

Editors: Raymond D. Aller, MD, & Dennis Winsten

Leveraging lab analytics: from requesting data to implementing tools

January 2024—When the medical microbiology laboratory at Yale-New Haven Hospital makes operational changes, it uses data analytics to monitor their impact. Yet the process of implementing laboratory analytics can be challenging.

“I’ve come across bad data sets, incomplete data, and situations where a simple data query is not going to be adequate for addressing the question at hand,” says David Peaper, MD, PhD, director of Yale-New Haven’s microbiology lab and associate professor of laboratory medicine at Yale School of Medicine.

As a microbiologist without formal training in data science, Dr. Peaper developed and honed his technical skills and techniques for laboratory data analysis on the job while creating and implementing data-driven tools for assessing lab operations. The lessons he learned in the process, sometimes by trial and error, have provided insight into Yale-New Haven’s laboratory operations.

“We are subject matter experts in our disciplines, so we need to develop data-extraction and -analysis skills because we are best suited to have at least this basic understanding of what is going on in our laboratory,” Dr. Peaper said during a presentation about creating a laboratory analytics program, conducted at the 2023 annual meeting of the Association for Diagnostics and Laboratory Medicine.

During the planning phase of such undertakings, the laboratory should develop a structured way to ask data analysts for information that will serve as the basis of the project, Dr. Peaper explains.



Dr. Peaper

Using turnaround time as an example, he says the laboratory must determine what aspect or portion of TAT it wants to analyze using this common metric for evaluating lab efficiency. Turnaround time could refer to the preanalytical part of the process, stretching from when a test is ordered until it is collected. It could also refer to the time from when a specimen is collected until it is received in the lab or the time from specimen collection until results release. Therefore, the laboratory should include in its data request the time stamps that define the portion of the turnaround time process that it will analyze and, more importantly, why it wants to analyze that portion of the TAT process, Dr. Peaper says.

“Maybe you are really interested in the time from specimen receipt until Gram stain for the pediatric emergency department or the pediatric ICU,” he adds. “That is another level of detail that you need to convey.”

Sharing details is important in part because most data analysts do not have deep technical knowledge of lab practices, Dr. Peaper says. Therefore, they may not know, for instance, that microbiology laboratories often receive a specimen, stain it, and produce an initial result within an hour but that it may take anywhere from a day to a few weeks to receive additional test results. Consequently, “the more specific and detailed your request for information can be, the more likely you are to get useful data earlier.”

Sometimes laboratories need data on patients who are located in a particular building or floor in the hospital complex, he notes. If the subunit of pediatric patients that the laboratory is studying is housed in North Pavilion 12, for example, providing that specific location could help data analysts verify that they are retrieving the desired information.

"There is a tendency to say, 'Just give me everything and I'll sort it out,'" Dr. Peaper says. But that approach is not only less efficient but less safe, he adds. Most laboratory data contain protected health information, and receiving a larger-than-needed data file could potentially put a greater amount of patient information at risk of being compromised in the event of a cyberattack.

Once the lab receives the data it requested, Dr. Peaper recommends double- and triple-checking it for accuracy. "No one knows your lab data like you," he says.

In reviewing data files obtained for analytics purposes, Dr. Peaper has found data that are missing critical elements, such as birth dates, and uncovered proficiency test data integrated with regular test results. He has also seen data with impossible values, such as results for a 30-minute test that have time stamps suggesting the test took far shorter or far longer than the expected time.

Depending on the intended use for the data, Dr. Peaper may exclude contaminating data or data with impossible values from further analysis. He recommends carefully documenting which data are being excluded, in the same way data parameters are defined in academic research papers.

Whether the data are to be presented "to my laboratory team, or my department chair, or external stakeholders, the more formal the process you use, the better," he says.

Dr. Peaper initially used tools embedded in Microsoft Excel for analyzing data extracts, which, he says, works well for ad hoc analysis.

Comma-separated values files or Excel files can be dropped into Excel quickly, and the program's pivot tables make it easy for users to sort, group, and rearrange data for analysis. However, Excel has limitations, he notes. For example, the program is capped at 1,048,576 rows of data, and the workflow processes for working with data are manual.

One project that Dr. Peaper says "bumped up against the limits of what is reasonably accomplished in Microsoft Excel" involved assessing the clinical and laboratory impact of a new polymerase chain-reaction testing protocol for the flu. The laborious project linked laboratory report data with clinical emergency department report data via encounter IDs to evaluate the impact of the testing protocol on time to PCR result, length of emergency department stay, and time to Tamiflu administration. The project demonstrated that the new protocol resulted in faster PCR result TAT, shortened emergency department stays, and a shorter time to Tamiflu prescription.

"We're talking about linking multiple events, including clinical parameters, in a particular encounter across time, and there is a tremendous opportunity for error," Dr. Peaper says, in explaining the technical challenges of the project.

To address the difficulties of working with multiple sets of data and increasingly sophisticated rules, Dr. Peaper learned Python. This was a worthwhile undertaking, he says, because "when you start trying to combine data or filter data, there is a lot of opportunity to corrupt your data. The more you do things manually, the more likely this is to happen."

With Python, rules, such as for exclusion criteria, are written in code, which not only automates processes but makes workflows easier to adjust. Therefore, if a laboratory decides to change its exclusion criteria for a project, anyone with knowledge of Python can alter that part of the code related to exclusion criteria. Making that type of change in Excel would be tantamount to starting over, Dr. Peaper says.

While Dr. Peaper acknowledges that the learning curve can be long for determining how to logically ask questions

of data for coding purposes, he considers Python to be a valuable tool for pathology laboratories.

“By combining our subject matter expertise as laboratorians with being able to extract and analyze data, we breed collaboration with our colleagues and our institutions,” Dr. Peaper concludes. “These collaborations, and the insights that lab data provides, can help pathologists demonstrate their value to their institutions.”

—*Renee Caruthers*

Secondary uses of data focus of Joint Commission certification program

The Joint Commission has introduced the voluntary Responsible Use of Health Data certification program for U.S. hospitals and health care organizations.

The program will provide guidance on how to safely use data for purposes other than clinical care, such as for improving quality and operations, developing algorithms, and advancing artificial intelligence. It will also recognize health care organizations that establish policies and procedures to protect health record data.

“The certification will provide an objective evaluation as to whether an organization is committed to utilizing best practices in its secondary use of data and promoting responsible use of data,” according to a press release from the nonprofit standards-setting and accreditation organization.

RUHD certification is based on principles set forth by the Health Evolution Forum, a collaboration of medical industry CEOs, other senior executives, and thought leaders interested in making voluntary, industry-led improvements to the health care marketplace. RUHD certification requirements address the data de-identification process, data controls, limitations on data use, algorithm validation, patient-focused transparency about data use practices, and data oversight structure.

“As more health care organizations are leveraging clinical data for secondary purposes, there have been increased calls to assure responsible data stewardship,” said Jonathan Perlin, MD, PhD, Joint Commission Enterprise president and CEO, in a press statement. “We believe our Responsible Use of Health Data certification will help health care organizations use data responsibly to improve the safety, quality, and equity of care; develop new technologies; and discover new therapies benefitting all patients.”

Hospitals that achieve RUHD certification will be recognized publicly for establishing an objective and rigorous process for meeting the necessary requirements to protect health record data.

Paige and Nuance collaborate on pathology consultation network

Paige and the Microsoft subsidiary Nuance have teamed up to develop the largest digital consultation network in pathology, according to the companies.

Through their collaboration, Paige will embed its FullFocus digital slide viewer and FullFolio artificial intelligence-powered image manager in Nuance’s PowerShare image-sharing network, which is used by more than 14,000 facilities to share medical images.

Merging Paige’s digital pathology tools with the PowerShare network allows laboratories with any type of digital capability to obtain expert pathology consultations from top-tier hospitals and health care systems, the companies jointly reported.

[Nuance](#), 800-654-1187

NYU Langone Health to adopt Philips technologies

Amsterdam-based Royal Philips and NYU Langone Health have announced an eight-year strategic partnership that will give the New York-based academic medical center access to numerous Philips technologies that support real-time patient-care collaboration.

Under the deal, NYU Langone will adopt Philips' IntelliSite digital pathology system, enterprise informatics solutions, and artificial intelligence-enabled diagnostic imaging technologies, such as the Philips Lumify handheld ultrasound. NYU Langone will also use the Philips Patient Information Center iX central monitoring system and Philips Capsule Medical Device Information Platform to simplify the medical center's workflows.

"MDIP is designed to bring together disparate data from medical devices and systems into a single unified view," according to a Philips press release. "When deployed with Philips PIC iX in a predictable enterprise monitoring-as-a-service subscription model, the combined solutions are intended to provide seamless integration across the network."

[Royal Philips](#), 800-229-6417

Sonic Healthcare acquires PathologyWatch

The Sydney, Australia-based anatomic pathology provider Sonic Healthcare has signed binding agreements to acquire the Murray, Utah-based digital dermatopathology solutions provider PathologyWatch, which connects physician offices with dermatopathologists worldwide for case consultations and related interactions.

"The PathologyWatch acquisition is attractive to Sonic as it creates multiple return opportunities" by accelerating Sonic's transition to digital pathology, providing the company with a competitive advantage in the dermatopathology marketplace, and capitalizing on significant synergies with Franklin.ai, Sonic's artificial intelligence-platform partner," according to a Sonic press release.

The founders, management staff, and other employees of PathologyWatch will join Sonic's U.S. and global teams.

Sonic is uniquely positioned to extend PathologyWatch's reach into countries in which the anatomic pathology provider has a significant presence, including Australia, Switzerland, and Germany, according to a press statement from Pathology-Watch.

The acquisition was expected to close at CAP TODAY press time.

[PathologyWatch](#), 877-463-5434

Proscia to digitize glass slides for CellNetix

The independent pathology service provider CellNetix Pathology and Laboratories, Tukwila, Wash., has signed an agreement for Proscia to digitize a portion of the approximately 1 million glass slides it generates annually.

"In modernizing its practice with Proscia, CellNetix will achieve greater flexibility along with productivity and efficiency gains in its routine workflows," according to a press statement from Proscia. "It will also streamline collaboration among the 60-plus pathologists distributed across its regional network."

[Proscia](#), 215-608-5411

Tribun Health to install digital pathology platform at Spire Healthcare

Paris-based Tribun Health has contracted for its digital pathology platform with Spire Healthcare, a United Kingdom-based independent health care group.

As part of the partnership, Tribun will install its CaloPix image-management system, CaloPix storage and archive-management functionality, and forthcoming artificial intelligence applications at Spire Healthcare hospitals and clinics across England, Wales, and Scotland.

Corista pathology solution gets CE mark under IVDR

Concord, Mass.-based Corista has reported that its DP3 image-management system for digital pathology is now CE marked under the European In Vitro Diagnostic Medical Devices Regulation for use in routine diagnosis in the European Union and United Kingdom. The company also announced a strategic partnership with Mortsel, Belgium-based Agfa HealthCare.

"We are extremely pleased to extend our services into the EU, and obtaining this CE mark is a crucial milestone toward our delivery of integrated pathology for health care providers across the EU," said Corista CEO and cofounder Elizabeth Wingard.

Corista will integrate its DP3 technology into Agfa HealthCare's Enterprise Imaging for Pathology platform, which is available in the European Union and United Kingdom.

[Corista](#), 978-287-6188

Dr. Aller practices clinical informatics in Southern California. He can be reached at rayaller@gmail.com. Dennis Winsten is founder of Dennis Winsten & Associates, Healthcare Systems Consultants. He can be reached at dennis.winsten@gmail.com.