Decision Support to Enhance Automated Laboratory Testing by Leveraging Analytical Capabilities

J. Mark Tuthill, MD

INTRODUCTION

Business analytics, clinical decision support, machine learning, artificial intelligence. How do these buzz words impact running a modern laboratory? The answer is, profoundly. Without analytical capabilities, a laboratory is effectively deaf, dumb, and blind. It is through the use of data analytics that the modern laboratory is most effectively and efficiently managed. This includes impact on daily operations and laboratory performance, and more sophisticated opportunities to predict in advance not only health system business needs and laboratory performance requirements, but new opportunities to impact patient care in novel ways. As computational power continues to increase, and software is developed to analyze and synthesize data into information, knowledge is created, and ultimately a higher order of information, wisdom. Although this may be hyperbole to a degree, what is clear is that laboratories that leverage the data they create on a daily basis to change the way they work and support clinical

Disclosure: The author has nothing to disclose.

Henry Ford Health System, 2799 W. Grand Boulevard, K-6 Pathology, Detroit, MI 48202, USA

E-mail address: Mtuthil1@hfhs.org

KEYWORDS

- Business analytics
- Clinical decision support
- Laboratory automation
- Dashboards
- Artificial intelligence
- Learning health systems

KEY POINTS

- Understand data science and the four types of data analytics.
- Recognize that laboratory clinical decision support leverages several types of data analytics to support business processes.
- Describe the practical tools that support laboratory clinical decision support, such as real-time dashboards, autovalidation, and reflex testing.
- Visualize future possibilities for clinical decision support leveraging new developments in artificial intelligence leading to learning systems.
practitioners are more efficient, successful, and more deeply valued by hospital leadership.

Henry Ford Health System (HFHS) Pathology and Laboratory Medicine (PALM) has long used data generated from routine testing to support not only decisions about laboratory operations but to communicate to clinical departments the value and impact of laboratory testing on their patient services. Because hospitals are tightly connected departmental ecosystems, it is well recognized that laboratory testing is key to effective patient care. Whether the impact of turnaround time (TAT) on emergency room throughput or infectious disease reporting to monitor nosocomial infection rates, nearly all modern laboratories provide analytical information that impacts patient care. In fact, some of this reporting is required by government and regulatory agencies. Good examples of this are requirements for reporting cancer; infectious disease rates for certain organisms; and analytes, such as lead levels.

These primary analytical reporting requirements have continued to grow over time and can consume a significant amount of time and effort by laboratories, but are actually a limited starting point compared with future opportunities. Imagine, rather than reacting to nosocomial infection outbreaks after the fact, that analytics could predict such outbreaks proactively or least provide real-time alerts to an increasing trend. Further imagine that panels of laboratory tests could be analyzed using artificial intelligence algorithms to predict other values that were not directly tested. It is beginning to be seen that such algorithms are not only plausible, but possible and clinically effective. Such algorithms save time and money by eliminating expensive reference laboratory testing.

The application of laboratory data analytics will be far reaching as more sophisticated tools become available and easier to use. Some of these tools allow integration of data from disparate systems to be leveraged in unique ways. Rather than just looking at daily laboratory operations, such analytical paradigms allow for laboratory data to be used to look at provider utilization and performance, patient outcomes, care modeling, and financial forecasting to mention a few. This is the next generation of laboratory analytics.

This article addresses the historical, current, and future state of laboratory analytics using examples, and offering a framework to organize thinking around analytical capabilities. This helps the reader understand where they are in the continuum of data analysis, suggesting steps toward the future.

UNDERSTANDING AND ORGANIZING ANALYTICS CONCEPTS: DATA SCIENCE

Business analytics or data analytics has become increasingly formalized. This has resulted in professionals with specialization in data science at the undergraduate and post-graduate level. As this field develops, the lexicon of data analytics has become more formalized using recognized nomenclature that should be understood by laboratorians specifically and health care professionals in general. The primary driver for the growth of data science has been the increase in computing power and data storage, and the development of sophisticated vendor solutions and open source technology. This includes leveraging cloud-based solutions, Internet technology, robust databases, and data warehouses. This technology, in conjunction with increasing pressure on laboratories to address costs and efficiency in the current regulatory environment, has created a desire for laboratories to better understand their work, and to improve their workflow. Whereas clinical decision support (CDS) in the clinical arena typically focuses on efficacious use of laboratory services by health care practitioners, within the laboratory such support systems are directed at preanalytical, analytical, and postanalytical aspects of diagnostic testing.
Data science has recognized four basic types of analytics\(^3\):

- **Descriptive analytics**: What happened?
- **Diagnostic analytics**: Why it happened?
- **Predictive analytics**: What will happen?
- **Prescriptive analytics**: How can we make it happen?

Each of these types of analytics has increasing sophistication in technical requirements and impact. Higher order analytics will lead to the deployment of applications that leverage artificial intelligence and machine learning culminating in “learning systems” that are the outcome of effectively deployed prescriptive analytics. To date most laboratory CDS is driven by descriptive analytics in combination with diagnostic analytics, but predictive and prescriptive analytics are beginning to be developed.

**CLINICAL DECISION SUPPORT IN THE AUTOMATED LABORATORY: OPERATIONAL IMPACTS**

Leveraging data analytics in the laboratory provides the tools that support operations and best practices (Table 1). This is different than how one views CDS in the clinical arena. Looking at such basic data, such as laboratory testing, TAT can provide an end-to-end example. When one looks at TAT at the level of descriptive analytics, one uses the raw data in combination with expected best practices to determine whether the desired outcome was met. That is, what happened? If the desired TAT metric was not achieved, diagnostic analytics are used to determine why the goal failed to be achieved. For example, preanalytical factors, such as sample integrity (eg, hemolysis or a clotted sample) or transport time, can be identified. Using predictive analytics, one may model how changes can improve the process. Finally, in a finely tuned system leveraging artificial intelligence and machine learning such an analytical system could suggest the answer. For example, a recommendation to add a new courier route at a given time of day. Perhaps even automatically identifying a practitioner with a higher number of preanalytical defects allowing just in time education to be provided that is directly related to the type of defect experienced.

In the preceding simple example, the operational requirement is to capture the appropriate data elements that can be processed by a pipeline leading from raw description to sophisticated responses and interventions, that is, prescription. To date this has been challenging for laboratories. Although most laboratories leverage descriptive analytics in their CDS processes, the data leveraged are typically static and historic. That is, one is assessing and reacting to yesterday’s information today. With such design approaches one can leverage diagnostic analytics to demonstrate

<table>
<thead>
<tr>
<th>CDS System</th>
<th>Data Type</th>
<th>Analytical Type</th>
</tr>
</thead>
<tbody>
<tr>
<td>TAT dashboard</td>
<td>Static</td>
<td>Descriptive</td>
</tr>
<tr>
<td>TAT dashboard</td>
<td>Dynamic</td>
<td>Descriptive, diagnostic</td>
</tr>
<tr>
<td>Volume dashboard</td>
<td>Static</td>
<td>Descriptive</td>
</tr>
<tr>
<td>Volume dashboard</td>
<td>Dynamic</td>
<td>Descriptive, diagnostic, predictive, prescriptive</td>
</tr>
<tr>
<td>Critical value call back</td>
<td>Static or dynamic</td>
<td>Descriptive, diagnostic</td>
</tr>
<tr>
<td>Reflex testing</td>
<td>Dynamic</td>
<td>Descriptive, diagnostic, prescriptive</td>
</tr>
<tr>
<td>Autoverification</td>
<td>Dynamic</td>
<td>Descriptive, diagnostic, predictive, prescriptive</td>
</tr>
<tr>
<td>Learning systems</td>
<td>Dynamic</td>
<td>Descriptive, diagnostic, predictive, prescriptive</td>
</tr>
</tbody>
</table>
a transport delay or clotted specimen, and may even be able to model how changes to
the system could be implemented, but there is little hope of a dynamic, prescriptive
intervention, never mind a learning system. For this to occur, data from the automated
laboratory must be continuously analyzed in conjunction with information from other
systems such as inbound clinical orders, data from the LIS (Laboratory Information
System), instruments, and even the automation line control software. As a result of
the reliance on static, historic data, few laboratories have achieved an end-to-end an-
alytics pipeline. However, as practitioners become more mature and sophisticated in
applying data analytics, the beginnings of such a pipeline are able to be conceptual-
ized. A key constraint to this work is not only the ability to capture the required data
points, but also the human resources to focus on these efforts.

REFLEX TESTING AND TESTING CASCADES

Although the TAT example presented is a common day-to-day CDS monitor, one can
imagine more sophisticated applications of CDS within and outside the laboratory.
These include such behaviors as reflex testing, leading to testing cascades where
the results of one test may lead automatically to the performance of another. The
classic example is the reflex ordering of a ferritin in the case of microcytic anemia
to rule in iron deficiency. In the automated core laboratory, this may result in a sample
being automatically retrieved from a storage stockyard for additional testing, versus
the manual approach. This is a real, albeit simple, example of laboratory CDS.
Leveraging the concepts of data science, it is not difficult to imagine more sophisti-
cated examples, such as alerting a clinician that the antibiotic resistance of a bacterial
isolate is not congruent with a patient’s current medication. A learning system may
even suggest the most appropriate therapy for a given patient taking into account mul-
tiple variables, such as kidney function, allergy status, and the patient’s problem list.

One can imagine many such scenarios, but most still fall into descriptive and diag-
nostic analytics, with predictive analytics and modeling being carried out offline by a
human. As more data are gathered from different areas of the patient care environ-
ment, analyzed and learned from, the game will change. Imagine, for example, that
the correct initial antibiotic could be predicted based on multiple historical data points
with a statistical degree of confidence. Although this is the purpose for publishing the
antibiotic resistance profile for organism in an environment, such information is
descriptive or a best diagnostic. Only when multiple factors are taken into account
can such models become increasingly predictive and prescriptive.

THE USE OF ANALYTICS AND CLINICAL DECISION SUPPORT BY HENRY FORD
HEALTH SYSTEM PATHOLOGY AND LABORATORY MEDICINE

Static Data and Dashboards

Historically, PALM has long used data to help make clinical decisions around labora-
tory testing, defect elimination, and process improvement. Leveraging static TAT and
defect reports the laboratory has been able to work with clinical customers to support
their needs. One early example was monitoring TAT for markers of cardiac damage for
samples coming from the emergency room; previously creatine kinase-myocardial
band, and currently cardiac troponin. It is well recognized that TAT of these tests
directly impacts emergency room wait time and throughput. TAT metrics were typi-
cally produced as daily reports showing the prior day’s success or failure in meeting
a TAT threshold for a given analyte. These reports were produced by extracting the
prior day’s laboratory data and descriptively analyzing them for the outcome: pass
or fail. If there was a failure to meet the desired TAT threshold, additional data could
be analyzed to determine the reason for failure. Often times this required additional information to be generated or queries into other systems. Because these data were static and a day old, no direct intervention was possible at the time of failure. More recently this information is presented on more refined paper dashboards distributed by email, but effectively it is still historical and static.

**Autoverification**

The use of autoverification or autovalidation is one of the earliest and most impactful applications of CDS in the clinical laboratory. It is not well recognized by most outside the laboratory what is required to release results downstream. Basic result reporting requirements, such as valid quality assurance for a given testing run and the relationship of test results to a reference range, have long required manual intervention to review test results to determine the suitability of releasing them. Autoverification mitigates this effort by allowing rules to be established for the automatic release of results.\(^6\) For example, if a result is within the normal range, and the quality assurance was valid, the result can be released directly. This is a powerful form of laboratory CDS that significantly impacts laboratory resources and improves TAT. Imagine having to review thousands of normal test results before releasing them. The autoverification process is based on descriptive analytical algorithms and is simple to implement because most LIS have this functionality built-in. Autoverification has direct impact on the resources required for testing. By removing “normal results” from the review stream, laboratory personnel are better able to focus on important, abnormal, and impactful values. Autoverification rules can be quite sophisticated going beyond quality assurance to include delta checks, completion of reflex testing, and results of additional testing that may be a required part of a testing cascade. Not only does this eliminate errors of omission, but it ensures that clinicians get fully completed information at the time results release, eliminating iterative review of the medical record.

**Critical Values**

Laboratory testing that results in values that are of high clinical impact for patient care are one of the most disruptive, challenging aspects of providing laboratory services. So called “critical values” not only are important for quality patient care, but also have significant requirements from regulatory bodies impacting laboratory accreditation, and justly so. The failure to report a critical result to a provider in a timely fashion can negatively impact patient care.\(^7,8\) Examples of this abound. Systems that support a closed loop communication process between laboratory and clinicians have been difficult to implement. The process actually requires significant CDS. From the identification of a critical value, to validation of the result, identification of a provider, contacting the provider, communicating the details for the result and its critical nature, documenting the callback, and significant rework for any step in the process, the process can take hours (Box 1). If one were to follow the paradigms of data analytics, one might see the problem in a completely different light.

Recognizing the critical value is a descriptive analytical process; the callback process, diagnostic analytics; and leaping to the prescriptive aspects of the process, and the business intelligence required for the intercommunications, it is no wonder the problem has been so difficult to solve. If such results could be anticipated earlier in the testing cycle the volume of calls might drop using predicative analytics to adjust rules for particular patient populations, and yet identify truly emergent results.

Recently, process improvements at HFHS PALM have led to significant reworking of the callback process. We discovered key aspects of this process are highly amenable to CDS. For example, the volume of testing is directly related to critical value
generation. Therefore, volume dashboards can be leveraged to anticipate staffing and the complexity of communication. Using descriptive diagnostics, we can recognize the reasons we have critical values and the details around their origins. Predicting and modeling support requirements led us to new efficiencies with prescriptive knowledge leading to better patient outcomes and more efficient provider communication.

Dynamic Dashboards and Expanded Data Feeds

More recently, HFHS PALM has begun to develop dynamic data feeds to an analytical pipeline that is continuously updated. This allows for deviations in the preanalytical or analytical processes to be recognized in near real time allowing for more rapid intervention. Although these dashboards need to be monitored, they can be viewed on screen, with dynamic refreshing. Furthermore, without any additional query, the details behind a summary dashboard can be drilled into to get to the details. As designed, the details can include additional data elements that are helpful at the level of diagnostic analytics. This includes data from third-party systems that not only aid understanding as to why a deviation occurred, but may aid in its resolution. Although we are still short of predictive or prescriptive models, dynamic data feeding our pipeline is the first step.

Yet another example related to the TAT is our use of ePending logs or TAT outlier monitor dashboards. This information is displayed dynamically on large screens that effectively function as electronic Andon boards. When a given analyte from a location exceeds its threshold, it can be visually highlighted on the TAT outlier monitor indicating a problem. Because this is the same pipeline providing the overall TAT report the data can be drilled into to visualize and diagnose the reason for the threshold failure. These TAT outlier monitor dashboards now include detailed information coming from the automation line, such as current sample location or testing that shares tubes. This aids in resolving deviations in real time.

Dynamic Volume Reports

However, the real value of the dynamic data pipeline derives from the ability to view data continuously and over time (currently 1 month). This allows one to relate laboratory performance dynamically and across time without the need to generate new queries to the LIS. This also allows us to look at the data iteratively using filters and refinements within the pipeline. The ability to visualize trends begins to allow predictive analytics as a CDS tool in the laboratory. Even a simple dashboard, such as volume of orders by location, for current versus historic orders is profound, facilitating improved allocation of resources, shift planning, reagent ordering, and so

---

**Box 1**

Steps in the critical value call back process

1. Identify a critical value
2. Validate the truth of the result
3. Identify the provider to contact
4. Contact the provider
5. Document the communication with read back by the provider
6. Iteratively rework failed steps

---
forth. For example, if one notes a current high volume of orders from a given site compared with its historic normal, the laboratory can react prospectively to support service requirements.

**UTILIZATION**

Laboratory test utilization is another example where moving from purely descriptive to prescriptive analytics has an impact on the efficient use of laboratory services allowing for better CDS. Many of the articles in this issue describe the impact of CDS tools, such as alerts and ordering cascades or best practices to appropriately direct use of laboratory services. Detailed, real-time analysis of utilization can help to identify interventions that could be considered for future CDS activities. For example, capturing and storing laboratory test orders, results, patient demographics, and International Classification of Diseases information not only allows the laboratory to monitor services, but can also provide alerts and summary information. An example of this is monitoring testing that can directly impact patients. For example, monitoring the lack of hemoglobin A1c testing (HbA1c) in a patient with an International Classification of Diseases code of diabetes for a given calendar year. When this is taken to the level of prescriptive analytics such a system would alert the clinician, or perhaps even the patient, that important laboratory testing has not been done. And this is just the beginning. Perhaps the lack of HbA1c testing could be determined to be a surrogate marker for other aspects of missing patient care. Thus, a missing HbA1c is predictive, and potentially prescriptive for other forms of intervention. Perhaps this is not a medical problem but a social one, such as a patient having lost insurance, or being unable to afford medication, thus the lack of HbA1c is predictive of a health care delivery problem, not just a missing test. As another example, imagine the power of monitoring the care events that follow a positive pregnancy test, particularly if the patient never comes back. In fact, such systematic intervention and data analytics to the level of prescriptive analytics are the unpinning of what has become to be referred as “Lab 2.0.”

**FUTURE VISION: LEARNING HEALTH SYSTEMS**

Taken in toto, an analytics pipeline that goes beyond descriptive and diagnostic analytics and begins to leverage predictive and prescriptive combined with machine learning and artificial intelligence achieves a new level of sophistication: a learning system. A learning health system is defined by the Institute of Medicine as one in which: “science, informatics, incentives, and culture are aligned for continuous improvement and innovation, with best practices seamlessly embedded in the delivery process and new knowledge captured as an integral by-product of the delivery experience.”

The process cycle of such systems has been described as including the following steps:

- Assemble
- Analyze
- Interpret
- Feedback
- Change

CDS as viewed through the prism of a learning system leads to a vision of a spontaneous, dynamic environment for modeling and implementing CDS that may have been otherwise unappreciated. Could such systems someday provide for sophisticated CDS that exceed the limits of human observation?
SUMMARY

CDS in the laboratory derives from the data analytical process model. Such analyses, although dependent on descriptive and diagnostic analytics, will ultimately leverage more sophisticated analytics including predictive and prescriptive analytics. The culmination of the effective implementation of such systems will lead to the development of learning health systems within the clinical laboratory. Leveraging this information is essential to the efficient operation of an automated core laboratory. Examples of CDS in the laboratory have been discussed in the context of data analytics models that allow the reader to organize the implementation of such processes and CDS tools.

REFERENCES
