



Have you heard about the (other) major regulatory change this year?

CLIA changed its proficiency testing criteria — and almost no one is talking about it

By Sten Westgard

If you got lost in last year's labyrinth of laboratory-developed tests (LDTs), or caught up in the latest CLIA changes to personnel requirements, you can be forgiven if you missed the update to the CLIA proficiency testing criteria.¹ In a quirk of timing, the new criteria became effective on July 11, 2024, but were not yet enforced. The actual teeth of the new rules didn't bite until January 1, 2025, when the proficiency testing programs must implement the changes.

Why is CLIA changing PT criteria?

Proficiency (PT) testing criteria were established in 1992, as part of the implementation of the CLIA 1988 regulations. For the last 33 years, those PT criteria have stood unchanged, even as everything else in the laboratory rapidly evolved.

As instruments improved, the 1992 PT goals became easier to hit, so much so that criticisms began to be voiced about their relevance. The '92 goals were too broad, no longer reflective of modern instrumentation, and more importantly, no longer reflected modern use and interpretation of test results. CLIA PT began to seem less like a true test of laboratory performance and more like a rubber stamp. Belatedly, laboriously, CLIA updated the goals. The new goals are meant to reflect the method performance and clinical use of the 21st century. It seems uncontroversial that instruments from this century should not be judged by the standards of the last century.

Even the implementation of these changes has rolled out in a kind of slow motion. After a proposed set of criteria were issued in 2019, the official changes were only announced on July 11, 2022 — with a full two years allocated for labs to prepare for the changes. By the time these criteria go into effect, it will have been almost two and a half years since they were announced.

Earning CEUs

See test online at <https://ce.mlo-online.com/courses/have-you-heard-about-the-other-major-regulatory-change-this-year/>.

Passing scores of 70 percent or higher are eligible for 1 contact hour of P.A.C.E. credit.



Scan code to go directly to the CE test.

LEARNING OBJECTIVES

Upon completion of this article, the reader will be able to:

1. List the reasons why CLIA is changing PT criteria.
2. Discuss the changes that may make a difference in PT results.
3. Describe the Sigma metric and its utility in PT success.
4. Discuss the reasons for PT failures in the newly applied CLIA rules.

Changes that will make no difference?

For many laboratories, proficiency testing has been pro-forma, a routine activity without risk of failure, a motion to go through to maintain compliance, but never something reflective of true performance or quality. For years, it's been so easy to pass the surveys that laboratories have been assuming that, even with new changes, they are still guaranteed to pass.

But look over the size of the changes that will go into effect on January 1st:

Analyte or test	Old criteria	New criteria	Difference
Hematocrit	6%	4%	33.33% reduction
Potassium	+/- 0.5 mmol/L	+/- 0.3 mmol/L	40% reduction
Magnesium	25%	15%	40% reduction
Uric acid	17%	10%	41% reduction
Hemoglobin	7%	4%	42.8% reduction
C3 complement	3 SD	15%	varies by method
T3 Uptake	3 SD	18%	varies by method
Triiodothyronine	3 SD	30%	varies by method
HbA1c	NDR: 20% or 3 SD (API)	8%	60% reduction
CA 125	NDR: 3 SD (CAP) / 2 SD (API)	20%	varies by method
Estradiol	NDR: 3 SD (CAP) / 2 SD (API)	30%	varies by method

Figure 1. Proficiency testing goals.

Key abbreviations. CAP: College of American Pathologists. API: American Proficiency Institute.

A full list of these changes can be found in the *Federal Register*,¹ as well as an annotated version on Westgard Web.² There are three main kinds of changes:

- Directly regulated tests where defined goals are being tightened. In the table, hematocrit, potassium, magnesium, uric acid, and hemoglobin are examples. For these, we can calculate how much smaller the new goals will be.
- Directly regulated tests without fully defined goals, now have newly defined goals. In the table, C3 complement, T3 uptake, and triiodothyronine are examples. It's not possible to state how large the reduction is, because each method was previously judged by a method-specific group SD. For a method that has a large group SD, the reduction will be more significant than a method that has a small group SD.
- Previously unregulated methods now being directly regulated with defined goals. In the table, HbA1c, CA 125, and estradiol are examples. It's even harder to determine the size of the reduction because not only were the previous goals method-specific, the previous goals were also PT-provider-specific. For unregulated methods, each PT provider was able to set their own goals.

HbA1c is a curious example of the impact of CLIA 2024 requirements. The new CLIA requirement is 8%, and the old goal for API was 20%, while the goal for CAP, following the National Glycohemoglobin Standardization Program (NGSP) program, was a goal of 6%. In this one instance, CAP users may find it easier to pass proficiency testing if they adjust to the CLIA requirement.

Why haven't we heard more about this?

While CMS published all of these rule changes in the *Federal Register*, the proficiency testing providers haven't aggressively publicized them. While alerts have been sent out, and notifications have been posted on various websites, no one has offered laboratories a detailed preview of the impact of these goals on their future survey results.

Overburdened, understaffed laboratories — the most common type of laboratory in the United States — have not had the bandwidth to contemplate their future failures or successes in proficiency testing. The LDT controversy and CLIA personnel changes have drowned out the PT notices.

Conventional wisdom is that these changes aren't going to impact labs significantly. But if CMS reduces a goal by 60%, where is the concrete evidence that this won't impact at least a few laboratories? Why change goals if they have no impact at all? Surely some labs are going to feel a new pain come January 1st. The question is, will it be felt by your laboratory?

What predicts your PT future? The analytical Sigma metric

One simple technique to predict PT failure is to calculate the analytical Sigma metric of your method. This calculation has been around for over 20 years and doesn't require any additional resource or study. You simply tap into the performance data you are already collecting — and many of the major control vendors have already built it into their QC or peer comparison software.

The Six Sigma approach has been around even longer. While it's long past the peak of its popularity, when it was closer to a cult than a management technique (Remember black belts, master black belts, and champions?), the core utility from Six Sigma is still there. A universal scale of zero to six, where Six Sigma is the most desirable outcome.

There are three ingredients to the analytical Sigma metric:

- Your cumulative imprecision — something like intermediate reproducibility, basically a few months of your QC performance.
- Your bias — which you can obtain several ways. Using peer comparison software (the difference between your cumulative mean and the peer group cumulative mean) is a convenient way. But you could also use the very PT surveys that we're concerned about — again finding the difference between your survey results and those of your peer group. If you're lucky and well-funded, you might be able to benchmark your method in an accuracy-based survey, or test some reference materials, or compare against a reference method — these ways will give you a more traceable, "true" bias.
- The allowable total error (TEa). This is not something from your laboratory — it's from CLIA. These are the requirements that just changed.

These variables get arranged in this order:

$$\text{Analytical Sigma metric} = [\text{TEa}\% - |\text{bias}\%|] / \text{CV}\%$$

[You can also work this equation out in units instead of percentages, where the CV is replaced by the SD.]

Think of it this way: CLIA gives you a target to hit (TEa), and your imprecision and bias determine whether or not you hit the bullseye or miss the target completely. If you achieve Six Sigma, you've hit the bullseye and achieved world class performance and there's no danger of PT failure in your future. Five Sigma means excellent performance and also no worry of PT failure. Four Sigma is good performance, with PT failure

unlikely. Three Sigma is considered the minimum acceptable performance by theory, but if you achieve it, your PT failure worries are still very low. Because CLIA allows a 20% failure rate, the Sigma metric you worry about doesn't emerge until 2.3 or lower. Sigma metrics of 2.3 and lower indicate higher likelihood of future PT failures.

Who will suffer most under the new CLIA PT rules?

Since 2022, when the changes were announced, Westgard QC has analyzed real-world data from hundreds of instruments across the United States. What we've found confirms what every laboratory suspects: methods, instruments, and laboratories are not created equally.³ While it's convenient for those at administrative and executive levels to pretend that all diagnostic manufacturers are the same, and analytical quality is simply a commodity (thereby justifying the selection of the cheapest box), those closer to the bench level know the ground truth: some instruments are better than others, some methods are more imprecise than others, and some laboratories are plagued by QC and PT failures more than others.

The pain of the tighter CLIA PT requirements will not be felt equally across laboratories. If you choose the wrong method or instrument, you'll see an increase in PT failures in 2025. However, if you've selected a method or instrument with high Sigma quality, these new PT criteria may not even make you blink.

For poor-quality methods, the suffering (and PT failures) will stretch out over time. It will manifest as a slow-motion disaster, where in any given survey, the low Sigma method might be spared failure because the PT samples didn't test the vulnerable part of the range, where one failure in PT is not automatically followed by another PT failure in the next survey, because the samples and levels tests change with each survey. Labs with poor instruments and methods have one management approach for PT: hoping for good luck.

For high-quality methods, there's an extra bonus: knowing the analytical Sigma metric allows you to streamline your QC. Using Westgard Sigma Rules,⁴ a 2019 evolution of the

Westgard Rules introduced back in 1981, you can reduce the number of rules, levels, and even (if you're in the mood for advanced strategies) frequency of QC, as your analytical Sigma metric increases (See Figure 2). If you find your assays are achieving Six Sigma, you can even stop using Westgard Rules altogether and rely on something as simple as a 1:3s control rule.

One last question: Are you ready? Or are you just hoping to be lucky?

All the tools to help you predict your PT failures are available free. Most of the proficiency testing providers allow you to review your old survey data and manipulate it to impose different performance specifications. But if you don't enjoy resurrecting old data to run a simulation, you can use your current data to estimate the Sigma metric today.

Remember, any Sigma metric over 2.3 will be unlikely to face more failures in PT surveys. But while you breathe a sigh of relief over the surveys, remember that you can take advantage of higher Sigma metrics to reduce your overall QC effort.

This year's PT constriction doesn't have to become a crisis: it can also be the opportunity for your laboratory to bring its QC practices into the 21st century.

At Westgard QC, we've spent the last two years accumulating real-world data from hundreds of instruments across the diagnostic landscape. We have identified which instruments are going to be happy on January 1st and which ones are going to have a bad hangover. If you're interested in participating in our national performance database (and seeing how your results compare to the national spectrum), don't hesitate to reach out to benchmarking@westgard.com. 📧



Scan code to go directly to the CE test.

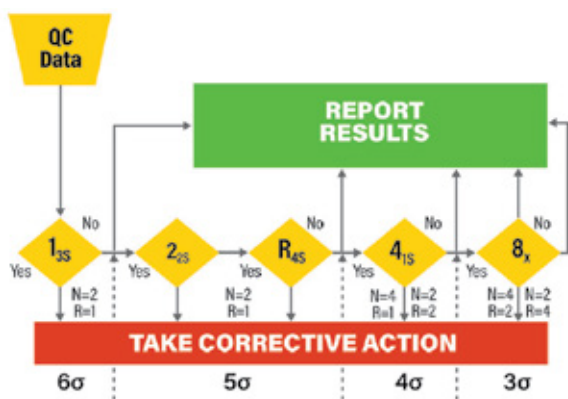


Figure 2. In this Westgard Sigma Rules flowchart, the traditional Westgard Rules are cut into sections by dashed vertical lines, each one designated at the bottom by a Sigma metric. For example, the 6 Sigma vertical line means that only the 1:3s rule is necessary for QC. As the Sigma metric declines, more Westgard Rules are needed, until 3 Sigma is reached. At that point the minimum acceptable quality (3 Sigma) requires the maximum Westgard Rules. Below 3 Sigma? The Westgard Rules can't save you — you'll need all of them and more — and those are also the assays that are vulnerable to future PT failure.

References

- Centers for Medicare & Medicaid Services. Clinical Laboratory Improvement Amendments of 1988 (CLIA) proficiency testing regulations related to analytes and acceptable performance. *Federal Register*. Published July 11, 2022. Accessed November 6, 2024. <https://www.federalregister.gov/d/2022-14513>.
- Westgard S. 2024 CLIA Acceptance Limits for Proficiency Testing - Westgard. Westgard.com. Accessed November 6, 2024. <https://westgard.com/clia-a-quality/quality-requirements/2024-clia-requirements.html>.
- Westgard S. Beckman Coulter Dxl 9000 immunoassay analyzer assay performance meets new CLIA 2024 PT goals with highest proportion of Six Sigma performance amongst manufacturers. Presented at: ADLM 2024. https://www.researchgate.net/publication/383058141_Beckman_Coulter_Dxl_9000_immunoassay_analyzer_assay_performance_meets_new_CLIA_2024_PT_goals_with_highest_proportion_of_Six_Sigma_performance_amongst_manufacturers.
- Westgard JO, Westgard SA. Establishing Evidence-Based Statistical Quality Control Practices. *Am J Clin Pathol*. 2019;1;151(4):364-370. doi:10.1093/ajcp/aqy158.



Sten Westgard is Director of Client Services and Technology for **Westgard QC**. Sten has presented at conferences and workshops worldwide on topics such as the "Westgard Rules," IQCP, Six Sigma, method validation, and quality management, and received an Outstanding Speaker award from the AACC for 2008 through 2012, and 2016. Sten is an adjunct faculty member at the Mayo Clinic School of Health Sciences in Rochester, Minnesota; an adjunct faculty member at the University of Alexandria, Egypt; an adjunct visiting faculty member at Kastruba Medical College of Manipal University, Mangalore, India; and an honorary visiting professor at Jiao Tong University, Shanghai.



STRECK

The lab moves pretty fast.



If you don't stop and think about the accuracy of your ESR testing, you might miss important results.

With the Diesse MINI-CUBE, sed-rate testing is simple. Load your samples in the instrument. While it's (accurately) measuring the sed-rate, you do your other work. It seems too easy, but then again, shouldn't it be?



Learning more couldn't be easier!
Watch my video and request a demo.

streck.com

