OPEN ACCESS – ORIGINAL ARTICLE

Quantifying Total Allowable Error Violations in Serum-Sodium Quality Control: A Computer Simulation Experiment of Two- to Six-Sigma Processes

Mark Angelo Ang and Karen Cybelle Sotalbo

Department of Pathology, College of Medicine, University of the Philippines Manila

ABSTRACT

Background. Serum-sodium reporting tolerates a total allowable error (TEa) of only ±4 mmol/L, yet many laboratories continue to operate at the marginal three-sigma level because the quantitative benefit of additional sigma capability is poorly characterized.

Objectives. The study aims to translate sigma metrics into clinically intuitive risk estimates by (1) quantifying the proportion of QC results that exceed the TEa at five sigma levels (2 – 6 σ) and (2) determining whether successive sigma gains produce statistically significant reductions in error.

Methodology. Five (5) hypothetical assays were parameterized with a common mean of 140 mmol/L and CVs corresponding to 2-, 3-, 4-, 5- and 6-sigma performance. For each assay, 1,000 Monte-Carlo iterations were run, each iteration simulating 36,500 QC results (assuming 100 runs/day for 365 days) drawn from $N(\mu = 140, \sigma = \mu \times CV)$. The error rate (the proportion of results outside ±4 mmol/L) was recorded per iteration. Distributions were summarized (mean, range, SD); differences were evaluated with one-way ANOVA followed by Tukey's HSD.

Results. Mean (±SD) error rates declined significantly with increasing sigma: Assay A (2 σ): 0.0456 ± 0.0011; Assay B (3 σ): 0.00270 ± 0.00027;Assay C (4 σ): $6.3 \times 10^{-5} \pm 4.1 \times 10^{-5}$; Assay D (5 σ): $5.8 \times 10^{-7} \pm 8.0 \times 10^{-7}$; and Assay E (6 σ): $2.0 \times 10^{-7} \pm 3.1 \times 10^{-7}$. The maximum single-iteration error rate fell from 0.0505 at 2 σ to 1.1×10^{-4} at 4 σ. The 5 σ and 6 σ processes produced zero TEa violations in ≥96 % of iterations. ANOVA confirmed a global difference ($p < 0.001$); all pairwise contrasts were significant ($p < 0.001$) except between 5 σ vs 6 σ ($p = 0.62$).

Conclusions. Each one-sigma gain yields an order-of-magnitude reduction in TEa violations until a plateau is reached at ≥5 σ, where residual analytical risk is negligible. These simulations support the recommendation that laboratories operating serum-sodium assays below 4 σ should prioritize precision improvements or enhanced QC strategies, whereas ≥5 σ assays may safely adopt less intensive QC without compromising patient safety.

Key words: Rstudio, total allowable error, TEa, quality control, six sigma, sigma metrics

ISSN 2507-8364 (Online)
Printed in the Philippines.

Received: 7 July 2025.
Accepted: 22 July 2025.
Published online first: 23 October 2025.
https://doi.org/10.21141/PJP.2025.08

Corresponding author: Assoc. Prof. Mark Angelo C. Ang, MD
Email: mcang1@up.edu.ph
ORCiD: https://orcid.org/0000-0003-1292-9493

INTRODUCTION

Reliable measurement of serum sodium is critical for the diagnosis and management of a broad spectrum of disorders, from hyponatraemia-related cerebral edema to hypernatremic dehydration.[1] Analytical error leading to misclassification of sodium concentrations can precipitate inappropriate therapeutic interventions with substantial patient harm.[2,3] Contemporary laboratory quality management therefore mandates a stringent total-allowable-error (TEa) for sodium of ±4 mmol/L, one of the narrowest limits applied in clinical chemistry.[4,5]

The sigma metric has gained prominence as a unifying index that relates analytical imprecision (coefficient of variation, CV) and bias to the governing TEa.[6,7] A higher sigma value denotes fewer defect opportunities per million test results and provides operational guidance on the intensity of internal-QC (IQC) protocols. Westgard and co-workers have formalized "risk-based QC" frameworks in which the frequency and complexity of QC rules are scaled

to an assay's sigma capability.[8–10] However, despite wide acceptance, many laboratories continue to operate sodium assays below or just at the marginal 3-sigma threshold, probably partly because the quantitative benefit of additional sigma gains has not been described in clinically intuitive terms such as expected TEa violations over an annual workload.[11–13]

Computer simulation presents a practical, resource-efficient approach to closing this evidence gap. By repeatedly sampling QC data under controlled distributional assumptions, one can translate abstract sigma metrics into tangible estimates of patient-level risk—namely, the proportion of results that would fall outside TEa over time.[14] Many earlier simulation studies have primarily examined single-sigma scenarios or incorporated multirule algorithms whose complexity can make it more challenging to clearly discern the underlying relationship between sigma level and error frequency.[15–18]

The present work employs a large-scale Computer Simulation (Monte-Carlo) design to model five hypothetical serum-sodium assays spanning 2- to 6-sigma process performances, while holding systematic bias at zero to reduce complexity in analysis. Our primary objective was to estimate, for each sigma process level, the yearly proportion of QC results breaching the ±4 mmol/L TEa for the hypothetical serum sodium assays. Secondary aims were (1) to visualize the changing dispersion pattern by means of representative Levy–Jennings plots and (2) to quantify and demonstrate statistical differences in error rates across sigma levels. These data provide simulated but empirically grounded and informationally familiar targets that can assist laboratories that are seeking to balance assay-selection costs against patient-safety gains.

METHODOLOGY

Study design

A Monte-Carlo computer assisted simulation experiment using the R programming language, in RStudio integrated development environment (IDE), was performed to characterize the analytical-error behavior of five hypothetical serum-sodium assays that differ only in imprecision (coefficients of variation, CVs) and hence in sigma performance.

For each assay, 1,000 independent annual Levy-Jennings (LJ) runs were simulated, mimicking, for example, 1,000 different laboratories using the same platforms. For each of the 1,000 simulated Levy-Jennings, 36,500 quality-control (QC) results were generated, corresponding to 100 QC measurements per day (which is an average case load per day for a low to medium case-load laboratory) for 365 days. All simulations were executed in the R programming language, inside Rstudio[19,20] with the random-number generator initialized to a fixed seed to ensure full reproducibility.

Analytical specifications

The total allowable error (TEa) for serum sodium was fixed at ±4 mmol/L, in accordance with CLIA recommendations. The process mean for every assay was set to 140 mmol/L to mimic a normal level QC material. Five CV values were pre-specified to emulate 2-, 3-, 4-, 5- and 6-sigma processes (Table 1). For every simulated run the within-run standard deviation (SD) was calculated as SD = CV × mean.

Table 1. Simulated Serum Sodium Assays with CV based on assumed Sigma Metric

Assay	Sigma metric	CV
A	2 σ	0.014286
B	3 σ	0.009524
C	4 σ	0.007143
D	5 σ	0.005714
E	6 σ	0.004762

Data-generation procedure

For each assay in every iteration, 36,500 QC results were sampled from a normal distribution $N(\mu = 140, \sigma = \mu \times \text{CV})$. A result was classified as "out-of-bounds," "breach," or a "Violation" when it lay outside the symmetric TEa limits (≤136 or ≥144 mmol/L). The iteration-specific error rate was defined as

$$\text{Error rate}_{ij} = \frac{\text{Count of } |x - \mu| \geq 4}{36\,500}$$

where i denotes assay and j denotes iteration.

Visualization of a single iteration

To illustrate typical LJ behavior, one representative iteration was plotted, using R code, for all five assays (Figures 1 to 5).

Outcome measures

Across the 1,000 iterations per assay, the following statistics were computed: mean error rate, minimum, maximum, and standard deviation. Iteration-level error-rate distributions were visualized with box-and-whisker plots sharing a common y-axis, displaying distribution of the proportion of the QC results that fall beyond ±TEa.

Statistical analysis

Differences among the error rates of the different sigma-metric assays were assessed with one-way analysis of variance (ANOVA). After a significant global F-test ($\alpha = 0.05$, two-sided), pairwise differences were explored by Tukey's honestly significant difference (HSD) test, which controls the family-wise error rate. All statistical analysis procedures were performed in the R programming language inside RStudio.

Software environment

All simulations were implemented in the R programming language. A copy of the entire code is available as supplementary data to this manuscript.

Software tools employed included the R programming language[19] for simulation coding and statistical analysis, RStudio[20] as the integrated development environment, Microsoft Excel[21] for data management, Microsoft Word[22] for manuscript preparation, and ChatGPT[24] for code debugging support and manuscript language assistance and refinement.

All conceptual and creative development, experimental design, analysis implementation decisions, and manuscript drafting were solely and independently initiated by the co-primary authors. ChatGPT was utilized solely to assist in identifying coding inconsistencies, debugging, and enhancing manuscript language clarity, without substantial contributions to the study's intellectual content or experimental framework. The co-primary authors affirm full intellectual ownership and responsibility for the entirety of the content presented in this manuscript.

Ethical considerations

As the study did not involve any actual human or animal subjects, tissues, or personal information, ethical approval from the Institutional Review Board (IRB) or Institutional Animal Care and Use Committee (IACUC) was not sought.

RESULTS

Error-rate distributions across sigma levels

A total of 1,000 Monte-Carlo iterations, each comprising 36,500 quality-control (QC) results, were generated for every assay. The proportion of QC results that exceeded the total allowable error (TEa = ±4 mmol/L) is displayed as the "*error rate*." Summary statistics for the iteration-level error rates are presented in Table 2.

The error-rate distribution widened markedly as analytical precision deteriorated. Assay A (2-sigma) exhibited error rates centered on 4.6%, whereas Assays D and E (≥5-sigma) produced out-of-bounds results only sporadically. More than 95% of their iterations contained no (zero) TEa violations.

Figure 1 visually displays the distribution of the 1,000 iteration-level error rates for each assay. The width of the boxes (inter-quartile range, IQR) and the overall whisker length decreased as the sigma metric improved:

Assay A (2 σ). The median error rate was 4.56% and the IQR spanned 4.48%–4.64%. Whiskers extended to 4.17% (minimum) and 5.05% (maximum). No individual iteration fell within the allowable error limits for all 36 500 QC results, underscoring the clinical inadequacy of a 2-sigma process.

Assay B (3 σ). The distribution contracted sharply: the median error rate was 0.27% with an IQR of 0.25%–0.29%. Although outliers were still present, every iteration's error rate was <0.35%, representing an 18-fold reduction compared with Assay A.

Assay C (4 σ). The median error rate approached zero (6.1×10^{-5}), and 75% of iterations exhibited ≤1 TEa violation. Outliers—iterations with two or three violations—appeared as isolated points above the upper whisker.

Assay D (5 σ) and Assay E (6 σ). Both boxplots were heavily compressed at the zero line. For Assay D, 959 iterations (95.9%) recorded zero violations; for Assay E the figure was 992 iterations (99.2%). The few non-zero iterations for these high-sigma assays manifested as single-point outliers with error rates below 0.0005%.

The boxplots confirm the clear drop in the number of TEa violations with incremental gains in sigma performance. The graph further illustrates that the performance gap between 5- and 6-sigma assays may be considered negligible in routine operations (100 runs per day case load).

Statistical comparison

One-way analysis of variance (ANOVA) demonstrated a highly significant difference in mean error rates across the five assays (*p value* <0.001). Post-hoc Tukey testing (Table 3)

Table 2. Mean error rates of serum assays across different sigma metrics

Assay (sigma)	Mean CV	Mean error rate*	Minimum	Maximum	SD
A (2 σ)	0.014 286	0.045600	0.041700	0.050500	0.001100
B (3 σ)	0.009 524	0.002700	0.001920	0.003400	0.000270
C (4 σ)	0.007 143	0.000063	0.000000	0.000110	0.000041
D (5 σ)	0.005 714	0.000001	0.000000	0.000005	0.000001
E (6 σ)	0.004 762	0.000000	0.000000	0.000001	0.000000

*Mean of 1,000 iterations; error rate expressed as a proportion of 36,500 observations per iteration.

Table 3. Tukey HSD Post-hoc tests

Assay Comparisons	Difference in mean error rates	Lower CI	Upper CI	Adjusted *p* value
B - A	-4.28E-02	-4.2906E-02	-4.2779E-02	<0.000001
C - A	-4.55E-02	-4.5536E-02	-4.5409E-02	<0.000001
D - A	-4.55E-02	-4.5601E-02	-4.5474E-02	<0.000001
E - A	-4.55E-02	-4.5601E-02	-4.5475E-02	<0.000001
C - B	-2.63E-03	-2.6934E-03	-2.5665E-03	<0.000001
D - B	-2.69E-03	-2.7580E-03	-2.6312E-03	<0.000001
E - B	-2.70E-03	-2.7586E-03	-2.6317E-03	<0.000001
D - C	-6.47E-05	-1.2809E-04	-1.2273E-06	0.043187
E - C	-6.52E-05	-1.2864E-04	-1.7752E-06	0.040410
E - D	-5.48E-07	-6.3978E-05	6.2882E-05	1.000000

confirmed that every pairwise contrast involving Assay A or Assay B differed significantly (p <0.001). In contrast, the comparison between the two highest-sigma assays (D vs E) was non-significant (p = 0.62), indicating indistinguishably low residual error at ≥5-sigma performance.

Plots of a representative iteration of Levy Jennings charts across sigma metrics

A representative set of Levy–Jennings charts (Figures 2 to 6), from a single iteration, support our findings. All simulated LJ charts share an identical y-axis scale bounded by the

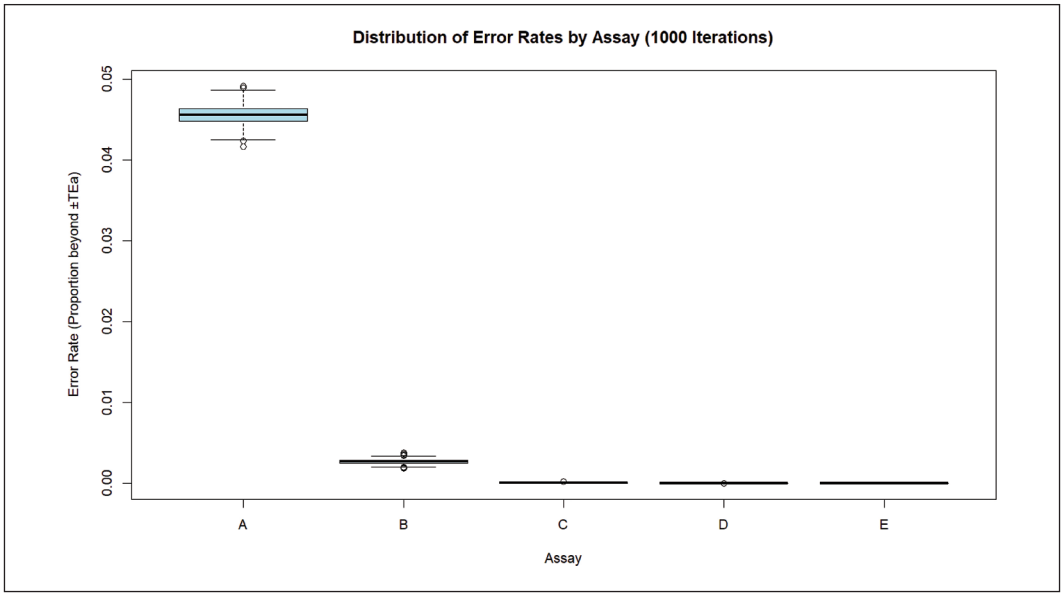

Figure 1. Box and Whiskers plot of the error rate of the 5 different simulated assays across 1000 simulations.

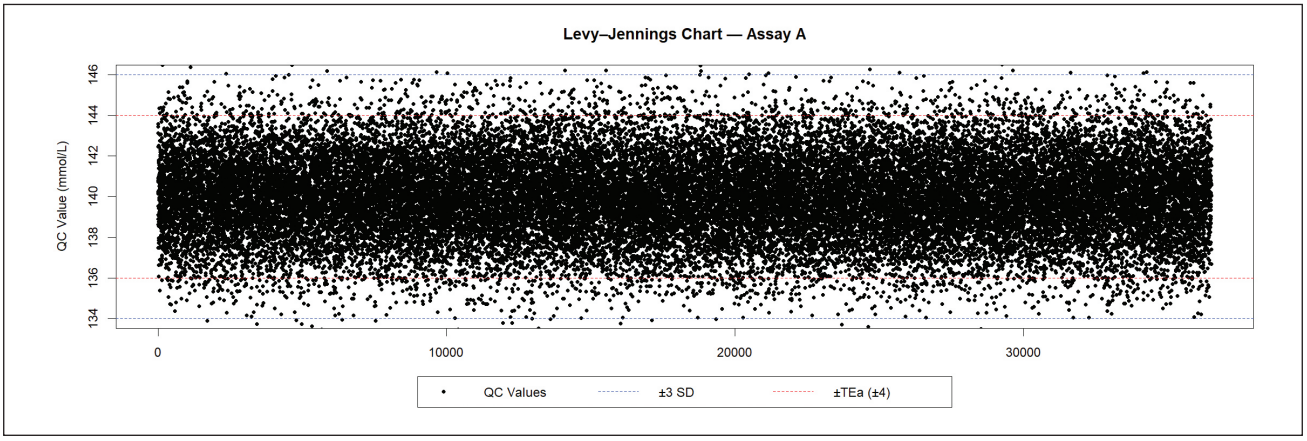

Figure 2. A representative iteration of an LJ chart for serum sodium with a 2-sigma process, showing 36,500 data points.

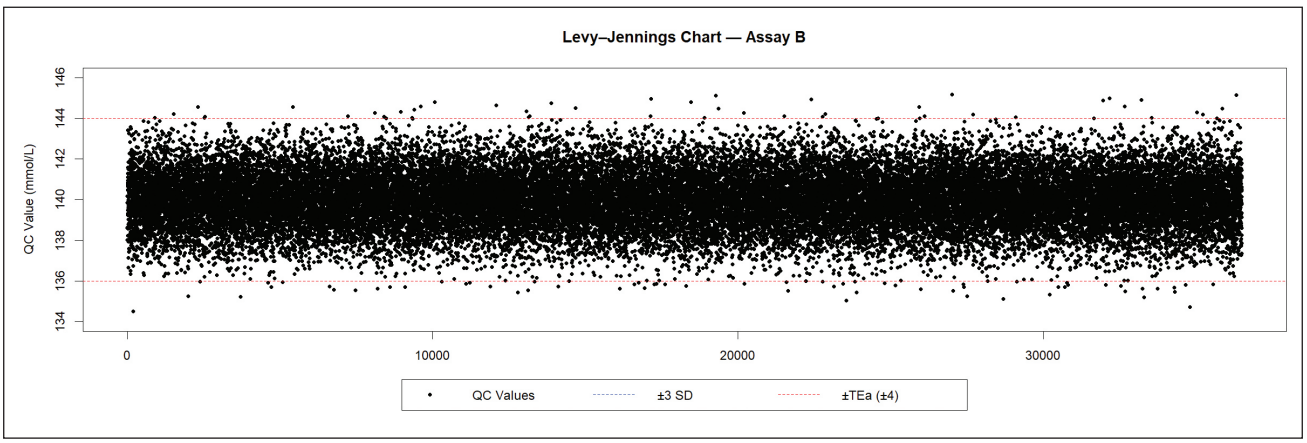

Figure 3. A representative iteration of an LJ chart for serum sodium with a 3-sigma process, showing 36,500 data points.

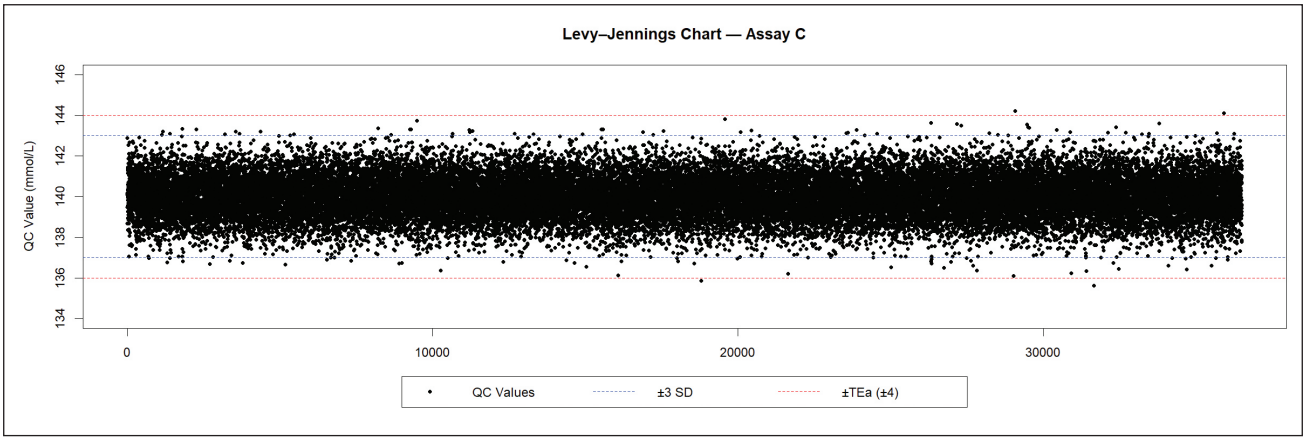

Figure 4. A representative iteration of an LJ chart for serum sodium with a 4-sigma process, showing 36,500 data points.

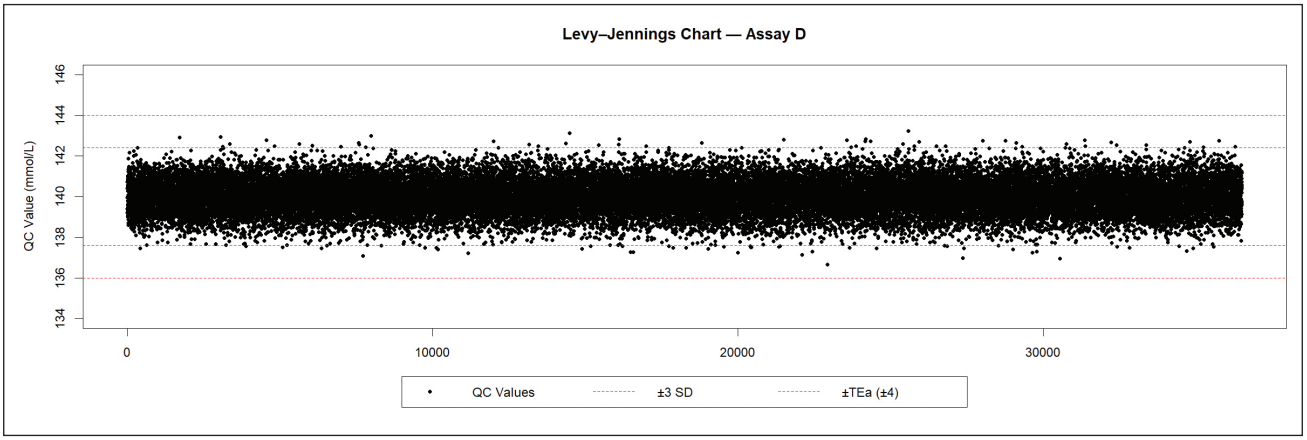

Figure 5. A representative iteration of an LJ chart for serum sodium with a 5-sigma process, showing 36,500 data points.

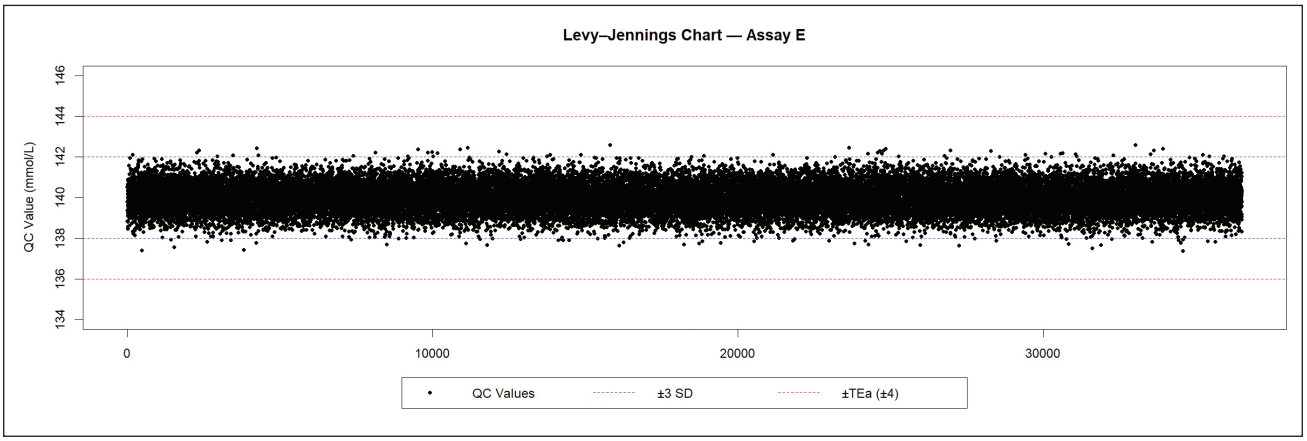

Figure 6. A representative iteration of an LJ chart for serum sodium with a 6-sigma process, showing 36,500 data points.

larger of the ±3 SD limits for Assay A or the regulatory TEa (±4 mmol/L), facilitating direct visual comparison.

In the 2-sigma chart (Assay A) QC points are widely dispersed. Frequent clusters breach both the blue dashed ±3 SD lines and the red dashed TEa boundaries, with prolonged sequences of consecutive violations evident.

In the 3-sigma chart (Assay B) most observations lie within ±3 SD. Only isolated points cross the TEa limits and visual "strings" of excursions or clusters are rare.

For the 4-sigma process (Assay C) all values are confined to a narrow band around the process mean. Occasional single-point outliers touch, but seldom exceed, the TEa lines.

The 5- and 6-sigma charts (Assays D and E) show an almost solid black band at the mean. QC points rarely approach the ±3 SD lines and no point in the depicted iteration exceeds the TEa limits.

The progressive collapse of the data cloud toward the process mean from Assay A to E provides an immediate qualitative illustration of the exponential reduction in clinically relevant error as sigma capability increases.

Collectively these results illustrate the exponential decline in TEa violations or breaches as assay precision improves from 2- to 6-sigma capability. Our results also show that once a process attains ≥5-sigma performance, further gains in CV may yield negligible additional reduction in clinically relevant errors.

DISCUSSION

This computer assisted simulation analysis experiment demonstrates a precipitous, exponential, tightly graded relation between an assay's sigma capability and its propensity to yield clinically unacceptable results for serum sodium.

Across 1,000 independent annual simulations per assay, the mean TEa violations (error rates) precipitously fell from 4.6% for the 2-sigma process to <0.0001% for the 4-sigma process and to zero for ≥5-sigma processes. One-way ANOVA confirmed that these gradients were statistically robust, and Tukey comparisons showed that each stepwise gain of one sigma level translated into a statistically, and, meaningful reduction in error, with the sole exception of the 5- versus 6-sigma contrast, where residual risk was already vanishingly small.

Representative Levy–Jennings (LJ) charts vividly illustrated this pattern. Whereas the 2-sigma assay generated frequent "runs" that fall outside of both the ±3 SD and ±Tea QC LJ lines, the 5- and 6-sigma assays plotted as a narrow ribbon around the process mean, with no violations in the illustrated iteration.

Our simulation experiment results align with Westgard's theoretical error grids, which predict ~45 000 defect opportunities per million (DPM) for a 2-sigma process and <4 DPM for a 5-sigma process when TEa is set at 1.65 SD.[9,10,14,24] While absolute rates differ because our TEa is fixed in concentration units rather than multiples of SD, the simulated decline in error as sigma rises is congruent with analytic-quality models employed by CLSI EP23-A2 and IFCC.[7,9,25] Moreover, the finding that a 3-sigma process still yields error in the order of 1×10^{-2} corroborates several external-quality-assessment reports showing that 3-sigma sodium assays fail 'total-error' proficiency challenges four- to five-times more often than 5-sigma assays.[26]

Clinical and operational implications

From a risk-management perspective, any test with ≥5-sigma capability for sodium essentially meets a "six-sigma" quality target after accounting for biological variation (≈0.5%).[2,4,5] Such performance permits less intensive QC scheduling, potentially reducing reagent waste, technologist time and instrument down-time without compromising patient safety. Conversely, a 2-sigma process is categorically inadequate. Even under idealized random-error conditions, it would deliver approximately 1,600 out-of-specification results per year at the simulated workload. A 3-sigma process, though markedly better, still produces about 1 error every four days, underscoring the need for either tighter imprecision goals or supplementary QC rules (e.g., Westgard multirules, moving averages) if this level of performance cannot be improved.

The 4-sigma scenario presents an inflection point: the average laboratory might find its residual error tolerable if additional patient-based QC[27] is in place. However, the tail of the error-rate distribution (maximum 1.1×10^{-4}) implies that sporadic TEa violations remain possible and must be weighed against the clinical consequences of sodium misclassification.

Strengths and limitations

A major strength of this work is the large simulation size—36,500 observations per iteration and 1,000 iterations per assay—providing tight confidence around the error-rate estimates.

The framework is transparent, fully reproducible and parameterized directly in concentration units of serum sodium, enhancing its practical relevance, understandability, and direct translatability to actual bench experience - features that would support use by laboratory staff and managers.

Many limitations stem from the model simplification choices we made in the simulations. First, only imprecision was varied. Other key factors such as systematic bias, drift, reagent lot effects and matrix interferences were not modelled, to keep the simulation tractable and feasible using standard computational resources. In real-world quality control, bias significantly impacts sigma metrics and is often the reason assays fall short of performance expectations despite good precision. The exclusion of bias in our model assumes idealized conditions and may thus lead to an overestimation of QC performance. Therefore, future studies should incorporate bias to more accurately reflect operational laboratory scenarios and better guide QC protocol design. These are relevant sources of variation that can and should be modelled using grounded assumptions.

Second, the assumption of normality may underestimate distribution tail behavior for real-world assays with non-Gaussian error distributions. The assumption of a normal distribution in simulating control values may underrepresent the frequency of extreme deviations. In practice, assay result distributions can exhibit skewness or heavy tails due to pre-analytical or biological variability, especially in low-sigma processes. This may lead to underestimation of error rates and overconfidence in assay reliability. Thus, caution must be exercised in applying these findings directly to routine QC planning without empirical distribution analysis. Follow-up simulations using statistical models with "fatter" tails may be informative.[28]

Third, the upper and lower TEa lines were treated as a fixed symmetric bounds, yet in practice different regulatory bodies can specify asymmetric or clinically variable goals.[5,29] This study used CLIA guidelines as the source of TEa due to their regulatory relevance and widespread use in clinical laboratory practice. However, it is important to note that using TEa derived from biological variation or percentage-based goals could alter sigma estimations. Wider TEa ranges would likely yield higher sigma values, potentially resulting in fewer TE violations. This underscores the importance of context-specific TEa selection when interpreting QC performance.

Finally, the analysis focused on a specific analyte only - serum sodium, which has a measurand with stringent TEa (±4 mmol/L). Direct extrapolation to analytes with wider or percentage-based TEa requires caution.

FUTURE DIRECTIONS

Incorporating systematic shifts and trending errors would refine understanding of QC frequency requirements, particularly for mid-tier (3- to 4-sigma) assays. Future work should also evaluate the combined effect of sigma performance and multirule QC algorithms to identify cost-efficient QC strategies under resource-constrained settings. Expanding the simulation to additional measurands with biologically driven TEa (e.g., potassium, calcium) would generalize the conclusions to a broader clinical-chemistry portfolio.

CONCLUSIONS

This simulation confirms that sigma metrics provide a powerful, quantitative gauge of analytical reliability. A 5-sigma sodium assay is effectively "zero-defect" under routine workloads, validating reduced QC burdens, whereas a 3-sigma assay, although showing a significant improvement from 2-sigma assay, still carries a measurable risk of clinically significant error and warrants enhanced QC oversight. Laboratories should therefore prioritize analytical-imprecision improvements to at least 4-sigma, and preferably at least 5-sigma, when selecting or validating serum-sodium methods.

STATEMENT OF AUTHORSHIP

The authors certified fulfillment of ICMJE authorship criteria.

AUTHOR DISCLOSURE

The authors declared no conflict of interest.

DATA AVAILABILITY STATEMENT

The data that support the findings of this study are available from the corresponding author upon reasonable request.

FUNDING

None.

REFERENCES

1. Spasovski G, Vanholder R, Allolio B, et al. Clinical practice guideline on diagnosis and treatment of hyponatraemia. Eur J Endocrinol. 2014;170(3):G1-47. PMID: 24569125 DOI: 10.1530/EJE-13-1020
2. Tormey WP, Carney M, Cuesta M, Sreenan S. Reference Change values for sodium are ignored by the American and European treatment guidelines for hyponatremia. Clin Chem. 2015;61(12):1430-2. PMID: 26152751 DOI: 10.1373/clinchem.2015.242602
3. Al Dhaybi O, Bakris GL. Factitious acidosis and severe hypoalbuminemia caused by unsuspected assay interference. Clin Kidney J. 2021;14(3):1023-4. PMID: 34840734 PMCID: PMC8612138 DOI: 10.1093/ckj/sfaa074
4. Freitas F. Analytical quality specifications in the clinical laboratory – limiting analytical error with allowable error Limits; 2013.
5. Oosterhuis WP, Coskun A, Sandberg S, Theodorsson E. Performance specifications for sodium should not be based on biological variation. Clin Chim Acta. 2023;540:117221. PMID: 36640931 DOI: 10.1016/j.cca.2023.117221
6. Guo X, Zhang T, Gao X, et al. Sigma metrics for assessing the analytical quality of clinical chemistry assays: a comparison of two approaches. Biochem Medica. 2018;28(2):020708. PMID: 30022883 PMCID: PMC6039159 DOI: 10.11613/BM.2018.020708
7. Luo Y, Yan X, Xiao Q, et al. Application of Sigma metrics in the quality control strategies of immunology and protein analytes. J Clin Lab Anal. 2021;35(11):e24041. PMID: 34606652 PMCID: PMC8605144 DOI: 10.1002/jcla.24041
8. Westgard SA, Bayat H, Westgard JO. Selecting a Risk-Based SQC procedure for a HbA1c total QC plan. J Diabetes Sci Technol. 2018;12(4):780-5. PMID: 28905657 PMCID: PMC6134308 DOI: 10.1177/1932296817729488
9. Westgard JO, Westgard SA. Quality control review: implementing a scientifically based quality control system. Ann Clin Biochem Int J Lab Med. 2016;53(1):32-50. PMID: 26150675 DOI: 10.1177/0004563215597248
10. Westgard JO. Managing quality vs. measuring uncertainty in the medical laboratory. CLin Chem Lab Med. 2010;48(1):31-40. PMID: 19919330 DOI: 10.1515/CCLM.2010.024
11. Bozkaya G, Aksit M, Zeytinli Aksit M. Evaluation of clinical chemistry tests in emergency laboratory by sigma metrics. Turk J Biochem. 2018;43(1):9-14. DOI: 10.1515/tjb-2017-0143

12. Moodley N, Gounden V. Assessment of sigma metrics for routine chemistry testing in 4 laboratories in Kwa-Zulu Natal, South Africa. J Appl Lab Med. 2022;7(3): 689-97. PMID: 34636901 DOI: 10.1093/jalm/jfab117
13. Feldhammer M, Brown M, Colby J, Bryksin J, Milstid B, Nichols JH. A survey of sigma metrics across three academic medical centers. J Appl Lab Med. 2021; 6(5):1264-75. PMID: 34060592 DOI: 10.1093/jalm/jfab028
14. Westgard JO. Simulation and modeling for optimizing quality control and improving analytical quality assurance. Clin Chem. 1992;38(2):175-8. PMID: 1540997
15. Boyd JC, Bruns DE. Quality specifications for glucose meters: assessment by simulation modeling of errors in insulin dose. Clin Chem. 2001;47(2):209-14. PMID: 11159768
16. Boyd JC, Bruns DE. Effects of measurement frequency on analytical quality required for glucose measurements in intensive care units: assessments by simulation models. Clin Chem. 2014;60(4):644-50. PMID: 24430017 DOI: 10.1373/clinchem.2013.216366
17. Parvin CA. Quality-control (QC) performance measures and the QC planning process. Clin Chem. 1997;43(4):602-7. PMID: 9105260
18. Gernand W. How Reliable are critical error calculations? Clin Chem. 2006;52(5):895-6. PMID: 16638957 DOI: 10.1373/clinchem.2006.067991
19. R Core Team. R: A Language and environment for statistical computing [computer program]. Published online 2024. https://www.R-project.org/.
20. Posit Team. RStudio: Integrated development environment for R [computer program]. Published online 2025. http://www.posit.co/.
21. Microsoft Corporation. Microsoft Excel [computer program]. Published online 2025. https://www.microsoft.com.
22. Microsoft Corporation. Microsoft Word [computer program]. Published online 2025. https://www.microsoft.com.
23. OpenAI. ChatGPT [large language model]. Version: June 2025. Published online 2025. https://chat.openai.com/.
24. Jones GRD. Reevaluation of the power of error detection of Westgard multirules. Clin Chem. 2004;50(4):762-4. PMID: 15044336 DOI: 10.1373/clinchem.2003.025585
25. Liu Q, Bian G, Chen X, et al. Application of a six sigma model to evaluate the analytical performance of urinary biochemical analytes and design a risk-based statistical quality control strategy for these assays: A multicenter study. J Clin Lab Anal. 2021;35(11):e24059. PMID: 34652033 PMCID: PMC8605169 DOI: 10.1002/jcla.24059
26. Weykamp C, Siebelder C. Evaluation of performance of laboratories and manufacturers within the framework of the IFCC model for quality targets of HbA1c. J Diabetes Sci Technol. 2018;12(4):747-52. PMID: 29145735 PMCID: PMC6134299 DOI: 10.1177/1932296817741320
27. Badrick T, Bietenbeck A, Cervinski MA, et al. Patient-based real-time quality control: review and recommendations. Clin Chem. 2019;65(8):962-71. PMID: 31263035 DOI: 10.1373/clinchem.2019.305482
28. Taleb NN. Statistical consequences of fat tails: real world preasymptotics, epistemology, and applications; 2020. DOI: 10.48550/ARXIV.2001.10488
29. Lyon AW, Lyon ME. Evaluation of precision and bias specifications required to achieve the 2018 FDA guidance criteria for glucose meter performance using simulation Models. J Diabetes Sci Technol. 2020; 14(3):513-8. PMID: 31752535 PMCID: PMC7576955 DOI: 10.1177/1932296819889639