NATFORLAB 2026

Risk-Based Internal Quality Control and Patient-Based Quality Control

Hikmet Can Çubukçu

MD, PhD, MSc, EuSpLM, Assoc. Prof. in Medical Biochemistry

Chair of EFLM Committee on Laboratory Error Database
Member of IFCC Committee on AI in Laboratory Medicine
Expert/Consultant in EFLM Committee on Accreditation and ISO/CEN standards
Department of Medical Biochemistry, Sincan Training and Research Hospital
Department of Rare Diseases, General Directorate of Health Services, Turkish Ministry of Health

NATFORLAB 2026 · 19 September 2026

Conflict of interest

None declared.

Outline

1. Background: analytical performance and conventional internal QC
2. QC practice in clinical laboratories
3. Risk-based internal QC
4. Patient-based QC
5. Take-home messages

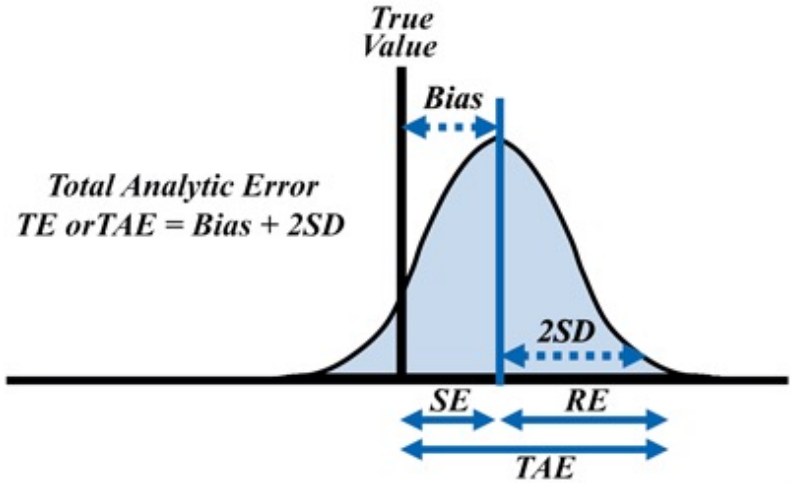
1

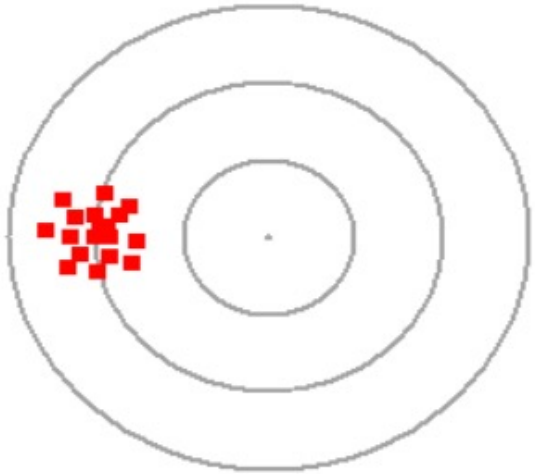
Background

Analytical performance and conventional internal QC

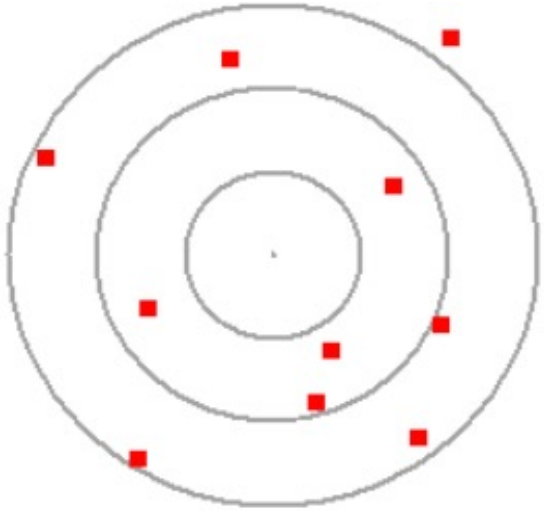
Types of errors

- **Systematic error (bias)**
  - Constant systematic error
  - Proportional systematic error
- **Random error (imprecision)**
  - Measured as SD or CV





Systematic error



Random error

Total analytic error (TE or TAE) = Bias + 2SD

Analytical performance specifications (APS): the Milan models

2014 – 'Defining analytical performance goals 15 years after the Stockholm Conference on Quality Specifications in Laboratory Medicine' (1st EFLM Strategic Conference)

Model	Advantage	Disadvantage
Model 1 – based on the effect of analytical performance on clinical outcomes	The most appropriate target for the intended clinical use of the test	Studies and targets are insufficient
Model 2 – based on components of biological variation of the measurand	Targets have been defined for a large number of analytes	For some analytes, the targets may not be achievable?
Model 3 – based on state-of-the-art	More achievable targets	Being technically achievable does not necessarily mean it meets the clinical need

Sandberg S, Fraser CG, Horvath AR, et al. Defining analytical performance specifications: consensus statement from the 1st Strategic Conference of the EFLM. Clin Chem Lab Med 2015;53:833–5.

Where can we find analytical performance specifications?

Model 2 · Biological variation

EUROPEAN FEDERATION OF CLINICAL CHEMISTRY AND LABORATORY MEDICINE

EFLM Biological Variation Database

https://biologicalvariation.eu

Formulas for desirable APS:

Imprecision: $CV_A < 0.5 \times CV_I$

Bias $< 0.25 \times (CV_I^2 + CV_G^2)^{½}$

MAU $< 2 \times 0.5 \times CV_I$

TEa $< 1.65 \times (0.5 \times CV_I) + 0.25 \times (CV_I^2 + CV_G^2)^{½}$

Model 3 · State of the art

Data Innovations
Revolutionizing the lab. Transforming patient care.

Total Allowable Error Table
datainnovations.com/allowable-total-error-table

WESTGARD QC

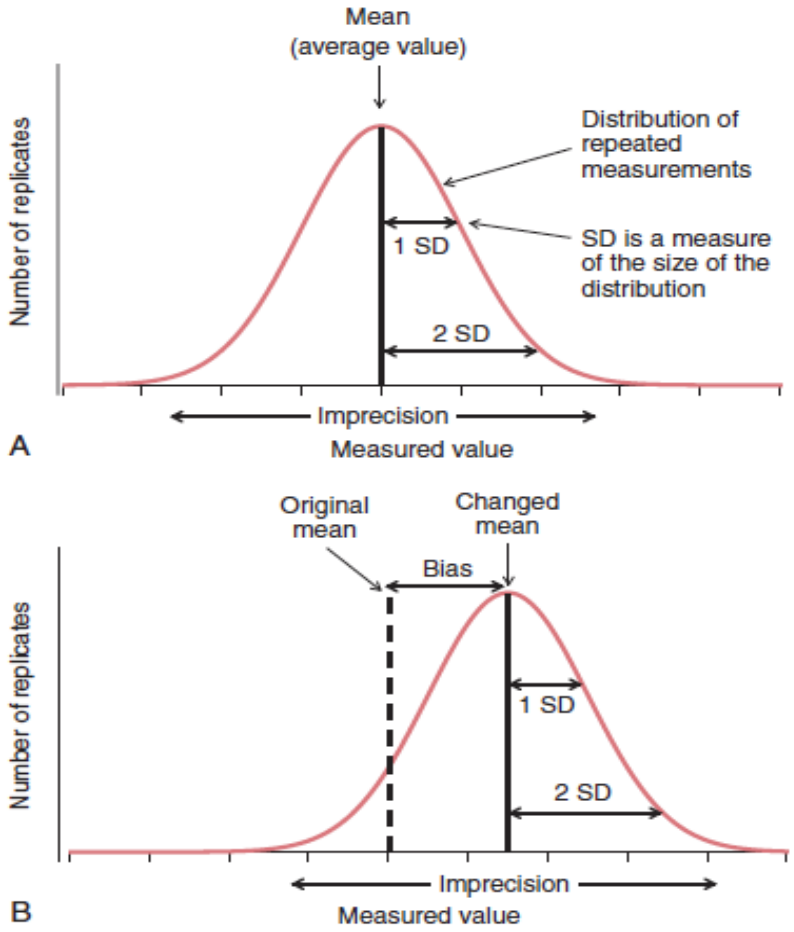
CLIA 2024 requirements
westgard.com/2024-clia-requirements.htm

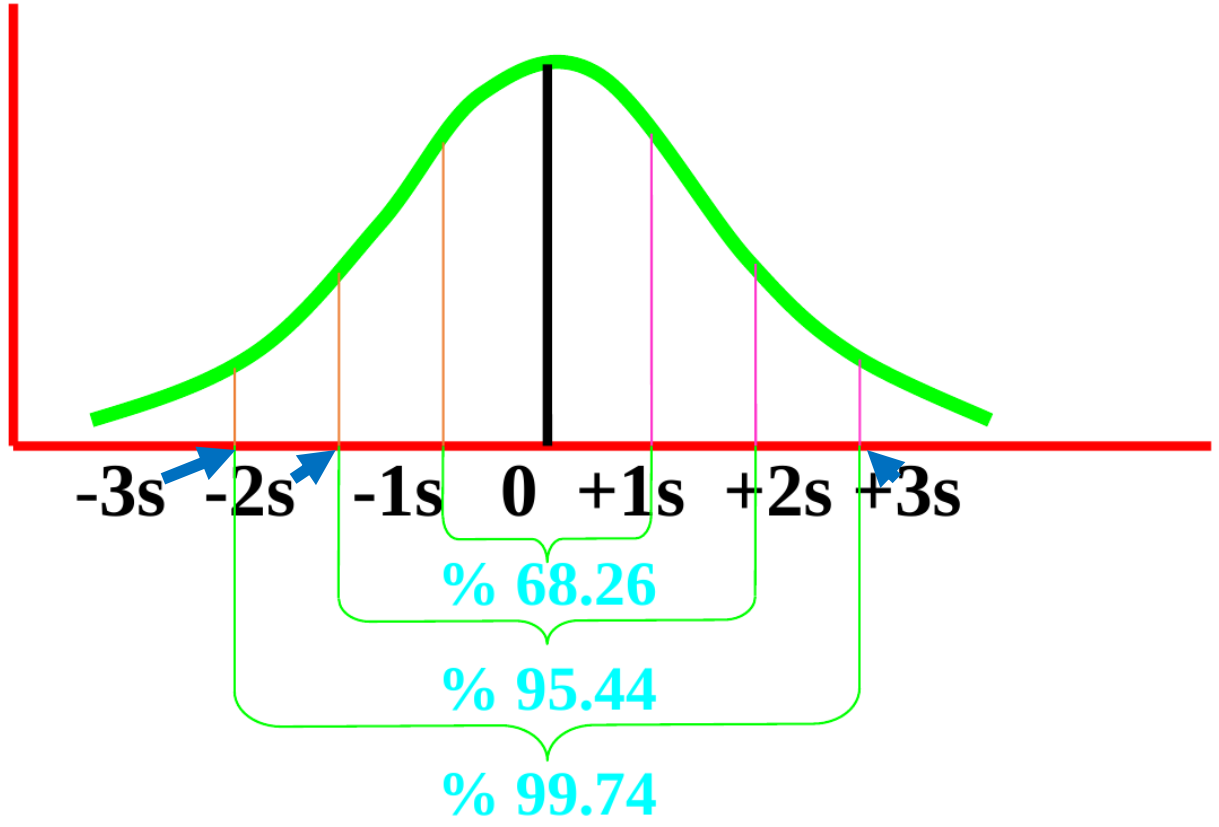
RCPAQAP
The Royal College of Pathologists of Australasia Quality Assurance Programs

RCPAQAP (Australia) – analytical performance specifications
rcpaqap.com.au/resources

Aarsand AK, Fernandez-Calle P, Webster C, et al. The EFLM Biological Variation Database. https://biologicalvariation.eu

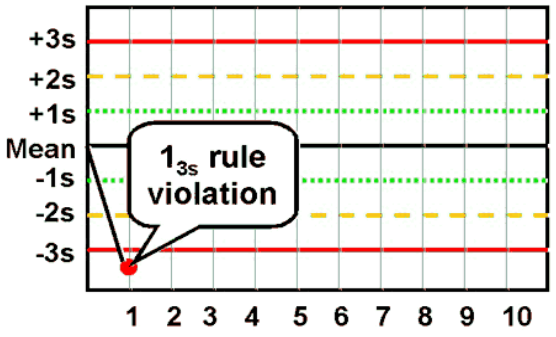
Internal QC basics: the distribution of control results





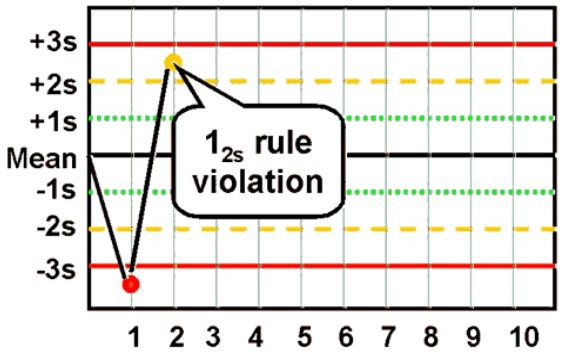
Mean ± 1s: 68.26 % · Mean ± 2s: 95.44 % · Mean ± 3s: 99.74 % of control results

Westgard rules (1)



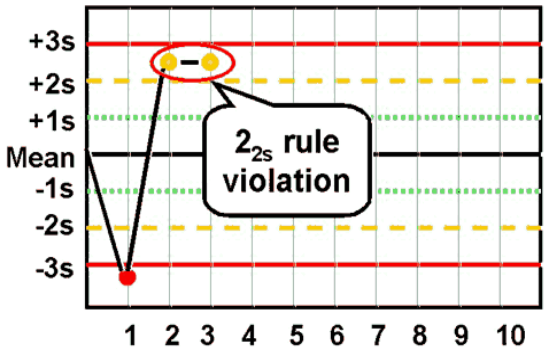
1-3s rule ⇒ Bias or large imprecision

When a single control measurement exceeds the mean plus 3s or the mean minus 3s control limit.



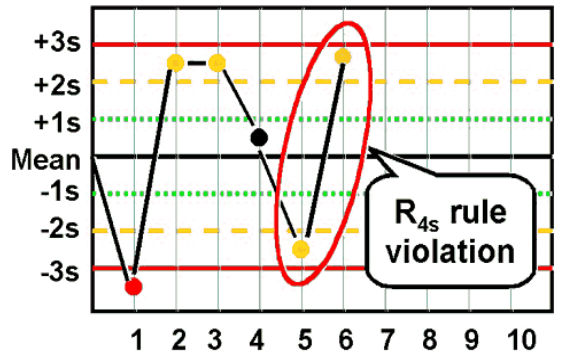
1-2s rule ⇒ Bias or imprecision (warning rule)

A warning rule to trigger careful inspection of the control data by the following rejection rules.



2-2s rule ⇒ Bias

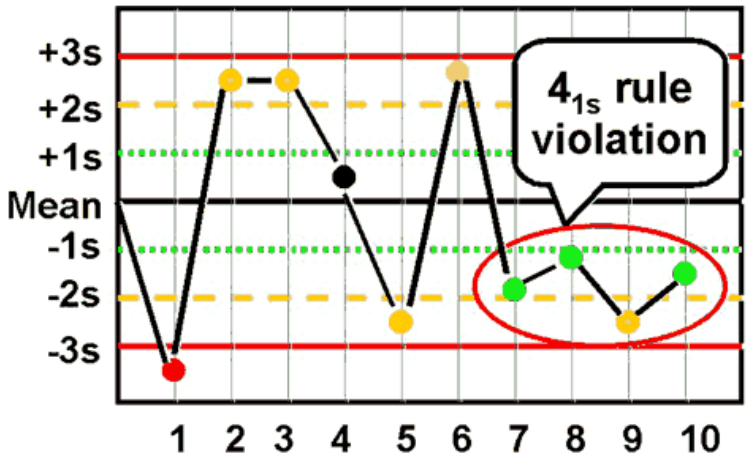
When 2 consecutive control measurements exceed the same mean plus 2s or the same mean minus 2s control limit.



R-4s rule ⇒ Imprecision

When 1 control measurement in a group exceeds the mean plus 2s and another exceeds the mean minus 2s.

Westgard rules (2)

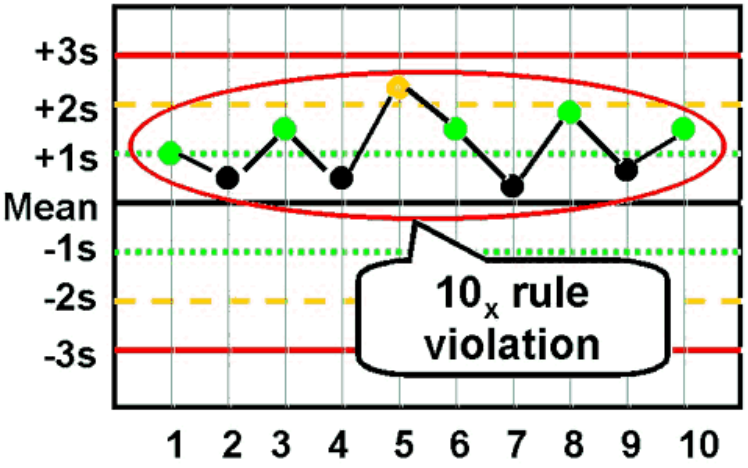


4-1s rule ⇒ Bias

When 4 consecutive control measurements exceed the same mean plus 1s or the same mean minus 1s control limit.

2 of 3s rule ⇒ Bias

When 2 out of 3 control measurements exceed the same mean plus 2s or mean minus 2s control limit. When 3 QC materials are run.



10x rule ⇒ Bias trend (not recommended)

When 10 consecutive control measurements fall on one side of the mean.

2

QC practice in clinical laboratories

Where we are now

QC practice in the US

AJCP / Original Article

Quality Control Practices for Chemistry and Immunochemistry in a Cohort of 21 Large Academic Medical Centers

Quality Control Rules

Hospital	Chemistry	Immunochemistry
A	±2 SD	±2 SD
B	±2 SD	±2 SD
C	±2 SD	±2 SD
D	±2 SD, some ±2.5 SD and ±3 SD	±2 SD
E	±2 SD	±2 SD
F	±2 SD	±2 SD
G	±2 SD	±2 SD
H	±2 SD	±2 SD
I	Variable (1-3S, 1-3.5S, 1-4S, 1-5S, 2-2S, 2 of 3-2S, R-4S, 3-1S, 4-1S, or N-x, 1-2S, 1-2.5S)	Variable (1-3S, 1-3.5S, 1-4S, 1-5S, 2-2S, 2 of 3-2S, R-4S, 3-1S, 4-1S, or N-x, 1-2S, 1-2.5S)
J	±2 SD	±2 SD
K	±2 SD	±2 SD
L	±2 SD	±2 SD
M	±2 SD	±2 SD
N	±2 SD	±2 SD
O	±2 SD	±2 SD
P	±2 SD	±2 SD
Q	3 SD	3 SD
R	±2 SD	±2 SD
S	2-2S, 1-3S, R-4S, 10x	2-2S, 1-3S, R-4S, 10x
T	2 or 3 SD	Some 2 or 3 SD; others Westgard rules
U	±2 SD	±2 SD

Hospitals Responding in Full

Barnes-Jewish Hospital/Washington University, St Louis
Brigham and Women's Hospital
Cedars-Sinai Medical Center
Duke University Hospital
Hospitals of the University of Pennsylvania–Penn Presbyterian
Houston Methodist Hospital
Johns Hopkins Hospital
Massachusetts General Hospital
Mayo Clinic
Mount Sinai Hospital
New York-Presbyterian Columbia
New York-Presbyterian Cornell
Northwestern Memorial Hospital
Ronald Reagan UCLA Medical Center
Stanford Health Care–Stanford Hospital
The Cleveland Clinic
Tisch Hospital, NYU Langone Health
University of California, San Francisco Medical Center
University of Colorado Hospital
University of Michigan Hospitals and Health Centers
UPMC Presbyterian Shadyside, Pittsburgh

Quality Control Frequency

Hospital	CHEM	IM	Stat IM	QC Events/d: CHEM	IM	Tn	hCG
A	2-3 Lv qd (electrolytes q8h)	2-3 Lv qd	2-3 Lv qd, negative QC q8h for Tn and hCG	3	1	3	3
B	2-3 Lv q8h	2-3 Lv q8h	2-3 Lv q8h	3	3	3	3
C	1 Lv (alternating) q2h	2-3 Lv q8h	cTn 4 Lv qd, 2 Lv alternating q2h, hCG 2 Lv q8h	12	3	12	3
D	2 Lv q12h	2 Lv q12h	2 Lv q8h	2	2	3	3
E	2 Lv q8h	2 Lv qd	2 Lv qd	3	1	1	1
F	2 Lv q8h	2 Lv qd	2 Lv qd	3	1	1	1
G	High control q12h, low control qd	2-3 Lv qd (medium Lv q12 for Lv 3 tests) or high Lv q12h/low Lv qd	Tn/CKMB/NT-proBNP: 2 Lv q12h; hCG: high q12h, low qd	2	2	2	2
H	2-3 Lv q12h	2-3 q12h	2-3 q12h	2	2	2	2
I	2-3 Lv q8h or qd	2-3 Lv qd	2-3 Lv qd	3	1	1	1
J	2 Lv q8h (q4h for electrolytes)	2-3 qd, 1 Lv (alternating) second/third shift	2-3 Lv day shift, 1 Lv (alternating) second/third shifts	6	3	3	3
K	2 Lv qd	2-3 Lv qd	Tn/hCG: 3 Lv qd	1	1	1	1
L	2 Lv q6h, 2 Lv startup/shutdown	2-3 Lv q12h	5 Lv q8h; hCG 2-3 Lv q12h	6	2	3	2
M	2 Lv at startup, 1 Lv (q4h alternating)	2-3 Lv at startup, 1 Lv at shutdown	Tn/hCG 2/3 Lv at startup, 1 Lv (alternating) q4h	7	2	7	7
N	3 Lv startup/shutdown, then 2 levels QC q2h	3 Lv startup/shutdown, plus 2 levels q8h	3 Lv q8h	12	4	3	3
O	3 Lv qd	2-3 Lv qd	Core laboratory: Tn/hCG 3 Lvq12h/q24 hours; ED laboratory: Tn/hCG 2 Lv qd	1	1	2	1
P	2-3 Lv q12h	2 Lv q8h	2 Lv q8h	2	3	3	3
Q	3 Lv day shift, 2 Lv other shifts	3 Lv day shift, 2 Lv other shifts	3 Lv qd, 2 Lv q12h	3	3	2	2
R	2 Lv q12h	2 Lv q12h	3 Lv q12h	2	2	2	2
S	2-3 q12h	q12h, certain tests once at start up	q12h, some qd	2	2	2	2
T	2 Lv qd	2 Lv qd	2 Lv qd	1	1	1	1
U	3 Lv q8h	3 Lv qd	3 Lv qd	3	1	1	1

CHEM, chemistry; CKMB, creatine kinase MB isoenzyme; cTn, cardiac troponin; ED, emergency department; hCG, human chorionic gonadotropin; IM, immunochemistry; Lv, level; NT-proBNP, NT-pro B-type natriuretic peptide; QC, quality control; QC Event, one or more QC levels run consecutively at a specified time on an analyzer; qd, per day; q(x)h, every x hours; Tn, troponin.

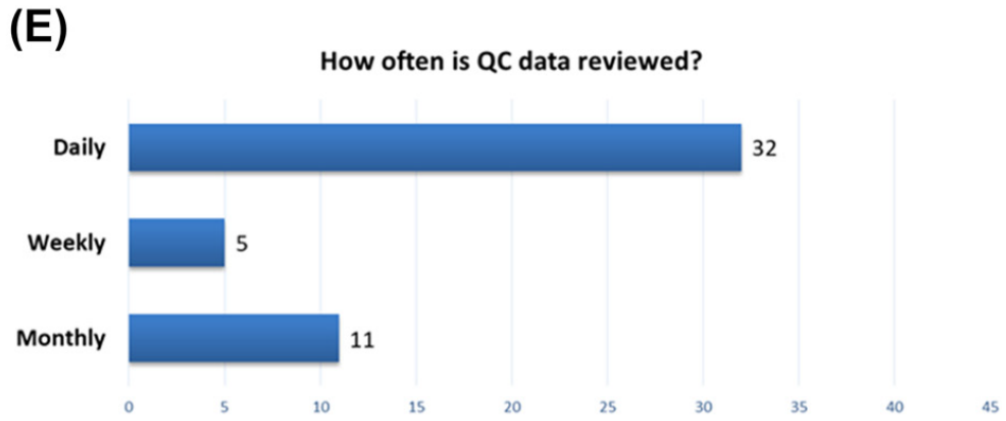
Rosenbaum MW, et al. Am J Clin Pathol 2018;150:96–104.

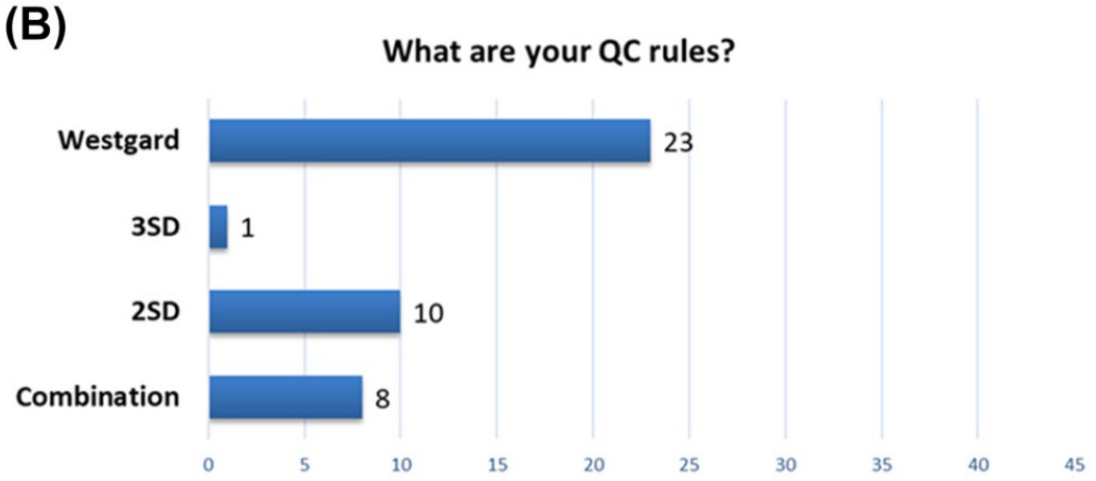
QC practice in the world

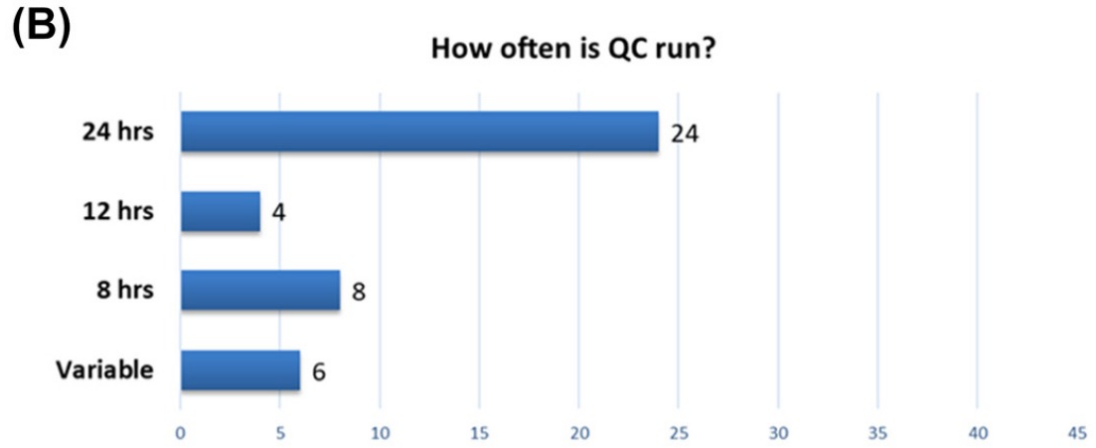
Sarah E. Wheeler, Ivan M. Blasutig, Pradeep Kumar Dabla, Jean-Marc Giannoli, Anne Vassault, Ji Lin, Kandace A. Cendejas, Armand Perret-Liaudet, Renze Bais, Annette Thomas, Egon P. Amann* and Qing H. Meng*

Quality standards and internal quality control practices in medical laboratories: an IFCC global survey of member societies

IFCC global survey of member societies: “What are your QC rules?”, “How often is QC run?”, “How often is QC data reviewed?”







Wheeler SE, et al. Clin Chem Lab Med 2023;61:2094–101.

Where we are now — and what we should be

Where we are now

What we should be

3

Risk-Based Internal Quality Control

Selecting the QC rule and QC frequency to mitigate patient risk

Risk-based internal QC: the goal

To select appropriate:

Quality control rule

Frequency of QC measurements
(number of patient samples between QC events)

… to mitigate patient risk

Parvin CA. Planning statistical quality control to minimize patient risk: it's about time. Clin Chem 2018;64:249–50.

Risk-based QC approaches

MaxE(Nuf) and Pfr-based plan: We can calculate maximum run size and acceptability of QC rule scheme

CLSI EP23 ed2-based plan: We can be informed about whether the QC strategy has managed risk or not

Risk-based QC tools

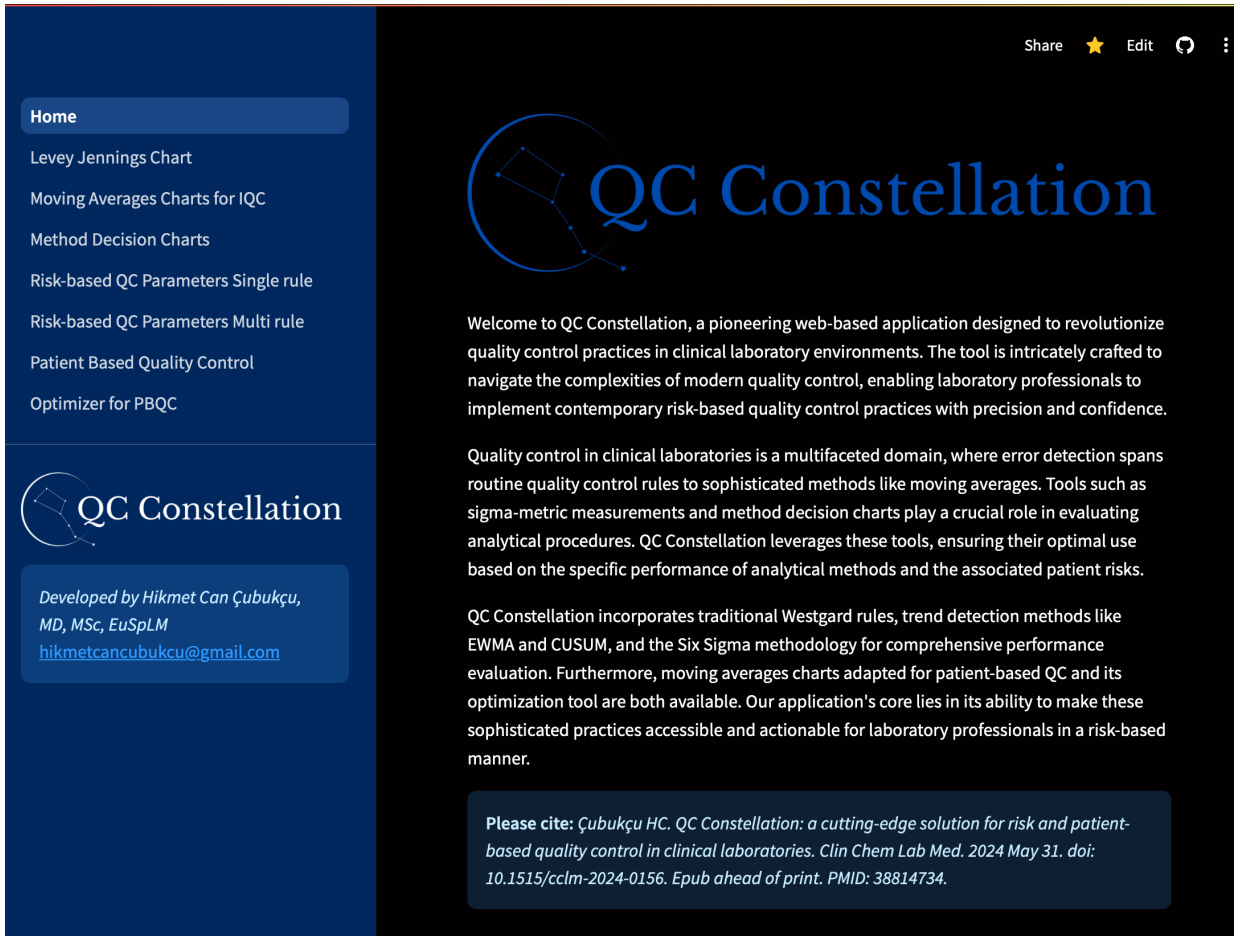
To manage the IQC plan with a risk-based approach:

- Sigma-metric
- Method decision charts
- Risk-based QC parameters

To track IQC results:

- Levey-Jennings charts
- Moving averages charts

A solution for risk and patient-based QC: QC Constellation



DE GRUYTER Clin Chem Lab Med 2024; 62(11): 2185–2197

Hikmet Can Çubukçu*

QC Constellation: a cutting-edge solution for risk and patient-based quality control in clinical laboratories

https://doi.org/10.1515/cclm-2024-0156
Received January 31, 2024; accepted April 30, 2024; published online May 31, 2024

Abstract

Objectives: Clinical laboratories face limitations in implementing advanced quality control (QC) methods with existing systems. This study aimed to develop a web-based application to addresses this gap, and improve QC practices.
Methods: QC Constellation, a web application built using Python 3.11, integrates various statistical QC modules. These include Levey-Jennings charts with Westgard rules, sigma-metric calculations, exponentially weighted moving average (EWMA) and cumulative sum (CUSUM) charts, and method decision charts. Additionally, it offers a risk-based QC section and a patient-based QC module aligning with modern QC practices. The codes and the web application links for QC Constellation were shared at https://github.com/hikmetc/QC_Constellation, and http://qcconstellation.com, respectively.
Results: Using synthetic data, QC Constellation demonstrated effective implementation of Levey-Jennings charts with user-friendly features like checkboxes for Westgard rules and customizable moving averages graphs. Sigma-metric calculations for hypothetical performance values of serum total cholesterol were successfully performed using allowable total error and maximum allowable measurement uncertainty goals, and displayed on method decision charts. The utility of the risk-based QC module was exemplified by assessing QC plans for serum total cholesterol, showcasing the application's capability in calculating risk-based QC parameters including maximum unreliable final patient results, risk management index, and maximum run size and offering risk-based QC recommendations. Similarly, the patient-based QC and optimization modules were demonstrated using simulated sodium results.
Conclusions: In conclusion, QC Constellation emerges as a pivotal tool for laboratory professionals, streamlining the management of quality control and analytical performance monitoring, while enhancing patient safety through optimized QC processes.

Keywords: performance evaluation; quality control; risk-based quality control; patient-based quality control; risk management; sigma metric

***Corresponding author: Hikmet Can Çubukçu**, MD, EuSpLM, General Directorate of Health Services, Rare Diseases Department, Turkish Ministry of Health, Bilkent Yerleskesi, 6001. Cadde, Universiteler Mahallesi 06800, Çankaya/Ankara, Türkiye,
E-mail: hikmetcancubukcu@gmail.com. https://orcid.org/0000-0001-5321-9354

Introduction

Quality control practices in clinical laboratory environments encompass a broad spectrum of error detection methods, ranging from routine quality control rules to moving averages. Additionally, various applications such as sigma-metric measurements and method decision charts are employed for evaluating the performance of analytical procedures. The appropriate use of these quality control tools, tailored to the performance of the analytical method and the risks of errors on patients, is now feasible with contemporary risk-based quality control practices.

In statistical quality control (SQC) applications, internal quality control results are typically evaluated using traditional Shewhart charts alongside Westgard rules [1]. While effective, Shewhart charts exhibit limitations in detecting small shifts or trends in the data. To address this, trend detection methods like Exponentially Weighted Moving Average (EWMA) and Cumulative Sum (CUSUM) charts have been introduced [2]. EWMA, obtained by calculating the weighted average of current and previous data, is superior to the Shewhart chart in detecting small shifts from the target value when a low weighting factor (0.05–0.2) is used [3, 4]. Notably, EWMA's application extends beyond internal SQC, finding utility in patient-based quality control programs as well [5]. On the other hand, CUSUM charts, another moving average technique, visually represent deviations from the target value [6]. Similar to EWMA, CUSUM charts offer improved detection of small shifts compared to Shewhart charts [7]. While its effectiveness in internal SQC is well-established [8], there is currently limited research exploring its implementation in patient-based quality

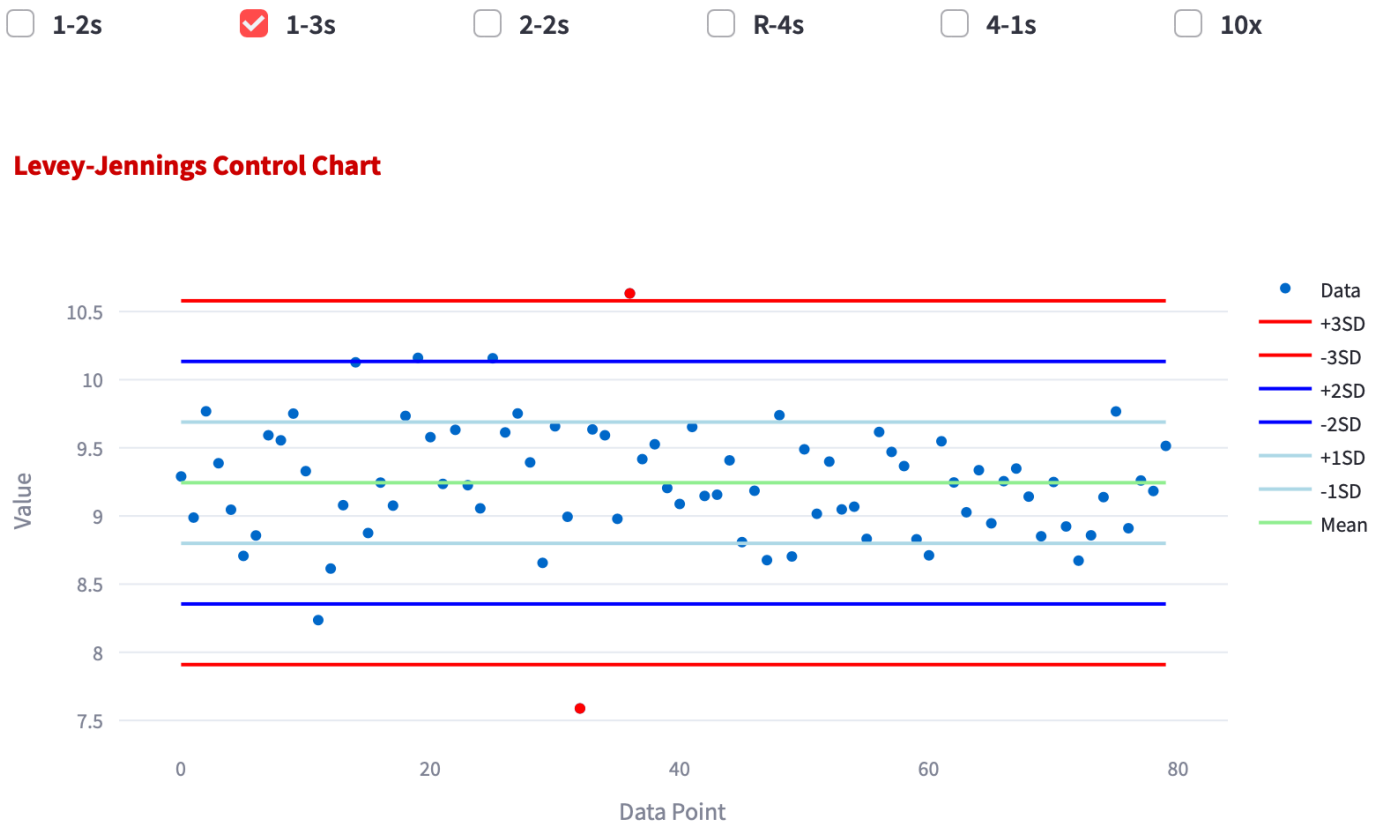
Scan for the app

App link: https://hikmetc-qcconstellation.streamlit.app

Çubukçu HC. Clin Chem Lab Med 2024; doi:10.1515/cclm-2024-0156.

Levey-Jennings control chart and Westgard rules

- Common practice in laboratory medicine.
- IQC results are evaluated using traditional Shewhart charts alongside Westgard rules.



Moving average charts

- Superior to the Shewhart chart in detecting small shifts in internal QC.
- Implementation in patient-based QC as a cost-effective alternative.

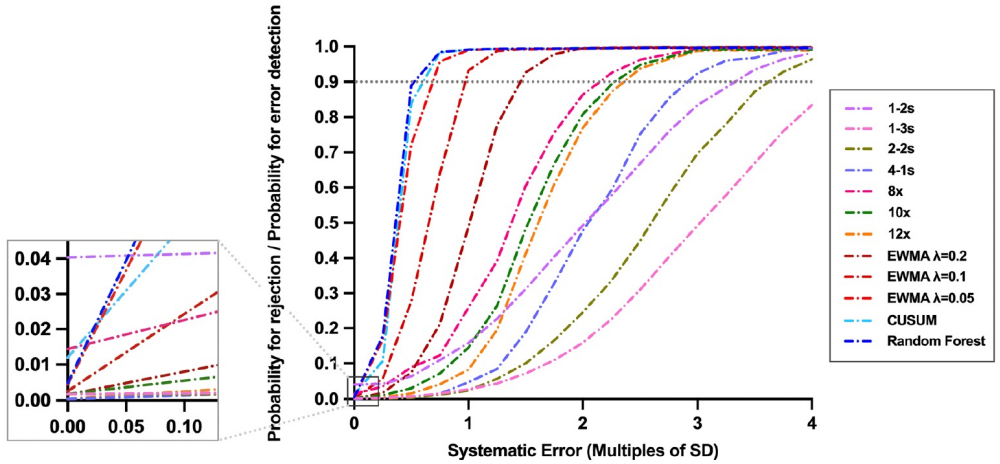
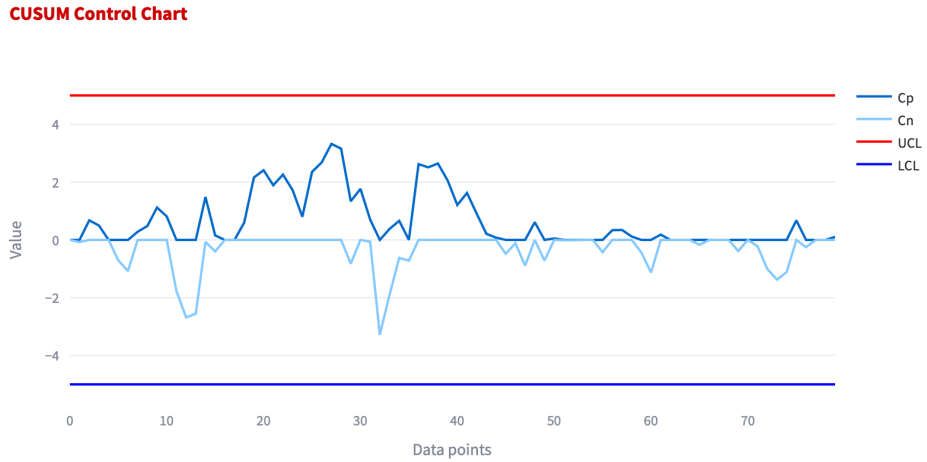
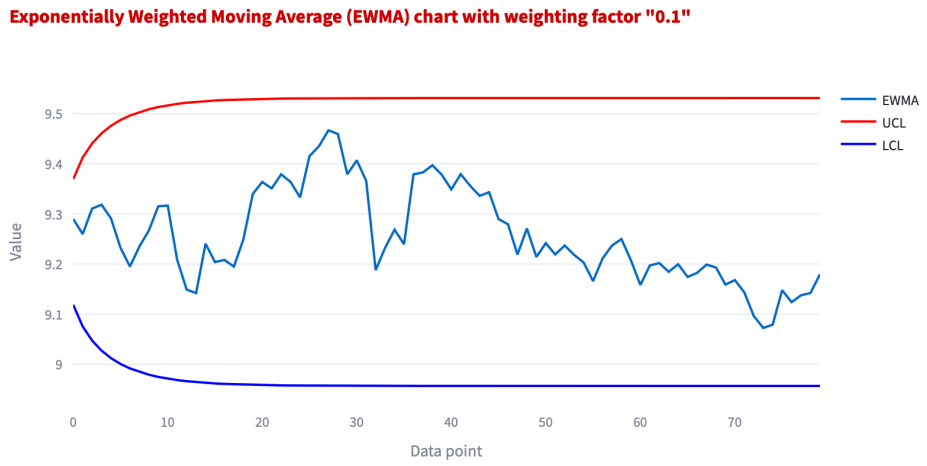


Figure 2: Power graph of EWMA, CUSUM, IQC rules and random forest model.





Çubukçu HC. Performance evaluation of internal quality control rules, EWMA, CUSUM, and the novel machine learning model. Turk J Biochem 2021;46:661–70.

Risk factors considered in the management of the IQC plan

Analytical performance

(Sigma-metric value)

Clinical impact of the error

(Severity of harm category of the measurand)

Six sigma practice

$$\text{Sigma-metric} = \frac{\text{TEa (\%)} - \text{Bias (\%)}}{CV_A\ (\%)} \quad *$$

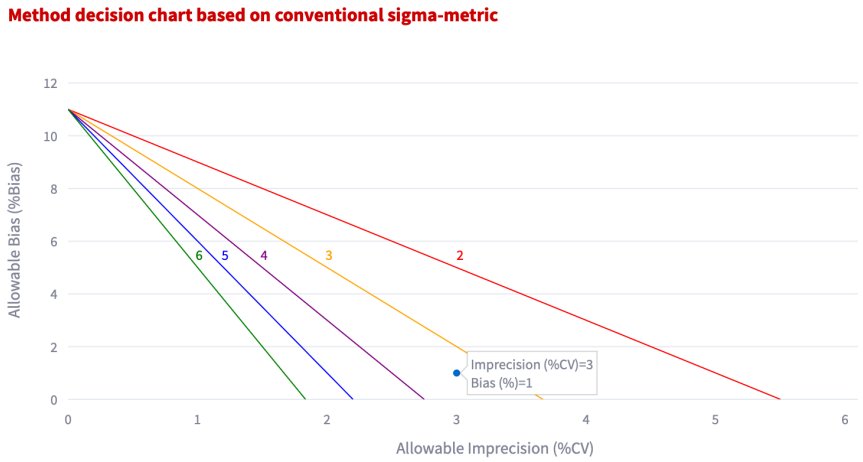
$$\text{Sigma-metric} = \frac{CV_I\ (\%)}{CV_A\ (\%)} \quad **$$

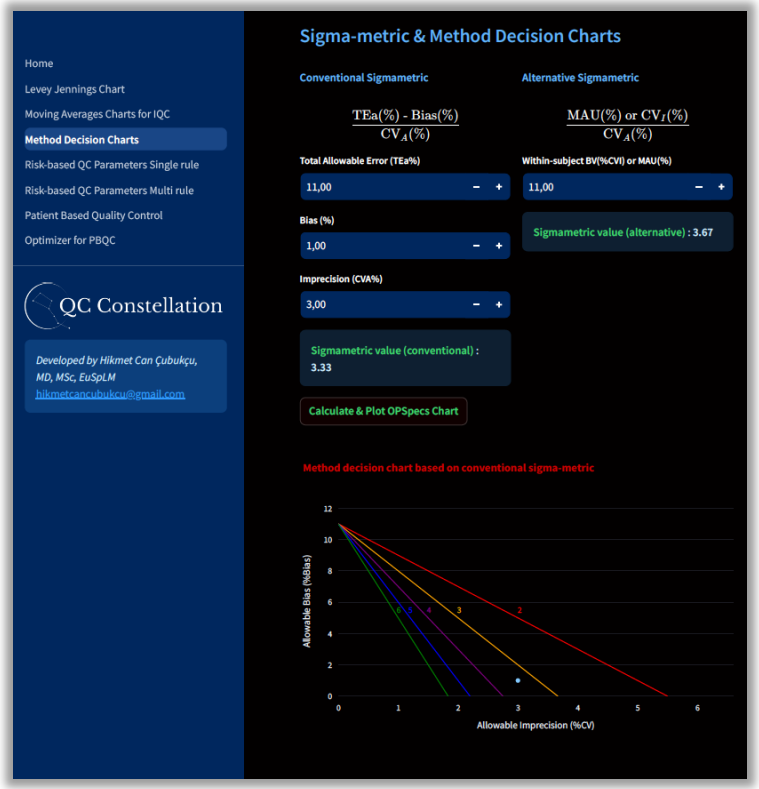
$$\text{Sigma-metric} = \frac{\text{MAU (\%)}}{CV_A\ (\%)} \quad ***$$

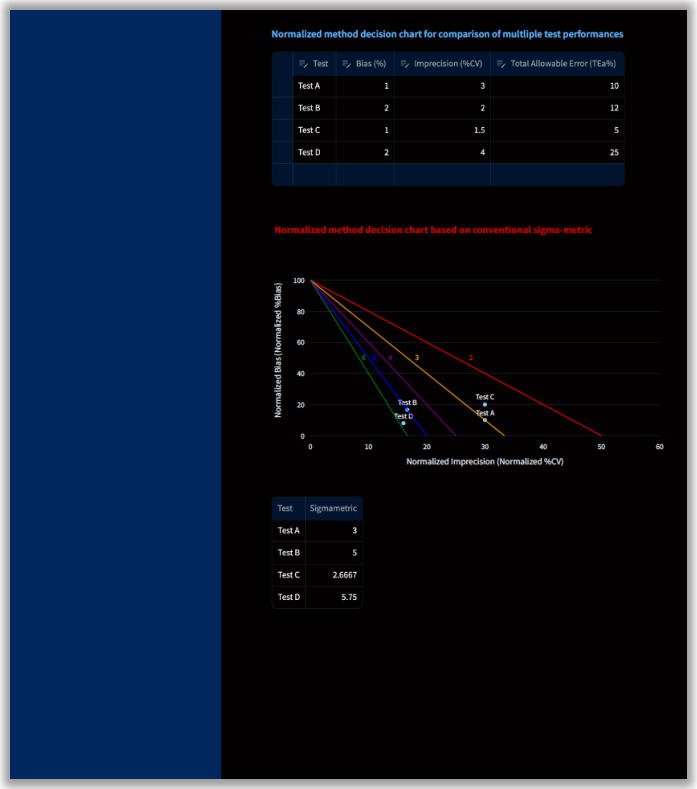
Sigma-metric degree	Description
< 2-sigma	Unacceptable
≥ 2-sigma – < 3-sigma	Poor
≥ 3-sigma – < 4-sigma	Marginal
≥ 4-sigma – < 5-sigma	Good
≥ 5-sigma – < 6-sigma	Excellent
> 6-sigma	World class

*Westgard JO, Westgard SA. Six Sigma quality management system and design of risk-based statistical quality control. Clin Lab Med 2017;37:85–96. **Oosterhuis WP, Coskun A. Sigma metrics in laboratory medicine revisited: we are on the right road with the wrong map. Biochem Med (Zagreb) 2018;28:020503. ***Çubukçu HC. Clin Chem Lab Med 2024; doi:10.1515/cclm-2024-0156.

Method decision charts





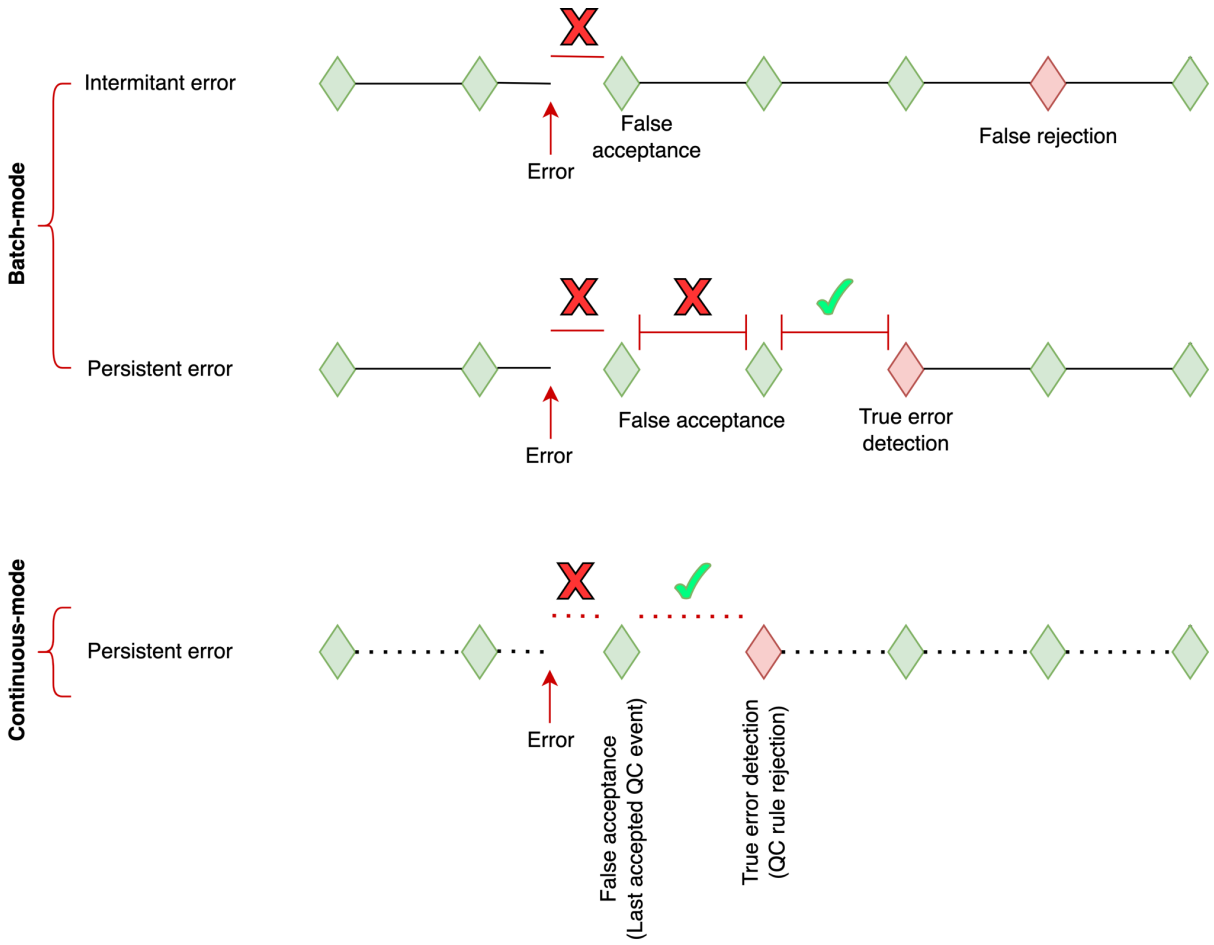


- y-axis: allowable bias (%); x-axis: allowable imprecision (%)
- The measurement procedure is plotted as a point (CVA on the x-axis, bias on the y-axis); coloured diagonal lines represent sigma-metric thresholds

QC Constellation: sigma-metric & method decision charts (conventional and normalized)

Çubukçu HC. Clin Chem Lab Med 2024; doi:10.1515/cclm-2024-0156.

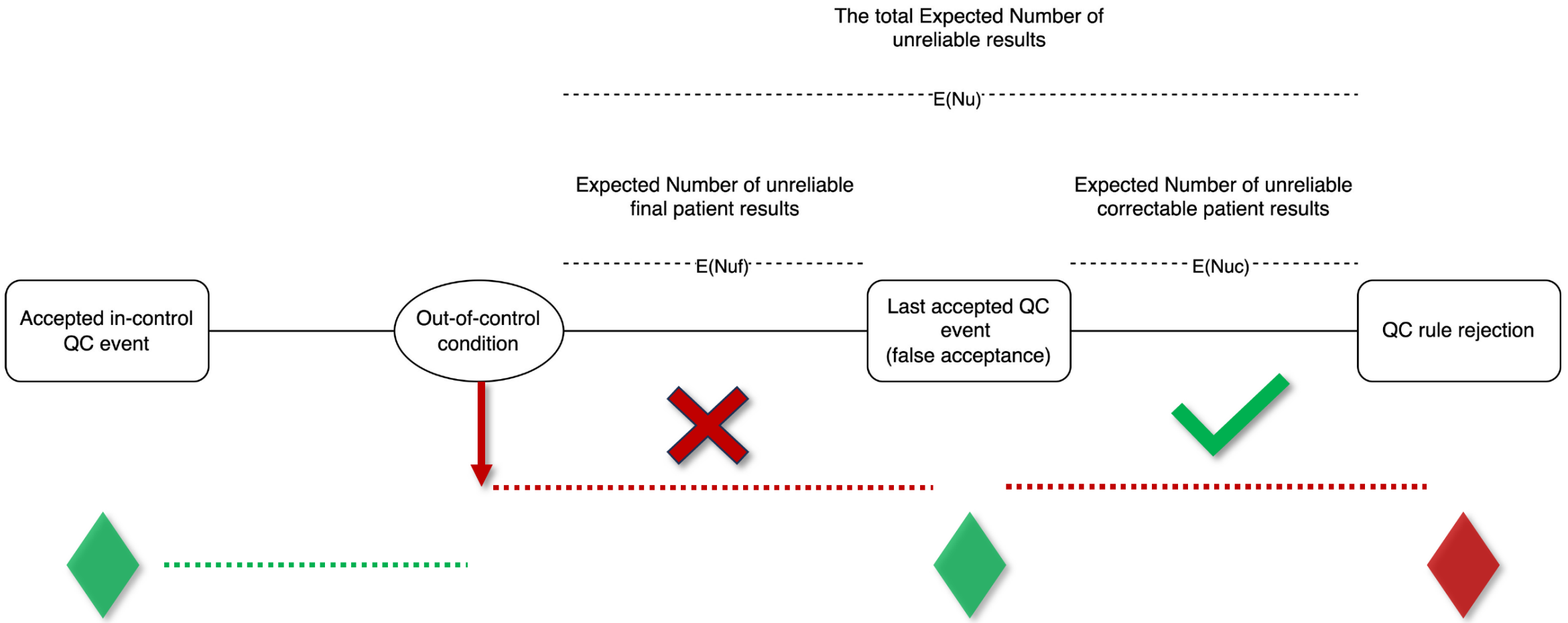
Laboratory testing modes



- **Batch mode:** the process operates in batch units; a batch is either in-control or out-of-control.
- **Intermittent error:** affects only a single batch. **Persistent error:** continues to affect subsequent batches until detected.
- **Continuous mode:** patient specimens are tested in a continuous stream; QC events occur periodically and an out-of-control condition might occur at any point in the testing stream.

Adapted from: Parvin CA. Assessing the impact of the frequency of quality control testing on the quality of reported patient results. Clin Chem 2008;54:2049–54.

Continuous mode in focus



MaxE(Nuf) and Pfr-based plan: probabilistic calculations

- **E(NB):** the expected number of patient samples between QC events.
- **E(N0):** the expected number of patient test results generated between the time an out-of-control condition occurs and the next QC event. E(N0) = E(NB)/2
- **E(NP):** the average number of patient samples processed from the start of an out-of-control condition to QC detection. E(NP) = E(N0) + E(NB) × (1/Ped − 1)
- **ΔPE:** increase in the probability of generating a result that exceeds the TEa (or MAU) due to the presence of an out-of-control error condition.
- **E(Nu):** the expected number of unreliable results. E(Nu) = E(NP) × ΔPE

E(Nuf) = ΔPE × [(ARLed − 1) × E(NB) − (1 − Ped) × (E(NB) − E(N0))]

MRS (maximum run size) = E(NB) / MaxE(Nuf) × (6 − severity of harm score)

Çubukçu HC. Clin Chem Lab Med 2024; doi:10.1515/cclm-2024-0156. · Parvin CA. Clin Chem 2008;54:2049–54.

Severity of harm scores (ISO 14971:2019 & ISO 24971:2020)

Score	Definition
1 – Negligible	Results in inconvenience or temporary discomfort
2 – Minor	Results in temporary injury or impairment not requiring medical or surgical intervention
3 – Serious / Major	Results in injury or impairment requiring medical or surgical intervention
4 – Critical	Results in permanent impairment or irreversible injury
5 – Catastrophic / Fatal	Results in death

AJCP | ORIGINAL ARTICLE

Measuring the impact

Severity of harm from laboratory errors in 195 tests

Hikmet Can Çubukçu, MD, MSc,[1] Murat Cihan, MD,[2] Hamit Hakan Alp, PhD,[3] Serkan Bolat, MD,[4] Oğuzhan Zengi,[5] Kamil Taha Uçar, MD,[5] Deniz İlhan Topcu, MD, PhD,[6] Muhammed Fevzi Kılınçkaya, MD,[7] Habib Özdemir, MD,[8,9] Murat Gülşen, MD,[1] Hayri Canbaz, MD,[10] Doğan Yücel,[11] Muhittin Abdulkadir Serdar, MD[12]

[1]Rare Diseases Department, General Directorate of Health Services, Turkish Ministry of Health, Ankara, Türkiye; [2]Ordu University Education and Research Hospital, Ordu, Türkiye; [3]Department of Biochemistry, Faculty of Medicine, Van Yüzüncü Yıl University, Van, Türkiye; [4]Medical Biochemistry Department, Sivas Cumhuriyet University, Sivas, Türkiye; [5]Department of Medical Biochemistry, Health Science University, Istanbul Başakşehir Çam and Sakura City Hospital, Istanbul, Türkiye; [6]Department of Medical Biochemistry, İzmir City Hospital, İzmir, Türkiye; [7]Department of Medical Biochemistry, Helse Møre og Romsdal, Molde, Norway; [8]Department of Medical Biochemistry, Faculty of Medicine, Manisa Celal Bayar University, Manisa, Türkiye; [9]Türkiye Health Data Research and Artificial Intelligence Applications Institute, Health Institutes of Türkiye, Istanbul, Türkiye; [10]Department of Emergency Medicine, University of Yıldırım Beyazıt, Yenimahalle Training and Research Hospital, Ankara, Türkiye; [11]Department of Medical Biochemistry, Lokman Hekim University, Ankara, Türkiye; [12]Department of Medical Biochemistry, Acibadem Mehmet Ali Aydinlar University, Ankara, Türkiye

ABSTRACT

Objectives: This study aimed to objectively assess the potential severity of harm associated with erroneous results in 195 laboratory tests by surveying 514 specialist physicians and medical biochemistry experts.

Methods: The survey obtained participants' (75 medical biochemists, 439 clinicians) opinions on severity of harm for the erroneous results of 195 tests. The comprehensive list of errors and their effects on test results were obtained from the literature, and then matched with severity of harm scores, from 1 (negligible effect) to 5 (life-threatening injury/death), obtained from the survey responses.

Results: Participants perceived tests such as cardiac biomarkers, blood gases, coagulation parameters (activated partial thromboplastin time, prothrombin time, international normalized ratio, and dimerized plasmin fragment D), critical ions (potassium, sodium), toxic trace elements (lead, mercury), and specific serum drug levels (lithium, digoxin) to have a greater potential for patient harm in case of errors. Medical biochemistry specialists assigned higher severity scores to some laboratory tests, including total bilirubin, pseudocholinesterase, platelet indices, and some drug levels (cyclosporine, methotrexate, vancomycin).

Conclusions: A substantial agreement (91%) was observed between medical biochemists and clinicians in terms of the most frequently chosen severity of harm score. The study provided objective severity scores and identified high-risk tests for targeted quality improvement.

KEY WORDS

laboratory errors; severity of harm; patient safety; risk; risk management

Am J Clin Pathol XXXX 2024;XX:1-0
HTTPS://DOI.ORG/10.1093/AJCP/AQAE144

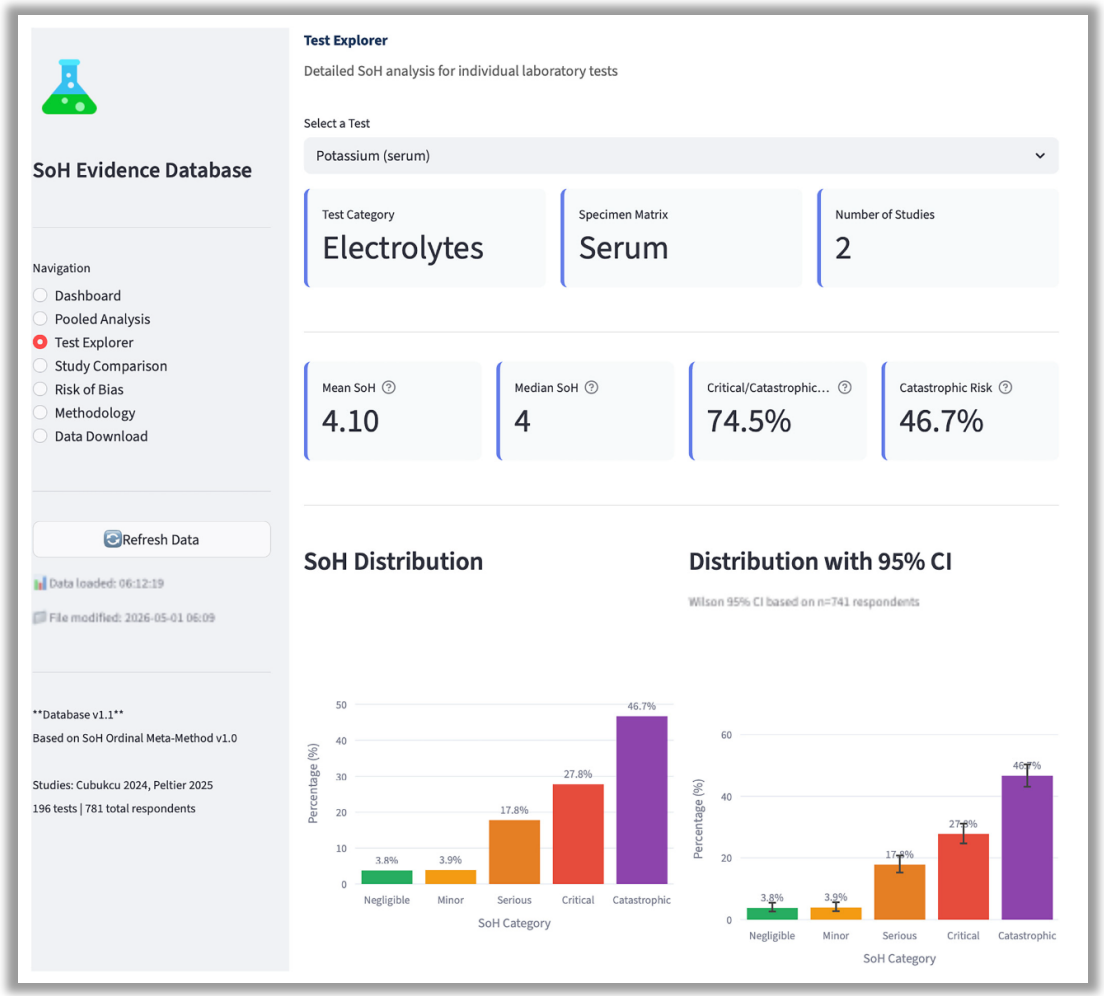
Received: August 7, 2024
Accepted: September 27, 2024

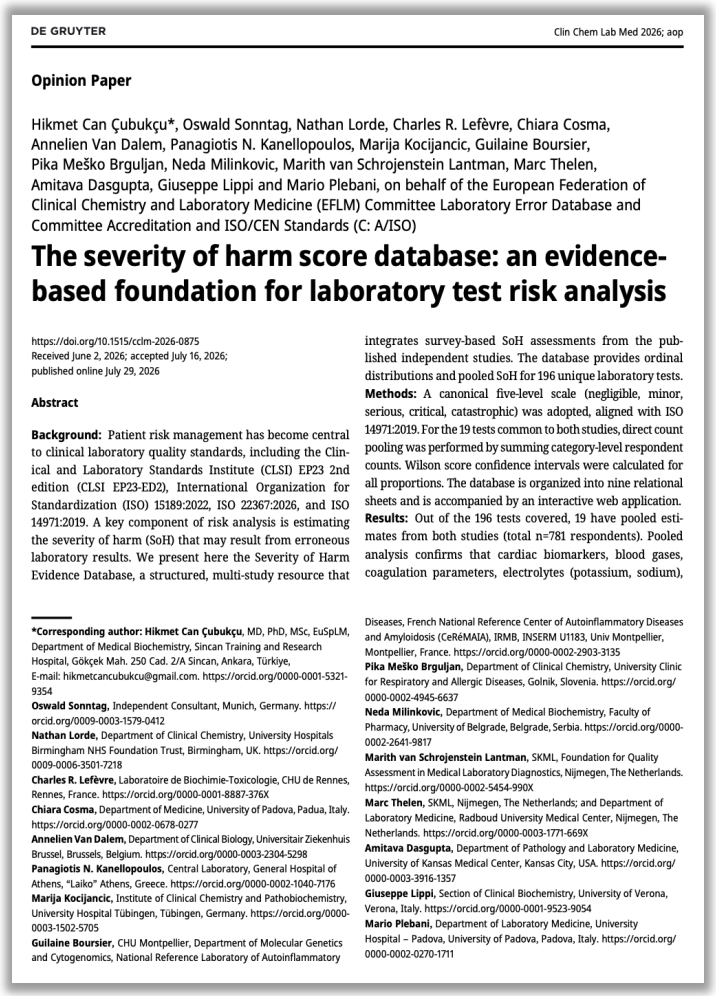
Corresponding author: Hikmet Can Çubukçu; hikmetcancubukcu@gmail.com

Funding: This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

Çubukçu HC, et al. Measuring the impact: severity of harm from laboratory errors in 195 tests. Am J Clin Pathol 2024; doi:10.1093/ajcp/aqae144.

Severity of harm score database (ISO 14971:2019 & ISO 24971:2020)





Scan for the SoH database

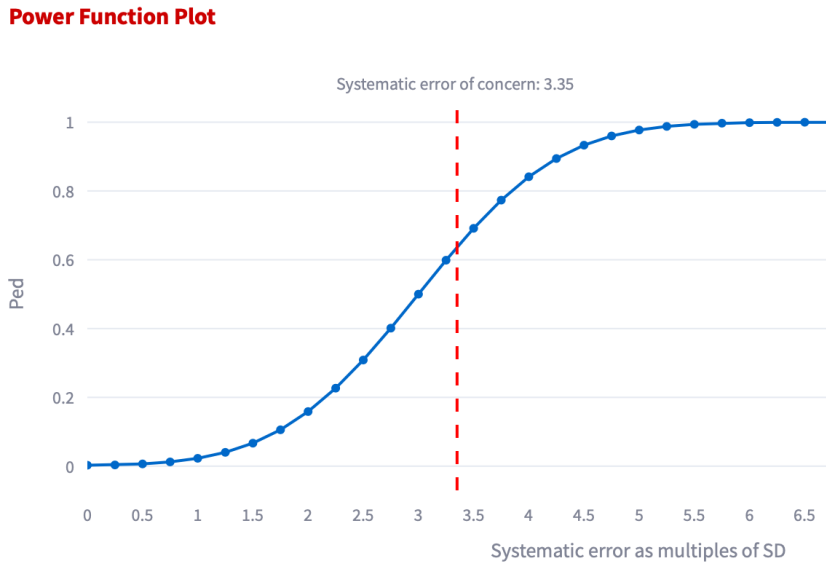
Çubukçu et al.: The severity of harm score database. CCLM 2026; doi:10.1515/cclm-2026-0875

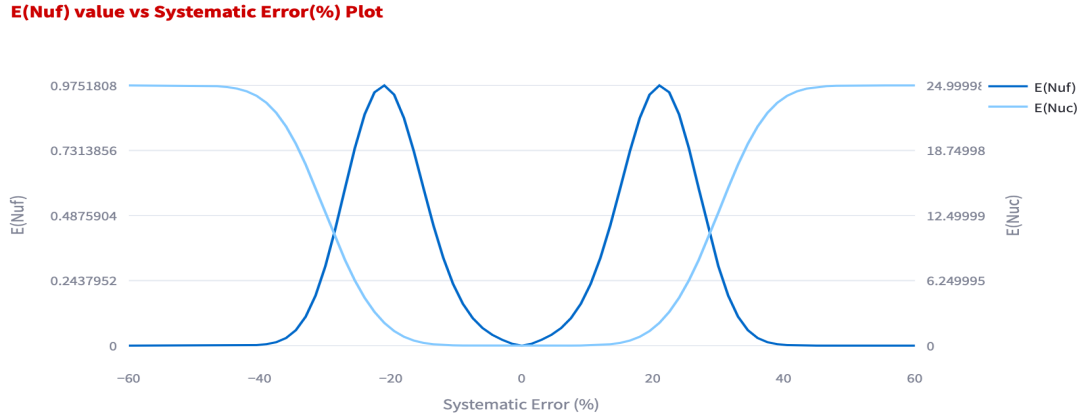
MaxE(Nuf) and Pfr-based plan: troponin example

Risk based QC Parameters

Parameter	*Value*
Probability of false rejection (%)	0.27
Sigma-metric value	5.0
Systematic error (%) at MaxE(Nuf)	21.0
MaxE(Nuf)	0.98
Max run size	51

hs-troponin
(catastrophic severity of harm)
Pfr 0.27 %
sigma 5.0
MaxE(Nuf) 0.98
max run size 51





Interpretation

- **If Pfr > 5%:** try a different multirule or reduce the number of QC measurements.
- **If MaxE(Nuf) > (6 – severity of harm score):** reduce the number of patient samples between QC events or try a multirule with more QC measurements.

Çubukçu HC. Clin Chem Lab Med 2024; doi:10.1515/cclm-2024-0156.

CLSI EP23 ed2-based plan: terms

- **PE:** the probability of producing an erroneous result across the range of out-of-control conditions.*
- **P(TEa in-control):** the probability of a result exceeding TEa (or MAU) when the analytical process is in-control.
- **ARLed:** the average run length to error detection.
- **MPBF:** the mean number of patients between failures. This value is calculated by multiplying the average number of days between instrument failures by the average number of patients tested per day.***
- **Ph|u:** the probability of harm given an unacceptable result, which represents the conditional probability of an erroneous result harming a patient.**
- **Acceptable PH** for negligible, minor, serious, critical, and catastrophic severity of harm categories is 0.01, 0.001, 0.0001, 0.00001, and 0.000001, respectively.*

*CLSI. EP23-ED2:2023 – Laboratory quality control based on risk management. **Parvin CA, Baumann NA. Assessing quality control strategies for HbA1c measurements from a patient risk perspective. J Diabetes Sci Technol 2018;12:786–91. ***Karnutsch D, Occhipinti F, Tumiatti D, Mueller T. Evaluation of the impact of changing quality control rules and frequency on the risk management index. Lab Med 2021;52:211–18.

CLSI EP23 ed2-based plan

$$P_E = P(\text{TEa in-control}) + \frac{E(\text{Nuf})}{\text{MPBF} + \text{ARLed}}$$

$$\text{Predicted } P_H = P_E \times P_{h|u}$$

$$\text{RMI} = \frac{\text{Predicted } P_H}{\text{Acceptable } P_H}$$

RMI: risk management index

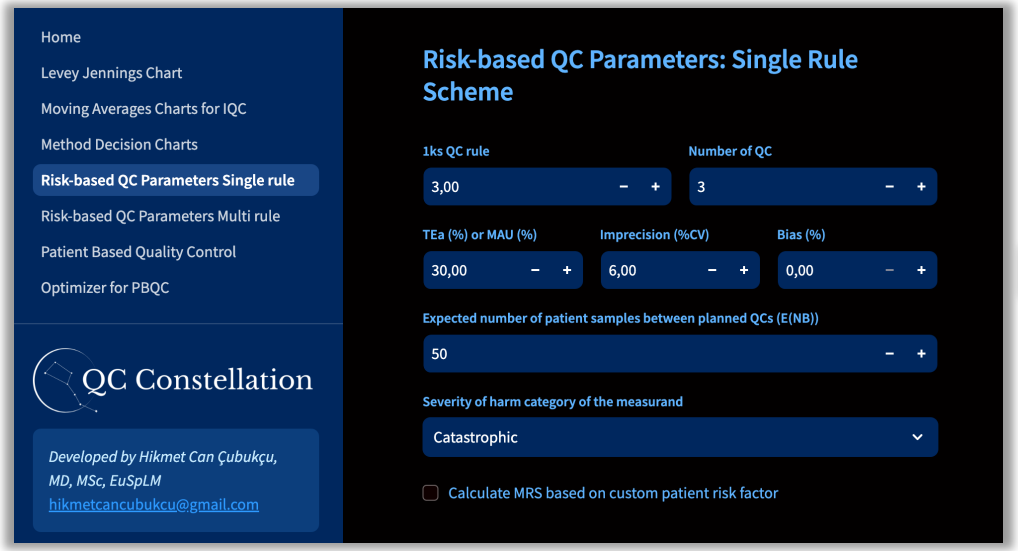
Interpretation: If RMI < 1, the QC strategy is regarded to have managed risk.

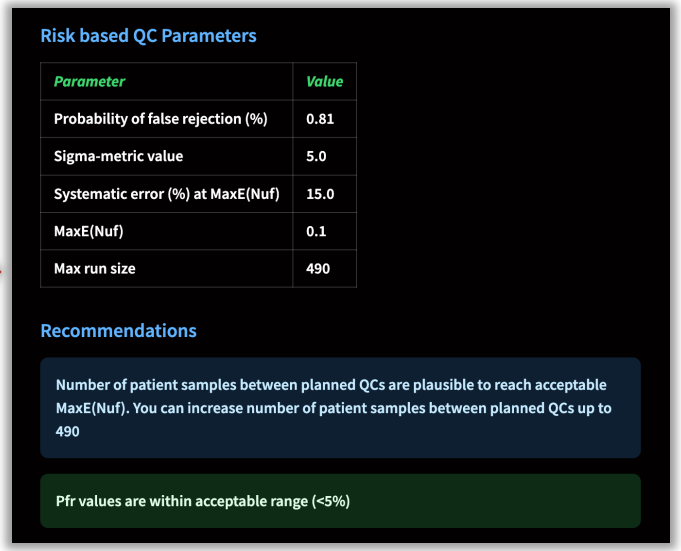
Severity of harm	Acceptable PH
Negligible	0.01
Minor	0.001
Serious	0.0001
Critical	0.00001
Catastrophic	0.000001

CLSI. EP23-ED2:2023 – Laboratory quality control based on risk management. 2023. · Parvin CA, Baumann NA. J Diabetes Sci Technol 2018;12:786–91. · Karnutsch D, et al. Lab Med 2021;52:211–18.

Both plans in QC Constellation

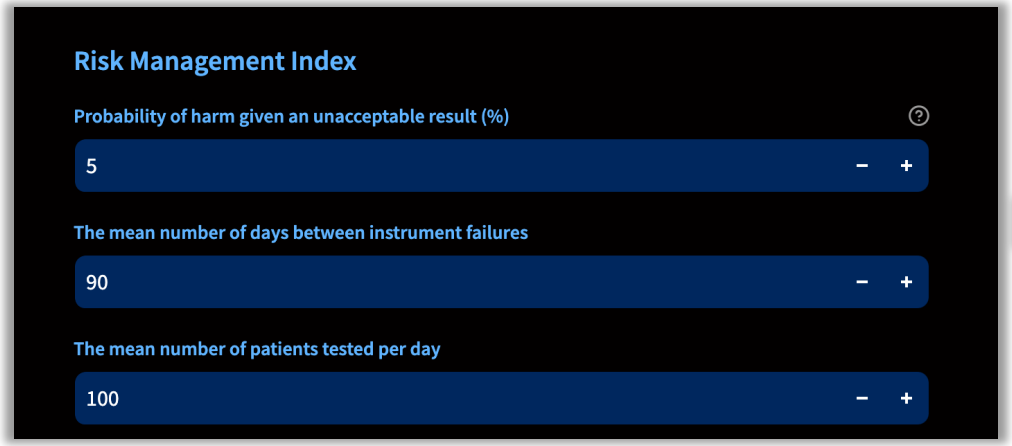
MaxE(Nuf) and Pfr-based plan





Risk-based QC parameters and recommendations (maximum run size, Pfr acceptability)

CLSI EP23 ed2-based plan



Risk management index (RMI) and its acceptability (≤ 1)

Disadvantages of both approaches

CLSI EP23 ed2:

- The document does not explicitly elucidate the intricate calculations of probabilities.
- EP23 ed2 does not directly assist in determining the maximum run size and proper QC rule.*

MaxE(Nuf) and Pfr-based method:

- MaxE(Nuf) serves as a surrogate metric for patient risk.
- The probability of harm to patients due to the inaccurate results reported for a particular test is not based only on the quantity of reported erroneous results, but also on the associated risk factors, such as the probability and severity of harm associated with an erroneous result for that specific test in relation to a patient.**

*Çubukçu HC. Clin Chem Lab Med 2024; doi:10.1515/cclm-2024-0156. **Yago M. Risk-based statistical quality control planning should be based on patient risk. J Appl Lab Med 2018;2:970–1.

Advantages of QC Constellation

Combines the EP23 ed2 approach with the MaxE(Nuf)–Pfr-based approach to present a more comprehensive and eclectic method.

Includes the severity of harm category of the measurand (clinical value of the measurand) in QC rule and frequency (or MRS) selection.

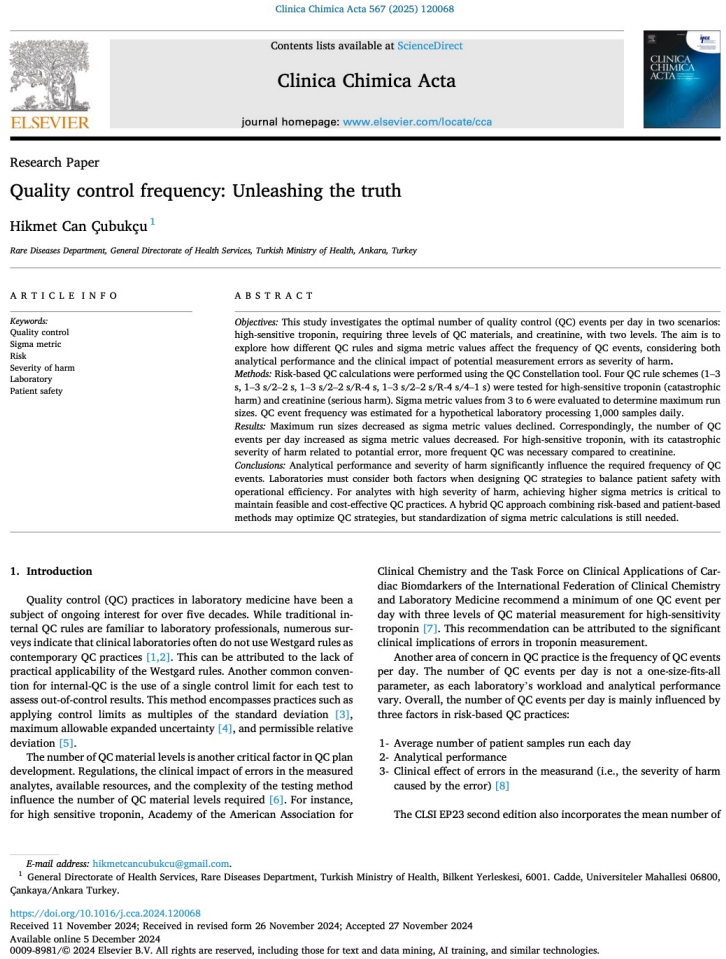
QC Constellation

hikmetc-qcconstellation.streamlit.app

Çubukçu HC. QC Constellation: a cutting-edge solution for risk and patient-based quality control in clinical laboratories. Clin Chem Lab Med 2024; doi:10.1515/cclm-2024-0156.

Irregular errors call for continuous monitoring of patient results

- **Irregular errors, may not be detected by routine risk-based internal QC strategy**, such as sudden instrument malfunctions or sample-specific interferences.
- Therefore, **continuous monitoring of patient results** is necessary to detect such errors on time, which requires the application of **patient-based quality control**.
- A **hybrid QC approach** combining risk-based and patient-based QC methods may optimize QC strategies, but **standardization of sigma metric calculations is still needed**.



Clinica Chimica Acta 567 (2025) 120068

Contents lists available at ScienceDirect

Clinica Chimica Acta

journal homepage: www.elsevier.com/locate/cca

Research Paper

Quality control frequency: Unleashing the truth

Hikmet Can Çubukçu [1]

Rare Diseases Department, General Directorate of Health Services, Turkish Ministry of Health, Ankara, Turkey

ARTICLE INFO

Keywords:
Quality control
Sigma metric
Risk
Severity of harm
Laboratory
Patient safety

ABSTRACT

Objectives: This study investigates the optimal number of quality control (QC) events per day in two scenarios: high-sensitive troponin, requiring three levels of QC materials, and creatinine, with two levels. The aim is to explore how different QC rules and sigma metric values affect the frequency of QC events, considering both analytical performance and the clinical impact of potential measurement errors as severity of harm.
Methods: Risk-based QC calculations were performed using the QC Constellation tool. Four QC rule schemes (1–3 s, 1–3 s/2–2 s, 1–3 s/2–2 s/R-4 s, 1–3 s/2–2 s/R-4 s/4–1 s) were tested for high-sensitive troponin (catastrophic harm) and creatinine (serious harm). Sigma metric values from 3 to 6 were evaluated to determine maximum run sizes. QC event frequency was estimated for a hypothetical laboratory processing 1,000 samples daily.
Results: Maximum run sizes decreased as sigma metric values declined. Correspondingly, the number of QC events per day increased as sigma metric values decreased. For high-sensitive troponin, with its catastrophic severity of harm related to potantial error, more frequent QC was necessary compared to creatinine.
Conclusions: Analytical performance and severity of harm significantly influence the required frequency of QC events. Laboratories must consider both factors when designing QC strategies to balance patient safety with operational efficiency. For analytes with high severity of harm, achieving higher sigma metrics is critical to maintain feasible and cost-effective QC practices. A hybrid QC approach combining risk-based and patient-based methods may optimize QC strategies, but standardization of sigma metric calculations is still needed.

1. Introduction

Quality control (QC) practices in laboratory medicine have been a subject of ongoing interest for over five decades. While traditional internal QC rules are familiar to laboratory professionals, numerous surveys indicate that clinical laboratories often do not use Westgard rules as contemporary QC practices [1,2]. This can be attributed to the lack of practical applicability of the Westgard rules. Another common convention for internal-QC is the use of a single control limit for each test to assess out-of-control results. This method encompasses practices such as applying control limits as multiples of the standard deviation [3], maximum allowable expanded uncertainty [4], and permissible relative deviation [5].

The number of QC material levels is another critical factor in QC plan development. Regulations, the clinical impact of errors in the measured analytes, available resources, and the complexity of the testing method influence the number of QC material levels required [6]. For instance, for high sensitive troponin, Academy of the American Association for Clinical Chemistry and the Task Force on Clinical Applications of Cardiac Biomdarkers of the International Federation of Clinical Chemistry and Laboratory Medicine recommend a minimum of one QC event per day with three levels of QC material measurement for high-sensitivity troponin [7]. This recommendation can be attributed to the significant clinical implications of errors in troponin measurement.

Another area of concern in QC practice is the frequency of QC events per day. The number of QC events per day is not a one-size-fits-all parameter, as each laboratory's workload and analytical performance vary. Overall, the number of QC events per day is mainly influenced by three factors in risk-based QC practices:

1- Average number of patient samples run each day
2- Analytical performance
3- Clinical effect of errors in the measurand (i.e., the severity of harm caused by the error) [8]

The CLSI EP23 second edition also incorporates the mean number of

E-mail address: hikmetcancubukcu@gmail.com.
[1] General Directorate of Health Services, Rare Diseases Department, Turkish Ministry of Health, Bilkent Yerleskesi, 6001. Cadde, Universiteler Mahallesi 06800, Çankaya/Ankara Turkey.

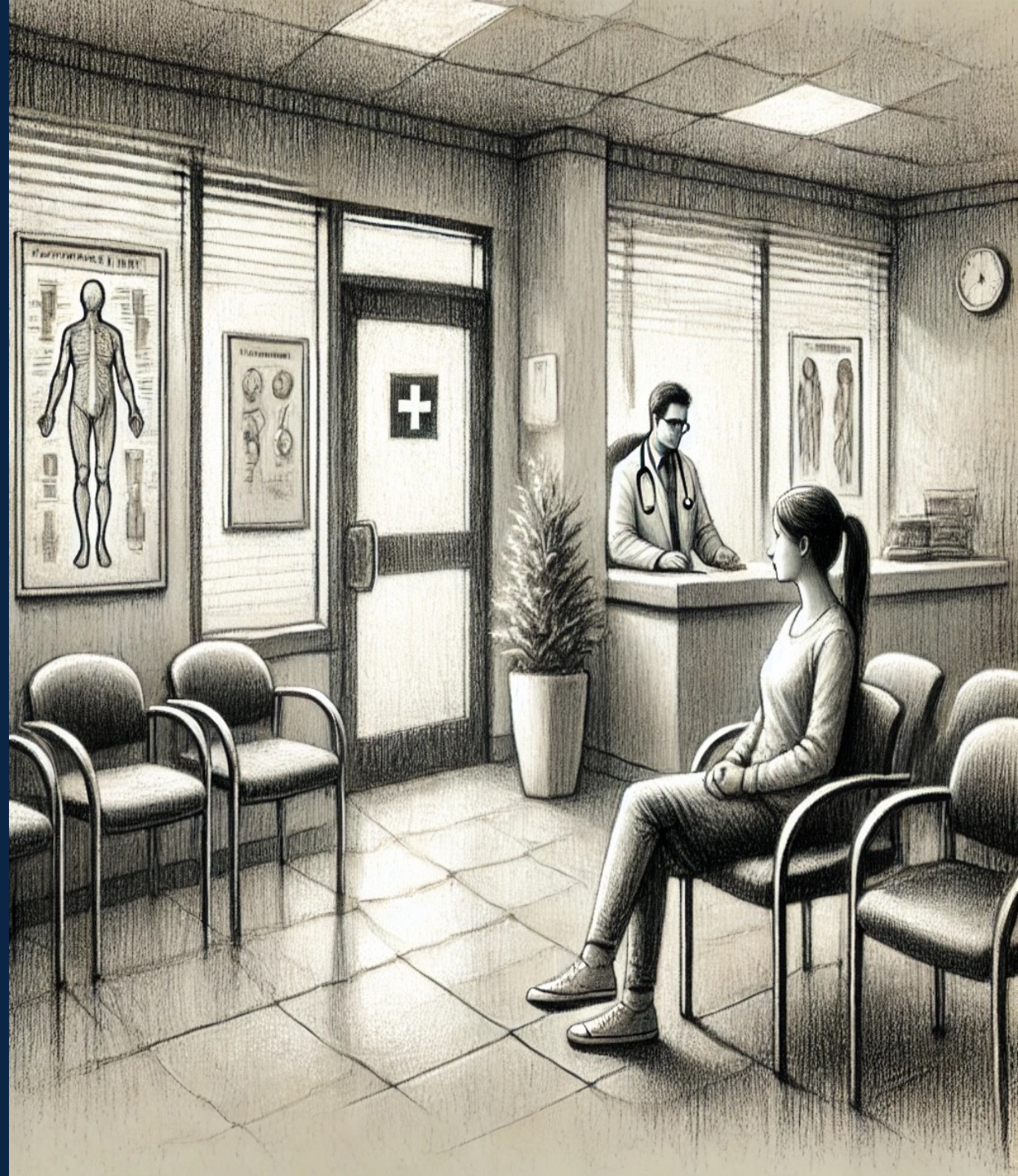
https://doi.org/10.1016/j.cca.2024.120068
Received 11 November 2024; Received in revised form 26 November 2024; Accepted 27 November 2024
Available online 5 December 2024

Çubukçu HC. Quality control frequency: unleashing the truth. Clin Chim Acta 2025;567:120068.

4

Patient-Based Quality Control

Using patient results to monitor analytical performance in real time



ISO 15189:2022(E)

Better: quality requirements are not met

c) If appropriate IQC material is not available, the laboratory shall consider the use of other methods for IQC. Examples of such other methods may include:

1) trend analysis of patient results, e.g. with moving average of patient results, or percentage of samples with results below or above certain values or associated with a diagnosis;

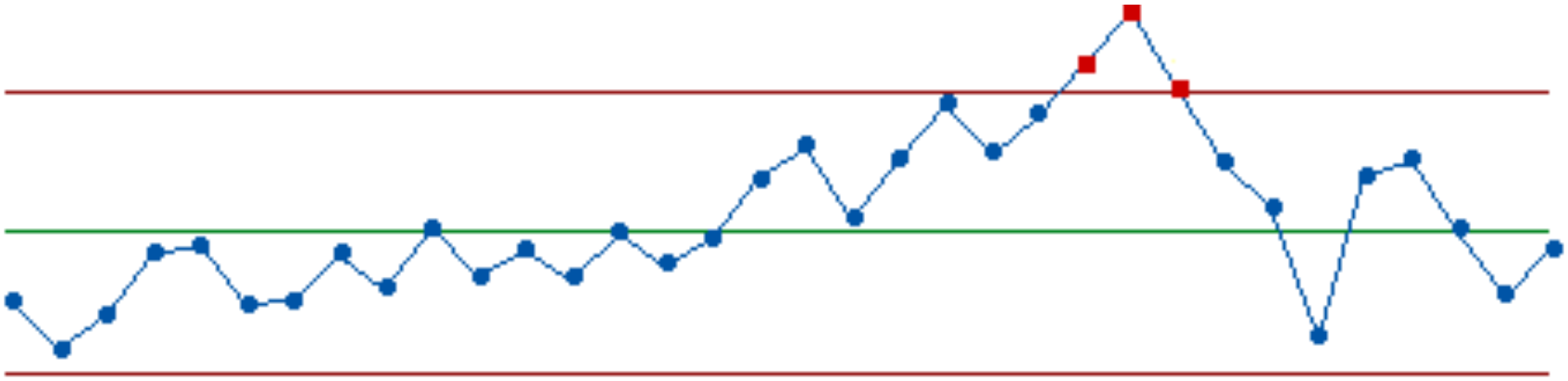
2) comparison of results for patient samples on a specified schedule to results for patient samples examined by an alternative procedure validated to have its calibration metrologically traceable to the same or higher order references as specified in ISO 17511;

3) retesting of retained patient samples.

Trend analysis of patient results (e.g. moving average of patient results) is listed among the other methods for IQC when appropriate IQC material is not available.

History of patient-based quality control

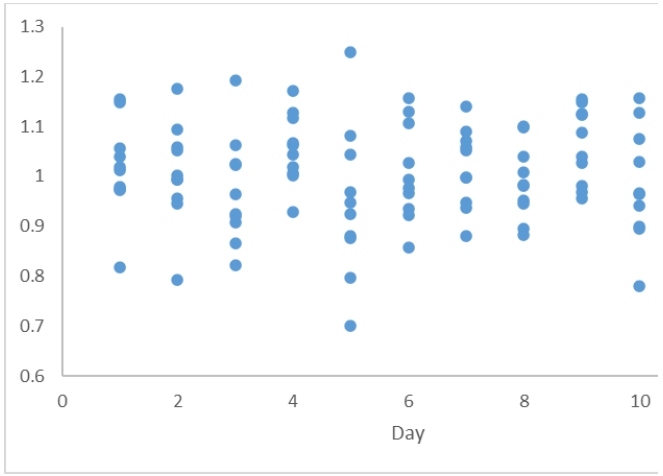
- First introduced by Dorsey in 1963. The average of patient of mean corpuscular hemoglobin as means of QC for Coulter Counter.
- In 1965, Hoffman and Waid proposed that the average of patient's test values are taken at the end of each day.
- Alternative to IQC **?**

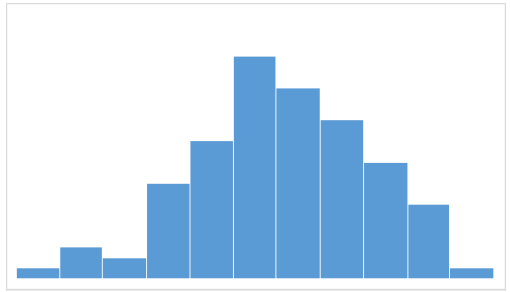


Crit Rev Clin Lab Sci 2020;57(8):532–47.

What is patient-based quality control?

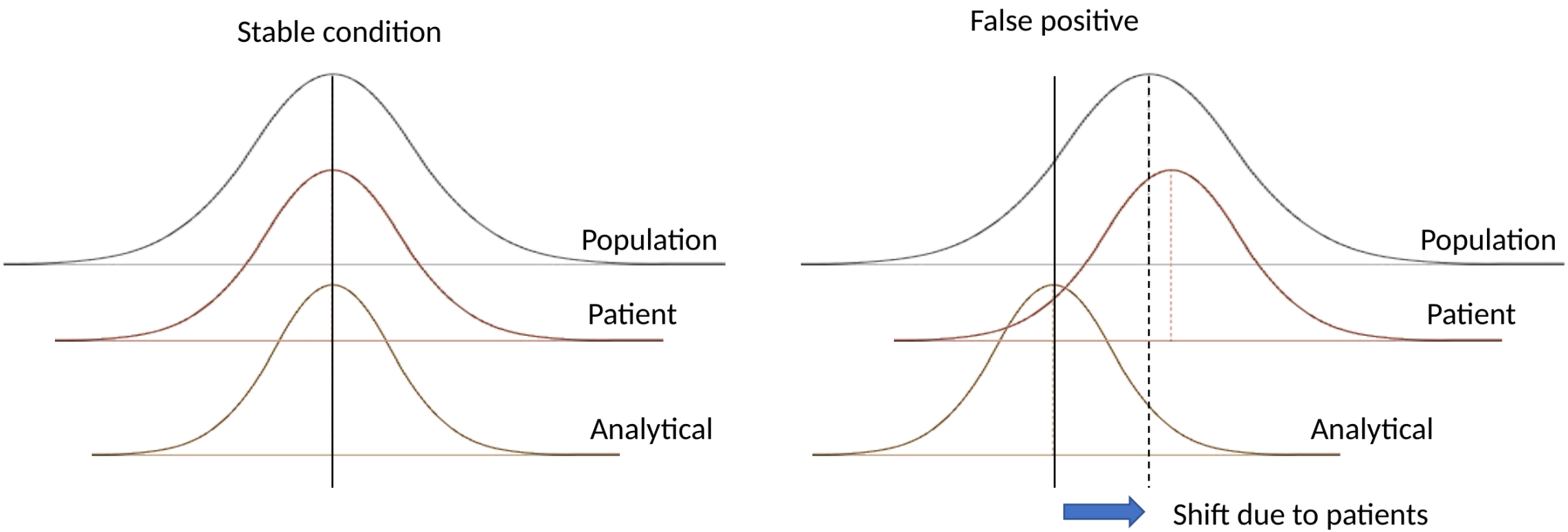
- QC that uses patient results to monitor analytical performance
- Patient result calculated as moving statistics (e.g. moving average)
- Predefined control limits to indicate an out-of-control condition
- Monitor in ‘real-time’ (as new result generated)



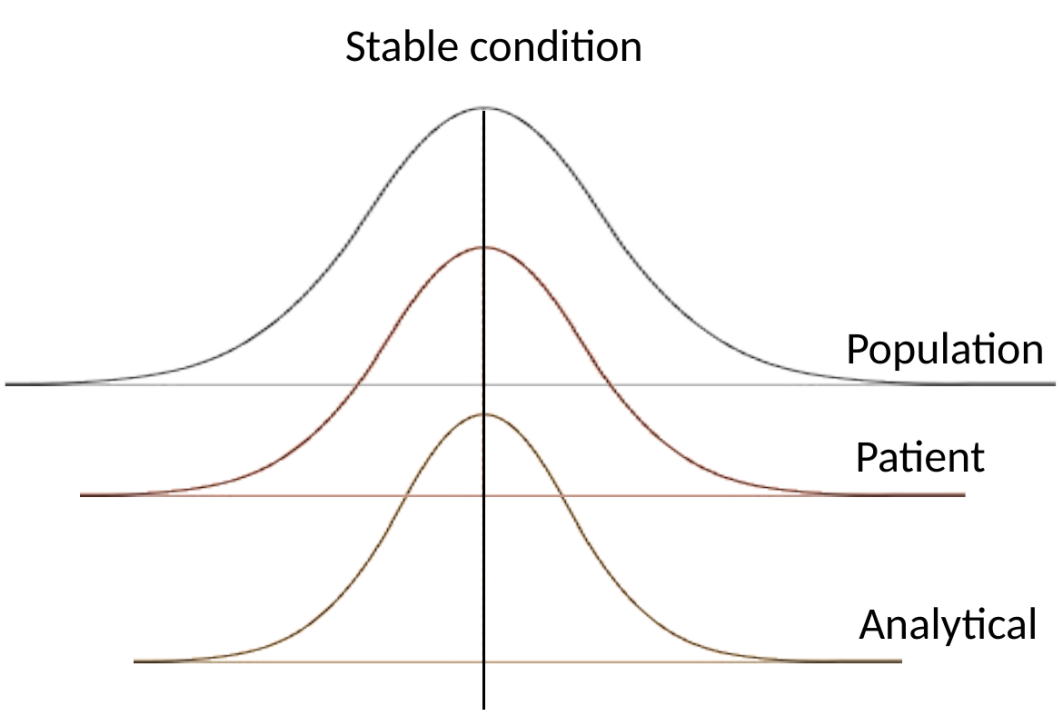


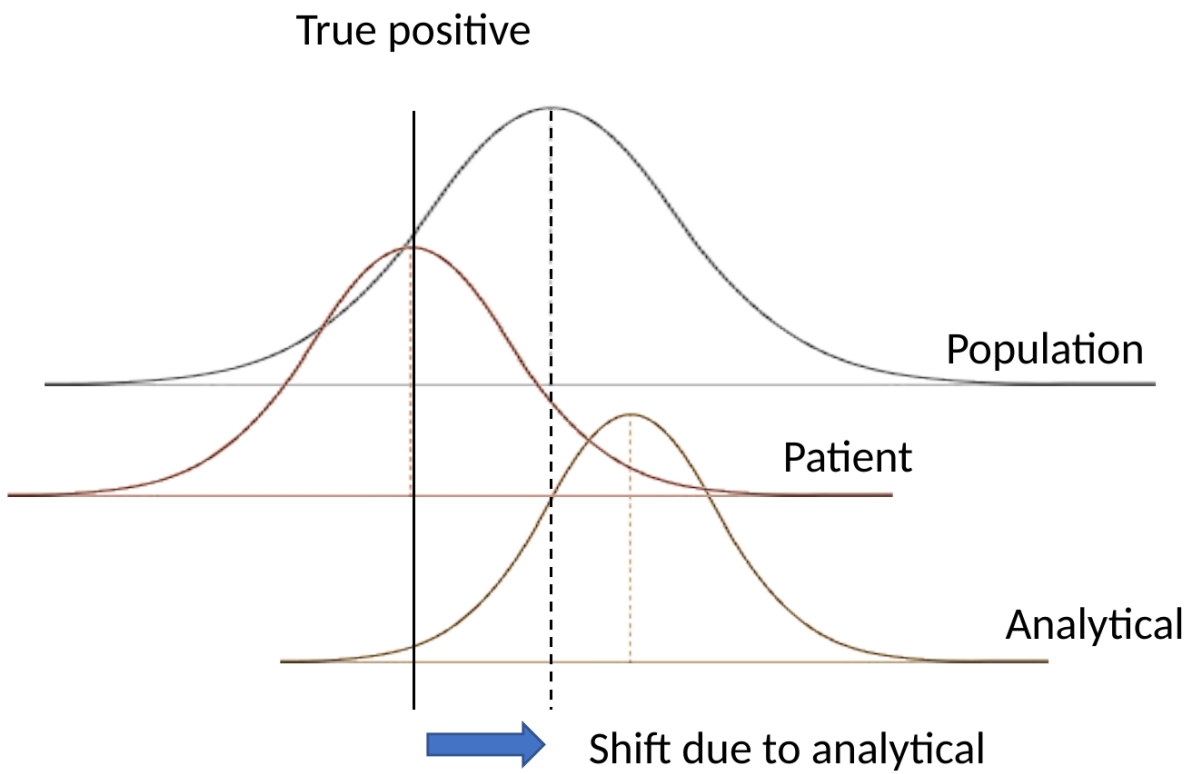
Clin Chem 2020;66(8):1072–83.

Shift due to patients: false positive



Shift due to analytical error: true positive





Moving average

- Average: the sum of the values divided by number of values
- Moving average continuously recalculates the average within a fixed block of patient results when a new patient result is produced
- Also known as rolling average or rolling mean
- The number of values in the block (block size), n, is user-defined
- Influenced by patient characteristics and analytical performance – need to minimise patient variation

$$\text{average} = \frac{\text{sum of values}}{\text{number of values}}$$

Crit Rev Clin Lab Sci 2020;57(8):532–47.

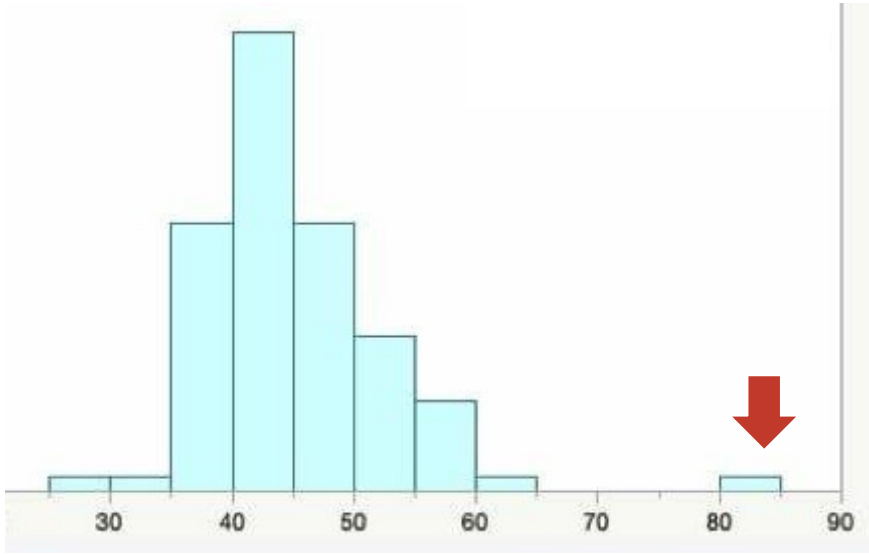
Moving average example (block size = 3)

Value	Moving average
5	–
6	–
7	(5+6+7)/3 = 6
8	(6+7+8)/3 = 7
6	(7+8+6)/3 = 7
15	**(8+6+15)/3 = 9.7**
9	10.0
5	9.7
8	7.3
3	5.3

- **Advantage:** reduce noise / ironing out variation
- **Drawback:** affected by extreme values

Outlier removal

- Outlier is an observation data that is distant from other observations
- Outliers should be removed / minimised
- Reduce distortion of the overall distribution
- Improve performance of model
- Drawback: reduced data, power of detection



Truncation

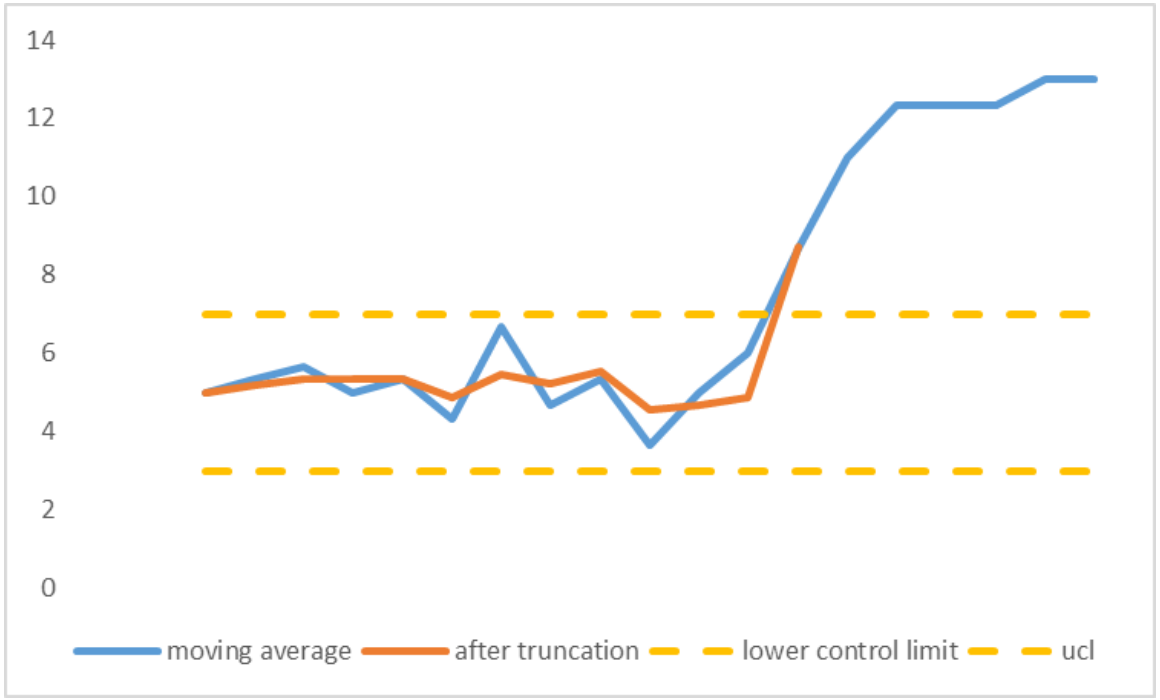
- Exclude the value (outliers) when it exceeds truncation limits
- Truncation limits can be defined: Kth percentile
- E.g. truncation lower and upper limits are 2 and 10
- Reduce the variability of the results / effects of extreme values
- Data is discarded

Value	MA (without truncation)
5	–
6	–
7	6.0
12	8.3
8	9.0
6	8.7
4	6.0
9	6.3
1	4.7
5	5.0
8	4.7
3	5.3

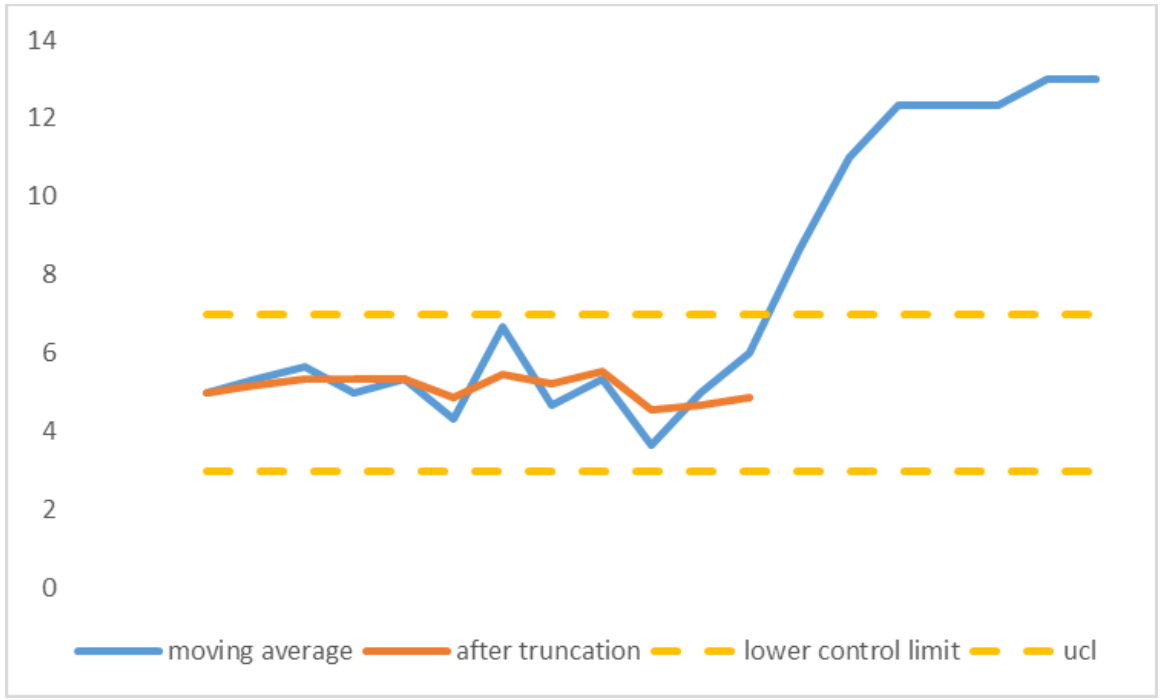
Value	MA (with truncation)
5	–
6	–
7	6
8	7
6	7
4	6
9	6.3
5	6
8	7.3
3	5.3

Truncation (examples)

Upper truncation limit = 10

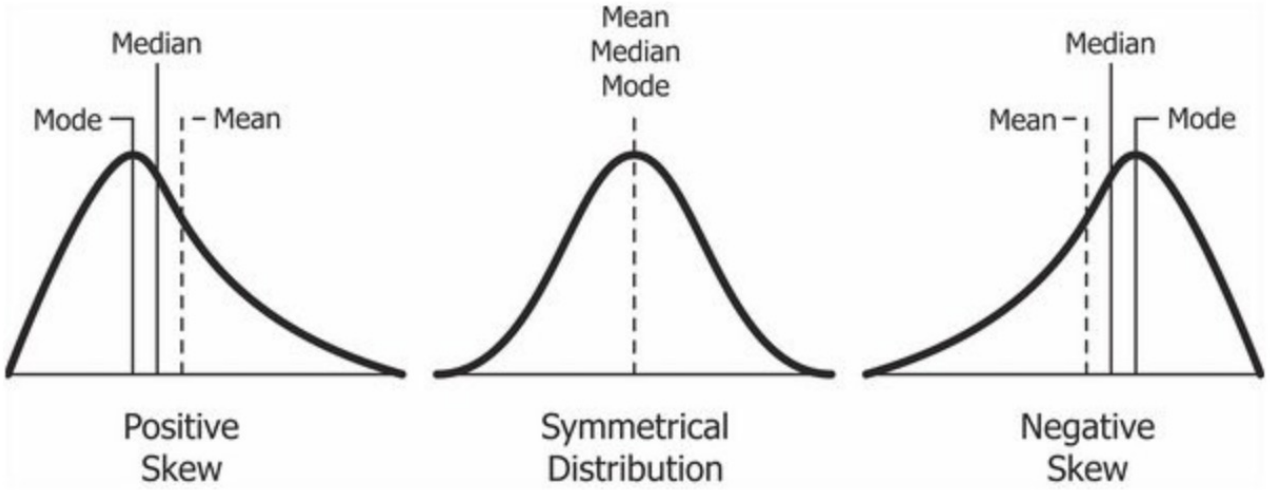


Upper truncation limit = 6.5 (bad example)



Transformation

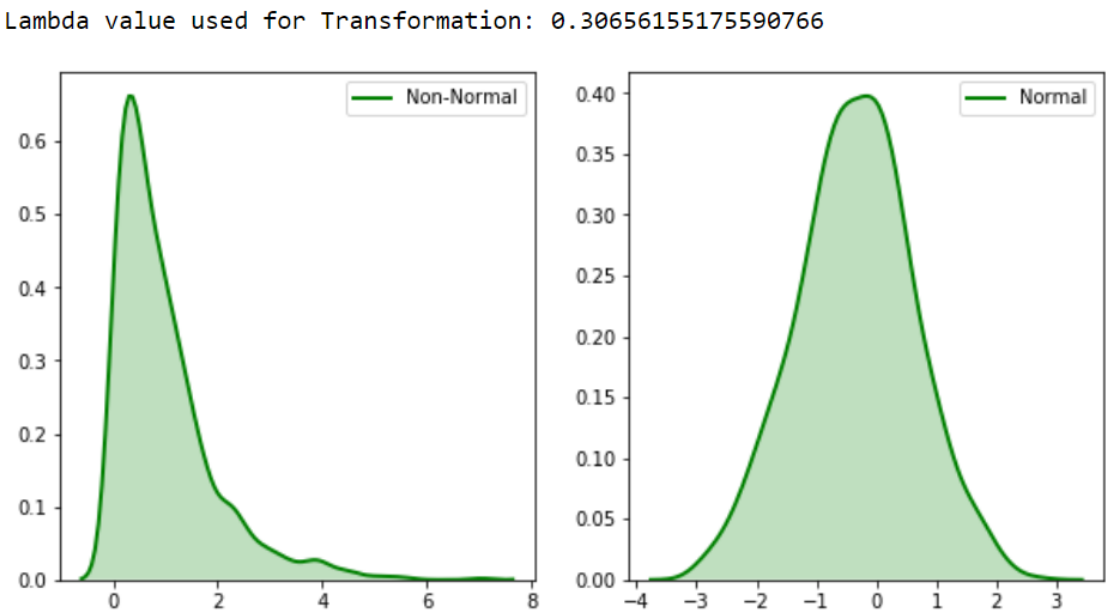
- Many clinical chemistry laboratory data are positively skewed
- Skewed distribution shifts mean value and has more extreme values
- Skewed distribution reduces the performance of PBQC
- Transformation helps better approximate a normal distribution



Box-Cox transformation

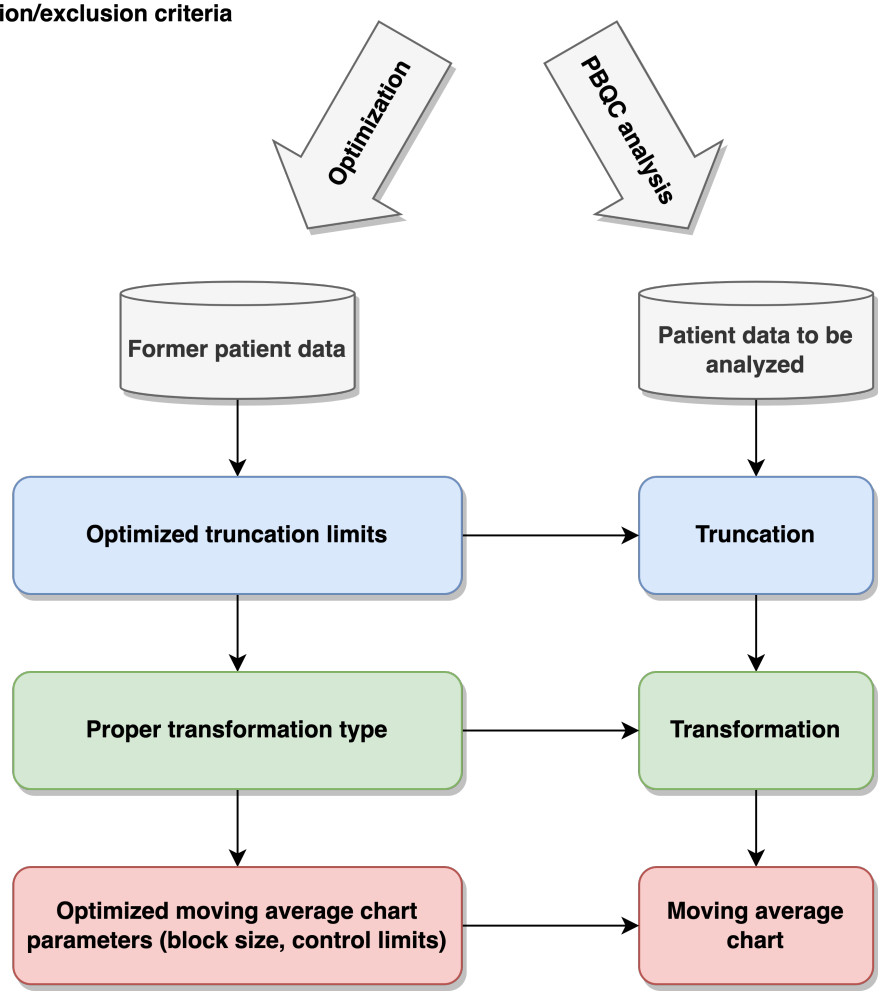
- Transforms non-normal values to approximate normal distribution
- The lambda parameter is determined by maximizing the transformed data to normal distributed data

$$y_i^{(\lambda)} = \begin{cases} \dfrac{y_i^{\lambda} - 1}{\lambda} & \text{if } \lambda \neq 0, \\ \ln(y_i) & \text{if } \lambda = 0, \end{cases}$$



PBQC summary

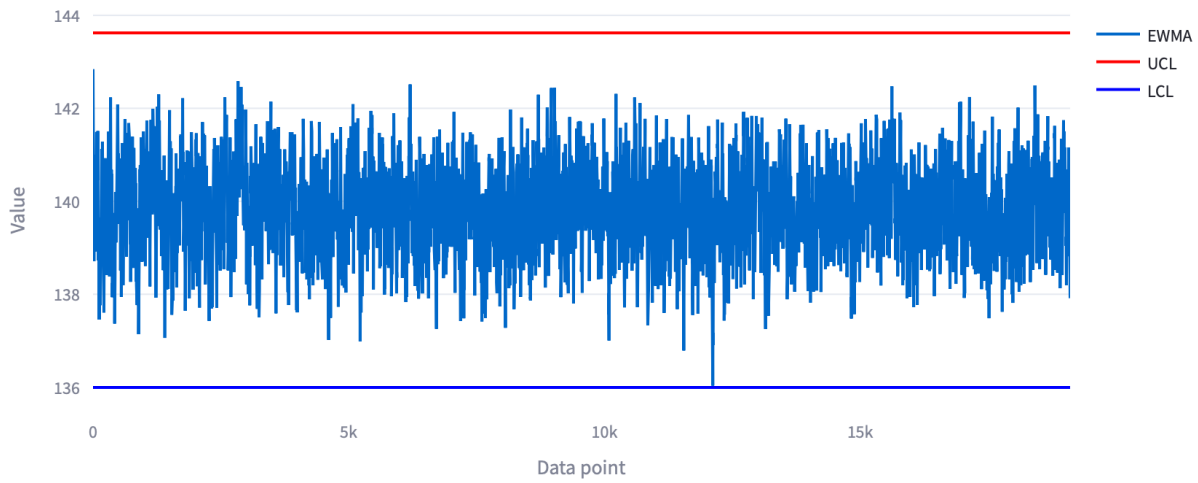
- **PBQC:** Continuous process monitoring with a commutable and cost-effective approach.*
- **Optimization:** Simulation to determine optimal parameters for moving average charts.**

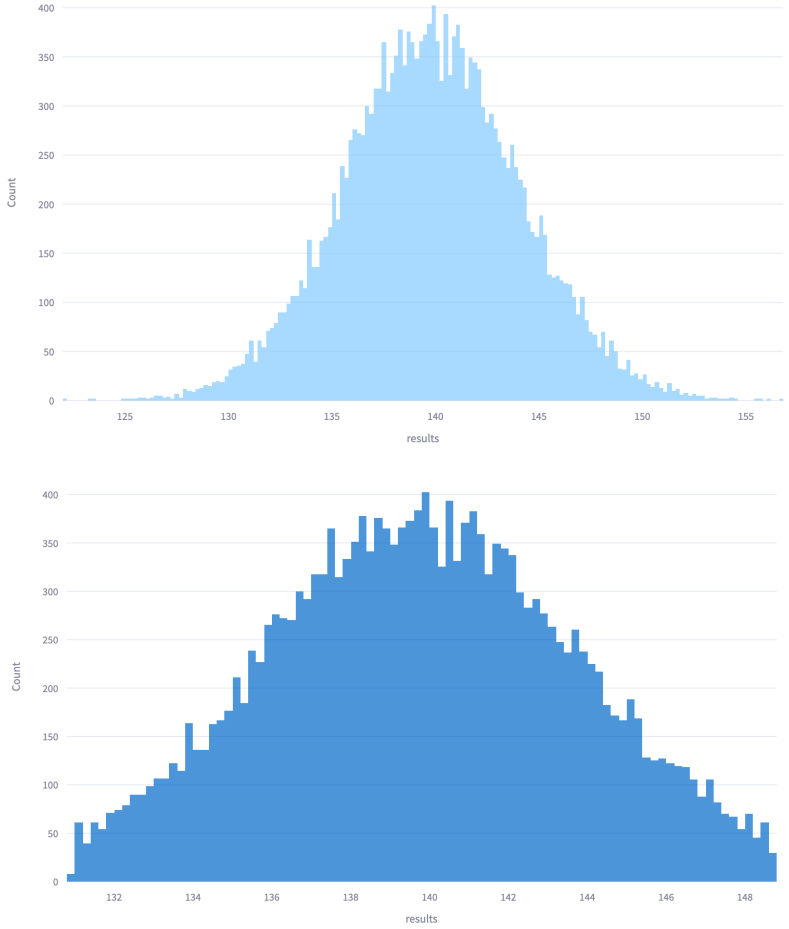


*Badrick T, et al. Clin Chem 2019;65:962–71. **Çubukçu HC. Clin Chem Lab Med 2024; doi:10.1515/cclm-2024-0156.

PBQC parameters: truncation limits and control limits

- **Truncation limits:** to reduce population noise and outliers.*
- **Control limits:** to achieve optimal sensitivity and specificity for PBQC.*

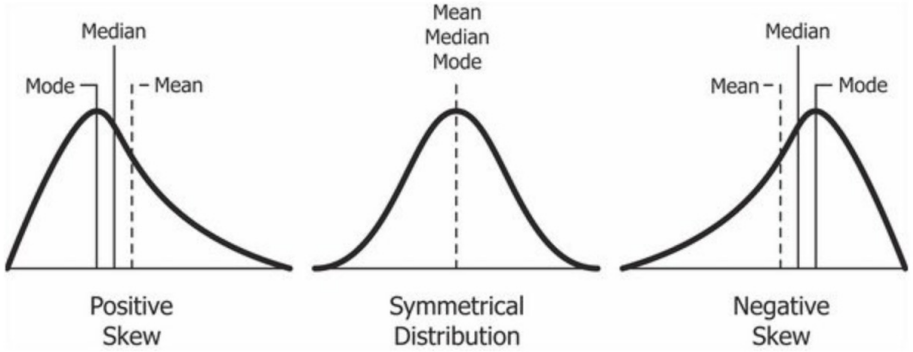




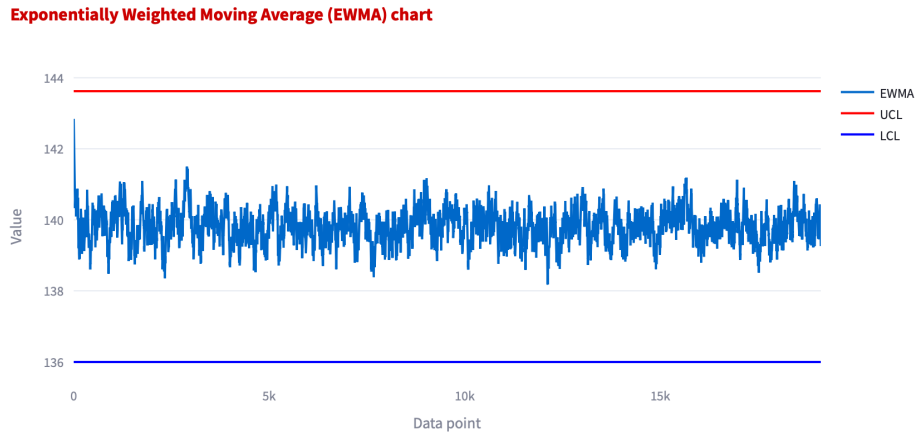
*Badrick T, et al. Clin Chem 2019;65:962–71. **Çubukçu HC. Clin Chem Lab Med 2024; doi:10.1515/cclm-2024-0156.

PBQC parameters: block size and transformation

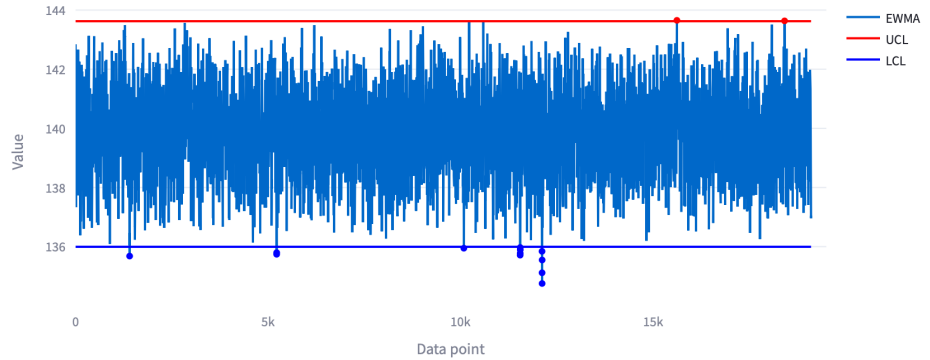
- **Block size:** lower block size faster error detection; but may increase false rejection; e.g.: for EWMA lower block size higher weighting factor, higher block size lower weighting factor.
- **Transformation:** box-cox or log-transformation, to reduce population noise.



Block size: 70



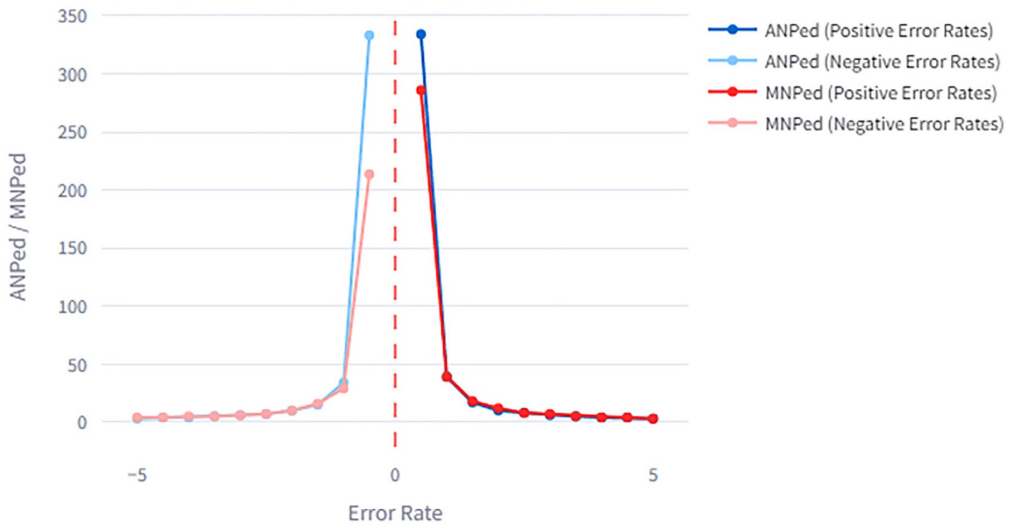
Block size: 10



*Badrick T, et al. Clin Chem 2019;65:962–71. **Çubukçu HC. Clin Chem Lab Med 2024; doi:10.1515/cclm-2024-0156.

PBQC optimization

- Optimization procedures involve varying factors such as **truncation limits, control limits, and block sizes** (for EWMA scheme etc.).*
- The simulations evaluate **different parameter combinations with different degrees of error** and **identify the set** that yields the highest performance (lowest ANPed, MNPed, and false positive rate; highest sensitivity, specificity, and Youden index).**



CUSUM scheme parameters optimized based on the highest youden index

truncation limit	transformation status	control limit (h)
131.74076778485463-148.02505373556048	Box-Cox Transformed Data	10

Performance

Metric	Value
Sensitivity (True Positive Rate)	1
Specificity	1
False Positive Rate	0
Youden Index	1
ANPed	3
MNPed	4

ANPed / MNPed: average / median number of patient samples affected from onset of the error to detection.

*Çubukçu HC. Clin Chem Lab Med 2024; doi:10.1515/cclm-2024-0156. **Topcu DI, Çubukçu HC. Clin Chim Acta 2022;534:50–6.

Optimizer for PBQC: a desktop app

Codes: https://github.com/hikmetc/aps-calculator/releases

Key notes from PBQC

PBQC

- Calculate moving statistic using selected algorithm when new patient result is produced

Common algorithms

- Moving average
- Moving median
- Exponentially weighted moving average
- Moving standard deviation
- Moving sum of positive

Model performance can be optimized by

- Block size
- Winsorisation or truncation
- Box-Cox transformation
- Selection of control limits

Take-home messages

1. **Risk-based internal quality control (IQC)** should be guided by both the **analytical performance** of the test and the **potential clinical impact of errors (severity of harm scores)**.

2. **Implementing risk-based IQC becomes practical** with a foundational understanding of its principles and freely available tools.

3. **Patient-based quality control (PBQC)** helps us to **continuously track the patient results** to **timely detect an error**.

4. **Optimization of PBQC scheme** to **achieve optimal parameters** is a necessary step before initiating PBQC.

Let’s carry our responsibility

APS Calculator

QC Constellation

Verify My Lab
Method Verfication Suite

APS Calculator

QC Constellation

Verify My Lab

QC Constellation (video)

Contact: hikmetcancubukcu@gmail.com · hikmet@live.hk

Codes and other projects: https://github.com/hikmetc

Thank you for your attention

NATFORLAB 2026