

In lab QC, how much room for improvement?

Anne Paxton

October 2014—The debut of the CMS' new quality control option, IQCP, has sharpened the focus on QC in the laboratory and raised hopes that risk management concepts can make QC more robust. But one of the most highly regarded quality control experts in the U.S. voices skepticism about the impact of IQCP—and indeed, about U.S. quality control standards in general.

As a voluntary, customizable QC option under CLIA, IQCP or Individualized Quality Control Plan is expected to give labs greater flexibility in achieving QC compliance. However, the CLIA QC standards, unchanged since 2003, will remain the same—and that's a problem, says James Westgard, PhD, who spoke about QC weaknesses at the 2013 Lab Quality Confab presented by The Dark Report. He believes that CLIA's sluggish evolution on QC has nurtured a nationwide attention deficit on the subject of meaningful quality management.

[hr]An interview with James Westgard, PhD[hr]

From his standpoint, the state of the practice in QC falls far short of perfect, and even well short of sufficient. "My opinion is that CLIA has kind of frozen quality management practices to the early 1990s," Dr. Westgard says.

A co-founder of Westgard QC, in Madison, Wis., Dr. Westgard is author of several books on laboratory quality management, including *Six Sigma Risk Analysis: Designing Analytic QC Plans for the Medical Laboratory* (2011). He has more than 40 years of experience in laboratory quality management.

Dr. Westgard, an emeritus professor at the University of Wisconsin in Madison, was the first chairman of the Evaluation Protocols Area Committee in CLSI, the Clinical and Laboratory Standards Institute (then known as NCCLS). Trained as an analytical chemist, "I got started in the laboratory about the time automation was just making big inroads, in 1968." He spent considerable time dealing with methods validation protocols and the best statistical analyses to use. For him, the big question has always been: How do you decide whether a method of testing is actually acceptable or not?

Clinical laboratory professionals used to depend on their own analytic skills to answer that question. "It was really with the introduction of automation in the late '60s and early '70s that QC got a strong push," he says. "That's because you may have a lot of confidence in your individual skills for performing a test, but once you have a machine doing the test, how do you know it's right?"

The standard practice of setting control limits at the mean, plus or minus two standard deviations, worked fine for QC—until multi-test instruments came on the scene. "Everyone knows that about one out of 20 results is expected to be outside two SD limits; that's the false rejection rate inherent with these two-standard-deviation limits. But that's based on one test. Once you start doing six, 12, or 20 tests with a multi-test system, you have a multiplier effect on that false rejection rate," Dr. Westgard says.

In those days, with the simultaneous-batch-type analyzer, you couldn't just pick one test to run, he explains. "You used up the capacity of the system every time you had to do a repeat test. So we soon got to the point where, because of the number of tests being run, you had about a 50 percent chance with each run that at least one test was out of control." That problem stimulated his work on QC and led to the development of a multi-rule QC procedure that is commonly known as Westgard Rules and became a standard of practice in laboratories in the 1980s.

The Clinical Laboratory Improvement Amendments of 1988 were supposed to extend the practices of QC to all laboratories in 1992, when CLIA '88 regulations took effect. "The regulations themselves described what your QC procedures are supposed to be able to do: monitor precision and accuracy of the system and detect medically important errors." The promise was that "manufacturers would make claims for their precision and accuracy and

what a lab should do as far as QC, and if the FDA approved that claim, then all the lab had to do was follow the manufacturer's directions."

But that part of CLIA—"QC clearance"—was never implemented. "First there was resistance by the manufacturers. Then later on, the Food and Drug Administration decided they had enough on their hands, and they didn't want to deal with it either." Every two years, from 1992 on, there would be a new final rule putting off the effective date of the FDA clearance of QC—until 2003, Dr. Westgard says. "Then they declared they didn't need the FDA clearance of QC anymore because the analytic systems had gotten so much better."

At that time, many manufacturers were arguing that the new test systems and point-of-care devices had built-in controls and labs didn't need to be running external controls. Under CLIA's final rule in 2003, the CMS compromised by establishing the EQC, or Equivalent QC procedures. "That allowed labs, instead of doing two levels of controls per day, to go to two levels per week or even two levels per month, if they provided certain validation data. One protocol stipulated that the lab should run controls for a 10-day period. If everything was okay, then you were qualified to reduce QC to two levels a month. So if you were stable for 10 days, then you wouldn't have to test until 30 days went by. Obviously, that's not how stability testing should work."

The problem with the whole approach was that the validation protocols that the CMS prescribed were not scientifically valid, Dr. Westgard says. But in spite of that argument having been made up front with CMS, "I think they were just stuck with having to adopt EQC to accommodate the POC test systems that were being widely used."

In the absence of the FDA's clearing QC, the CMS' default minimum—two levels of control per day—became the standard practice in labs. It was the least amount of QC that needed to be done. "There's no scientific basis for that practice, but because CLIA said so, the minimum became a maximum over time. That's the nature of regulation." As a result, most labs fall back on what the regulations require, he says. "You are only required to run controls. The regulation doesn't require that you run the right QC."

Despite this low bar, many laboratories are very good, he notes. "But as you increase the workload in labs, and people have less and less time to think about what they're doing and they're just trying to keep up with the workload, then there are certain things that fall by the wayside. And that, unfortunately, is what happens with quality practices."

IQCP was devised, in part, to resolve the EQC problem by offering a risk assessment approach. "In theory, it's a good approach," Dr. Westgard says. "It's just that in practice, laboratories have never done formal risk analysis, and formal risk analysis is not a trivial undertaking." Laboratories face a steep learning curve in understanding and correctly applying quality practice plans based on risk, in his view.

It's true that people make judgments of risk all the time. "You look to see what the weather is and decide whether you need an umbrella or not. That's risk assessment. But that's not the same thing as looking at the potential for harm in a lab test result and figuring out, 'Can I risk this or not?'" Even though CLSI has a guideline called EP23A, it is qualitative, it is subjective, and most people don't understand it, he says.

Unfortunately, no one actually knows how confident we can be in laboratory test results, he says. "We all hope for the best."

QC should always begin with the question: How good does this test need to be? Dr. Westgard emphasizes. "Then we look at how good we are. And if the test is much better than it needs to be, then it doesn't take much QC. If it's only 'close to' as good as needed, then you have to monitor the test much more carefully."

He compares the process to budgeting. "When you set a financial budget, you know that if you spend too much in one place, you run over your budget. The analogy for lab tests is an 'error budget.' We know we have certain sources of error like precision and inaccuracy. How much of the budget gets spent by different error sources, and is it possible that we will overrun the budget if a problem occurs?"

“Well, we have information on that, in the form of what quality is required for the test. If you define that in your budget, then you can measure the errors for your methods in the laboratory to be sure they fit within the budget. That should be an ongoing part of quality management: keeping track of how big these errors are and how they relate to the amount of error that is allowable.”

Dr. Westgard concurs that laboratory test quality has improved because of the technology of the diagnostics industry. But the quality demands for any one test change when the use of the test changes. He cites HbA1c as a good example. In the past 10 or 15 years, it has gone from being a test used to monitor diabetes to becoming the basis for diagnosing diabetes and monitoring treatment.



Dr. Westgard: “Formal risk analysis is not a trivial undertaking.”

“HbA1c is now, perhaps, the most critical test we run. If you look at just the CAP proficiency testing requirements five years ago, the test needed to be correct within 15 percent. This year, the requirement is that it needs to be correct within six percent. As doctors have used the test more critically, the quality required gets to be more demanding. And the real question is: Is the improvement keeping up with what is required?”

This is another area in which CLIA standards have not kept pace, he notes. “CLIA started with a list of 70 to 80 tests that were required to participate in proficiency testing—three events per year, five samples per event, with criteria defined for acceptable performance.”

But the CMS has never updated this list of regulated tests, which does not include HbA1c or prostate specific antigen. As a result, standards for those tests have been left up to private standard-setting organizations (in the case of HbA1c, the National Glycohemoglobin Standardization Program or NGSP, as well as the CAP). “So it’s CAP that is really setting the standards. If you ask CMS, they will say they have a plan to update the list, and it’s ‘going to happen.’ But they’ve been talking about that for the last five years.”

Larger academic institutions tend to be better at QC, Dr. Westgard says. “People recognize that CLIA is just a minimum. But the biggest problem is, the smaller the lab, the more pressure there is for production. At a certain level, like POC testing, you may have operators with little or no understanding of the laboratory testing process itself.”

It’s a mistake to think that analytical quality is no longer a problem, he warns. “That’s based on the idea that there are so many more errors occurring in the preanalytic phase than in the analytic phase, that we don’t need to worry about the analytic phase anymore. People almost universally believe that.”

But the most harm to the patient is caused by analytic errors, he emphasizes. He points to a well-known study by Italian researchers Mario Plebani and Paolo Carraro, "Lab Errors and Patient Care," as evidence (Clin Chem. 2007;53:1338-1342). The study found that in 160 confirmed laboratory errors, 46 led to inappropriate patient care, and 24 of those were analytical errors—meaning that more than 50 percent of the time, analytical errors were the major cause of inappropriate patient care. "That is the same paper that is cited to justify that there are fewer analytic errors than preanalytic and postanalytic errors, but many people overlook the fine print that says analytic errors are the most serious problem related to patient care," Dr. Westgard says.

Unlike preanalytic errors, which are more easily discovered, many analytic errors go undetected, he adds. "When you get a result, can you recognize it's an error? Most clinicians can't do that. So if you're relying on them to tell you about analytic quality, the fact that you don't have any complaints does not mean quality is okay." While there are algorithms that make it possible to look at the relationship between test results, "all of them are relatively insensitive, even when we reduce them to a mathematical algorithm."

Private-sector standard-setting groups like the CAP could find implementation of IQCP challenging, Dr. Westgard believes. "The risk assessment approach is actually very difficult to do in an objective way. The practice guidelines that are coming out describe a very qualitative and subjective process. From my perspective, the result will be that 'any QC will do,' as long as you make it look like you are going through the process, because there's no objective way to monitor the process." The challenge will be deciding how to inspect so that the QC plan actually works, he predicts.

Laboratories need to get over their fear of statistics, he contends. "You just have to say the word 'statistics,' and people throw up their hands and say 'okay, now you've lost me.' And that's a real problem in labs." But statistics are not as frightening as people think, he insists. "They are a way of summarizing the data in order to manage the question you're addressing."

Calculations of quality on a sigma scale work the same way. "The value of six sigma is that it gives you a methodology to define how good a test is. If you know for your method in the laboratory how good the test needs to be, the precision, and the bias, those three numbers can be combined to calculate a sigma metric to characterize quality on the sigma scale."

He advises every laboratory to determine their quality on the sigma scale. "The sigma scale gives you kind of a universal way of looking at how good you are—six sigma being the goal for world-class quality, and three sigma recognized in industry as a marginal level of quality. If you go below that, you shouldn't even be in production." Laboratories that have a test with a two sigma rating should stop performing the test, he says.

People in the laboratory understand how precision and bias affect the quality of their method. But they often don't know which methods perform the best, he says. This is a particular problem with point-of-care testing, particularly waived tests. "For these waived tests, there's no proficiency testing requirement. We have no idea what the quality of that testing is, and that is really a serious problem." A 2014 article in Clinical Chemistry showed that, in a study of seven devices waived by the FDA and certified by the NGSP as being equivalent, three did not "meet generally accepted analytical performance criteria" and the data show they do not perform up to three sigma standard (Lenters-Westra E, et al. Clin Chem. 2014;60:1062-1072). Another article in that same issue shows that six out of seven methods used in central laboratories provide less than three sigma quality (Woodworth A, et al. Clin Chem. 2014;60:1073-1079). "That means that most of the HbA1c testing that is occurring today is not good enough for the intended clinical use of the test."

However, sigma scale ratings of commercial diagnostic tests are unlikely to get circulated widely, he says, because of the starkness of attaching a sigma number to every method. "As a laboratorian, I would know immediately that I should avoid certain methods. So the manufacturers would never put up with it."

As more and more hospitals complete the upgrade to the electronic health record, he thinks the EHR is going to drive quality to improve but will not necessarily have a direct influence on QC. The EHR will bring some of QC's weaknesses more out into the open, and he sees that as a good thing.

“When you start getting test results and different methods from different labs put into the same record, you start to see the inconsistencies. Each manufacturer of a method typically has some patent, has done something different or unique to obtain intellectual property rights to the method and prevent other manufacturers from using it.” But for quality patient care, “all labs need to get the same result. So there’s an inherent problem there.”

“We can’t standardize the methodology to be sure we get the same results. Instead we must ‘harmonize’ the results by assigning different calibration or correction factors.” That then requires programs like the NGSP to evaluate and certify that results from different methods are comparable or equivalent, he says. “And that is very costly, which then limits our efforts to ensure comparability of test results from different methods.”

Physicians don’t understand how different results can be from one method to another, Dr. Westgard says. “There is this assumption that with all this high technology we use, we must be able to do these measurements right.”

Another perennial issue for laboratories is that the benefits of great QC, and the harms from poor QC, are external failure costs, Dr. Westgard points out. “You could scrimp on QC to keep costs down, and if you generate a bunch of bad results, that just means more nursing time or more physician time, and that’s over in someone else’s budget.”

A few studies have tried to estimate what the cost of quality amounts to. But hospitals don’t have a line item for misdiagnoses in their budgets, and they’re not likely to have one anytime soon, he says. “In principle, everyone understands that one bad test result and mistreatment of a patient in the ER—that’s a huge amount of money. But knowing that does not in any way help the laboratory.”

Quality control practices in Europe are ahead of those in the United States, Dr. Westgard maintains. “One of the major differences in Europe is that the standards for practice are coming out of the International Standards Organization or out of the professional organizations. ISO 15189 was developed in the early 2000s, and we now have an update as of 2012. CLIA was developed between 1988 and 1992, and my opinion is that CLIA is kind of stuck there, whereas in Europe they have continued to evolve with ISO 15189.”

The CAP’s 15189 accreditation program is, of course, based on ISO 15189. And that’s a telling fact, Dr. Westgard says. “That makes it pretty clear that what we have in CLIA is not the standard of practice in the world anymore. ISO 15189 is really the global standard, and U.S. labs are for the most part almost ignorant of it.” Despite the hopes for IQCP, he does not believe it would stand up to ISO accreditation standards.

The contribution of CLIA in the QC arena, he summarizes, has been one-size-fits-all QC minimums, leading to equivalent quality control, and now risk-based QC, which Dr. Westgard says will tax laboratories with negotiating highly complex QC procedures. The net effect, in his view, is that in the U.S. the quality of laboratory tests may suffer. Professional groups and clinical laboratories themselves, he believes, will have to assume more responsibility to ensure QC meets the challenge of detecting medically important errors before they cause patient harm.

[hr]

Anne Paxton is a writer in Seattle.