

Westgard Rules versus Sigma Metrics: Which Better Reflects Analytical Quality? — A Comparative Review

Anush Nazar¹, Simi William², Rajesh Thomas³, Jasmin Anush⁴

¹Anush Nazar, Faculty in MLT, SNVHSS Cherthala, Kerala India-688524, Email: anushnazar@gmail.com

²QC Officer, Maternity Hospital, Al Sabab, MOH, Kuwait. Email: simijohnson777@gmail.com

³General Manager (Laboratory), G G Hospital, Pattom, Thiruvananthapuram, Kerala India -695004 Email: rajeshbteekoy@gmail.com

⁴QC Manager, Sanjeevani Labs, Cherthala, Kerala India-688524, Email: jasmianush@gmail.com

Vol. 1 | Issue 3 | 2026 | pp. 17-23

Received: 08-Jan-2026

Accepted: 15-Mar-2026

Published: 23-Jul-2026

DOI: <https://doi.org/10.5281/zenodo.22078706>

P-ISSN: 3117-8065 | E-ISSN: 3117-8073

Address for Correspondence

Anush Nazar,
Faculty in MLT, SNVHSS Cherthala,
Kerala India-688524,
Email: anushnazar@gmail.com

How to cite this article: Nazar A, William S, Thomas R, Anush J. Westgard Rules versus Sigma Metrics: Which Better Reflects Analytical Quality? — A Comparative Review. ICMEJLM. 2026;1(3) p17-23. DOI: 10.5281/zenodo.22078706

ABSTRACT

Background: Internal quality control (IQC) underpins the analytical reliability of laboratory results. Westgard multirule QC and Six Sigma metrics are the two dominant frameworks used to evaluate analytical quality, yet they are frequently, and inaccurately, presented as competing alternatives rather than as tools that answer fundamentally different questions.

Objective: To critically compare Westgard Rules and Sigma Metrics as reflections of analytical quality, and to determine whether either framework, alone, adequately captures it.

Main findings: Across precision, accuracy, total analytical error, error-detection capability, false rejection, cost, automation and applicability, the two frameworks address complementary rather than overlapping dimensions of quality. Sigma metrics quantify the inherent analytical capability of a method relative to a defined total allowable error (TEa), enabling cross-laboratory benchmarking and evidence-based allocation of quality-control resource, but cannot themselves detect an out-of-control run. Westgard Rules provide real-time, run-level monitoring with a well-characterised statistical basis, but applied uniformly across a test menu they are statistically inefficient, over-controlling high-capability methods and under-controlling poor ones. Published evidence, spanning clinical chemistry, haematology, coagulation and blood-gas analysis, consistently shows that Sigma-based selection of Westgard rules, control frequency and control numbers improves the efficiency and defensibility of IQC programmes, while unresolved variability in TEa sources continues to limit cross-study comparability.

Conclusion: Neither Westgard Rules nor Sigma Metrics alone fully reflects analytical quality. Sigma metrics evaluate what a method is capable of; Westgard Rules monitor whether that capability is being sustained from day to day. The strongest evidence, and current international guidance, supports their integration within a risk-based quality control plan rather than treating either as a stand-alone indicator of quality.

Keywords: Six Sigma; Westgard rules; internal quality control; total allowable error; analytical quality; risk-based quality control

1. INTRODUCTION

Analytical quality is the foundation on which clinical decision-making in laboratory medicine rests: an estimated two-thirds of major clinical decisions draw directly on laboratory results, so undetected analytical error carries a disproportionate capacity for harm [1,2]. Internal quality control (IQC) is the mechanism by which laboratories detect such error before results are released, and the design of an IQC programme — which rules to apply, how many

control materials to run, and how often — is one of the most consequential, and most debated, decisions in laboratory quality management [3].

Two frameworks dominate this debate. Westgard multirule QC, introduced by Westgard, Barry, Hunt and Groth in 1981, combines several statistical control rules into a single run-acceptance algorithm designed to maximise error detection while limiting false rejection [3,4]. Six Sigma, first applied to laboratory medicine by Nevalainen

and colleagues in 2000, instead expresses a method's analytical capability as a single benchmarkable number — the Sigma metric — derived from its total allowable error, bias and imprecision [5,6]. Both frameworks are taught extensively, embedded in laboratory information systems and quality-management software, and referenced side by side in continuing-education material, which has fostered a persistent — and often unproductive — framing of the two as rival philosophies of quality [7,8].

This framing matters practically as well as conceptually. Laboratories that select Westgard rules without reference to Sigma performance risk two simultaneous failures: unnecessarily strict monitoring, and consequent waste of control material and staff time, on methods that are already highly capable; and inadequate monitoring of methods that are not [7,9]. Conversely, laboratories that rely on Sigma metrics as a stand-alone quality indicator overlook that a Sigma value describes long-term capability and provides no mechanism for detecting an out-of-control run occurring today [7,10]. Compounding the difficulty, the total allowable error (TEa) values on which Sigma metrics depend are drawn from at least four incompatible sources — regulatory (CLIA), biological-variation-based (Ricos/EFLM), state-of-the-art, and outcome-based — which frequently disagree for the same analyte, undermining the comparability that makes Sigma metrics attractive in the first place [11,12,13,14]. International guidance, including ISO 15189:2022 and IFCC practical recommendations, requires laboratories to justify IQC design through documented risk assessment, yet offers no single

Table 1. Westgard Rules versus Sigma Metrics: conceptual comparison

Attribute	Westgard Rules	Sigma Metrics
Purpose	Accept/reject a specific analytical run	Quantify long-term analytical capability relative to TEa
Statistical basis	Control-chart theory; power functions (Ped, Pfr) [17,18]	Process-capability index; Sigma = (TEa – Bias)/CV [5,6]
Underlying philosophy	Process monitoring — is the process behaving as validated, today?	Process capability — can the process, in principle, meet the requirement?
Primary quality objective	Prevent release of results from an out-of-control run	Benchmark and prioritise methods for QC intensity and improvement
Time horizon	Real time, run-by-run	Longitudinal (reviewed periodically, e.g. per lot/quarter)
What a poor result means	This run cannot be trusted today	This method may never reliably meet its quality requirement without redesign

These are not two measurements of the same thing expressed differently; they are answers to two different questions that happen to be asked of the same analytical process. This distinction is the organising principle for the comparative evaluation that follows.

3. COMPARATIVE EVALUATION OF ANALYTICAL QUALITY

This section evaluates, parameter by parameter, how well each framework reflects analytical quality, drawing on evidence published predominantly since 2015.

Precision and imprecision

Sigma metrics incorporate imprecision directly, as the denominator of the Sigma equation; a method's CV, estimated from long-term IQC data, is one of only two inputs (with bias) that determine its Sigma classification [5,6]. Westgard Rules do not quantify imprecision at all — they detect its consequences (an R4s violation signals excess random error) but provide no numerical estimate of the method's underlying imprecision. For evaluating precision as a property of a method, Sigma metrics are unambiguously the more informative tool; Westgard Rules are a downstream detector, not a measure, of imprecision [5,7].

Accuracy, bias and trueness

prescriptive answer to how Westgard Rules and Sigma Metrics should be combined in that assessment [15,16].

This review critically compares Westgard Rules and Sigma Metrics in evaluating analytical quality and discusses whether they should be viewed as competing or complementary quality tools.

2. CONCEPTUAL DIFFERENCES

Westgard Rules are a set of statistical decision criteria — 1s, 2s, R4s, 41s, 10x̄, and the 12s warning rule — applied to control observations within and across runs to decide whether a run should be accepted or rejected [3,17]. Their statistical basis is control-chart theory: each rule, and each combination of rules, has a characterisable probability of error detection (Ped) and false rejection (Pfr) that can be plotted as a power function [17,18]. Their objective is run acceptance: is this specific set of results, produced today, trustworthy enough to release?

Sigma Metrics instead express analytical performance on the industrial Six Sigma scale via $\text{Sigma} = (\text{TEa} - |\text{Bias}|) / \text{CV}$, where TEa is the allowable total error for the analyte, Bias is systematic deviation from a reference or peer-group target, and CV is long-term imprecision [5,6]. Their statistical basis is process-capability analysis, not control-chart theory. Their objective is capability assessment: is this method, in principle, good enough to be used for its intended clinical purpose, and how much monitoring does it need?

The same asymmetry holds for bias. Sigma metrics require an explicit bias estimate, typically from external quality assessment (EQA)/proficiency-testing peer-group comparison, and the Quality Goal Index ($\text{QGI} = \text{Bias} / (1.5 \times \text{CV})$) built on this framework tells a laboratory whether bias or imprecision is the dominant driver of a low Sigma value [18,19]. Westgard Rules detect bias only indirectly, through sustained shifts captured by 2s, 41s or 10x̄ rules, and cannot distinguish a run-to-run shift from a persistent method bias without longitudinal review [3,17]. Sigma metrics therefore better reflect accuracy as an intrinsic method property; Westgard Rules better reflect whether bias has changed since the method was last characterised.

Total analytical error and measurement uncertainty

Sigma metrics are built directly on the total error (TE) model, and inherit its central controversy: TE combines bias and imprecision into a single tolerance limit, whereas the internationally harmonised measurement uncertainty (MU) framework treats identified bias as correctable through traceability rather than an error component to be tolerated [20,21,22]. Panteghini and Sandberg have argued the two frameworks can be reconciled through traceability-based bias correction [20,23]; Oosterhuis and Theodorsson maintain that TE, while pragmatically dominant, lacks the rigour and harmonisation of

the MU model [21]; Kallner has questioned whether combining trueness and precision into a single expression is conceptually sound at all [22]. Westgard Rules are agnostic to this debate entirely — they operate on raw control observations and do not depend on any particular error model. In this narrow sense, Westgard Rules are more theory-neutral, but this is also a limitation: they offer no mechanism for expressing how much error a result may plausibly contain, only whether control observations have crossed a limit.

Error-detection capability and false rejection

Error-detection probability (Ped) and false-rejection probability (Pfr) are native, well-characterised properties of Westgard rule combinations, calculable via power-function analysis for any given N (controls per run) and R (control levels) [17,18]. Sigma metrics have no direct Ped/Pfr equivalent — but they determine, indirectly, how demanding the Ped/Pfr trade-off needs to be: a high-Sigma method can tolerate simple rules with modest Ped because genuine errors of clinically important size are rare, while a low-Sigma method requires rule combinations with high Ped even at the cost of increased Pfr, because the margin for error is small [7,9]. Evidence from Sigma-metric implementation studies consistently shows that matching rule stringency to Sigma performance reduces unnecessary false rejections on high-performing analytes while tightening detection on poor performers — an efficiency gain neither framework can achieve alone [7,9].

Patient risk and clinical decision-making

Neither framework measures patient risk directly, but each contributes a different risk signal. Sigma metrics identify which analytes carry the greatest inherent risk of undetected error before any error occurs, supporting proactive risk stratification consistent with CLSI EP23-A risk-based QC planning [24,25]. Westgard Rules provide the operational safeguard that prevents an identified out-of-control run from reaching a clinician [3]. Blood-gas evidence illustrates the danger of relying on either signal alone: analysers that passed conventional total-error acceptance criteria for pH nonetheless showed markedly poor Sigma performance, meaning a method could appear "in control" by simple accuracy checks while still carrying substantial capability-related risk that only a Sigma calculation would reveal [26].

Table 2 (summary). Net comparative assessment by parameter

Parameter	Better reflected by
Precision (imprecision, CV)	Sigma Metrics — direct numerical input
Accuracy / bias / trueness	Sigma Metrics — explicit bias term; QGI diagnostic
Total analytical error / measurement uncertainty	Sigma Metrics quantify it; Westgard Rules are theory-neutral
Error-detection capability (Ped)	Westgard Rules — native, calculable property
False rejection (Pfr)	Westgard Rules — native, calculable property
Patient risk stratification (proactive)	Sigma Metrics
Patient risk mitigation (reactive, daily)	Westgard Rules
Cost-effectiveness at scale	Sigma-informed Westgard QC (integrated) > either alone
Ease of implementation	Westgard Rules — lower infrastructure requirement
Suitability, small laboratories	Westgard Rules more robust; Sigma feasible with standardisation
Suitability, high-volume/reference laboratories	Sigma Metrics add substantial efficiency value

No single row favours one framework across every relevant dimension of quality, and several — cost-effectiveness and patient-risk management in particular — are best served only by their combination. This is the empirical basis for the integration argument developed in Section 5.

Table 3 synthesises 17 key comparative and applied studies published largely since 2018. The pattern that emerges is consistent: studies that evaluate Sigma metrics alone confirm their value for benchmarking and prioritisation but do not, by themselves, address run-level error

Cost-effectiveness

Sigma-based QC design has repeatedly been shown to reduce control material consumption and staff time by concentrating intensive monitoring on genuinely poor-performing analytes rather than applying uniform multirule QC across an entire menu [7,9]. Multi-country comparative studies report substantial reductions in unnecessary control runs once Sigma-tiered QC replaces uniform Westgard protocols [27,28]. Westgard Rules, applied without Sigma stratification, are the less cost-effective approach at scale, though they remain the only mechanism actually executing that (Sigma-informed) monitoring once designed.

Ease of implementation and automation

Westgard multirule logic is built into virtually all laboratory information systems and analyser middleware, requiring minimal additional infrastructure [3,17]. Sigma-metric calculation requires reliable long-term CV data, a defensible bias estimate from EQA/PT participation, and an explicit, disclosed TEa source — a higher implementation bar, though increasingly met by commercial Sigma-dashboard tools that automate the calculation from existing IQC/EQA data streams [7,10]. For laboratories without such tools, manual Sigma calculation and rule selection remains feasible but labour-intensive [27,29].

Applicability in small versus high-volume laboratories

High-volume and reference laboratories are well suited to Sigma-based approaches: large data volumes support robust CV and bias estimation, and the resource savings from Sigma-tiered QC scale favourably with test volume [7,9]. Small laboratories face a harder trade-off — Westgard Rules remain implementable regardless of volume, but limited EQA/PT participation options and shorter IQC data histories can make Sigma estimates statistically unstable [27,29,30]. Several comparative studies from resource-limited settings have nonetheless found Sigma metrics feasible and valuable even at modest volumes, provided TEa source and calculation method are explicitly standardised across the network [27,29,30].

4. EVIDENCE FROM PUBLISHED LITERATURE

detection; studies that evaluate integrated Sigma-based Westgard QC report efficiency gains (fewer unnecessary control runs, better-targeted rule stringency) without loss of error-detection performance; and multi-source TEa comparisons consistently demonstrate that Sigma classification is sensitive to the TEa source chosen, regardless of laboratory setting or country.

Table 3. Summary of key published comparative studies (2018–2026)

Author, year	Design / setting	Main finding	Favours
Westgard, Bayat, Westgard, 2018 [6]	Methodological review	Sigma Method Decision Charts and QGI formalise capability-based QC design	Integrated use
Guo et al., 2018 [31]	Comparative methodology, clinical chemistry	Two Sigma-calculation approaches yield materially different results for the same data	Sigma Metrics (with caution)
Kumar & Mohan, 2018 [32]	Single-centre, India, chemistry	Sigma metrics identify IQC performance gaps not evident from Westgard-rule compliance alone	Sigma Metrics
El Sharkawy et al., 2018 [28]	Multi-centre, Egypt (4 labs)	Substantial inter-laboratory Sigma variation for identical assays	Integrated use (harmonisation needed)
Westgard & Bayat, 2025 [10]	Reply/methodological clarification, immunoassay	Sigma-based rule recommendations are not expected to alter underlying method performance, only monitoring intensity	Integrated use
van Heerden, George, Khoza, 2022 [27]	Retrospective, South Africa, 19 chemistry analytes	Sigma classification varied from 21–53% "acceptable" analytes depending on TEa source (Ricos/RCPA/CLIA/EFLM)	Neither alone (TEa harmonisation needed)
Ozdemir & Ucar, 2022 [29]	Haematology analyser, CBC parameters	Sigma values for the same analyser differ materially across TEa sources; calls for CBC-specific harmonisation	Neither alone
Othmani et al., 2026 [33]	Prospective, Tunisia, paediatric chemistry	TEa source variability produces inconsistent QC-strategy recommendations from the same IQC data	Integrated use (with harmonisation)
Erdogan-Doventas et al., 2024 [19]	Coagulation laboratory	QGI alongside Sigma identifies whether bias or imprecision limits performance, directing improvement	Sigma Metrics
Hamid et al., 2025 [26]	Comparative, 3 blood-gas analysers	Passing total-error acceptance criteria did not guarantee acceptable Sigma performance for pH	Sigma Metrics (as risk signal)
Ilardo et al., 2024 [34]	Method verification workflow	Sigma metrics embedded within a risk-based method verification process, applied prospectively	Integrated use
Liu et al., 2026 [35]	Multi-centre, serum lipids	Six Sigma model used to design differentiated QC strategy across analytes and centres	Integrated use
Pham, 2025 [36]	Single-analyser evaluation, biochemistry	Sigma-based quality indicator evaluation on a routine chemistry analyser	Sigma Metrics
Rinaldi et al., 2024 [30]	Comparative, bias-source variation	Sigma values shift meaningfully depending on the bias source used in calculation	Neither alone
Cervinski et al., 2023 [37]	Methodological review, PBRTQC	Patient-based real-time QC detects between-run error inaccessible to interval-based Westgard QC	Integrated use (with PBRTQC)
Ma et al., 2025 [38]	Tertiary hospital, China, immunoassay	PBRTQC flagged a reagent-lot problem that conventional IQC had not detected	Integrated use (with PBRTQC)
Giannoli et al. (IFCC TF-GLQ), 2025 [16]	Guidance/consensus	IFCC practical recommendations for IQC under ISO 15189:2022 endorse risk-based, capability-informed rule selection	Integrated use

No study in this set reports Westgard Rules alone, applied without reference to method capability, as superior to a Sigma-informed or integrated approach on any efficiency or defensibility metric; conversely, no study reports Sigma metrics alone as an adequate substitute for run-level monitoring. The evidence base, while still dominated by single-centre and cross-sectional designs, is directionally consistent with integration rather than competition.

5. INTEGRATION RATHER THAN COMPETITION

The comparative evidence in Sections 3 and 4 converges on a single structural point: Sigma Metrics evaluate analytical capability — what a method can, in principle, achieve relative to its quality requirement — while Westgard Rules monitor analytical performance — whether that capability is being sustained on a given day [5,7,10]. A capability assessment cannot substitute for daily monitoring, and daily monitoring designed without reference to capability is statistically inefficient at best and inadequate at worst [7,9]. This is not

a compromise position reached for want of a clear winner; it is the specific relationship described by Westgard's own Method Decision Chart / OPSpecs framework, in which a method's Sigma-determined operating point directly specifies the rules, N and R required to monitor it appropriately [6,7].

Sigma-based QC planning

Sigma-based Westgard QC assigns rule combinations, control numbers and control frequency according to each analyte's demonstrated Sigma performance rather than a uniform default: analytes at Sigma ≥ 6 can be monitored with minimal rules and few controls; analytes at Sigma 4–5 warrant standard multirule QC; analytes below Sigma 3 require the full multirule scheme, increased control frequency, and — where Sigma remains persistently low despite optimisation — consideration of method replacement rather than ever-more-intensive QC of an inherently inadequate method [6,7,10].

Table 4. Recommended QC strategy by Sigma performance tier

Sigma tier	Suggested Westgard rule(s)	Typical N (controls/run)	QC intensity
≥ 6	1 _{3s} only	N = 2	Minimal — infrequent QC may suffice
5 – <6	1 _{3s} / 2 _{2s}	N = 2	Standard review each run

4 – <5	1s _s / 2s _s / R4s	N = 2–3	Standard multirule QC
3 – <4	1s _s / 2s _s / R4s / 4s	N = 3–4	Increased frequency; scrutinise trend patterns
<3	Full multirule + 10x̄/8x̄	N ≥ 4, every batch	Prioritise root-cause investigation and method redesign over QC intensification alone

Table 4. Synthesised from Sigma-based Westgard QC methodology [6,7,10]. Exact N, R and rule combinations should still be verified against the laboratory's own power-function or OPSpecs analysis.

Risk-based quality control and current guidance

CLSI EP23-A formalises a structured risk-management process for developing an individualised quality control plan (QCP), incorporating manufacturer risk information, regulatory requirements and the specific clinical setting — Sigma performance is one legitimate input into that broader risk assessment, not a stand-alone substitute for it [24,25]. ISO 15189:2022 requires laboratories to define acceptable limits for control results and apply statistically valid techniques to monitor performance, without prescribing a single method — a deliberate flexibility that IFCC's 2025 practical recommendations interpret as endorsing exactly this kind of capability-informed, risk-based rule selection [15,16]. Westgard's own current position, articulated most recently in reply to critics who expected Sigma-based rule recommendations to improve underlying method performance, is unambiguous: Sigma-based tools are designed to match monitoring intensity to demonstrated capability, not to substitute for the calibration, maintenance or method change actually required to raise that capability [10].

Algorithm: Sigma-guided selection of Westgard multirules

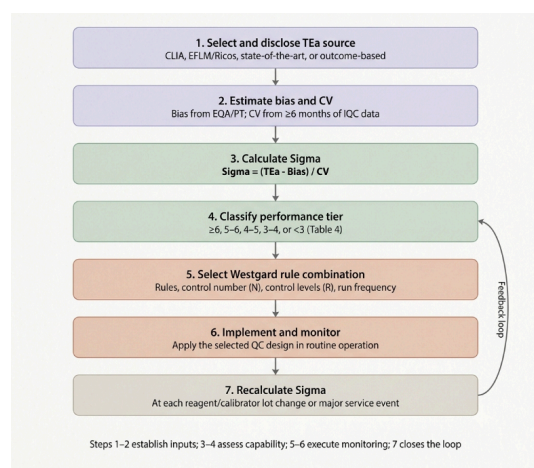


Figure 1 — Summarises the practical workflow: (1) calculate Sigma using a disclosed TEa source, current bias (EQA/PT) and long-term CV; (2) classify performance using Table 4; (3) select the corresponding rule combination, N and R; (4) implement and monitor; (5) recalculate Sigma at each reagent/calibrator lot change or significant service event, and adjust QC design accordingly. This loop operationalises the integration argument as a repeatable laboratory process rather than a one-off classification exercise.

6. CURRENT CHALLENGES

TEa source variability: CLIA, Ricos/EFLM biological-variation, RCPA and state-of-the-art specifications frequently disagree, and multiple recent studies show the same IQC/EQA data can be classified from "acceptable" to "unacceptable" depending on which is used [11,14,27,29,30].

Variability in Sigma calculation methodology itself: differing approaches to bias estimation and equation formulation can materially change reported Sigma values even before TEa source is considered [31,29].

Misinterpretation of Sigma values: a Sigma value calculated without disclosure of TEa source, bias-estimation method or CV

observation window is not meaningfully comparable across laboratories or over time [7,33].

Excessive dependence on Westgard Rules applied uniformly: a fixed rule set applied across an entire test menu without Sigma stratification over-controls high-capability analytes and under-controls poor ones [7,9].

Lack of harmonization: repeated calls across the literature for standardised, analyte-specific TEa specifications and standardised Sigma-calculation methodology remain, to date, only partially answered [11,14,28,27,29].

Resource-limited laboratories: constrained EQA/PT participation, shorter IQC data histories and limited LIS connectivity make robust Sigma estimation and automated Sigma-based rule selection harder to achieve, even though several studies from such settings demonstrate feasibility with explicit standardisation [27,29,30].

7. FUTURE PERSPECTIVES

Patient-based real-time quality control (PBRTQC) uses statistics derived from patient results themselves — moving averages, medians or the average-of-normal method — to provide near-continuous monitoring at negligible reagent cost, filling the between-run blind spot inherent to interval-based Westgard QC [37]. Recent case evidence shows PBRTQC detecting reagent-lot problems missed by conventional IQC, supporting its role as a complement to, not a replacement for, Sigma-informed Westgard QC [38]. Moving-average and average-of-normal approaches, in particular, offer a low-cost adjunct for electrolytes and other high-volume analytes where control material cost or supply is limiting [37]. Machine-learning approaches to PBRTQC algorithm optimisation are emerging, with IFCC working-group recommendations emphasising validation rigour and transparency before deployment [39]. Risk-based QC, as formalised in CLSI EP23-A and reflected in ISO 15189:2022, is likely to remain the overarching framework within which Sigma metrics, Westgard Rules and PBRTQC are jointly deployed, with digital quality-management dashboards increasingly automating the

Sigma-recalculation and rule-adjustment loop described in Section 5 [24,15,25,16].

8. CONCLUSION

Which better reflects analytical quality? The comparison developed throughout this review does not support a categorical answer, and the evidence does not point toward one. Sigma Metrics more directly reflect precision, bias and total-error-related capability, because these are their explicit numerical inputs; they provide a benchmarkable, proactive signal of which analytes carry the greatest inherent risk of undetected error. Westgard Rules more directly reflect error-detection performance and false-rejection behaviour in daily operation, because these are the properties for which their statistical basis was purpose-built; they provide the only mechanism, of the two, capable of stopping a specific erroneous run from reaching a clinician today.

Sigma Metrics primarily evaluate the analytical capability of a method — what it is fundamentally able to achieve relative to a defined quality requirement. Westgard Rules primarily monitor analytical performance during routine operation — whether that capability is being sustained from run to run. Neither, used alone, fully reflects analytical quality as required under ISO 15189:2022's risk-based framework: a method deemed capable by Sigma metrics still requires run-level monitoring, and a run judged acceptable by Westgard Rules provides no assurance that the underlying method is fit for its clinical purpose in the first place.

The evidence synthesised here — spanning clinical chemistry, coagulation, blood-gas and immunoassay applications, and consistent across single-centre and multi-centre studies from 2018 to 2026 — supports integrating Sigma Metrics with Westgard Rules within a documented, risk-based quality control plan, rather than treating either as a stand-alone indicator of quality or as competitors for the same role. This conclusion should not be overstated: the underlying evidence base remains dominated by cross-sectional, single-centre designs, and unresolved variability in TEa sources continues to limit how confidently Sigma metrics can be compared across laboratories. What can be said with reasonable confidence is narrower but still practically important: laboratories that select and calibrate Westgard rule intensity according to demonstrated Sigma performance, within a formal risk-based QC plan, appear consistently better positioned — on cost, efficiency and error-detection grounds alike — than laboratories that rely on either framework in isolation.

REFERENCES

1. Plebani M. Errors in clinical laboratories or errors in laboratory medicine? *Clin Chem Lab Med*. 2006;44(6):750-759.
2. Sciacovelli L, Panteghini M, Lippi G, Sumarac Z, Cadamuro J, Galoro CADO, et al. Defining a roadmap for harmonizing quality indicators in laboratory medicine: a consensus statement on behalf of the IFCC Working Group "Laboratory Errors and Patient Safety" and EFLM Task and Finish Group "Performance specifications for the extra-analytical phases". *Clin Chem Lab Med*. 2017;55(10):1478-1488.
3. Westgard JO, Barry PL, Hunt MR, Groth T. A multi-rule Shewhart chart for quality control in clinical chemistry. *Clin Chem*. 1981;27(3):493-501.
4. Westgard JO, Klee GG. Quality management. In: Burtis CA, editor. *Textbook of Clinical Chemistry*. 2nd ed. Philadelphia: WB Saunders; 1994. p.548-592.
5. Nevalainen D, Berte L, Kraft C, Leigh E, Picaso L, Morgan T. Evaluating laboratory performance on quality indicators with the six sigma scale. *Arch Pathol Lab Med*. 2000;124(4):516-519.
6. Westgard S, Bayat H, Westgard JO. Analytical Sigma metrics: a review of Six Sigma implementation tools for medical laboratories. *Biochem Med (Zagreb)*. 2018;28(2):020502. doi:10.11613/BM.2018.020502
7. Westgard S, Bayat H, Westgard JO. Special issue on Six Sigma metrics — experiences and recommendations. *Biochem Med (Zagreb)*. 2018;28(2):020301.
8. Westgard JO, Westgard SA. The quality of laboratory testing today: an assessment of sigma metrics for analytic quality using performance data from proficiency testing surveys and the CLIA criteria for acceptable performance. *Am J Clin Pathol*. 2006;125(3):343-354.
9. Koch DD, Oryall JJ, Quam EF, Feldbruegge DH, Dowd DE, Barry PL, Westgard JO. Selection of medically useful QC procedures for individual tests done on a multi-test analytical system. *Clin Chem*. 1990;36(2):230-233.
10. Westgard S, Bayat H. Reply to Cristelli et al: implementation of new Westgard Rules suggested by the Westgard Advisor. *Biochem Med (Zagreb)*. 2025;35(2):020401.
11. Sandberg S, Fraser CG, Horvath AR, Jansen R, Jones G, Oosterhuis W, et al. Defining analytical performance specifications: consensus statement from the 1st Strategic Conference of the European Federation of Clinical Chemistry and Laboratory Medicine. *Clin Chem Lab Med*. 2015;53(6):833-835.
12. Ricos C, Alvarez V, Cava F, Garcia-Lario JV, Hernandez A, Jimenez CV, et al. Current databases on biologic variation: pros, cons and progress. *Scand J Clin Lab Invest*. 1999;59(7):491-500.
13. Aarsand AK, Fernandez-Calle P, Webster C, Coskun A, Gonzales-Lao E, Diaz-Garzon J, et al, on behalf of the EFLM Working Group on Biological Variation. The EFLM Biological Variation Database [Internet]. Available from: <https://biologicalvariation.eu>
14. Clinical Laboratory Improvement Amendments (CLIA). Acceptable performance criteria for proficiency testing. Federal Register. 2024.
15. International Organization for Standardization. ISO 15189:2022 Medical laboratories — Requirements for quality and competence. 4th ed. Geneva: ISO; 2022.
16. Giannoli JM, Vassault A, Carobene A, Perret-Liaudet A, Blasutig IM, Dabla PK, et al, for the IFCC Task Force on Global Lab Quality. Ensuring internal quality control practices in medical laboratories: IFCC recommendations for practical applications based on ISO 15189:2022. *Clin Chim Acta*. 2025 (Epub ahead of print). PMID:40064457.
17. Westgard JO, Groth T, Aronsson T, Falk H, de Verdier CH. Performance characteristics of rules for internal quality control: probabilities for false rejection and error detection. *Clin Chem*. 1977;23(10):1857-1867.

18. Westgard JO, Barry PL. Cost-Effective Quality Control: Managing the Quality and Productivity of Analytical Processes. Washington DC: AACC Press; 1986.
19. Erdogan-Doventas Y, Ozemir-Anayurt E, Erdogan H, Kulan P, Koldas M. Performance evaluation of a coagulation laboratory using sigma metrics and quality goal index. *Clin Lab*. 2024;70. doi:10.7754/Clin.Lab.2024.240619
20. Panteghini M. Application of traceability concepts to analytical quality control may reconcile total error with uncertainty of measurement. *Clin Chem Lab Med*. 2010;48(1):7-10. doi:10.1515/CCLM.2010.020
21. Oosterhuis WP, Theodorsson E. Total error vs. measurement uncertainty: revolution or evolution? *Clin Chem Lab Med*. 2016;54(2):235-239. doi:10.1515/cclm-2015-0997
22. Kallner A. Is the combination of trueness and precision in one expression meaningful? On the use of total error and uncertainty in clinical chemistry. *Clin Chem Lab Med*. 2016;54(8):1291-1297. doi:10.1515/cclm-2015-0975
23. Panteghini M, Sandberg S. Total error vs. measurement uncertainty: the match continues. *Clin Chem Lab Med*. 2016;54(2):195-196. doi:10.1515/cclm-2015-1036
24. Clinical and Laboratory Standards Institute. EP23-A: Laboratory Quality Control Based on Risk Management; Approved Guideline. Wayne, PA: CLSI; 2011.
25. Nichols JH. Developing risk-based quality control plans: an overview of CLSI EP23-A. 2013. PMID:23331726.
26. Hamid H, Al-Wafi M, Ahmida MH, Elmansoury AM, Najah M. Beyond total error: an observational comparative study of sigma metrics in long-term arterial blood gas analyzer performance. *Cureus*. 2025;17(1):e77236. doi:10.7759/cureus.77236
27. van Heerden M, George JA, Khoza S. The application of sigma metrics in the laboratory to assess quality control processes in South Africa. *Afr J Lab Med*. 2022;11(1):a1344. doi:10.4102/ajlm.v11i1.1344
28. El Sharkawy R, Westgard S, Awad MA, AbdelKarem OI, El Hadidi I, Gaballah A, Shaheen E. Comparison between Sigma metrics in four accredited Egyptian medical laboratories in some biochemical tests: an initiative towards sigma calculation harmonization. *Biochem Med (Zagreb)*. 2018;28(2):020711. doi:10.11613/BM.2018.020711
29. Ozdemir S, Ucar F. Determination of Sigma metric based on various TEa sources for CBC parameters: the need for Sigma metrics harmonization. *J Lab Med*. 2022;46(2):133-141. doi:10.1515/labmed-2021-0116
30. Rinaldi SF, Ibadurrahmah AA, Ridwanna S, Harianto H. Comparison of the sigma metrics using the total error allowable algorithm with variation of bias source. *Indones J Med Lab Sci Technol*. 2024;6(1):27-35. doi:10.33086/ijmlst.v6i1.4930
31. Guo X, Zhang T, Gao X, Cheng X. Sigma metrics for assessing the analytical quality of clinical chemistry assays: a comparison of two approaches. *Biochem Med (Zagreb)*. 2018;28(2):020708. doi:10.11613/BM.2018.020708
32. Kumar BV, Mohan T. Sigma metrics as a tool for evaluating the performance of internal quality control in a clinical chemistry laboratory. *J Lab Physicians*. 2018;10(2):194-199. doi:10.4103/JLP.JLP_102_17
33. Othmani M, Amri Y, Chelbi S, Hadj Fredj S, Messaoud T, Dabboubi R. Discrepancies in sigma metrics driven by total allowable error variability: implications for QC strategy and laboratory efficiency. *J Appl Lab Med*. 2026;11(1):48-60. doi:10.1093/jalm/jfaf177
34. Ilardo C, Lamarti C, Al Muhanna B, Bastelica M, Benailly N. Sentinel testing, analytical sigma metrics, and a risk management approach as part of a simplified method verification/validation process. *Scand J Clin Lab Invest*. 2024 (Epub ahead of print). doi:10.1080/00365513.2024.2442512
35. Liu Q, Lin Y, Yang F, Dai Y, Hang H, Wang M, Hu M, Yang F. Application of the six sigma model to evaluate the analytical performance of serum lipid analytes and design quality control strategies: a multi-centre study. *Ann Clin Biochem*. 2026;63(1):3-13. doi:10.1177/00045632251350503
36. Pham VD. Six sigma in the evaluation of quality indicators using Roche Cobas c501 biochemistry analyzer. *J Lab Physicians*. 2025. doi:10.25259/JLP_285_2025
37. Cervinski MA, Bietenbeck A, Katayev A, Loh TP, van Rossum HH, Badrick T. Patient-based real-time quality control. *Adv Clin Chem*. 2023;117:223-261. doi:10.1016/bs.acc.2023.08.003
38. Ma C, Zhang Q, Hu Y, Ding W, Xia L, Qiu L. Application and insights on patient-based real-time quality control: detecting undetected errors in internal quality control through daily antibody positivity rate analysis. *Biochem Med (Zagreb)*. 2025;35(2):020801. doi:10.11613/BM.2025.020801
39. International Federation of Clinical Chemistry and Laboratory Medicine (IFCC) Working Group on Artificial Intelligence. Recommendations for best practice in the use of machine learning in laboratory medicine. 2023 [summarised in: *Diagnostics (Basel)*. 2024;14(16):1808. doi:10.3390/diagnostics14161808].