not be immediate. Financial barriers and competing priorities to hospital information technology and security teams must be overcome. Another limitation of our study is that prioritized recommendations were the results of an informal consensus building process; recommendations should vetted using a rigorously developed survey with a broader group of stakeholders in future work.

The benefits of standards for genetic information is that as EHRs data are increasingly shared, the information can be transmitted to outside EHR users and used by clinicians who serve multiple institutions. The Displaying and Integrating Genetic Information Through the EHR Action Collaborative (DIGITizE AC), is an ad hoc activity under the auspices of the Institute of Medicine Roundtable on Translating Genomic-Based Research for Health that is convening key stakeholders from health information technology vendors, academic health centers, government agencies, and other organizations to work together to examine how genomic information can be uniformly represented and integrated into electronic health records in a standards-based format.[40] This and other similar activities may address some of the priorities identified by our group. Without careful attention to the integration of genetic information into the EHR, the promise of personalized, genetic medicine at the individual patient and population level may not be fully realized.

COMPETING INTERESTS

SUPPLEMENTARY MATERIAL

Supplementary material is available online at http://jamia.oxfordjournals.org/.

REFERENCES

1. Hsiao C-J, Hing E. Use and characteristics of electronic health record systems among office-based physician practices: United States, 2001–2013. NCHS data brief, no 143. Hyattsville, MD: National Center for Health Statistics; 2014.
2. Lautenbach DM, Christensen KD, Sparks JA, Green RC. Communicating genetic risk information for common disorders in the era of genomic medicine. *Ann Rev Genomics Hum Genet.* 2013;14:491–513.
3. Lynch J, Whatley A, Uchendu US, Ibrahim SA. Race and genomics in the Veterans Health Administration. *Am J Public Health.* 2014;104 (Suppl 4): S522–S5224.
4. Kannry JL, Williams MS. The undiscovered country: the future of integrating genomic information into the EHR. *Genet Med.* 2013;15(10):842–845.
5. Kannry JL, Williams MS. Integration of genomics into the electronic health record: mapping terra incognita. *Genet Med.* 2013;15(10):757–760.
6. Kho AN, Rasmussen LV, Connolly JJ, et al. Practical challenges in integrating genomic data into the electronic health record. *Genet Med.* 2013; 15(10):772–778.
7. Ury AG. Storing and interpreting genomic information in widely deployed electronic health record systems. *Genet Med.* 2013;15(10):779–785.
8. Marsolo K, Spooner SA. Clinical genomics in the world of the electronic health record. *Genet Med.* 2013;15(10):786–791.
9. Farrugia G, Weinshilboum RM. Challenges in implementing genomic medicine: the Mayo Clinic Center for Individualized Medicine. *Clin Pharmacol Ther.* 2013;94(2):204–206.
10. Starren J, Williams MS, Bottinger EP. Crossing the omic chasm: a time for omic ancillary systems. *JAMA.* 2013;309(12):1237–1238.
11. 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):419–422.
12. Overby CL, Kohane I, Kannry JL, et al. Opportunities for genomic clinical decision support interventions. *Genet Med.* 2013;15(10):817–823.
13. Hoffman MA. The genome-enabled electronic medical record. *J Biomed Inform.* 2007;40(1):44–46.
14. Welch BM, Eilbeck K, Fiol GD, Meyer LJ, Kawamoto K. Technical desiderata for the integration of genomic data with clinical decision support. *J Biomed Inform.* 2014;51:3–7.
15. National Human Genome Research Institute. *Clinical Sequencing Exploratory Research (U01).* [cited 2014. http://grants.nih.gov/grants/guide/rfa-files/RFA-HG-10-017.html.
16. Gottesman O, Kuivaniemi H, Tromp G, et al. The Electronic Medical Records and Genomics (eMERGE) Network: past, present, and future. *Genet Med.* 2013;15(10):761–771.
17. Tarczy-Hornoch P, Amendola L, Aronson SJ, et al. A survey of informatics approaches to whole-exome and whole-genome clinical reporting in the electronic health record. *Genet Med.* 2013;15(10):824–832.
18. Institute of Medicine (IOM). The Interpretation of Genomic Data. In: Olson S, Beachy S, Giammaria C, Berger A, eds. *Integrating Large-Scale Genomic Information into Clinical Practice: Workshop summary.* USA: National Academies Press; 2012.
19. Hoffman JM, Haidar CE, Wilkinson MR, et al. PG4KDS: a model for the clinical implementation of pre-emptive pharmacogenetics. *Am J Med Genetics Part C, Semin Med Genet.* 2014;166c(1):45–55.
20. Bell GC, Crews KR, Wilkinson MR, et al. Development and use of active clinical decision support for preemptive pharmacogenomics. 2014;21(e1):e93–e99.
21. Weitzel KW, Elsey AR, Langaee TY, et al. Clinical pharmacogenetics implementation: approaches, successes, and challenges. *Am J Med Genetics Part C, Semin Med Genet.* 2014;166c(1):56–67.
22. Goldspiel BR, Flegel WA, DiPatrizio G, et al. Integrating pharmacogenetic information and clinical decision support into the electronic health record. *JAMIA.* 2014;21(3):522–528.
23. Laerum H, Bremer S, Bergan S, Grunfeld T. A taste of individualized medicine: physicians' reactions to automated genetic interpretations. *JAMIA.* 2014;21(e1):e143–e146.
24. Neri PM, Pollard SE, Volk LA, et al. Usability of a novel clinician interface for genetic results. *J Biomed Inform.* 2012;45(5):950–957.
25. Deshmukh VG, Hoffman MA, Arnoldi C, Bray BE, Mitchell JA. Efficiency of CYP2C9 genetic test representation for automated pharmacogenetic decision support. *Methods Inform Med.* 2009;48(3):282–290.
26. Bielinski SJ, Olson JE, Pathak J, et al. Preemptive genotyping for personalized medicine: design of the right drug, right dose, right time-using genomic data to individualize treatment protocol. *Mayo Clinic proceedings.* 2014; 89(1):25–33.
27. Peterson JF, Bowton E, Field JR, et al. Electronic health record design and implementation for pharmacogenomics: a local perspective. *Genet Med.* 2013;15(10):833–841.
28. Gottesman O, Scott SA, Ellis SB, et al. The CLIPMERGE PGx Program: clinical implementation of personalized medicine through electronic health records and genomics-pharmacogenomics. *Clin Pharmacol Therapeutics.* 2013; 94(2):214–217.
29. Aronson SJ, Clark EH, Varugheese M, Baxter S, Babb LJ, Rehm HL. Communicating new knowledge on previously reported genetic variants. *Genet Med.* 2012;14:713–719.
30. ClinVar. [cited March 4, 2015]. http://www.ncbi.nlm.nih.gov/clinvar/.
31. ClinGen: Clinical Genome Resource. 2015 [cited February 26, 2015]. http://www.clinicalgenome.org.
32. Feblowitz J, Henkin S, Pang J, et al. Provider use of and attitudes towards an active clinical alert: a case study in decision support. *App Clin Inform.* 2013;4(1):144–152.
33. Relling MV, Klein TE. CPIC: Clinical Pharmacogenetics Implementation Consortium of the Pharmacogenomics Research Network. *Clin Pharmacol Therapeutics.* 2011;89(3):464–467.
34. CPIC. CPIC Informatics Working Group. 2015 [cited February 6, 2015]. https://http://www.pharmgkb.org/page/cpicInformatics.

RESEARCH AND APPLICATIONS

35. Scheuner MT, Sieverding P, Shekelle PG. Delivery of genomic medicine for common chronic adult diseases: a systematic review. *JAMA.* 2008;299(11): 1320–1334.

36. Klitzman R, Chung W, Marder K, *et al.* Attitudes and practices among internists concerning genetic testing. *J Genetic Counseling.* 2013;22(1):90–100.

37. Freedman AN, Wideroff L, Olson L, *et al.* US physicians' attitudes toward genetic testing for cancer susceptibility. *Am J Med Genetics Part A.* 2003; 120a(1):63–71.

38. Green RC, Berg JS, Grody WW, *et al.* ACMG recommendations for reporting of incidental findings in clinical exome and genome sequencing. *Genet Med.* 2013;15(7):565–574.

39. Index for Excerpts from the American Recovery and Reinvestment Act of 2009 (ARRA) [cited February 20, 2015]; Revised 2/19/09 ONC Pre-Decisional Draft. http://www.healthit.gov/sites/default/files/hitech_act_excerpt_from_arra_with_index.pdf

40. (IOM) IoM. Developing Guiding Principles for Integrating Genomic Information Into the Electronic Health Record Ecosystem. [cited February 20, 2015]. http://www.iom.edu/Activities/Research/GenomicBasedResearch/Innovation-Collaboratives/EHR.aspx.

AUTHOR AFFILIATIONS

[1]Department of Laboratory Medicine, University of Washington, Seattle, WA, 98195, USA

[2]Department of Medicine, Division of Medical Genetics, University of Washington, Seattle, WA, USA

[3]Personalized Medicine, Partners Healthcare, Boston, MA, USA

[4]Department of Pediatrics, Columbia University Medical Center, New York, NY, USA

[5]Department of Medicine, Harvard Medical School, Boston, MA, USA

[6]Dana-Farber Cancer Institute, Boston, MA, USA

[7]National Human Genome Research Institute, NIH, Rockville, MD, USA

[8]Department of Genome Sciences, University of Washington, Seattle, WA, USA

[9]Department of Pediatrics, Baylor College of Medicine, Houston, TX, USA

[10]Department of Internal Medicine, University of Michigan, Ann Arbor, MI, USA

[11]Department of Biomedical Informatics and Medical Education, University of Washington, Seattle, WA, USA

[12]The Broad Institute of MIT and Harvard, Cambridge, MA, USA

[13]Department of Pathology and Laboratory Medicine, University of North Carolina at Chapel Hill, Chapel Hill, NC, USA

[14]Department of Genetics, University of North Carolina at Chapel Hill, Chapel Hill, NC, USA

[15]Department of Health Sciences Research, Division of Biomedical Statistics and Informatics, Mayo Clinic, Rochester, MN, USA

[16]Department of Biomedical and Health Informatics, The Children's Hospital of Philadelphia, Philadelphia, PA, USA

[17]Group Health Research Institute, Group Health Cooperative, Seattle, WA, USA

[18]Biomedical Informatics Research Center, Marshfield Clinic Research Foundation, Marshfield, WI, USA

[19]Department of Biomedical Informatics, Vanderbilt, Nashville, TN, USA

[20]Department of Preventive Medicine, Division of Health and Biomedical Informatics, Northwestern University Feinberg School of Medicine, Chicago, IL, USA

[21]Genome Medicine Institute, Geisinger Medical Center, Danville, PA, USA

[22]Department of Medicine, Program for Personalized and Genomic Medicine and Center for Health-Related Informatics and Bioimaging, University of Maryland School of Medicine, Baltimore, MD, USA

RESEARCH AND APPLICATIONS