Clinical Chemistry

Patient Privacy and Clinical Laboratory Data

Article:
Jason Y Park, et al.
Patient Privacy and Clinical Laboratory Data.
Clin Chem 2017;63:1321-25.
http://clinchem.aaccjnls.org/content/63/8/1321

Guest: Dr. Jason Park is an Associate Professor at the University of Texas Southwestern Medical Center in the Department of Pathology and the Eugene McDermott Center for Human Growth and Development. He is also the Medical Director of the Advanced Diagnostics Laboratory at the Children's Medical Center Dallas.

Bob Barrett: This is a podcast from *Clinical Chemistry,* sponsored by the Department of Laboratory Medicine at Boston Children's Hospital. I am Bob Barrett.

In the United States, the past decade has been a period of rapid expansion of electronic patient records. In 2008, only 9% of hospitals and 17% of physicians had electronic health records. As of 2015, 96% of hospitals and 78% of physicians used electronic health records. The goal in the era of electronic medical records is to achieve improvements in areas such as patient safety and operational efficiencies.

In the August 2017 issue of *Clinical Chemistry*, a Q&A article titled, "Patient Privacy and Clinical Laboratory Data" examines the current state of patient privacy with respect to data warehouses and health information exchanges. The moderator of this *Clinical Chemistry* article is Dr. Jason Park, Associate Professor at the University of Texas, Southwestern Medical Center in the Department of Pathology, and the Eugene McDermott Center for Human Growth and Development. He is also the Medical Director of the Advanced Diagnostics Laboratory at the Children's Medical Center in Dallas. Dr. Park is our guest in this podcast.

So doctor, first, what do electronic medical records and data warehouses have to do with the clinical laboratory?

Dr. Jason Park: Well Bob, electronic medical records are detailed clinical records that are generated, stored, and accessed on electronic media at the location of patient care. Electronic medical records are updated in real time and contain not only clinical laboratory data, but also information such as patient-specific demographics, pharmacy information, radiology, and physician notes.

Data warehouses in contrast, aggregate electronic medical records from multiple patients seen at multiple healthcare facilities. These virtual warehouses integrate not only clinical information, but also financial and operational data.

By integrating datasets into a single source, information can be searched and analyzed to improve clinical care, examine operational efficiencies, negotiate with insurers, and perform research. There are now numerous case studies which have demonstrated the value of clinical warehouses in improving the clinical care of patients, as well as increased operational efficiency of a health system.

As an example of the scale of the clinical data warehouse, in the recent *Clinical Chemistry* Q&A article, Dr. Toby Cornish from the University of Colorado School of Medicine described Data Compass, which is the name of the clinical data warehouse for the University of Colorado School of Medicine and UCHealth. Clinical laboratory data was added to this warehouse in 2016 and initially included over 369 million patient lab orders and results from UCHealth and over 124 million patient lab orders and results from Children's Hospital of Colorado. In addition, Data Compass is also planned to house genetic single nucleotide polymorphism snip results from the Colorado Center for Personalized Medicine's Biobank Project, when that data becomes available.

So using the University of Colorado's Data Compass as an example, we can see that millions of patients' medical records, including hundreds of millions of laboratory results, can reside in a single clinical data warehouse.

Bob Barrett: So how do health information exchanges relate to electronic medical records and clinical data warehouses?

Dr. Jason Park: Health information exchanges are the methods or services of electronically transferring medical data between legally separate entities. If two separate entities, say a reference laboratory and a hospital, are part of the same health information exchange, then a patient's data can be shared electronically. A typical scenario is a patient has outpatient laboratory testing performed at a reference laboratory and then becomes admitted to a hospital. With the health information exchange, the hospital personnel taking care of a patient can access the previously performed laboratory data and immediately initiate care. Not only can health care be delivered more quickly, but repeat testing can sometimes be avoided.

Some key features for a high quality health information exchange is to correctly identify the same patient in both entities sharing information, and then to provide fast and secure data transfer. In an ideal setting, a patient's medical records are accessible regardless of where they are seen for healthcare. This could include entities such as outpatient primary care, pharmacy, radiology, laboratory, and hospital.

Health information exchanges currently connect entities at the regional, statewide, and national level.

An ideal future scenario for health information exchanges is for a patient visiting any emergency room in the country to have their full electronic medical record immediately available to their healthcare providers.

In simplistic terms, the data warehouse aggregates the data for many patients seen in the same health system, whereas the health information exchange enables the sharing of data between health systems and healthcare providers.

Bob Barrett: But I imagine privacy concerns arise from this information sharing.

Dr. Jason Park: Absolutely. Although there is a clear clinical advantage to information sharing, there is also an increased privacy risk by having a patient's medical record accessible for multiple entities. As mentioned in the Q&A article by Dr. Michael Hogarth, "It is not possible to have zero risk, however, it is possible to keep the risk low, such benefits of sharing the data dramatically exceeds the risks."

Bob Barrett: Well, let's talk about that. Are there ways to decrease the privacy risks from this information sharing?

Dr. Jason Park: There are established legal protections for patient privacy that are based on the 1996 Health Information Portability and Accountability Act, better known as HIPAA. This 1996 legislation was followed by the Health Information Technology for Economic and Clinical Health Act, known as HITECH, in 2009. These laws were passed to encourage the adoption of electronic medical records by regulating the privacy of patients' medical information. HIPAA provides a foundation for the use and external sharing of clinical data. One of the key protections by HIPAA is that it defines what is considered de-identified patient information. Patient information which is de-identified according to HIPAA can mitigate the privacy risks and still be useful in clinical data warehouses.

However, de-identification also introduces new risks, since de-identified health information is no longer subject to requirements of HIPAA.

Bob Barrett: All right, well let's get this straight. Are you saying that the risk to privacy can be decreased by de-identification, but at the same time, de-identified data poses new risks?

Dr. Jason Park: Yes. Patient data which has been de-identified according to HIPAA still has information which is unique to a specific patient. The HIPAA privacy rule defined 18 specific identifiers, including the patient's name, address, telephone numbers, fax numbers, email addresses, Social Security number, medical record number, et cetera. Once these 18 identifiers have been removed from a dataset, then the information is considered de-identified and is no longer subject to HIPAA and most state privacy laws. However, by combining de-identified healthcare data with other datasets, it is possible to re-identify individuals.

This process of re-identification is not currently prohibited by federal law. There have been multiple studies that have demonstrated that using a combination publicly accessible data bases such as phonebooks and voter registration, can be used to re-identify patient information which was the de-identified according to HIPAA.

Indeed, a study from Vanderbilt University published in 2012 in the *Journal of The American Medical Informatics Association*, showed that in a model using 8.5 million laboratory results from 61,280 patients, any results from a complete blood count or chemistry seven panel was 99% unique among the 61,280 patients. Thus, by knowing the complete blood count for a patient at a certain date of testing, one could then re-identify that patient's medical record from a de-identified dataset of tens of thousands of patients.

Similar re-identification studies have been performed using genetic data. In general, data de-identified according to the requirements of HIPAA poses less of an immediate privacy risk. However, one should not complacent and consider the dataset to be without risk of re-identification. As some of the contributors to the Q&A article indicate, de-identification should be used whenever possible within a data warehouse. In addition, clinical laboratories in organizations should consider protecting de-identified datasets in the same manner they would protect fully identifiable patient data.

Additional protections include the use of agreements when sharing data to specifically exclude the re-identification of patients in the use of that data.

Bob Barrett: So are there any aspects of data warehouses which can decrease the risk to patient privacy?

Dr. Jason Park: Yes, there are certainly are aspects of data warehouses and health information exchanges which can decrease the risk to patient privacy. When we consider how data is stored and protected, there are some key advantages to the use of a centralized data warehouse. Although we may consider it risky to centralize data into a single physical structure which can be attacked by a hacker or be subject to natural disasters, the consolidation of data also creates an opportunity to consolidate resources to enhance security and protect patient data.

A data warehouse may be housed in a location that has hardened security measures to the physical structure of a building, or enhanced security for electronic access to data. In addition, consolidated data can have devoted resources for redundancy or backup in the event of a natural disaster, fire, or hardware failure.

In the absence of data warehouses or health information exchanges, the default is to store data on portable electronic media or paper, and transfer that information between healthcare providers by physical mail. It is important to consider the security of patient data within the context of a centralized data warehouse versus the security of traditional patient records distributed among thousands of healthcare providers.

Bob Barrett: Well finally Dr. Park, any final words of advice for laboratory professionals regarding the privacy of patient laboratory data?

Dr. Jason Park: I would like to end with a quote from Tim Berners-Lee, the inventor of the worldwide web. While describing how customers need to be in control of their data, he said: “Data is a precious thing and will last longer than the systems themselves.” As laboratory professionals, we need to be aware of the permanence of patient data and ensure protection of patient laboratory data not only when it is used for current clinical needs, but also when it is sent to the patient’s electronic medical record, stored within data warehouses, and shared with healthcare partners across health information exchanges.

Bob Barrett: Dr. Jason Park is an Associate Professor at the University of Texas Southwestern Medical Center in the Department of Pathology and the Eugene McDermott Center for Human Growth and Development. He is also the Medical Director of the Advanced Diagnostics Laboratory at the Children’s Medical Center Dallas. He’s been our guest in this podcast from *Clinical Chemistry*.

I’m Bob Barrett. Thanks for listening!