An electronic health record (EHR) log analysis shows limited clinician engagement with unsolicited genetic test results


Department of Medicine, Division of Nephrology, Columbia University, New York, New York, USA.

†Jordan G. Nestor and Alexander Fedotov are co-first authors.

*Corresponding Author: Jordan G. Nestor, MD, Department of Medicine, Division of Nephrology, Columbia University, New York, New York, USA; jgn2108@cumc.columbia.edu

Received 25 November 2020; Revised 21 January 2021; Editorial Decision 5 February 2021; Accepted 12 February 2021

ABSTRACT

How clinicians utilize medically actionable genomic information, displayed in the electronic health record (EHR), in medical decision-making remains unknown. Participating sites of the Electronic Medical Records and Genomics (eMERGE) Network have invested resources into EHR integration efforts to enable the display of genetic testing data across heterogeneous EHR systems. To assess clinicians’ engagement with unsolicited EHR-integrated genetic test results of eMERGE participants within a large tertiary care academic medical center, we analyzed automatically generated EHR access log data. We found that clinicians viewed only 1% of all the EHR-integrated eMERGE results. We also identified several patterns of user engagement with the EHR’s Genetics section and with eMERGE genetic test results based on the user’s clinical role, practice setting, and the context for their involvement in a participant’s care.

LAY SUMMARY

To evaluate whether clinicians at our institution viewed eMERGE study participants genetic test results entered in the EHR, we analyzed users’ EHR access logs. We found clinician users viewed only 1% of all the EHR-integrated eMERGE results. We also identified several patterns of user engagement with the EHR’s Genetics section and with eMERGE genetic test results based on the user’s clinical role, practice setting, and the context for their involvement in a participant’s care.
opportunities to use EHR log data as a performance metric to more precisely inform ongoing EHR-integration efforts and decisions about the allocation of informatics resources in genomic research.

Key words: clinical engagement with genomic results, electronic health records, log analysis

INTRODUCTION

The success of Precision Medicine initiatives relies, in large part, on understanding how clinicians utilize genomic data. Several NIH-funded initiatives, such as the Clinical Sequencing Evidence-generating Research (CSER) consortium,1 Implementing GeNomics in practice (IGNITE),2 and the Electronic Medical Records and Genomics (eMERGE) Networks,3 have set out to understand how clinicians engage with medically actionable genomic findings. Within these initiatives, members of the eMERGE Network—a NHGRI-funded consortium of academic and integrated health systems established in 2007 (https://emerge-network.org)—have invested institutional resources into electronic health record integration. To date, these efforts have primarily focused on the electronic transmittal of standardized genomic data from laboratories to healthcare providers.4 Despite dedicated efforts across these sites to display actionable genomic data in the EHR, it remains to be seen if clinicians utilize unsolicited genomic findings in their medical decision-making, and whether it impacts patient care. EHR access and audit logs offer unique opportunities to scale observational research on user behaviors in a disease-agnostic and minimally biased manner.5–7 To date, no attempts have been made to assess clinician engagement with genomic data using EHR access logs.8–14 Therefore, we set out to evaluate the breadth of clinicians’ engagement with unsolicited EHR-integrated genetic test results of eMERGE participants using automatically generated access logs within a large tertiary care academic medical center.

MATERIALS AND METHODS

The eMERGE Network’s phase III study set out to return medically actionable genomic findings to 25 380 biobank participants and integrate the data into the EHR to support personalized medicine efforts and to track long-term outcomes. Participants underwent clinical-grade genetic screening with a customized next-generation sequencing panel (ie, eMERGE-Seq platform) of 109 clinically relevant, and medically actionable genes15,16. Between 2016 and 2017, 1071 eMERGE participants were prospectively enrolled at New York-Presbyterian Hospital (NYP)/Columbia University Irving Medical Center.17,18 The study’s return of results phase began October 1, 2018–May 31, 2019. In total, 1021 participants with “negative” (ie, nondiagnostic) results were mailed a genetic counseling letter and a copy of the molecular pathology report; 50 participants with “positive” (ie, variants classified as Pathogenic or Likely Pathogenic, per the American College of Medical Genetics and Genomics,19 or a genetic risk variant deemed actionable) results were invited to return for an in-person genetic counseling. Genetic test results and counseling letters were entered (.pdf) in the Genetics section of a home-grown EHR (iNYP).20

This study did not include human subjects. This purely observational (ie, noninterventional) outcomes involved the analysis of materials (eg, data, documents, records or specimens) collected solely for nonresearch purposes (eg, medical treatment, diagnosis, etc.). The risk of harm to subjects exposed as a result of this research was no more than minimal.

Data sources and preprocessing

iNYP, a hospital-supported web-based platform, displays patients aggregated clinical information, and is accessed directly through its website and through both the hospital- (SCM Allscripts) and university-supported (Crown) EHR systems. iNYP was the platform selected to input eMERGE results because of its flexibility for developing test result display functions, popularity among users and its easy-to-use reputation.21,22

In the context of log analysis, we defined “clinician engagement” with genetic test results as any episode of patient record access by a clinician that involved an interaction with the Genetics section (and any document within this section) of the iNYP platform—for example, an attempt to access the Genetics section, an attempt to view the genetic test report, etc. A logged user session was defined as a recorded set of unique user’s interactions with iNYP that involved the Genetics section within a single platform login/logout (or timeout due to inactivity) period. The degree of clinician’s engagement with results was assessed based on the number and type of documents (ie, the pathology report and/or counseling letter) accessed in the Genetics section, along with the amount of time spent on this activity (also, see Data analysis).

A custom-programmed script was used to retrieve iNYP raw log session files of all hospital- and university-affiliated users who viewed any type of genetic data in iNYP from October 1st, 2018 to December 16th, 2019. Logs captured user access of iNYP’s Genetics section, including users who accessed it through the university- and hospital-supported EHRs. The following elements were captured in each log file: relevant search terms (eg, “genetics,” “emerge,” etc.),
date and time of access, user login, EHR platform used to access iNYP, patient medical record number, and when applicable—a unique eMERGE identification number and a timestamp of the genetic test result document viewed. Initially, 23,306 raw logs session files were retrieved (Figure 1) using this script. We then removed 437 files that contained “emergency.” The 22,869 files were of all users who accessed iNYP’s Genomics section. Next, we filtered out 22,690 files that did not contain “emerge” and the corresponding eMERGE identification number. The remaining 179 files were of users who accessed the Genomics section of an eMERGE participant. Then, 145 files were excluded as the users were study team members or nonclinicians (eg, IT employees, third-party contractors, administrators, etc.). Adjudication of nonclinician was based on their university/hospital-designated employee title and absence of a National Provider Identification number. For the remaining access logs, we manually identified each user and categorized each of them by gender, employer (eg, hospital or university), clinical role (ie, attending/fellow/resident-level physicians vs. nurse practitioners), specialty (ie, medicine vs. nonmedicine disciplines) using employee directories. Then, a clinician (JGN) performed in-depth chart reviews to determine the context in which each user accessed the EHR’s Genomics section of an eMERGE participant.

**Data analysis**

Clustering analysis, long recognized as an effective classification method for exploratory analysis of high-dimensional and heterogeneous EHR log data, was used to identify patterns of clinician engagement with EHR-integrated genetic results. Variables extracted from access logs and participants’ charts included in the final dataset used for the clustering analysis were as follows: time of Genomics section access, EHR platform used to access iNYP, practice setting, clinician categorical data, type of genetic test result document opened and clinical context of Genomics section access. Given the mixed-type data, along with the lack of a priori information about the number of clusters one should expect in our dataset, we chose a hierarchical clustering approach for this analysis that is known to work especially well with smaller datasets. Agglomerative complete linkage clustering of Gower’s coefficient distances was performed using R package “cluster.” Two independent and well-established methods of verifying the quality of clustering solutions were used. Multi-scale bootstrap resampling was performed to assess the uncertainty of the clustering solution using R package “pvclust.” The approximately unbiased (AU) $P$-value was computed with 1000 times resampling and clusters with an AU $P$-value larger than .95 were set as significant modules. The optimal number of clusters ($k$) was determined by using the Silhouette coefficient method. The amount of time spent by a unique user reviewing a section of 31 eMERGE participants ($n = 5$ participants with positive results; $n = 26$ participants with negative results). Of these 30 clinicians, only 13 went on to open at least one of the genetic test result documents ($n = 5$ participants with positive results; $n = 8$ participants with negative results), representing 1% of the 1071 study results that were uploaded onto iNYP.

**User characteristics and clinician workflows**

User characteristics and workflow features are summarized in Table 1. Users were mostly (67%) males and had the following roles: physicians at the level of attending ($n = 14$), fellow ($n = 4$), and resident ($n = 9$); nurse practitioners ($n = 3$). Nineteen (63%) users practiced internal medicine (ie, “medicine”): nephrology ($n = 8$); general internal medicine ($n = 5$); gastroenterology ($n = 2$); hematology/oncology ($n = 2$); cardiology ($n = 1$); rheumatology ($n = 1$). The 11 remaining users were in specialties other than medicine: anesthesiology ($n = 4$); surgery ($n = 3$); radiology ($n = 2$); neurology ($n = 1$); pathology ($n = 1$).

Seventy-seven percent of users ($n = 23/30$) accessed the EHR during “regular” workhours (ie, Monday–Friday, 06:00–18:00) compared with 20% who accessed it during “on-call” (ie, evenings and weekends) hours, and one clinician who accessed the EHR of more than one participant during “regular” and “on-call” hours. Most (63%) clinicians accessed iNYP using hospital-supported platforms—the iNYP website or SCM Allscripts ($n = 5$ and $n = 14$, respectively), while 37% accessed iNYP through the university-supported EHR (Crown). Clinicians accessed the Genomics section in the following practice settings: outpatient (60%); inpatient/hospitalized (37%); both (3%).

Users accessed the Genomics section in the following contexts: around the date and time of a scheduled outpatient appointment with a participant (37%); as part of the inpatient care team during a participant’s hospitalization (37%); as a participant’s established outpatient provider (17%); not ascertained (10%).

Clinicians specialized in medicine ($n = 19$) accessed the Genomics section of 20 eMERGE participants. Among them, 13 (68%) viewed at least one result document compared with 6 (32%) users who only accessed the Genomics section. Furthermore, all 5 positive results were viewed by a clinician specialized in medicine. In 2 cases, the clinician’s (ie, one nephrologist and a gastroenterologist) intervention was informed by the positive finding and documented as part of the participant’s outpatient encounter. The remaining 11 users, all from specialties other than medicine, accessed the Genomics section of 12 participants. None of these 11 users opened a genetic test result document, or accessed the Genomics section, for any of the participants with positive results.

**Engagement levels by user traits**

Hierarchical clustering revealed 5 main types of clinician engagement with the genetic results, summarized in Table 2. Types A and B represent clinicians who in absolute majority of cases fully engaged with the genetic results and opened at least one of the documents. These types are represented primarily by attending physicians, specialized in medicine. (A) who were all part of the inpatient care team during a participant’s hospitalization and viewed genetic results mostly “on-call” and mostly using the hospital-supported EHR or (B) mostly viewed genetic results around the time of a participant’s scheduled outpatient appointment, during “regular” workhours, using the university-supported EHR. These clinicians, on average, spent 151 and 93s reviewing genetic results, respectively. The
remaining 3 cluster types (ie, C, D, and E) represent clinicians whose engagement with the genetic results was transient and limited to only accessing the Genetics section before transitioning to the next recorded activity, without viewing any documentation there. These types were represented by attending physicians of nonmedicine specialties who spent on average 17, 9, or 3 s accessing the Genetics section, respectively. Nonparametric ANOVA test of the time that clinicians spent accessing genetic results indicated that observed differences are significant (P < .001).

**DISCUSSION**

Participating eMERGE sites have devoted tremendous resources in development of infrastructures to support EHR integration and display of genomic data. In this study, we describe one institutional experience assessing clinicians’ engagement with unsolicited genetic test results displayed in the EHR as a proof of concept for using objective EHR user logs. We found that clinicians viewed only 1% of the total number of eMERGE results that were integrated in the EHR. The 13 users that went on to view the genetic test results were more likely to be senior clinicians (ie, attendings-level physicians) specialized in medicine (vs. anesthesiology, surgery, etc.). By categorizing each user by role, practice setting, and context for their involvement in a participant’s care, we also identified several patterns of user engagement with the EHR’s Genetics section and with unsolicited genetic test results based on user characteristics. The analysis also allowed us to assess the outcome of 2 users’ engagement with the genetic test results through the capture of 2 documented instances where the positive results influenced physician medical decision-making and subsequent management.

Similar to other EHR log studies, we were able to observe the clinician behaviors across a diverse set of clinical specialties. The log data provided us with enough identifiable information about the user to guide deeper chart review. This in turn allowed us to observe what users were doing when they accessed participants’ EHR data, giving us insights into their different workflows and tasks. Overall, few users in our study went on to view participants unsolicited genetic test results in the EHR, regardless of whether the findings were positive or negative. Unlike recent network studies on physicians’ attitudes toward unsolicited genetic test results, which use surveys and qualitative interviews, our study used objective and disease-agnostic data to assess the scope of clinician engagement with unsolicited genomic data. Furthermore, we make granular level observations of individual users in our study and compare which users viewed the study results versus who merely accessed the Genetics section before moving on to another section of the EHR. The insights into users’ specialties, levels of training, practice settings, and workflows, shed light onto users’ needs (ie, who needs to see what and when). For example, an attending physician, in internal medicine, seeing nonacute patients in an outpatient setting, may rely on a broad range of data for medical decision-making. In contrast, a junior resident-level physician in an

### Table 1. User characteristics

<table>
<thead>
<tr>
<th>Characteristics</th>
<th>Users N = 30</th>
<th>User session logs N = 34&lt;sup&gt;a&lt;/sup&gt;</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>N = 20</td>
<td>n = 14</td>
</tr>
<tr>
<td>Clinical role</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Attending-level physician (most senior)</td>
<td>14 (47%)</td>
<td>10 (47%)</td>
</tr>
<tr>
<td>Fellow-level physician</td>
<td>4 (13%)</td>
<td>2 (15%)</td>
</tr>
<tr>
<td>Resident-level physician</td>
<td>9 (30%)</td>
<td>1 (29%)</td>
</tr>
<tr>
<td>Nurse practitioner</td>
<td>3 (10%)</td>
<td>1 (&lt;1%)</td>
</tr>
<tr>
<td>Clinical specialty</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Anesthesiology</td>
<td>4 (13%)</td>
<td>5</td>
</tr>
<tr>
<td>Surgery</td>
<td>3 (9%)</td>
<td>3</td>
</tr>
<tr>
<td>Radiology</td>
<td>2 (6%)</td>
<td>2</td>
</tr>
<tr>
<td>Neurology</td>
<td>1 (3%)</td>
<td>1</td>
</tr>
<tr>
<td>Pathology</td>
<td>1 (3%)</td>
<td>1</td>
</tr>
<tr>
<td>Medicine</td>
<td>19 (63%)</td>
<td>14</td>
</tr>
<tr>
<td>Time user accessed EHR</td>
<td></td>
<td></td>
</tr>
<tr>
<td>During “regular” workhours&lt;sup&gt;b&lt;/sup&gt;</td>
<td>23 (77%)</td>
<td>16</td>
</tr>
<tr>
<td>During “on-call” hours&lt;sup&gt;b&lt;/sup&gt;</td>
<td>6 (20%)</td>
<td>2</td>
</tr>
<tr>
<td>During “regular” and “on-call” hours&lt;sup&gt;b&lt;/sup&gt;</td>
<td>1 (3%)</td>
<td>2</td>
</tr>
<tr>
<td>Practice setting when user accessed EHR</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Outpatient/ambulatory</td>
<td>18 (60%)</td>
<td>10</td>
</tr>
<tr>
<td>Inpatient/hospitalized</td>
<td>11 (37%)</td>
<td>9</td>
</tr>
<tr>
<td>Both outpatient and inpatient settings</td>
<td>1 (3%)</td>
<td>2</td>
</tr>
<tr>
<td>Context when user accessed EHR</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Around the time of a scheduled outpatient clinical encounter</td>
<td>11 (37%)</td>
<td>5</td>
</tr>
<tr>
<td>Care team member of a hospitalized participants</td>
<td>11 (37%)</td>
<td>9</td>
</tr>
<tr>
<td>Established outpatient provider</td>
<td>5 (17%)</td>
<td>3</td>
</tr>
<tr>
<td>Not ascertained/unknown</td>
<td>3 (10%)</td>
<td>3</td>
</tr>
</tbody>
</table>

<sup>a</sup>Corresponding to 31 study participants.

<sup>b</sup>“Regular” weekday hours (ie, Monday–Friday, 06:00–18:00); “on-call” (ie, evenings and weekends) hours.

Totals do not equal 100% due to rounding.
esthesiology, performing pre-operative evaluations in acutely ill, hospitalized patients, may review EHR data in a more task-specific and narrow scope. These observations also give rise to questions that merit further investigation, such as the influence clinical role plays in the navigation and integration of patients’ EHR and genomic data, and the use of this information in users clinical decision-making.

The strengths of this study include the novel use of EHR access logs to assess clinicians’ engagement with unsolicited genomic data integrated in the EHR, at our institution. Our study provides valuable insights into the types of users who accessed the EHR’s Genet- ics section, but who may or may not have gone on to view the eMERGE genetic test results. Furthermore, this study is a proof of concept for the use of EHR log data as an objective assessment of the return on investment for integrating new types of clinical data into EHR platforms, and highlights the importance of addressing real-world barriers as part of EHR-integration efforts. An unexpected finding in our study was that despite iNYP’s reputed ease of use and clear clinical data displays, only a small portion of clinicians accessed the EHR-integrated eMERGE genetic test results. And, though it has been shown that certain types of EHR data goes unviewed by target users,33 our findings highlight an important, and still unresolved, question for genomic medicine—who are the clinicians that need to know about a patient’s actionable genomic findings? The scope of this problem is significant, and underscored by over a decade of work by national NHGRI-funded research networks, continuously focused on identifying sustainable ways to promote effective implementation of genomic findings into clinical care, through EHR-integration. Our findings underscore the urgency to identify these intended users as part of ongoing genomic implementation and EHR-integration research. Further study is needed into user characteristics, their role-specific informational

Table 2. Patterns of clinician engagement with EHR-integrated genetic test results

<table>
<thead>
<tr>
<th>Engagement type (number of unique user session logs)</th>
<th>Engagement pattern</th>
<th>Entry viewed</th>
<th>Average time spent accessing genetic testing results, s*</th>
</tr>
</thead>
<tbody>
<tr>
<td>“Inpatient attending physicians” (A, n = 5)</td>
<td>Mostly attending physicians; all of medicine specialties; all part of the inpatient care team during a participant’s hospitalization; mostly viewed genetic results during “on-call” hours; mostly using the hospital-supported EHRb</td>
<td>Mostly accessed the Genetics section and then clicked to view the genetic test results</td>
<td>151</td>
</tr>
<tr>
<td>“Outpatient attending physicians” (B, n = 11)</td>
<td>Mostly attending physicians; all of medicine specialties; mostly viewed results around the date and time of a scheduled outpatient appointment with a participant; mostly during “regular” hours; mostly using the university-supported EHRb</td>
<td>Mostly accessed the Genetics section and then clicked to view the genetic test results</td>
<td>93</td>
</tr>
<tr>
<td>“Non-medicine attending physicians” (C, n = 3)</td>
<td>All attending physicians; mostly from surgery; accessing the Genetics section for unknown reasons or as an established outpatient provider of a participant; all during “regular” hours; all using the hospital-supported iNYP website</td>
<td>Only accessed the Genetics section</td>
<td>17</td>
</tr>
<tr>
<td>“Outpatient trainees” (D, n = 7)</td>
<td>Mostly residents; from diverse clinical specialties; mostly accessing the Genetics section around the date and time of a scheduled outpatient appointment with a participant; all during “regular” hours; all using the hospital-supported EHRb</td>
<td>Mostly accessed the Genetics section</td>
<td>8</td>
</tr>
<tr>
<td>“Inpatient trainees” (E, n = 8)</td>
<td>Mostly residents; from anesthesiology; all part of the inpatient care team during a participant’s hospitalization; mostly accessing the Genetics section during “regular” hours; and mostly using the hospital-supported EHRb</td>
<td>Only accessed the Genetics section</td>
<td>3</td>
</tr>
</tbody>
</table>

a“Regular” weekday workhours (ie, Monday–Friday, 06:00–18:00); “on-call” (ie, evenings and weekends) hours.
bHospital-supported EHR is SCM Allscripts and the university-supported EHR is Crown.

*P < .001.
and workflow needs, and the factors that influence their engagement with clinical and genomic data presented in the EHR. This information will inform the design of EHR data displays and clinical-decision support tools that are accessible, easy-to-consume, usable across diverse practice settings,32–34 and effective in promoting clinicians’ engagement and utilization of the genomic data.

CONCLUSION

These data contribute important empirical knowledge on the application of objective EHR log data to evaluate clinician engagement with genomic data and with unsolicited findings displayed in the EHR, within a clinical context. It also provides insights into the informational and workflow needs of users and adds some valuable insights into types of activities that clinicians were engaged in when they accessed the unsolicited genetic test results. This study highlights opportunities to use EHR logs in ongoing EHR-integration efforts in both research and healthcare settings.

FUNDING

The project was supported by grants from the National Institutes of Health through grants TL1TR001875 (JGN), SU01HG008680-04 (eMERGE—Columbia University), and UL1TR001873 (CTSA) and the National Kidney Foundation’s Young Investigator Award (JGN).

AUTHOR CONTRIBUTIONS

CW, SS, SB, GH, AG, WC, JW, DF, AF, and JGN conceived the study. CW, AF, and JGN designed the study. JGN, AF, and DF collected the data. JGN and AF analyzed the data. JGN, AF, and CW wrote the manuscript. All authors contributed to and discussed the results and critically reviewed the manuscript.

DATA AVAILABILITY

The data underlying this article will be shared on reasonable request to the corresponding author.

CONFLICT OF INTEREST STATEMENT

None declared.

REFERENCES