COVID-19 Evidence Accelerator Collaborative

Diagnostics Evidence Accelerator #36

Thursday, *September 23, 2021, 12-1 PM ET*

Call Summary

Introduction to Diagnostics Evidence Accelerator Meeting #36

This week's Diagnostics Evidence Accelerator meeting consisted of 4 presentations:

1. Opening Remarks (Dr. Sara Brenner, Center for Devices and Radiological Health (CDRH), FDA)
2. Where Are We in The Data Flow (Gina Valo, FDA)
3. Quest Diagnostics Data Resources (Dr. William Meyer, Quest Diagnostics)
4. Q&A with Chris Stanley and Bill St. Pierre, Labcorp and Dr. William Meyer, Quest Diagnostics

As always, thank you to all of the analytic partners, strategic advisors, and scientific advisors that are participating in Project One of Diagnostics Evidence Accelerator.

Opening Remarks (Dr. Sara Brenner, Center for Devices and Radiological Health (CDRH), FDA)

CDRH is working across the spectrum to include diagnostic data quality, completeness, and utility by working closely with stakeholders across the data ecosystem. In order to guide an evidence based pandemic response at a local, state, and national level, capturing diagnostic data from COVID-19 is important. However, a lack of standards, interoperability, and connectivity by public health IT has led to fragmented, delayed and incomplete testing data which makes it difficult for the nation to fight the pandemic. Efforts that address data standards, IT infrastructure, and emerging technologies in policy are necessary to improve data quality upstream and its utility downstream, therefore the efforts taken to improve availability for post market or real-world evidence (RWE) and real-world data (RWD) will help FDA and other regulated industry evaluate the real-world performance of emergency use authorized (EUA) COVID-19 tests. It also helps the FDA better identify potential issues with tests that are under review for pre-market authorization and proactively corrected.

CDRH is also expanding the SHIELD program by leveraging high quality harmonized, post-market data to support and supplement the review and authorization of new diagnostic tests. By gathering RWE developers and regulators are able to both expedite free market review and the real-world performance of tests. Additionally, CDRH is working with working groups across agencies that focus on data collection, public health data systems, and innovation and public health data analytics which was highlighted under recent executive orders from the President focused on nations and state needs during the pandemic and beyond the pandemic in terms of preparation.

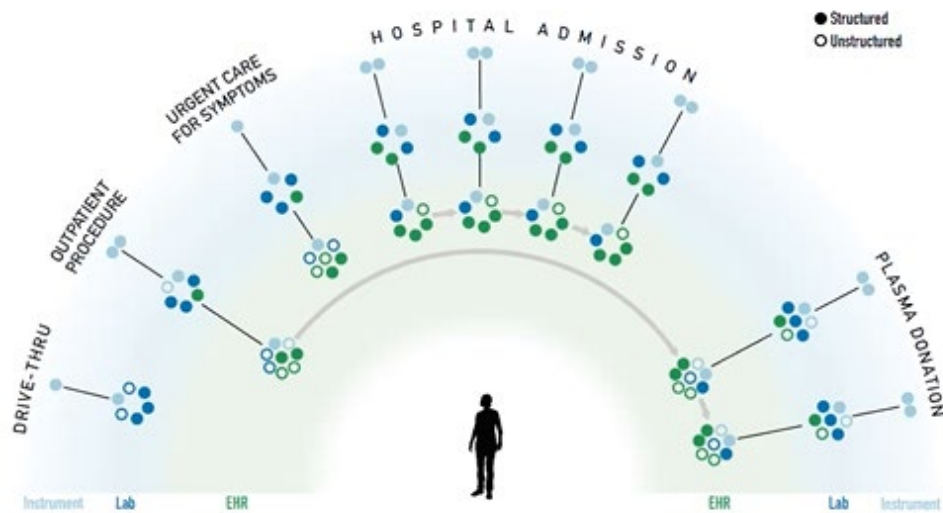## Where Are We in The Data Flow (Gina Valo, FDA)

Figure 1: Heidi the hypothetical patient.

The Heidi story (Figure 1) was a story of connecting lab data on many levels. In this space, we have been asking the question of how do we connect testing data to a patient, specifically asking what we knew about the patient before they took the test, what we know about the patient after they took the test and what we know about the test itself. This is important because it allows us to compare performance and patterns (i.e., regional patterns) across each test and pair this with patient-level data. Therefore, there is a need for collaboration across all stakeholders in this ecosystem, such as commercial labs.

Quest Diagnostics Data Resources (Dr. William Meyer, Quest Diagnostics)

Quest Diagnostics has an informatics system (Big Data Lake) that stores testing data after it is completed. The data lake consists of 10+ years of longitudinal lab data and 59+ billion test results. Yearly, Quest Diagnostics encounters 140 million unique patients resulting in approximately 4 billion results per year. They are serving approximately 50% of the physicians and hospitals in the US, have 470,000+ physicians connected to their Quanum portal, and have 675+ EMR interfaces.

The "Big Data Lake" is the primary storage that is used for data licensing. Figure 2 shows what data that is stored and not stored in data lake. The elements highlighted in yellow must be provided by the clients. Additionally, Quest Diagnostics Internal Unique Patient Identifier allows for retrospective and prospective tracking of individual patients. This data can be tokenized (de-identified) to allow for tracking across the system. There is a possibility to connect the streams that connect the Big Data Lake for licensing and the data that is not stored in the Big Data Lakes, however, this process is not an automatic connection. If there was a question regarding data in the Big Data Lake, then Quest Diagnostics can go into the other database to evaluate the question.

Figure 2: Data fields that are collected and stored in the "Big Data Lake".

Figure 3 shows the strengths and weaknesses that Quest Diagnostics has seen in the Informatic Big Data Lake.

Figure 3: Strengths and weaknesses that Quest Diagnostics has seen in the Informatic Big Data Lake.

The data are used by multiple stakeholders and reported to public health agencies electronically, however, each state public health agency has different requirements that must be taken into consideration when reporting. Additionally, researcher, health plans, health systems, and pharmaceutical firms and biotech firms use the clinical laboratory data.

Questions and Answers:

- Is there a possibility to connect the Big Data Lake for licensing and the data that are not stored in the Big Data Lakes?
  - There is a possibility to connect the streams that connect the Big Data Lake for licensing and the data that are not stored in the Big Data Lakes, however, this process is not an automatic connection. If there was a question regarding data in the Big Data Lake, then Quest Diagnostics can go into the other database to evaluate the question.
- Why are data such as manufacturer information not stored in the data lake? What would cause you to want to merge these?
  - The reason why the data that are on the right side is not stored in the data lake is because those are data that researchers or clients are typically not interested in, however, the data are stored on another database so that it is assessible.
- How is the information that goes to a clinician similar or different from the data in the Big Data Lake?
  - The information that goes to the clinician is medically significant information such as the test name, test results, and interpretation of results. These results have LOINC codes attached to it. However, all of the information that is stored in the Big Data Lake will be submitted to the clinician.
- When discussing the challenge of combining with data from other clinical laboratories, is this across Quest Diagnostics or different labs such as Labcorp.
  - This is combining with different labs, i.e., combining data generated by Quest labs and labs operated by other organizations. Different labs may use different LOINC codes for the same test. Connecting the data is a complex process.
- If a customer or agency asks to share certain data elements that are not part of the Big Data Lake, then what will it take to be able to share or add the element?
  - This will depend on the request and each request will have to be handled on its own. It is not possible to take the data elements and move them around. The technicality and effort of the addition of that request will have to be evaluated.

Q&A with Chris Stanley and Bill St. Pierre, Labcorp and Dr. William Meyer, Quest Diagnostics
Interviewed by Dr. Carla Rodriguez-Watson

Labcorp has a similar set and challenges up as Quest Diagnostics. The clinical and patient data are provided to the clinician, patient, and interested payor which is similar to Quest Diagnostic. The data that are not provided is manufacturer data; clients do not request that data. However, the data are stored in the lab information system (LIS).

Additionally, Labcorp has different capabilities to address the needs of different consumers such as public health agencies.

What is the "low hanging fruit" so that we can advance the interpretation or standardize these data across multiple labs?

Labcorp: Labcorp has a standardized platform across all of the labs where the same test codes, workstreams, and equipment is used. Each equipment, test, or reagent will have different standards and reference ranges, therefore aligning on one aspect such as LOINC codes will not be helpful in the standardizing process. Therefore, standardizing on the methodology, equipment, and codes may help, but there are other factors that need to be addressed too.

Quest: Additionally, a team is needed in order to interpret the results since there are many complexities in analyzing the data.

Is there a strategy to look at each test in order to standardize the test within the lab system and then across lab systems? One approach could be to take a high-volume test and standardize it across lab systems.

Labcorp: This will be a tedious task because there is a lot of data elements that will need to be standardized (i.e. discipline, procedure class, instruments). Labcorp has standardized across their system; however, it becomes more complex when other labs such as commercial and hospital labs are connected. Taking a high-volume test and standardizing it across lab systems could be a potential way to begin the standardizing process.

Standardizing on equipment allows for aligning based on equipment and methodology (almost a shortcut) - another alternative would be to standardize the requirements for interoperability that gets implemented on all equipment. This will not solve the test-to-test problem but would solve the connecting and cleaning problem. This would lead to standards for interoperability.

Labcorp: This is correct. There are different standards for each equipment and test interfaces. The HL7 standards is what Labcorp uses to send the data. However, every EHRs has a custom system that is made specifically for that EHR. This causes additional challenges.

On the Evidence Accelerator, we have been discussing the challenges for reporting data for reportable disease. There are a number of data fields that do not get reported, therefore what are the drivers to get reportable data (e.g. race and ethnicity and manufacturer) to flow?

Labcorp: The challenge for getting this type of data to flow is that the client may not be providing that specific data to the labs. The labs will have to receive the data in order to report it, however if it is a required field, then the field will be created in order to report it.

Next Steps

- Continue making data connections through the Evidence Accelerator and through www.EvidenceAccelerator.org.

Next Meeting: Thursday, October 7, 2021 12-1 pm ET