

Autoverification: lessons drawn from a core lab

Karen Titus

August 2022—Just as there is scant room in this world for pink “While You Were Out” notepads, paper checks, or the copying skills of Bartleby the Scrivener, laboratories would do well to leave manual result verification in the past.

But the need for such verification persists, with labs looking for ways to bring processes in line with automated analyzers and high-throughput testing, said Mayo Clinic’s Darci Block, PhD, DABCC, who talked about her experiences in a presentation at the Pathology Informatics Summit 2022 in May. It’s even more critical given the ongoing labor shortage in laboratories.

Result verification had become tedious and time-consuming when manual methods were applied in speeded-up laboratories. “And given the mundane nature of it, led to operator fatigue and potential for errors, and ultimately rendered itself impractical to do in this manner,” said Dr. Block, co-director of Mayo’s central clinical laboratory. In short, this older approach to the task has fallen into “I would prefer not to” territory.

Given the importance of result verification (a last opportunity to identify errors, a chance to prevent reporting inaccurate results, and ensuring compliance), labs would do well to ask if autoverification makes sense, Dr. Block said. The relevant CLSI guideline (“Auto15: Autoverification of Medical Laboratory Results for Specific Disciplines,” 1st ed., 2019) can help labs decide when it’s needed, including when (ahem) there’s a shortage of laboratory technologists, for compliance with quality requirements, and to meet turnaround time expectations.



Dr. Block

Autoverification is, at its heart, a tale of filters. Raw results from an instrument are sent through middleware software that basically replicates the manual work done by a technologist, comparing results against laboratory-defined acceptance parameters for a number of filters, such as quality control within allowable limits, absence of instrument flags (or similar errors), within allowable HIL limits (is the sample too hemolyzed, icteric, or lipemic?), and limit checks (is the result within the analytical measurement range? is dilution required? is it questionably low?).

As with all rules, they spring from the needs of the users, which can vary from one lab to the next. At Mayo Clinic, Dr. Block noted, “We do like when these rules hold things that are actionable. So it might mean that you have to pour the sample over into a false bottom tube, check for a clot, or repeat the testing with increasing or decreasing volume,” to name a few examples. Other commonly used rules address critical values and delta checks. (The latter may be of more historical significance to help identify misidentified samples that labs continue to evaluate in the advent of expanded positive patient ID system use.)

Her institution uses a number of logic/consistency rules. Among them:

- Is the total bilirubin greater than the direct bilirubin? “It shouldn’t be,” she said.

- Comparing I index to the total protein concentration. “You can identify interferences caused by paraproteins and other inconsistent results for further troubleshooting.”
- Anion gap.
- Metric of calcium times phosphorus.
- Metric of sodium divided by chloride.

At Mayo Clinic, Dr. Block said, “Our laboratory also uses what I would classify as process check rules. We’ll hold results where a sample treatment step such as polyethylene glycol precipitation or filtration was supposed to be performed, because in our highly automated and high-volume workflow, ensuring that manual processes such as these are performed prior to result release has the potential to be overlooked.

“So we choose to hold these results,” she continued, “and then verify that that processing was actually performed and final results are consistent with expected values,” versus autoverifying the results.

AST (U/L)	H – Index	Action
<100	50	Re-collect
100–<200	100	Re-collect
200–<300	200	Re-collect
≥300	500	Re-collect

Results falsely increased

Troponin T (ng/L)	H – Index	Action
<6–9	300	Re-collect
10–49	160	Re-collect
50–245	160	Call provider & re-collect
>245	265	Re-collect

Results falsely decreased

Courtesy of Nikola Baumann, PhD (AST), and Brooke Katzman, PhD (troponin T).

Fig. 1. Customizing hemolysis thresholds using analyte-specific middleware rules

Several other scenarios are especially ripe for machine learning and similar types of logic, she suggests. Among them: identifying IV fluid contamination from samples drawn from peripheral and central lines while medications are co-administered, EDTA contamination caused by order of draw errors, and pseudohyperkalemia caused by reasons other than hemolysis.

The general approach at Mayo, Dr. Block said, focuses on holding results where the technologist adds value to the work in process. “If we’re just holding things to hold,” she said, “it’s not helpful and possibly just delaying results.”

The central clinical laboratory is Mayo Clinic’s core lab, which tests some 8,000 specimens daily in chemistry, immunoassay, coagulation, and hematology. It includes three analytic lines that perform chemistry and immunoassay testing, as well as connected hematology automation. “We recently replaced aging Roche modular preanalytics with the Cobas Connection Modules. We have three p 612, p 471 modules that will centrifuge [and] do some limited aliquoting and sorting of specimens to these different destinations,” she said. The lab also has two p 701 refrigerated storage units that can store chemistry samples for the seven-day retention time.

Staff describe the lab as fast-paced, she said, and little wonder. The heaviest workflow occurs between 7 AM and 2 PM, Monday through Friday. Based on her own quick calculations, the lab delivers one result per second over a 24-hour period—but in reality the pace is much more intense.

Resource	Setup & support infrastructure	Design or review algorithms	Write/review rules	Test rules
IT	X		X	
MW, LIS, instrument vendors	X		X	
Instrument expert		X		X-Responsible
Supervisor and workflow experts		X		X-Responsible
LIS expert		X	X	X-Responsible
Laboratory director		X		X-Accountable
Project manager	Keeps project on timeline and within defined scope			

Courtesy of Darci Block, PhD.

Fig. 2. Middleware support plan

Go back to basics CLSI Auto15

The lab uses automated serum indices to assess specimen integrity. The middleware applies either the manufacturer's limits or lab-defined thresholds. Serum indices are performed for all specimens, "and then we've customized our rules and programmed them into the middleware to hold only results where the serum index threshold is exceeded," she said. If a threshold is exceeded, that result is not reported and a redraw request is submitted.

She offered the example of hemolyzed specimens. Middleware rules are programmed with analyte-specific thresholds. On a basic metabolic panel, if the potassium was identified as having hemolysis above threshold, this specific result would be held, while the other analytes would be autoverified and reported. The potassium would be re-collected with a potassium order. "With any luck, that sample will not be hemolyzed," and the lab can report the result.

The lab performed experiments to extend the hemolysis threshold for additional more-sensitive analytes, including direct bilirubin, as well as AST and troponin (**Fig. 1**). "We derive these rules based on measured concentrations to reduce our rejection rates," she explained. For example, previous thresholds led to too many needless re-collections for AST testing. "Hemolysis causes a falsely elevated result; therefore as AST concentration increases, the proportion that is falsely increased diminishes. So we were able to extend the H-index limit as the AST concentration went up," Dr. Block said.

The influence of hemolysis on troponin T was the exact opposite but just as impressive. Hemolysis causes troponin T measurement to be falsely decreased. The lab was getting false decreases based on previous H-index thresholds. The lab derived different thresholds that allow a slightly higher amount when results are low. "And then a fairly stringent H-index limit around the decision points where the reference interval lies," she said, "and where we expect to see fairly sensitive changing patterns in early MI."

What was the impact of these changes? For AST, rejection/re-collection rates dropped from 8.1 percent to 1.5 percent (an 82 percent reduction). That eliminated 708 specimen re-collections over roughly one year. For direct bilirubin, the lab calculated that rejection/re-collection rates dropped from seven percent to 2.7 percent—a 62 percent reduction—with the new H-index threshold, which eliminated re-collection of 306 specimens.

Autoverification testing strategies, like traveling with teenagers, have their low points as well as their rewards.

A good place to start, Dr. Block told her audience, is with CAP accreditation program checklist requirement GEN.43875, which addresses validation of the autoverification process initially and whenever there's a change to the system that could affect the autoverification logic.

"So it's important to have in mind what the middleware support model is for the laboratory or labs that are using this system," she said. "We usually think in terms of IT, which supports the application from a technical perspective, versus the lab expert [who] not only uses the system but can then help assist in testing and verifying

that it's functioning properly."

Beware of gray areas, she said. "In our case, we have a lab information system technical specialist [who] helps with configuration." The specialist also "serves as an arbiter between IT and the lab to help support configuration build and communicate other change requests."

Mayo Clinic's central clinical lab "has a long history of using middleware software," she said. A medical technologist built the first version in about 2000, and it evolved as instruments changed and automation software became more commonplace. The lab recently modified its current middleware rule build to accommodate a new automation and instrument upgrade; Dr. Block and colleagues rely on Infinity to help support the automation, and this added an extra layer of complexity to the workflow. Moreover, since a multiphase project was needed to replace instruments and automation in existing space, the project had many moving parts and interim states of partial automation. "We also layered on top of that a personnel reorganization and this other thing called COVID-19," she continued, "which meant there was much less on-site support and subsequent staff turnover. One or all of these challenges may have distracted us from making a robust testing plan."

Nevertheless, she reported, "We did it—we did well, considering all these obstacles. But eventually a very savvy tech in the laboratory noted that a test that should hold when hemolysis is present didn't hold as expected. After some forensic investigation, ultimately we discovered the cause was a different hemolysis message sent from a new instrument that wasn't updated within the middleware to hold." The lab was relying on existing rules that under normal circumstances would work. "However, the new instrument didn't utilize the same message, so it went unrecognized for three tests on the test menu. Ultimately we did perform root-cause analysis to determine what we could do to prevent this moving forward."

Having a testing scheme is crucial, said Dr. Block, who noted there are two basic ways that middleware logic gets tested.

One, the so-called dry-testing approach, involves a simulated environment that lets the lab test virtually all scenarios in a fairly robust way, she said. This approach can be facile and comprehensive. It also saves on the cost of reagents and having to test specimens. On the flip side, this method generates massive amounts of files and paperwork. "And you're blind to the manual inputs that you're assuming are happening in these various scenarios," Dr. Block said.

So-called wet testing, on the other hand, is a real-world scenario that involves ordering a test on a test patient, then following the flow of information through the whole system to check the behavior and confirm that the expected outcome has indeed occurred. The instrument flags and nuances are visible with this approach (and thus would have helped the central clinical lab identify its problem, she said), and it "provides a high amount of confidence that you've tested that algorithm well."

But this approach has its own disadvantages. "You can't test every single scenario, and some can't even be replicated," Dr. Block said. "And the cost of reagent and specimens is going to add up tremendously. It could actually be prohibitive."

In Mayo Clinic's case, Dr. Block said, "We decided we were missing a very explicit SOP," one that the technical specialists who support the middleware, as well as the testing lab, IT, and other stakeholders, could follow to produce a testing plan. The plan needed to describe, at a high level, what types of changes might need to be implemented in the middleware and whether it would be amenable to simulation testing, wet testing, or both.

"That was the first step," Dr. Block said, one that allowed all stakeholders to define their roles and responsibilities more clearly. With this plan, "We usually use the analytical SOP as a source of truth, and then that would be what we use during our downtime scenarios." The testing is done by the laboratory technical specialist, who signs off along with the medical director. "So the pearl is to make a middleware support plan that identifies the who, what, and which scenarios to apply—wet testing versus dry testing." (**Fig. 2**).

This well-defined plan was not the starting place for the central clinical lab, given that new software and instrumentation (including automation) were already in the works. Another hiccup occurred with turnover of some key staff who had built and maintained the existing middleware rules.

Still, neither the new instrumentation nor the data innovations were radically different. The lab also had plenty of vendor support, including training new staff in their jobs. In retrospect, that lulled the lab into a false sense of security, Dr. Block said. “We ultimately did not do as much of the full path wet testing with patient samples as we really needed and could have used.”

That experience brought Dr. Block to her second pearl of wisdom. “Once you have that plan, it’s important to engage the plan, sticking to it. At least don’t abandon it,” she said. If the project runs into hurdles—such as staff departures—the plan can make the handoff a little smoother, with less risk.

Finally, she said, lab leaders need to keep asking their team (and themselves) why they do what they do. The answers—or lack of good ones—can keep labs on the right path. There’s always room for labs to evolve and improve, she said. Autoverification, in other words, should not fall into the trap of running on autopilot.□

Karen Titus is CAP TODAY contributing editor and co-managing editor.