Fast and Efficient Data Utilization: The Benefits of Physician Access to Health and Laboratory System Databases

Parsa Hodjat, MD[1], S. Wesley Long, MD, PhD[1], Paul Christensen, MD[1]

[1]Departments of Pathology and Genomic Medicine, Houston Methodist Hospital, Houston, TX

Background

Massive volumes of data are recorded in Health (EHR) and Laboratory Information Systems (LIS) that can be effectively used for improving patient management. Unnecessary roadblocks in data acquisition can delay providing actionable information to questions raised by clinicians and hospital leadership. During the COVID-19 pandemic, we leveraged the direct LIS database access of our MD informaticists to provide quick and accurate answers, allowing us to keep up with the fast-paced evolution of institutional strategic planning. Three examples of such inquiries were:

- Possible increased rate of blood culture positivity in COVID-19 patients
- Change in incidence of rhinovirus and endemic coronaviruses during the COVID-19 pandemic
- Perceived increase in HSV-1 positivity in respiratory specimens of hospitalized patients

Methods

We created ad-hoc SQL queries to retrieve the necessary data elements from our LIS relational database (Oracle® DBMS). When necessary, such as for COVID-19 PCR test results, these queries were embedded in a Python script and added to a Linux cron schedule for automated updates. The query results were saved as comma-separated values (.csv) or Excel (.xlsx) files. Excel Power Query was used to join tables and process the data. Excel Pivot Tables and Pivot Charts were used to analyze the data.

Results

In all instances, the time spent on designing and executing the SQL queries was less than 5 hours. The data processing (configuring table joins, cleaning the data) and analysis (creating charts and summary tables) took 2.5 hours on average. In each of the instances we had the complete analysis within 8 hours.

Conclusions

Our experience with direct database access to the LIS database, with appropriate data governance, shows how physician training and appropriate access shorten the amount of time necessary to acquire actionable information. When relying on others for a data pull, there is risk of receiving suboptimal data in the initial queries and the need for multiple course corrections to achieve the best results, which further delays receiving the necessary information. We have found Excel Power Query and Excel Pivot Tables to be optimal tools for rapid, intermediate-level data processing, data exploration and data visualization.
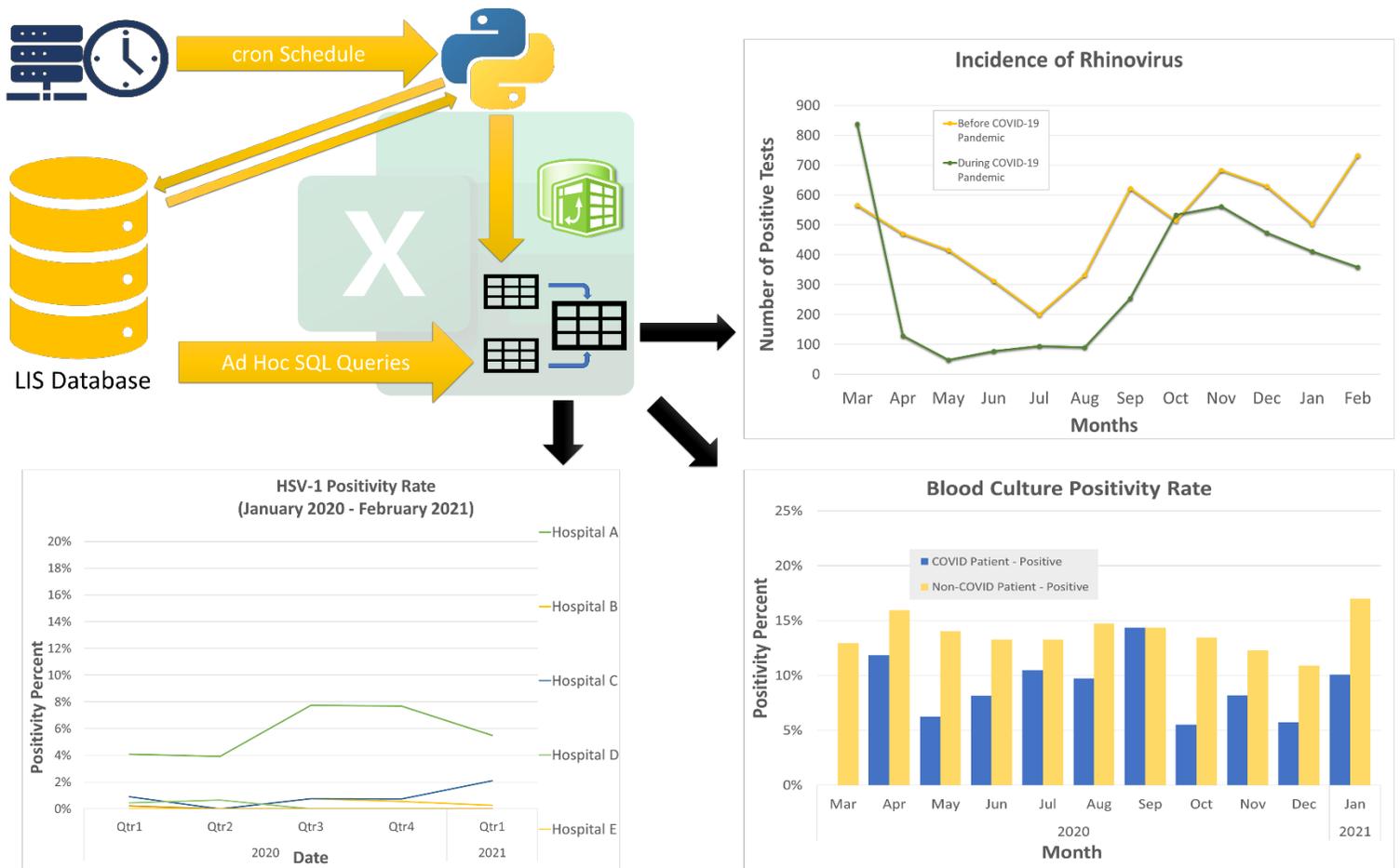
Figure 1. Fast and efficient report generation and data analysis with clinical informaticist's direct access to LIS database