Can EHRs Handle Genomics Data?

Save to myBoK

By Daniel DuBravec, CHTS, CEHRS

Virginia's Inova Fairfax Hospital recently announced plans to open a $245 million genomics research center as part of Inova's Center for Personalized Health. The research center will sequence nearly 20,000 whole genomes by the end of 2016. Whole genome sequencing, a process where a specialized laboratory device creates a full DNA map of a patient's genome, was once a rare procedure. But it is now conducted in nearly 240,000 labs across the country.[1]

Only a few years ago, researchers spent $10,000 to $50,000 outsourcing the process to specialized labs. Today, the whole process costs less than $1,200 and the cost continues to decline as personalized medicine becomes regular practice. In a report by the *Journal of the American Medical Informatics Association* (AMIA) on the future direction of electronic health records (EHRs), the authors recognized "with ever-decreasing costs of sequencing technology, patients' genomes are likely to be sequenced routinely in the course of clinical care in the not-so-distant future."[2]

With great amounts of data comes great responsibility. Laboratories and clinics are facing data security challenges in storing and integrating highly sensitive personal genomic data into their practices' EHRs. The growth of the precision medicine industry has created a need for nationally recognized genomic EHR standards and policies to protect individuals and communities, while local-level care centers require training on proper data storage and network safety.

Why Genomic Data is Important

At the 2015 Health Level Seven (HL7) Genomics Policy Conference in Washington, DC, physicians and genomic specialists stressed the importance of integrating genomic data with other types of healthcare data through hospital EHR platforms. Forecasting the future of personalized medicine, the presenters envision healthcare centers collecting and sequencing patients' genomic data and then generating analytical reports with EHRs which physicians would use to take note of genes with relationships to a multitude of hereditary diseases.

Five to 10 percent of all cancer is hereditary.[3] If hereditary life-threatening disease is identified early, personalized patient treatment can be administered, surgeries can be recommended to reduce the health risk to patients, healthcare providers can be alerted of potentially harmful prescription drug interactions based on a patients' genes, and family members can be notified of hereditarily passed conditions.

An estimated one in four people go undiagnosed with medical conditions that could be identified through whole genome sequencing.[4] Two recent studies on celiac disease explored how celiac patients live with the disease without knowing it for four years, on average. Women go undiagnosed even longer than men.[5] With whole genome sequencing, celiac genes can be identified early, letting the physician know to look at celiac disease as a potential health issue and to order additional testing.

By integrating the results of genetic testing with care centers' EHRs, automated flags can be generated, making automated linkages to hereditary diseases. The positive results of genomic integration are echoed by Molly Coye, MD, MPH, chief innovation officer at UCLA Health, stating that genomic data combined with EHRs could have "real clinical meaning" for patients.[6]

Challenges for Integrating Genomic Data with EHRs

Whole genome sequencing produces overwhelming amounts of raw data. Research scientists are still deciphering the full meaning of the data they are collecting. Scientists and physicians need to select specific data to care for a patient, integrate it into the care center's EHR, and produce actionable insights. Meanwhile, EHR vendors are still developing solutions for user interface, mapping personal genetic information to in-house or external data libraries, and enhancing search capabilities—all while improving privacy protection.

Using genomically integrated EHRs, physicians and genetic counselors will be able to interpret and identify relevant information using the EHR to better understand a patient's response to treatment based on their genetic information. An example of the effectiveness of incorporating genomic data into electronic health records is demonstrated at the Geisinger Health System. Staff members regularly update patient records in their health system's EHR with new genetic information, making clinicians aware of these changes and educating patients about their role in managing their own information.[7]

The major challenges the industry faces in integrating genomic data with EHR systems can be broken down into five categories. From the volume of data and the number of people, standardization of practices, and policies and privacy issues, to the actual storage and indexing of data that requires humans to garner actionable information, the challenges are big but not insurmountable.

1. What to do with all that data?

In an article on the technical needs for integration of genomic data into EHRs published in the *Journal of Biomedical Informatics*, the authors express valid concerns about the enormity of raw genomic data: "The large volume of each individual's DNA, protein and related data—hundreds of gigabytes to terabytes in its raw form—exceeds the capacity of commonly available network bandwidth and disk storage in healthcare settings."[8] Geneticists are predicting that the storage needs for patient genomic data at the capacity of 2-40 exabytes of storage, exceeding the overall data storage used by YouTube, as an estimated 100 million to two billion human genomes are sequenced by 2025.[9]

Rapid retrieval and analysis of this data is challenging for the typical database system used by most EHRs, leading to substantial performance issues. In the international weekly science publication *Nature*, authors Eric Green and Mark Guyer state in an article on charting the course of genomic medicine that "existing clinical informatics architectures are largely incapable of storing genome sequence data in a way that allows the information to be searched, annotated and shared across healthcare systems over an individual's lifespan."[10]

By storing the raw data outside of the EHR, users can transfer data sets into the system when they need them. The original raw data could be stored locally in the hospital in storage clusters or remotely in "the cloud." Cloud-based storage and applications companies such as Amazon and Google are currently developing the infrastructure to host the raw genomic data to alleviate the technical and infrastructure burdens from hospitals.[11] Security concerns over where the data is stored, whether in the cloud or in different clusters, remain and must be compliant with federal Health Insurance Portability and Accountability Act (HIPAA) Privacy and Security Rules.[12]

2. We're counting on standardization

Creating national standards for how individuals' genomic information is shared and stored is more difficult than it seems. EHR platforms must be tailored to incorporate the results of genetic tests, and then shared with healthcare professionals, who then share the data with their patients. This knowledge transfer requires health information technology (health IT) professionals and standards developing organizations (SDOs) to agree on genetic terminology, interoperability, and methods of transfer. Standard structures and language are required for genetic data to be exchanged between clinics and to drive clinical decision support tools.[13] The standards organization HL7 published an implementation guide outlining data standards for reporting genetic test results. Additionally, organizations such as the Institute of Medicine's Roundtable on Translating Genomic-Based Research for Health, the Clinical Genome Resource (ClinGen), and the National Center of Biotechnology, with the support of the National Institutes of Health (NIH), are all also tackling this issue so genomic data can eventually be used by health information exchanges to transmit data between different EHRs and other systems.

3. Privacy laws, policies, and practices are weak

Stolen genome data could have tragic consequences. Criminals who obtain access as a result of poor data usage policies, inadequate privacy standards, and ineffective legal protections can lead to discrimination, undisclosed paternity or ancestry data being made public, and other real-world implications of misused data. Prosaic issues such as targeted advertisements based on genome information are not far off, either.[14] Today's empowered patients expect their genomic data will have the most restrictive privacy policies and secure storage technology in place. However, many studies have shown that identifying individuals through their health data is not difficult. Using genealogy websites, users can be positively identified. Care centers

have a responsibility to explain the risks associated with sequencing, as well as their privacy policies and practices to patients. At the technical level, any stored genomic data needs to be encrypted with restrictive access controls.

Individuals will also be better protected when federal and state genomic laws are enacted and enforced. In a report investigating the ethical and practical challenges of sharing data from genome-wide association studies, the researchers concluded that the NIH should be challenged to create clearer guidance to ensure ethical safeguards are part of any institute's policy when sharing genomic data.[15] Encapsulating this concern over policies and protections, a report on "Patient Privacy in the Genomic Era" states that "research in the genomic privacy field is still in its infancy."[16] The Genetic Information Nondiscrimination ACT (GINA) provides federal protection from the misuse of genetic results and genetic discrimination in health insurance and employment, but offers no guidance in EHR integration.

The US Presidential Commission for the Study of Bioethical Issues released recommendations for ensuring patient privacy and data security while promoting scientific advancement through whole genome sequencing. But they are simply recommendations. At the state level, two examples of legislation being passed to protect patients' genetic information are Illinois, which requires patient consent before any genetic data may be disclosed under the Genetic Information Privacy Act (410 ILCS 513), and California, with the CA Genetic Information Non-discrimination Act (SB 559), which offers protections to individuals in the areas of health and life insurance coverage, housing, mortgage lending, employment, education, and elections.

4. Still in search of user-friendly systems

Thanks to browsing tabs, drop-down menus, and search fields, traditional data such as patient vitals and medical history are becoming more searchable in most EHRs. As clinicians seek pertinent information, the navigational design of the EHR provides search capabilities to easily find patient data. It even has the ability to create impromptu reports on the patient's condition. Most EHRs attempting to include genetic information provide a clinical notes section where free form text can be captured. This data is either automatically imported from laboratory systems or copied and pasted, where it lives undiscoverable and difficult to interpret. Another method used by healthcare providers is to attach a PDF as part of the patient record. Brad Strock, software developer for Epic Systems, expressed his frustration with this method at HL7's 2015 Genomics Policy Conference, saying that the largest challenge today is that the vast majority of genetic data is coming to providers as PDFs and faxes that sit on oncologists' desks. "It doesn't need to be that way," Strock said.[17]

Customizable EHR applications could help physicians create fields to categorize the data and make it searchable. Customizations, however, would deviate from EHR application standardizations and may cause interoperability issues.

5. Education for practitioners, physicians, and patients ongoing

Knowledge generated by genomics already exceeds the mental capacity of any given physician. Studies of primary care physicians show a poor understanding of genetics and dissatisfaction with support tools at the point of care.[18] The information contained within EHRs is pure data. If the information has not been reviewed and interpreted first, it has no practical use to a physician with little training in this area. When reviewing a patient's genomic information, care providers need to be able to clearly understand the story that the individual's genome tells. Health risks, family history, pre-existing conditions and genomic data can help the physician build a comprehensive care plan.

Genomic education remains a pressing need in clinical care. Genetic counselors can facilitate advances to ease interpretation. With fewer than 3,000 certified genetic counselors in the US, there aren't enough counselors to aid the interpretation of data. At the Genomic Medicine VIII meeting held in Rockville, MD this year, attendees talked about establishing a consensus within the genomic research community when interpreting genetic test results. Sharing experience and evidence will help in resolving discrepancies in interpretation.[19]

Work Has Begun to Solve Issues

Four studies have been launched within the Department of Veterans Affairs (VA) creating the nation's largest database linking genetic, clinical, lifestyle, and military exposure information involving more than 390,000 veterans. Information from this research will explore the role of genetics in obesity, diabetes, and abnormal lipid levels as drivers of heart disease.[20] As more and more data is being collected through studies like these, and as personalized medicine programs in local care centers

expand within the US, the challenge of how to use EHRs effectively and the standards for securing genomic data is an urgent issue.

In a recent survey conducted within the nation's hospital systems, the majority 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.[21]

As health IT professionals plan enhancements to their internal EHR applications, the key areas to focus on are how to safely store, smartly search, securely share, and accurately interpret this immense and vital information which will ultimately improve patient care and condense patient diagnosis times. These steps will require unprecedented cooperation between practitioners, healthcare providers, and state and local governments.

Notes

1. Reed, Tina. "Genomics 2.0: Putting Inova on the map." *Washington Business Journal*. February 28, 2014. www.bizjournals.com/washington/print-edition/2014/02/28/genomics-20-putting-inova-on-the-map.html.
2. Payne, Thomas H. et al. "Report of the AMIA EHR 2020 task force on the status and future direction of EHRs." *Journal of the American Medical Informatics Association*. May 29, 2015. http://jamia.oxfordjournals.org/content/early/2015/08/14/jamia.ocv066.
3. Haidle, Joy Larsen. "Family-Centered Clinical Genetics." Presented at the HL7 Genomics Policy Conference, Washington, DC, July 22, 2015.
4. Lee, Jaimy. "Seeking answers in a genetic code." *Modern Healthcare*. December 14, 2013. www.modernhealthcare.com/article/20131214/MAGAZINE/312149990.
5. O'Rourke, K. "Delays in Diagnosis of Celiac Disease Worry Experts." *Gastroenterology & Endoscopy News*. July 2015. www.gastroendonews.com/ViewArticle.aspx?d=In+the+News&d_id=187&i=July+2015&i_id=1207&a_id=32940.
6. Verel, D. "UCLA Health to integrate genomic data into EHR in pilot." MedCityNews. March 7, 2015. http://medcitynews.com/2015/03/ucla-health-integrate-genomic-data-ehr-pilot/.
7. Lee, Jaimy. "Seeking answers in a genetic code."
8. Masys, Daniel R. et al. "Technical desiderata for the integration of genomic data into Electronic Health Records." *Journal of Biomedical Informatics*. June 2012. www.j-biomed-inform.com/article/S1532-0464(11)00219-X/abstract.
9. "Genome data could outgrow YouTube, Twitter content by 2025 – report." RT. July 8, 2015. www.rt.com/news/272440-genome-data-youtube-twitter.
10. Green, Eric. D. and Mark S. Guyer. "Charting a course for genomic medicine from base pairs to bedside." *Nature*. February 09, 2011. www.nature.com/nature/journal/v470/n7333/full/nature09764.html.
11. Gebelhoff, Robert. "Sequencing the genome creates so much data we don't know what to do with it." *Washington Post*. July 7, 2015. www.washingtonpost.com/news/speaking-of-science/wp/2015/07/07/sequencing-the-genome-creates-so-much-data-we-dont-know-what-to-do-with-it/.
12. Shoenbill, Kimberly et al. "Genetic data and electronic health records: a discussion of ethical, logistical and technological considerations." *Journal of the American Medical Informatics Association*. January 1, 2014. http://jamia.oxfordjournals.org/content/21/1/171.
13. Ibid.
14. Telenti, Amalio, Erman Ayday, and Jean Pierre Hubaux. "On genomics, kin, and privacy." F1000 Research. March 31, 2014. http://f1000research.com/articles/3-80/.
15. McGuire, Amy. "Ethical and practical challenges of sharing data from genome-wide association studies: The eMERGE Consortium experience." *Genome Research*. June 1, 2011. http://genome.cshlp.org/content/21/7/1001.short#aff-2.
16. Raisaro, J. L., E. Ayday, and JP Hubaux. "Patient privacy in the genomic era." *Praxis*. May 7, 2014. www.ncbi.nlm.nih.gov/pubmed/24800770.
17. Raths, David. "DIGITizE: Displaying and Integrating Genetic Information Through the EHR." *Healthcare Informatics*. July 20, 2015. www.healthcare-informatics.com/article/digitize-displaying-and-integrating-genetic-information-through-ehr.
18. Nishimura, Adam A., Peter Tarczy-Hornoch, and Brian H. Shirts. "Pragmatic and Ethical Challenges of Incorporating the Genome into the Electronic Health Record." *Current Genetic Medicine Reports*. December 2014. http://link.springer.com/article/10.1007%2Fs40142-014-0051-9.

19. Jarvik, Gail, Heidi Rehm, and Dan Roden. “Consistency of Interpretation of Variants Across Expert Labs.” Presentation at Genomic Medicine VIII, June 8, 2015, Rockville, MD. www.genome.gov/Multimedia/Slides/GM8/Panel2.pdf.
20. Walsh, Beth. “VA studies to tap into Million Veteran database.” *Clinical Innovation + Technology*. July 9, 2015. www.clinical-innovation.com/topics/analytics-quality/va-studies-tap-million-veteran-database.
21. Shirts, Brian H. et al. “CSER and eMERGE: current and potential state of the display of genetic information in the electronic health record.” *Journal of the American Medical Informatics Association*. July 3, 2015. http://jamia.oxfordjournals.org/content/early/2015/07/02/jamia.ocv065.

Daniel A. DuBravec (ddubravec@lmi.org) is a senior consultant at LMI, a non-profit government consulting firm. DuBravec holds multiple EHR certifications and is a consultant for the government on EHR patient privacy and security standards.

Article citation:
DuBravec, Daniel A. "Can EHRs Handle Genomics Data?" *Journal of AHIMA* 86, no.11 (November 2015): 28-31.