ISSN 0973-2063 (online) 0973-8894 (print)

Bioinformation 22(10): 7069-7076 (2026)

©Biomedical Informatics (2026)

BIOINFORMATION
Discovery at the interface of physical and biological sciences

BIOMEDICAL INFORMATICS

www.bioinformation.net

Volume 22(10)

Research Article

Received October 1, 2026; Revised October 31, 2026; Accepted October 31, 2026, Published October 31, 2026

DOI: 10.6026/973206300227069

SJIF 2026 (Scientific Journal Impact Factor for 2026) = 8.478
2022 Impact Factor (2023 Clarivate Inc. release) is 1.9

Declaration on Publication Ethics:
The author's state that they adhere with COPE guidelines on publishing ethics as described elsewhere at https://publicationethics.org/. The authors also undertake that they are not associated with any other third party (governmental or non-governmental agencies) linking with any form of unethical issues connecting to this publication. The authors also declare that they are not withholding any information that is misleading to the publisher in regard to this article.

Declaration on official E-mail:
The corresponding author declares that lifetime official e-mail from their institution is not available for all authors

License statement:
This is an Open Access article which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly credited. This is distributed under the terms of the Creative Commons Attribution License

Comments from readers:
Articles published in BIOINFORMATION are open for relevant post publication comments and criticisms, which will be published immediately linking to the original article without open access charges. Comments should be concise, coherent and critical in less than 1000 words.

Disclaimer:
Bioinformation provides a platform for scholarly communication of data and information to create knowledge in the Biological/ Biomedical domain after adequate peer/editorial reviews and editing entertaining revisions where required. The views and opinions expressed are those of the author(s) and do not reflect the views or opinions of Bioinformation and (or) its publisher Biomedical Informatics. Biomedical Informatics remains neutral and allows authors to specify their address and affiliation details including territory where required.

Edited by Ritik Kashwani
E-mail: docritikkashwani@yahoo.com
Citation: Halder *et al.* Bioinformation 22(10): 7069-7076 (2026)

A prospective study of DMAIC-based six sigma implementation for analytical quality improvement in a clinical biochemistry laboratory

Mrityunjoy Halder[1], Priyanka Datta[2], Abhra Ghosh[3,*] & Manthan Dhingra[4]

[1]Department of Biochemistry, All India Institute of Medical Sciences, Patna, Bihar, India; [2]Department of Biochemistry, Bankura Sammilani Medical College & Hospital, Bankura, West Bengal, India; [3]Department of Biochemistry, Mata Gujri Memorial Medical College & Lions Seva Kendra Hospital, Kishanganj, Bihar, India; [4]Department of Internal Medicine, St. Francis Medical Centre, Lynwood, California, USA; *Corresponding author

Affiliation URL:
https://www.mgmmckishanganj.in/

ISSN 0973-2063 (online) 0973-8894 (print)
Bioinformation 22(10): 7069-7076 (2026)
©Biomedical Informatics (2026)

Author contact:
Mrityunjoy Halder - E-mail: mhaldernbmc@gmail.com
Priyanka Datta - E-mail: dattapriyanka12@gmail.com
Abhra Ghosh - E-mail: abhraghoshcmc@gmail.com
Manthan Dhingra - E-mail: manthandhingra2000@gmail.com

Abstract:
Analytical errors in clinical biochemistry can compromise patient safety and clinical decision-making. Therefore, it is of interest to evaluate a DMAIC-based quality improvement programme for 20 serum chemistry analytes over 11 months. Baseline and endpoint data were compared using CLIA 2025 allowable total error, IQC-derived coefficient of variation, EQA-derived bias and Quality Goal Index analysis. Unacceptable analytes decreased from 11 to 2, while world-class and excellent performers increased from 0 to 4 and 0 to 3, respectively. Thus, data shows the DMAIC-based Six Sigma programme improved analytical performance and supported risk-based quality control planning for sustainable laboratory quality improvement.

Keywords: Six sigma, DMAIC, sigma metrics, quality goal index, internal quality control, Westgard sigma rules, clinical biochemistry, analytical quality, patient safety

Background:
Clinical biochemistry laboratories play a critical role in diagnosis, treatment and disease monitoring, with laboratory results influencing approximately 60-70% of clinical decisions **[1].** The total testing process is vulnerable to pre-analytical, analytical and post-analytical errors, with pre-analytical errors accounting for a substantial proportion of laboratory failures **[2, 3]**. These errors may lead to delayed diagnosis, inappropriate treatment; repeat testing and compromised patient safety **[4]**. Six Sigma metrics provide a quantitative approach to assess analytical performance using allowable total error (TE), bias and coefficient of variation (CV), with ≥6σ indicating world-class performance and <3σ indicating unacceptable performance **[5].** Studies from India and other countries have demonstrated considerable variation in sigma performance among routine biochemical analytes, with poor performance commonly related to imprecision or inaccuracy **[6, 7].** The Quality Goal Index (QGI) helps identify whether poor sigma performance is primarily due to imprecision, inaccuracy, or both, thereby guiding appropriate corrective actions **[7]**. The DMAIC (Define, Measure, Analyze, Improve and Control) framework provides a structured approach for identifying root causes, implementing corrective actions and sustaining improvements in laboratory quality **[8]**. Lean Six Sigma approaches have also demonstrated improvements in laboratory efficiency, including reductions in turnaround time, unnecessary retesting and sample rejection **[9].** Sigma-based risk stratification further enables analyte-specific Westgard QC strategies, allowing QC frequency and rules to be matched with analytical risk **[10-12].** Therefore, it is of interest to evaluate the effect of a structured DMAIC-based Six Sigma programme on the analytical performance of 20 serum chemistry analytes and to evaluate the resulting risk-based quality control strategies using Westgard Sigma Rules.

Materials and Methods:
Study design and setting:
This prospective, longitudinal quality improvement study was conducted in the Department of Biochemistry, Mata Gujri Memorial Medical College & Lions Seva Kendra Hospital, Kishanganj. The quality improvement intervention was implemented over an 11-month period October 2024 to August 2025, following the DMAIC (Define, Measure, Analyze, Improve, Control) framework of Six Sigma methodology. Baseline analytical performance was assessed using retrospective internal quality control (IQC) data from July to September 2024 and was designated as Set 1. Endpoint performance was assessed using IQC data from June to August 2025 and was designated as Set 2. Twenty routine serum chemistry analytes were evaluated at both time points, allowing matched longitudinal comparison of sigma performance before and after the structured quality improvement intervention. A pre-specified no-demotion criterion was applied throughout the study period, whereby no analyte was permitted to regress to a lower sigma category during the intervention and sustainment phases.

Analyte panel and instrumentation:
All 20 analytes were assessed at both time points, ensuring a complete matched longitudinal comparison. The panel comprised: Albumin, Alkaline Phosphatase (ALP), Alanine Aminotransferase (ALT), Amylase, Aspartate Aminotransferase (AST), Calcium, Creatine Kinase, Creatinine, Glucose, HDL Cholesterol, Iron, LDL Cholesterol, Magnesium, Phosphorus, Total Bilirubin, Total Cholesterol, Total Protein, Triglycerides, Urea and Uric Acid. All analytes were processed on a fully automated clinical chemistry analyzer, HumaStar 600 from HUMAN Gesellschaft für Biochemica und Diagnostica mbH, Germany using photometric, potentiometric and enzymatic-kinetic detection modalities. Reagents, calibrators and quality control (QC) materials were sourced from the same manufacturer as the analyzer system.

Quality control data collection:
IQC data were collected over a 3-month period - for each dataset - July to September 2024 for Set 1 and June to August 2025 for Set 2. Routine QC was scheduled twice daily; however, all valid unscheduled QC measurements (*e.g.,* post-calibration, post-maintenance and post-troubleshooting verification runs) were

also included to capture the true operational variance of the analytical phase. To calculate the Coefficient of Variation (CV %), data points were excluded only if they were associated with a documented assignable cause (*e.g.,* verified calibration failure, mechanical error, or documented operator reconstitution error) that resulted in run rejection. Statistical outliers or rule violations that resolved upon simple rerun without an identifiable root cause were retained in the dataset to accurately reflect the assay's inherent random error. These approaches yielded a total of 213 and 199 QC measurements, per level per analyte, respectively, for Set 1 and Set 2. CV (%), representing random error (imprecision), was calculated as: CV (%) = (SD / Mean) × 100. Bias (%), representing systematic error (inaccuracy), was derived from EQA peer-group comparison as: Bias (%) = |((Lab mean − Peer group mean) / Peer group mean)| × 100. The Allowable Total Error (TEa, %) for each analyte was sourced from CLIA 2025 guidelines. The sigma metric (σ) was calculated as **[5, 11]**: Sigma (σ) = (TEa − Bias) / CV. Sigma was reported independently for IQC L1 and L2. Performance categories were: World Class (≥6σ), Excellent (5-6σ), Good (4-5σ), Marginal (3-4σ) and Unacceptable (<3σ). For analytes failing the threshold, the Quality Goal Index (QGI) was calculated to identify the dominant error type and direct corrective action **[7]**: QGI = Bias / (1.5 × CV) QGI <0.8 indicates imprecision-dominant failure (focus on CV reduction); QGI >1.2 indicates inaccuracy-dominant failure (focus on bias reduction); and QGI 0.8-1.2 indicates mixed error requiring both.

Intervention structure - four phases:
Based on sigma assessment and Ishikawa root cause analysis findings, interventions were organised into four sequential, cumulative phases:

Phase 1 - Recognition (Sept-Nov 2024):
Baseline sigma and Quality Goal Index (QGI) were calculated for all 20 analytes. An Ishikawa (fishbone) diagram was constructed across six causal domains - Manpower, Method/SOP, Machine/Analyzer, Material, Measurement/Monitoring and Environment/Workflow. Each failing analyte was assigned a primary error category (imprecision-dominant, inaccuracy-dominant, or mixed) based on QGI classification, IQC trend pattern analysis and root cause review. This domain assignment determined which corrective action layer was prioritised for each analyte.

Phase 2 - Standardization (Dec 2024-Feb 2025):
As operator-dependent imprecisions were identified as the dominant QGI finding across the panel, different interventions were introduced. These included weekly micro-training sessions covering QC vial reconstitution, Westgard rule interpretation and corrective action documentation; laminated QC startup checklists at each workstation; shift-wise sign-off logs with mandatory corrective action notes; and SOP revision incorporating analyte-specific CAPA flowcharts. For analytes persistently below 3σ, stricter Westgard multirules were applied, using the $1_{2}s$ as a warning rule and the 10_{x} rule for shift detection.

Phase 3 - Technical stabilization (Mar-May 2025):
Interventions in this phase addressed analyzer performance and material quality management. Preventive maintenance was decoupled from calendar scheduling and linked instead to sigma review findings, with mandatory calibration verification introduced after every service event. A lot-to-lot verification protocol was implemented for critical analytes, requiring a minimum of 20 parallel runs and sigma recalculation before any new reagent lot was approved for deployment. QC material handling was standardised through single-use 500 µL aliquots, a 48-hour open-vial discard policy and FEFO reagent rotation. Following a multi-analyte cluster imprecision event, reagent-grade water conductivity monitoring was added to the daily quality indicator checklist.

Phase 4 - Sustainment (Jun-Aug 2025):
A monthly sigma dashboard review meeting was established displaying analyte-level sigma colour-coded by category, QGI classification, CAPA status and month-on-month trend direction. Levey-Jennings review was restructured as a proactive activity: any analyte showing a 2-SD trend over five consecutive QC runs triggered pre-emptive investigation before a formal Westgard rule violation occurred. CAPA effectiveness was formally assessed at 4-8 weeks post-implementation. Seasonal trend tracking was conducted throughout to distinguish genuine system gains from environmental confounders.

Data analysis:
The primary outcome was change in sigma category distribution across the full 20-analyte matched panel between Set 1 and Set 2. Secondary outcomes included: number of analytes achieving world-class sigma (≥6σ) at endpoint; absolute CV% change for imprecision-dominant analytes; and adherence to the no-demotion criterion. Performance categories were: World Class (≥6σ), Excellent (5-6σ), Good (4-5σ), Marginal (3-4σ) and Unacceptable (<3σ), consistent with Westgard Six Sigma benchmarks. Sigma-based QC strategy assignment at endpoint was performed by mapping each analyte's conservative sigma (lower of IQC L1 and L2) to the Westgard Sigma Rules framework, determining the applicable QC rule set, number of controls per run (N) and maximum patient run size, the latter derived from the Sigma-Metric Run Size Nomogram targeting a maximum expected number of unreliable patient results per QC event (Max E (Nuf)) of less than 1, as recommended by CLSI C24-Ed4. Descriptive statistics were applied throughout; no inferential statistics were used.

Results:
Sigma assessment of all 20 analytes at baseline revealed widespread analytical quality deficiency. Based on a conservative assessment of IQC sigma values 11 analytes were Unacceptable (<3σ): Albumin, ALP, AST, Creatine Kinase, Creatinine, Iron, Magnesium, Phosphorus, Total Cholesterol,

Urea and Total Protein. Five analytes were Marginal (3-4σ): Glucose, HDL Cholesterol, Total Bilirubin, Triglycerides and Uric Acid. Four analytes were Good (4-5σ): Amylase, ALT, Calcium and LDL Cholesterol. The complete baseline dataset is presented in **(Table 1)**. QGI analysis classified the majority of failing analytes as imprecision-dominant (QGI <0.8), establishing operator-dependent CV variability as the primary driver of sigma failure. Creatinine was the worst performer (L1 sigma 1.70; CV 4.98%), followed by Albumin (L1 sigma 1.73), Creatine Kinase (L1 sigma 2.48), Phosphorus (L1 sigma 2.15), ALP (L1 sigma 2.66) and Iron (L1 sigma 2.86). AST showed discordant inter-level performance (L1 sigma 2.55 vs. L2 sigma 4.03), consistent with concentration-dependent CV behaviour in a kinetic enzyme assay. Fishbone analysis identified Manpower and Material as the two highest-priority root cause domains, with inconsistent QC startup technique across shifts and inadequate lot-to-lot reagent verification as the most operationally actionable findings. Since QGI analysis identified imprecision as the dominant failure mode for 15 of the 20 analytes, Phase 2 interventions were implemented first and targeted the human and procedural sources of CV generation. The QC startup checklist was among the highest-impact interventions in this phase. QC log review over the four weeks following its deployment showed a decline in unexplained single-level IQC failures, although this finding should be interpreted cautiously as other operator-directed interventions were introduced concurrently. These failures had not previously been systematically investigated and were broadly attributed to random analytical variation. Retrospective review of QC logs and operator records suggested they were substantially operator-driven, arising from inconsistent vial inversion technique prior to aspiration, variable equilibration time after refrigerator removal, differing startup sequences between shifts, failure to allow the analyzer to complete its self-diagnostic cycle before QC aspiration, use of the same QC vial across multiple shifts beyond the recommended open-vial period, inconsistent mixing of reconstituted lyophilized material and variable QC aspiration timing relative to the start of the analytical run. Collectively, these findings highlighted that a significant proportion of what had been recorded as random imprecision was in fact structured, operator-attributable variability amenable to standardization. The weekly micro-training programme addressed a specific competency gap identified during root cause analysis: junior operators could recognise a Westgard rule violation but could not distinguish between a 1_3s random error rejection and a 2_2s systematic error rejection in terms of the appropriate investigative response. Following eight training sessions, QC log analysis showed that error-type-directed corrective actions replaced reflexive, non-directed reruns as the predominant response to QC failures. Reflexive reruns had been masking imprecision rather than correcting it, producing apparent QC acceptability while underlying CV remained elevated. The analyte-specific CAPA flowcharts in the revised SOP directed operators toward CV-relevant variables first for imprecision-dominant analytes, producing faster stabilization after QC failures and reducing the recurrence rate of the same failure pattern within a 30-day window. By end of Phase 2, Total Protein L1 CV had fallen from 2.50% to approximately 1.4% and Creatinine L1 CV from 4.98% to approximately 3.1%. Phase 3 addressed residual imprecision and bias that the Phase 2 human standardisation layer could not resolve. Performance-linked maintenance produced the most visible gains in kinetic enzyme assays. The sigma dashboard in March 2025 identified deteriorating CV trends in ALP and AST despite stable operator practices, traced to cuvette reaction chamber temperature variance drifting 0.4°C above tolerance. A targeted service event reduced ALP L1 CV from 6.06% to 4.57% by endpoint, accounting for its entire sigma gain from 2.66 (Unacceptable) to 3.52 (Marginal). The same maintenance intervention drove Creatine Kinase L1 CV from approximately 3.8% (mid-Phase 2) to 2.29% at endpoint, enabling its transformation from Unacceptable to World Class. The lot-to-lot verification protocol simultaneously prevented a demotion and enabled a world-class transformation. For Phosphorus, the April 2025 incoming reagent lot produced a parallel-run CV of 2.9% versus the established 2.1%; the transition was deferred, directly preventing a Marginal-to-Unacceptable demotion. For Triglycerides, the incoming April 2025 lot produced a parallel-run CV of 1.9% versus the then-current 2.8%; deployment of this superior lot combined with Phase 2 CV gains drove Triglycerides from baseline 3.22 (Marginal) to 7.63 (World Class) - the largest single-analyte sigma gain in the study. A multi-analyte cluster CV event in March 2025 across Glucose, Calcium, Phosphorus and Total Protein simultaneously was traced to reagent-grade water conductivity drift; correction within 10 working days validated water system monitoring as a permanent control measure. Post-service calibration verification corrected previously unrecognised post-maintenance calibration drift in Urea and Total Protein, enabling Urea's transformation from Unacceptable (2.94) to World Class (7.17). The Phase 4 sigma dashboard identified early downward trends in AST, Phosphorus and Glucose during June-July 2025 before any formal Westgard rule violation had occurred. For AST, declining sigma (L1 4.1 in May to 3.9 in June) was correlated with a recent operator rotation; a competency verification session stabilised AST at endpoint sigma 3.77 (Marginal). For Phosphorus, mild CV drift following its April lot transition was identified as early calibration drift; recalibration recovered L2 sigma to 4.15 (Good) at endpoint. For Glucose, a downward L2 trend correlated with elevated ambient temperature in June 2025; reinforcement of environmental temperature controls recovered L2 sigma to 3.80 (Marginal). In all three cases, pre-emptive intervention directly preserved the no-demotion criterion. The root cause-directed corrective and preventive actions implemented during the DMAIC cycle, along with their monitoring metrics and endpoint outcomes, are summarized in **(Table 2).** At study endpoint, every analyte across all 20 parameters improved or maintained its conservative sigma category, with zero demotions. The complete endpoint data are presented in **(Table 3)** and the distributional comparison in **(Table 4)**. Unacceptable analytes fell from 11 to 2; Marginal analytes decreased from 5 to 4; World Class performers increased from 0 to 4; and Excellent analytes

increased from 0 to 3. CV reduction was the dominant mechanism: Total Protein L1 CV fell 57% (2.50%→1.07%), Triglycerides L1 CV fell 58% (4.24%→1.79%), Urea L1 CV fell 59% (2.71%→1.11%), Creatinine L1 CV fell 56% (4.98%→2.21%) and Creatine Kinase L1 CV fell 62% (5.95%→2.29%). Albumin remained unacceptable at both levels, while ALP remained unacceptable at L2 despite improving to marginal at L1. Albumin (L1 sigma 1.76) was constrained by a narrow TEa of 8% with persistent mixed-error QGI and ALP (L1 sigma 3.52) by residual kinetic assay temperature sensitivity and unresolved EQA bias. Both showed measurable improvement from baseline, satisfying the no-demotion criterion. Based on the August 2025 sigma values, each analyte was assigned a QC strategy per the Westgard Sigma Rules framework using its conservative sigma (lower of IQC L1 and L2). The strategy specifies three operational parameters for each analyte: the applicable Westgard rule set, the number of control measurements per run (N) and the maximum patient run size - the maximum number of patient samples that may be processed between two consecutive QC events while maintaining a Max E(Nuf) of less than 1 (*i.e.,* fewer than one erroneous patient result expected per QC event, per CLSI C24-Ed4 guidance). The complete QC strategy for all 20 analytes is presented in **(Table 5)**. The run size column made the patient safety implication of sigma improvement directly quantifiable. At baseline, the 11 analytes classified as Unacceptable were each restricted to a maximum run size of 45 samples per QC event - meaning QC was required approximately every hour and even then the probability of releasing at least one erroneous result per QC interval remained non-trivially high. At endpoint, Albumin remained unacceptable by both levels; ALP remained unacceptable at L2 but improved to marginal at L1. Four analytes permitted a run size of up to 1000 samples under a single 1_{3}s rule - the minimum QC burden consistent with patient safety for world-class processes. The aggregate shift in permissible run size across the panel represented a substantial reduction in the number of QC events required per working day while simultaneously providing greater protection against undetected analytical errors for the analytes remaining in lower sigma tiers, through the application of more sensitive multi-rule procedures. Two analytes required continued vigilance at endpoint. Albumin remained unacceptable at both IQC levels and was therefore restricted to a maximum run size of 45 patient samples with intensified twice-daily QC monitoring. Its persistently low sigma performance indicated the need for a dedicated matrix-specific calibration investigation, particularly targeting the EQA bias component. ALP also required continued monitoring. Although it improved to marginal performance at IQC Level 1, it remained unacceptable at IQC Level 2. Accordingly, under the conservative sigma approach, ALP was retained in the unacceptable category and restricted to a maximum run size of 45 patient samples, with duplicate control levels and an intensified multirule QC procedure applied twice daily. The no-demotion outcome across all 20 analytes confirmed that the layered intervention architecture produced durable, system-wide quality improvement rather than selective optimisation of individual analytes at the cost of others.

Table 1: Baseline sigma performance - September 2024 (Set 1), All 20 analytes

Analyte	CV L1%	CV L2%	Bias L1	Bias L2	Sigma IQC L1	Quality L1	Sigma IQC L2	Quality L2
Albumin	2.95	2.50	2.90	2.68	1.73	UNACCEPTABLE	2.13	UNACCEPTABLE
ALP	6.06	5.22	3.88	5.44	2.66	UNACCEPTABLE	2.79	UNACCEPTABLE
ALT	3.26	2.24	1.80	4.94	4.05	GOOD	4.49	GOOD
Amylase	2.82	2.38	8.05	7.99	4.24	GOOD	5.05	EXCELLENT
AST	5.28	3.40	1.54	1.30	2.55	UNACCEPTABLE	4.03	GOOD
Calcium	2.17	2.07	0.63	1.14	4.32	GOOD	4.28	GOOD
Creatine Kinase	5.95	5.72	5.24	5.07	2.48	UNACCEPTABLE	2.61	UNACCEPTABLE
Creatinine	4.98	4.16	1.53	1.10	1.70	UNACCEPTABLE	2.14	UNACCEPTABLE
Glucose	2.02	2.15	0.91	1.31	3.51	MARGINAL	3.11	MARGINAL
HDL Cholesterol	4.92	4.35	4.21	4.86	3.21	MARGINAL	3.48	MARGINAL
Iron	4.80	4.42	1.27	1.39	2.86	UNACCEPTABLE	3.08	MARGINAL
LDL Cholesterol	3.04	3.81	2.88	1.98	5.63	EXCELLENT	4.73	GOOD
Magnesium	4.05	3.82	3.05	2.93	2.95	UNACCEPTABLE	3.16	MARGINAL
Phosphorus	4.28	4.05	0.79	1.10	2.15	UNACCEPTABLE	2.19	UNACCEPTABLE
Total Bilirubin	5.10	4.76	3.78	3.72	3.18	MARGINAL	3.42	MARGINAL
Total Cholesterol	3.14	3.09	1.30	0.88	2.77	UNACCEPTABLE	2.95	UNACCEPTABLE
Total Protein	2.50	2.43	1.60	1.39	2.56	UNACCEPTABLE	2.72	UNACCEPTABLE
Triglycerides	4.24	4.20	1.35	1.48	3.22	MARGINAL	3.22	MARGINAL
Urea	2.71	2.36	1.03	1.14	2.94	UNACCEPTABLE	3.33	MARGINAL
Uric Acid	1.72	1.62	4.37	4.11	3.28	MARGINAL	3.64	MARGINAL

Quality categories: World Class (≥6σ); Excellent (5-6σ); Good (4-5σ); Marginal (3-4σ); Unacceptable (<3σ). CV: coefficient of variation; IQC: internal quality control.

Table 2: Root cause-directed corrective and preventive actions implemented during the DMAIC quality improvement cycle

Issue (Sept 2024)	Root Cause	Action Taken	Monitoring Metric	Status by Aug 2025
Multiple analytes in <3 sigma / marginal range	QC handled inconsistently across staff/shift	Short-cycle retraining on QC handling, Westgard interpretation and corrective action documentation	Monthly analyte-wise CV%, QC rerun frequency	Sustained improvement; no demotion observed
Same analyte problems recurring despite repeated reruns	Troubleshooting not standardized; individual judgment	SOP updated with analyte-specific decision flow	Recurrence rate of QC failures per analyte; CAPA closure time	Recurrence reduced; faster stabilization

QC failure actions not properly recorded	Documentation gaps; no shift accountability	Shift-wise QC sign-off plus mandatory action note in QC log	% QC events with complete action documentation	Documentation improved; easier trend review
Low-sigma analytes not prioritized	Uniform review of all analytes; no risk-based focus	Sigma-based QC review and analyte prioritization	Monthly sigma dashboard; count by sigma category	Major uplift in low performers without sacrifice to better analytes
Intermittent QC instability despite acceptable maintenance	Maintenance not linked to QC performance trends	Maintenance-QC correlation review; PM aligned to analyte-specific instability	QC shifts/trends before vs after PM	Stability improved, especially in enzyme/kinetic assays
Bias remained high despite acceptable IQC	Reagent/calibrator lot effects; incomplete lot verification	Lot-to-lot verification protocol for key analytes	EQA bias%, lot change review log, delta checks	Bias improved in several analytes over successive cycles
QC material handling variation	Inconsistent practices during busy shifts	Standardized QC aliquoting, labeling, storage and discard policy	QC vial handling audit; unexplained QC excursions	Day-to-day variability reduced
EQA review done mainly for compliance, not action	Data available but not used for prioritization	Monthly/bi-monthly sigma + QGI dashboard review meeting	Meeting records, QGI categorization, CAPA actions initiated	More proactive interventions; no analyte demotion
Drift recognized late (after control failure)	Reactive review culture	Levey-Jennings trend/shift review before rule violation escalates	Trend flags identified early; number of pre-emptive actions	Early detection improved sustainment
CAPA actions done but effectiveness not checked	Closure culture weak; no follow-up interval	Add CAPA effectiveness review date and recheck criteria	% CAPA with effectiveness verification completed	Better sustainment and less recurrence

Table 3: Endpoint sigma performance - August 2025 (Set 2), All 20 analytes

Analyte	CV L1%	CV L2%	Bias L1	Bias L2	Sigma IQC L1	Quality L1	Sigma IQC L2	Quality L2
Albumin	2.56	2.18	3.48	2.68	1.76	UNACCEPTABLE	2.44	UNACCEPTABLE
ALP	4.57	5.37	3.91	5.45	3.52	MARGINAL	2.71	UNACCEPTABLE
ALT	2.97	2.53	1.78	2.15	4.45	GOOD	5.08	EXCELLENT
Amylase	2.39	2.81	8.04	8.01	5.00	EXCELLENT	4.27	GOOD
AST	3.57	3.04	1.54	1.29	3.77	MARGINAL	4.51	GOOD
Calcium	1.95	2.29	0.62	0.24	4.81	GOOD	4.26	GOOD
Creatine Kinase	2.29	2.69	5.48	3.24	6.34	WORLD CLASS	6.23	WORLD CLASS
Creatinine	2.21	1.89	1.54	1.10	3.83	MARGINAL	4.71	GOOD
Glucose	1.50	1.76	0.90	1.31	4.73	GOOD	3.80	MARGINAL
HDL Cholesterol	3.01	2.57	4.20	4.63	5.25	EXCELLENT	5.98	EXCELLENT
Iron	2.46	2.89	1.62	2.31	5.44	EXCELLENT	4.39	GOOD
LDL Cholesterol	2.60	2.22	2.89	2.00	6.58	WORLD CLASS	8.11	WORLD CLASS
Magnesium	2.30	2.70	1.02	1.50	6.08	WORLD CLASS	5.00	EXCELLENT
Phosphorus	2.01	1.72	2.82	2.86	3.57	MARGINAL	4.15	GOOD
Total Bilirubin	3.25	3.82	2.78	2.77	5.30	EXCELLENT	4.51	GOOD
Total Cholesterol	2.15	1.84	1.29	0.87	4.05	GOOD	4.96	GOOD
Total Protein	1.07	0.92	1.61	1.40	5.97	EXCELLENT	7.17	WORLD CLASS
Triglycerides	1.79	2.10	1.34	1.48	7.63	WORLD CLASS	6.44	WORLD CLASS
Urea	1.11	1.31	1.04	1.14	7.17	WORLD CLASS	6.00	WORLD CLASS
Uric Acid	1.56	1.43	3.68	3.29	4.05	GOOD	4.69	GOOD

Quality categories as per **Table 1**

Table 4: Sigma category distribution - Set 1 (Sept 2024) vs. Set 2 (Aug 2025)

Sigma Category	Sept 2024 (n)	Aug 2025 (n)	Change
World Class (≥6σ)	0	4	+4
Excellent (5-6σ)	0	3	+3
Good (4-5σ)	4	7	+3
Marginal (3-4σ)	5	4	−1
Unacceptable (<3σ)	11	2	−9
Total	20	20	-

N: Categories were assigned using the lower sigma category between L1 and L2.

Table 5: Westgard sigma-based QC Strategy - August 2025 (Set 2)

Sigma Tier	Category	QC Rules	N per Run	Max Run Size	Analytes (August 2025)
≥6σ	World class	1_{3s}	N = 2 (L1 + L2, once daily)	≤ 1000 samples per QC event	Creatine Kinase, LDL Cholesterol, Triglycerides, Urea
5-6σ	Excellent	1_{3s} / 2_{2s} / R_{4s}	N = 2 (L1 + L2, once daily)	≤ 450 samples per QC event	Magnesium, Total Protein, HDL Cholesterol
4-5σ	Good	1_{3s} / 2_{2s} / R_{4s} / 4_{1s}	N = 2 (L1 + L2, once daily)	≤ 200 samples per QC event	Amylase, ALT, Calcium, Iron, Total Bilirubin, Total Cholesterol, Uric Acid
3-4σ	Marginal	$1_{2.5s}$ / 2_{2s} / R_{4s} / 4_{1s} / 10_{x}	N = 4 (L1×2 + L2×2, once daily)	≤ 90 samples per QC event	Creatinine, Phosphorus, AST, Glucose
<3σ	Unacceptable	1_{2s} / 2_{2s} / R_{4s} / 4_{1s} / 10_{x}	N = 4-6 (duplicate levels, twice daily)	≤ 45 samples per QC event	ALP, Albumin

N: number of QC measurements per run; Max Run Size: maximum patient samples between consecutive QC events targeting Max E(Nuf) <1 per CLSI C24-Ed4 and Westgard Sigma-Metric Run Size Nomogram. Rules: $1_{3}s$ = reject if any control >±3 SD; $2_{2}s$ = 2 consecutive controls >same ±2 SD; $R_{4}s$ = range between L1 and L2 >4 SD; $4_{1}s$ = 4 consecutive >±1 SD; 10_{x} = 10 consecutive same side of mean **[12, 13]**.

Discussion:

The prospective, longitudinal design of our study provides a crucial advancement in clinical laboratory quality management by implementing a 11-month continuous DMAIC framework constrained by a strict "no-demotion" criterion. Systematic analyses of the clinical laboratory literature reveal that while Six Sigma is a foundational concept in Total Quality Management (TQM) **[14, 15]**, the majority of published studies rely heavily on retrospective performance evaluations rather than tracking sustained, prospective quality improvement **[2, 7 and 16-18]**. Even among quasi-experimental multiphase studies that deploy new quality control (QC) rules, such as those by Choi *et al.* **[19]**, or the two-phase evaluation by Singh *et al.* **[20]**, post-intervention periods frequently witness the performance demotion of certain analytes. In our study, the enforcement of a no-demotion policy uniquely prevented this regression, establishing a highly durable analytical quality loop. While some analytes experienced minor fluctuations in raw sigma scores due to natural process variation, the layered interventions ensured that no analyte suffered a categorical demotion. A critical methodological decision in Sigma computation is the selection of the Total Allowable Error (TEa) target, which our study standardizes against CLIA 2025 guidelines **[11]**. Comparative investigations consistently highlight the profound variability introduced by utilizing different TEa and biological variation databases **[21]**. For instance, Ramachandra *et al.* demonstrated that calculating Sigma values against the highly stringent EFLM database frequently relegates a much higher proportion of routine chemistries to unacceptable tiers compared to CLIA targets **[22]**. Furthermore, extensive reviews by Westgard emphasize the operational confusion generated by conflicting international TE standards **[23]**. By maintaining standard CLIA targets, our study establishes the consistent methodology necessary for accurate benchmarking, successfully avoiding the target-shifting artefacts often observed in multi-analyser cohorts and multicenter plasma protein evaluations **[24]**. At baseline, our study identified several analytes, including AST, creatine kinase, creatinine, total cholesterol, urea and total protein, as unacceptable performers using the conservative sigma approach. QGI analysis attributed most of these failures predominantly to random error, indicating imprecision as the major operational weakness. This specific vulnerability profile is widely corroborated across international cohorts. Retrospective assessments from comprehensive specialized hospitals in Ethiopia **[2, 17]**, institutions in Turkey **[7]** and diverse clinical settings in India uniformly flag these identical routine assays for marginal or unacceptable performance **[6, 21 and 25]**. Our proactive translation of these QGI findings into targeted retraining and root cause analysis improved several poor performers, with some analytes achieving world-class performance, without merely masking errors through reflexive, non-directed reruns. The major translational value of our Six Sigma model lies in its integration with risk-based Statistical Quality Control (SQC) design. By dynamically assigning improved analytes to tailored Westgard rule sets and calculated patient run sizes, including capped run sizes of 45 samples for unacceptable performers and 90 samples for marginal performers between QC events, our study converted the complex mathematical risk models proposed by Westgard and Parvin into a practical, operational quality-control framework for routine clinical biochemistry testing **[13, 26 and 27]**. The clinical necessity of adjusting QC run frequencies to limit the maximum expected error per run is strongly echoed by Peng *et al.* **[5]** and Zhou *et al.* **[28]**, who similarly advocated for massive run sizes for world-class analytes alongside stringent, small run intervals for weak performers. The integration of such multistage bracketed SQC designs has been validated across various high-throughput settings **[10, 16]**, proving that tailored rules reduce unnecessary laboratory interventions while effectively mitigating patient risk. While this study rigorously isolates and optimizes analytical phase precision, the broader literature contextualises these analytical gains within the Total Testing Process (TTP) and economic efficiency. The implementation of optimised Sigma rules drastically curtails false rejection rates, yielding quantifiable financial savings in internal and external failure costs, as demonstrated in yearlong cost-benefit analyses **[29]**. Furthermore, modern error-detection models extend the Lean Six Sigma framework to extra-analytical phases, such as applying machine learning algorithms for predictive delta checks to catch sample misidentifications **[30]**, optimizing turnaround times for molecular assays like SARS-CoV-2 **[8]** and reducing pre-analytical vulnerabilities like sample rejection or hemolysis **[9]**. Finally, robust reporting mechanisms and post-analytical verifications complete this quality loop, showing that our analytical achievements fit seamlessly into the broader, system-wide paradigm of modern clinical quality management.

Conclusion:

The DMAIC-based quality improvement programme strengthened analytical performance and reduced the number of unacceptable analytes from 11 to 2. The increase in world-class analytes from 0 to 4, without category demotion, demonstrates measurable improvement in laboratory quality. Integrating sigma metrics, QGI analysis, CAPA and Westgard Sigma Rule-based QC planning offers a systematic approach to sustaining analytical reliability and enhancing patient safety.

Strengths and limitations:

The major strength of this study is its longitudinal DMAIC-based design, which assessed baseline sigma performance, implemented root cause-directed corrective actions and re-evaluated endpoint performance. Use of IQC-derived CV, EQA-derived bias, QGI analysis, conservative sigma classification and Westgard Sigma Rule-based QC planning made the approach practical and risk-oriented. The main limitations are its single-

centre design, use of one analytical platform, retrospective baseline IQC data and the simultaneous implementation of multiple interventions, which limited assessment of the independent effect of each corrective action.

References:

[1] El Karaaoui AK & Nada A. *Afr J Lab Med.* 2023 **12**:2044. [PMID: 37063602]
[2] Teshome M *et al. J Multidiscip Healthc.* 2021 **14**:125. [PMID: 33488088]
[3] Stulens S *et al. J Humanit Logist Supply Chain Manag.* 2021 **11**:369. [DOI: 10.1108/JHLSCM-08-2020-0072]
[4] Monday LM *et al. Glob J Qual Saf Healthc.* 2021 **4**:109. [PMID: 37261063]
[5] Peng S *et al. J Clin Lab Anal.* 2021 **35**:e23665. [PMID: 33270940]
[6] Goel P *et al. Pract Lab Med.* 2020 **23**:e00195. [PMID: 33392370]
[7] Çevlik T & Haklar G *et al. J Med Biochem.* 2024 **43**:43. [PMID: 38496028]
[8] Lagamayo D *et al. Philipp J Pathol.* 2023 **8**:27. [DOI: 10.21141/PJP.2023.03]
[9] Maripin N *et al. Rev Educ (Madr).* 2025 **408**:67. [DOI: 10.59671/Nr0mj]
[10] Costa-Pallaruelo M *et al. Biochem Med (Zagreb).* 2025 **35**:030704. [PMID: 41103687]
[11] https://westgard.com/clia-and-quality-regulation-requirements/quality-requirements/2024-clia-requirements.html
[12] Westgard JO *et al. Clin Chem.* 2018 **64**:289. [PMID: 29097516]
[13] Parvin CA. *J Appl Lab Med.* 2017 **1**:581. [PMID: 33379807]
[14] Rusanganwa V *et al. Glob Health Action.* 2024 **17**:2358633. [PMID: 38828509]
[15] Chaudhry AS *et al. J Clin Biochem Nutr.* 2023 **73**:108. [PMID: 37700849]
[16] Ramteke TD *et al. J Clin Diagn Res.* 2021 **15**:BC20. [DOI: 10.7860/jcdr/2021/47818.14722]
[17] Geto Z *et al. Indian J Clin Biochem.* 2022 **37**:285. [PMID: 35873618]
[18] Gadde R *et al. Bioinformation.* 2023 **19**:1043. [DOI: 10.6026/973206300191043]
[19] Choi H *et al. Pract Lab Med.* 2025 **47**:e00501. [PMID: 40978770]
[20] Singh S *et al. Asian J Med Sci.* 2024 **15**:70. [DOI: 10.3126/ajms.v15i12.70568]
[21] Panda CR *et al. Cureus.* 2023 **15**:e51386. [PMID: 38292960]
[22] Ramachandra SC *et al. EJIFCC.* 2025 **36**:546. [PMID: 41459165]
[23] Westgard S *et al. Biochem Med (Zagreb).* 2018 **28**:020502. [PMID: 30022879]
[24] Hu M *et al. J Clin Lab Anal.* 2025 **39**:e70080. [PMID: 40679174]
[25] Ghosh S *et al. Asian J Med Sci.* 2025 **17**:27. [DOI: 10.71152/ajms.v17i1.4907]
[26] Westgard JO *et al. Am J Clin Pathol.* 2019 **151**:364. [PMID: 30517600]
[27] Westgard JO & Westgard SA .*Clin Lab Med.* 2017 **37**:85. [PMID: 28153372]
[28] Zhou B *et al. J Clin Lab Anal.* 2020 **34**:e23126. [PMID: 31774217]
[29] Rathore M *et al. EJIFCC.* 2025 **36**:106. [PMID: 40584979]
[30] Seok HS *et al. Clin Chem Lab Med.* 2024 **62**:1421. [PMID: 38095534]

Caveat Emptor is applicable among the literate community where required and possible. The publisher, its journal, editors and the internal/external reviewers take adequate steps to check, evaluate, correct, edit, revise and improve content where possible and required.