Case closed: discrepant results at multiple sites

Kevin B. O'Reilly

October 2015—As hospitals are brought under single health systems, laboratory leaders are faced with the task of ensuring that their clinical lab results are comparable among various sites and instruments. But some have had more opportunity than most to investigate the mischief afforded by variations in instruments, reagents, and more.

Dr. Greene

Dina N. Greene, PhD, found herself in that position. She worked for four years at Kaiser Permanente Northern California, where she served as a clinical chemistry consultant for its 21 hospital laboratories in the area and directed hemoglobinopathy and myeloma testing for the system's regional laboratory.

"This is an increasing kind of problem with consolidation. As different universities acquire more hospitals and as hospital systems acquire other hospitals, it's going to be an increasing challenge," says Dr. Greene, now associate director of chemistry at the University of Washington Medical Center. She also is assistant professor in the Department of Laboratory Medicine at UW, which she joined in December 2014.

"You have to standardize your equipment—that's a fundamental part of this. Without standardizing the equipment, you just have so much more opportunity for wildly different results, especially if everything is going into the same electronic medical record," Dr. Greene says.

During a talk at this year's American Association for Clinical Chemistry meeting in Atlanta, Dr. Greene highlighted a puzzling case that she investigated along with Nikola Baumann, PhD, of the Mayo Clinic in Rochester, Minn. The case helps illustrate the complex challenge of aligning laboratory results across a health care network. A 29-year-old woman who was nine weeks pregnant presented to one of Kaiser's Northern California hospitals with severe nausea and vomiting. Nearly all of her laboratory values were unremarkable, except for an elevation of her aspartate aminotransferase, at 105 U/L. The reference range for that hospital laboratory was 14–36 U/L, using an Ortho Clinical Diagnostics Vitros analyzer.

The woman was diagnosed with hyperemesis gravidarum and treated with regular IV fluids and the antinausea medication ondansetron. Her AST peaked at 132 U/L, as measured by the hospital's analyzer, but by 20 weeks of gestation the symptoms had resolved. Her AST, measured on an outpatient basis this time using a Beckman Coulter AU5800 at Kaiser's regional laboratory, was at a normal 38 U/L given that instrument's reference range of 10–40 U/L.

At 33 weeks of gestation, the woman's abdominal symptoms recurred, but the outpatient AST results continued to be normal. However, a paired specimen evaluated at the hospital laboratory showed an elevated AST. By 36 weeks of gestation, the woman reported continuous pain in the right upper quadrant of her abdomen. Specimens evaluated stat at the hospital laboratory all showed an elevated AST, with bile acids mildly elevated. Yet all other liver and pancreatic markers were normal. At 37 weeks of gestation, the woman underwent an uneventful elective caesarean section.

“Mom and baby were just fine,” Dr. Greene told the AACC crowd.

While the outcome was good, the mystery of the discrepant AST results remained. At this time, Dr. Baumann, co-director of Mayo’s central clinical laboratory and director of central processing, was asked to consult on the case. She and Dr. Greene suspected that differences in reagent composition might be the root cause. Serum aliquots sent to the Mayo Clinic and another reference laboratory—both of which use Roche Cobas instrumentation—also returned with discrepancies. Mayo Clinic flagged the specimen as having an elevated AST of 243 U/L, while the other reference laboratory flagged the sample as having a low AST of 8 U/L.

So, this was no longer just a matter of different instruments and different laboratory sites yielding discrepant results. Now it was two outside laboratories using the same instrumentation yet reporting diametrically opposite results. Dr. Greene and her colleagues, however, were able to pinpoint how a combination of patient and laboratory factors combined to create the confusing results.

First, they noted differences in reagent composition. Mayo Clinic and the Kaiser hospital laboratory that reported elevated AST results both supplemented their reagent with pyridoxal-5-phosphate, the active form of vitamin B6, as a cofactor. But the Kaiser regional laboratory and the outside reference laboratory, which reported the patient’s results as low or normal, did not supplement their reagent with P5P. Meanwhile, it turned out that the patient was vitamin B6-deficient, which was discovered by measuring B6 vitamers in a fasting sample.

The Mayo Clinic laboratory also tested for and identified a rare macroenzyme of AST, termed macro AST, in the patient’s serum. Testing again, Dr. Greene and her colleagues found that without P5P, the patient’s AST was just 11 U/L and jumped 1,700 percent to 186 U/L in a sample where P5P was used in the reagent. Dr. Greene said she hypothesizes that the macroenzyme form of AST is more sensitive to B6 deficiency than “normal” AST. She added that Mayo’s own investigation of the discrepancy supports that view.

“More than anything, what this shows is that discrepant results from instrumentation really complicate clinical interpretations, and we have to understand the reagent composition, even in our FDA-cleared tests, in order to be able to solve these complicated things when they come up,” Dr. Greene said at the AACC session.

“If the patient just had unexplained elevated AST, we would have found the involved macroenzyme right away,” she added. “What we didn’t know was that the macroenzymes would be inconsistent between platforms, and that’s what this shows.”

The case is published in the October issue of *Clinical Chemistry* (Mills JR, et al. 2015;61[10]: 1241–1244).

To help head off these sorts of problems, one of Dr. Greene’s first jobs at Kaiser was to harmonize the hospital labs’ chemistry analyzers with the regional laboratory’s. The AU5800 was used at the regional laboratory, while the Vitros was used for chemistry at the hospital labs.

Dr. Greene and her colleagues examined the Beckman AU5800 and AU680, the Siemens Vista, and the OCD Vitros. They performed 40 serum tests, four urine tests, and two cerebrospinal fluid tests with each platform. They found no red flags with precision, linearity, or carryover as part of the basic validation. But when they did an interinstrument comparison to check on harmonization, they got interesting results.

After running about 100 samples on four different instruments within a 24-hour period, they found similar results between the AU5800 and the AU680. But when comparing the AU5800 with the Vitros, Dr. Greene and her colleagues found 12 analytes in which there was positive or negative bias of five percent or more. Similar discrepancies were seen between the AU5800 and the Vista.

“If you know anything about chemistry analyzers, you might think that this is because the Vitros is dry chemistry and the AU5800 is wet chemistry—that might explain these differences. But that’s actually not the case,” Dr. Greene said. “It has a lot more to do with specific reaction conditions, things like the substrate and the wavelength the instrument is monitoring at.”

For the lipase test, for example, there was a 133.1 percent bias in the Vitros relative to the AU5800. That bias occurred, Dr. Greene said, because "the Vitros method uses an unconventional substrate of questionable specificity for pancreatic cases."

Even comparing the AU5800 and the AU680, there was still interinstrument variation. In the end, Kaiser opted to transition all of its laboratories to the AU680. Undertaking such a change for the sake of laboratory harmonization requires a concrete, stepwise approach to validation, Dr. Greene said.

"At one site, you do a very extensive validation, then you figure out the kinks in the assays that are going to cause you problems," she said. "Active participation of knowledgeable lab staff at the first site ensures that issues will be identified and resolved before subsequent lab deployments."

The next steps are to build and test the interface, write procedures and train staff, optimize workflow changes, distribute technical bulletins to track reference range changes from previous instrumentation, and then go live. The process is repeated at each laboratory.

Kwong

A former colleague of Dr. Greene's, Shiu-Land Kwong, CLS, MT(ASCP), further details this multisite rollout.

"It's good for that first site to be really solid, so we have a good plan. It took over a year to get the first site ready. Behind the scenes, in addition to validation, there are LIS issues and other issues. There's all of this background work in trying to make sure we're able to duplicate it when we take off from the first site to the second one to all 21 sites," says Kwong, regional director of laboratory compliance and risk management in the Permanente Medical Group Laboratory System in Northern California.

"We've been able to cut down the implementation time to about eight weeks after the site construction is complete," Kwong says. "We're glad to have a package to be able to roll out at each site." The AU680 is live at eight of the Kaiser hospitals in Northern California, and another three are expected to go live by the end of the year.

Even with the same equipment, it remains a challenge to reduce or eliminate discrepancies. That requires understanding the clinically acceptable maximum imprecision and investigating matters when it is exceeded.

"It could be that calibration is overdue or there are issues with the instruments or reagents," Kwong says. "There's a lot of implications that this is something that needs to be looked at."

The initial validation and rollout is essential, she adds, because "it gives us an internal reference point in addition to what the manufacturers have published in their reference data."

And, yet, not all discrepancies can be eliminated.

"We can see this when we bring up all these instruments," Kwong says. "We can have two brand-new instruments, yet they're not the same. It would be great if the instrument-to-instrument variation would be minimized at the manufacturing level. However they put together the instruments, we may have a certain instrument consistently running on the negative or positive side. We do see that issue, even after we calibrate. That's just the way it is. There are constant biases between instruments. It is fine if the bias is within the clinically and analytically

acceptable limits defined by our laboratory system."

Reducing the variation among laboratory sites and instruments is especially important in the era of patients having direct access to test results, Kwong says.

"In our system, the patients have access to their information online and are able to trend results themselves," she says. "The patient doesn't know where the testing was performed. All they know is, 'This is my result.' Therefore, it's even more critical—when patients have electronic access to all laboratory results to look at them and trend them—that we have comparable results from all locations."

In her AACC talk, Dr. Greene offered up evidence of how faulty calibration can contribute to greater variation, even when all the tests are performed on the same platform. Comparing how 21 Kaiser sites used 42 Beckman Coulter Access 2 immunoassay systems to perform the Accu-TnI+3 troponin I assay, Dr. Greene and her colleagues found instrument-to-instrument variation was likely the result of calibration bias (*Clin Biochem.* 2015;48[4–5]:268–274). Sites that recalibrated and repeated the comparison saw the between-instrument bias shift.

All of this makes the argument for standardizing laboratory equipment across sites even stronger as a way to achieve greater harmonization and more reliable results, Dr. Greene says.

"That's what my data is showing. This is hard enough when everyone's on the same instrument. There's no way to do it when you're on different instruments," she says. "This also bolsters the argument for standardization of quality practices."

Dr. Greene no longer has the ability to see how a change in a reagent might play out across 42 instruments as she did in her previous position at Kaiser. But she still has three University of Washington laboratories within her purview, and Dr. Greene works with colleagues around the country to gather data on interesting cases that elucidate the challenges of lab harmonization.

"Other people may not think this work is sexy," she says, "but I find it fascinating."
[hr]

Kevin B. O'Reilly is CAP TODAY senior editor.