Original Paper

Investigating the Variable Component of the Systematic Error, a Neglected Error Parameter: Theoretical Reevaluation Study

Atilla Barna Vandra, MS

Spitalul Clinic Judetean de Urgenta Brasov, Brasov, Romania

Corresponding Author:

Atilla Barna Vandra, MS
Spitalul Clinic Judetean de Urgenta Brasov
Str. Berzei 2 Bl. B. ap 20
Brasov 500276
Romania
Phone: 40 722264666
Email: vandraattila@gmail.com

Related Articles:
Preprint (medRxiv): https://www.medrxiv.org/content/10.1101/2023.05.24.23290382v1
Peer-Review Report by Elvar Theodorsson (Reviewer C) : https://med.jmirx.org/2026/1/e88830
Peer-Review Report by Anonymous: https://med.jmirx.org/2026/1/e90221
Author's Response to Peer Review Reports: https://med.jmirx.org/2026/1/e88981

Abstract

Background: The existence of the variable component of the systematic error (VCSE) was known from the beginning. Still, it is a kind of taboo: it does not have a definition in the International Vocabulary of Metrology and is not present in equations, as it is considered transformed over time into random error.

Objective: This theoretical study aims to reevaluate the role and significance of the VCSE in quality control (QC).

Methods: Assuming three quintessential principles—(1) a parameter must be determined under the same conditions under which it is used, (2) a calibration cannot correct smaller biases than the calibration error, and (3) a constant cannot correct a variable—it was deduced that the source of the VCSE is bias drift caused by reagent instability and the shifts caused by human interventions. Both phenomena are mentioned in the literature. The two causes were confirmed by two series of computer simulations using 1000 normally distributed values with an SD of 1 to simulate random error and appropriately chosen bias values to simulate (1) drifts with different slopes and (2) variable shifts. Real-life examples from day-to-day QC, using Roche reagents on Cobas 6000 and Cobas PRO analyzers, confirmed the computer simulations.

Results: "The bias" is a definitional uncertainty because bias is time-variable. The causes of the cyclic variations are reagent instability and human intervention, confirmed by computer simulation and real-life QC data. Making a clear distinction between bias measured under repeatability and reproducibility within laboratory conditions, as in the case of SDs, and also separating constant and variable subcomponents of the systematic error, 2 sets of error parameters are obtained, each set being consistent with the measurement conditions. The link between them is the time-variable VCSE function. More properties of the VCSE(t) impose a distinction from random error component: predictability and corrigibility in the short term and non-Gaussian distribution. Its transformation into random phenomena is a myth based on confusion between random and variable error components. The accurate determination of the VCSE(t) function is possible, but it has an excessively high cost-effectiveness ratio. Because it is hidden in the bias measured in repeatability and in the SD in reproducibility within laboratory conditions, it helps us to avoid the redundant use in total measurement error and MU equations. Several false assumptions behind the Westgard rules were uncovered.

Conclusions: The new error model aims to serve as the foundation of a new QC system. Internal QC decisions are only consistent with graphs designed using SD measured in repeatability conditions; therefore, they are not consistent with the actual Westgard rules. Alarms should be avoided in cases of incorrigible biases. Immediately after calibration, constant biases, gradually increasing biases, and unexpected shifts in bias represent distinct situations, each requiring a unique strategy.

JMIRx Med 2026;7:e49657; doi: 10.2196/49657

Keywords: repeatability condition; reproducibility within laboratory condition, measurement; systematic error; clinical laboratory; quality control; bias; QC; statistical; statistics; mathematics; computer simulation

Introduction

The author was motivated to research and publish this study after observing several statistically impossible internal quality control (IQC) graphs designed with s_{RW} (SD measured under variable conditions, reproducibility within laboratory conditions), as recommended by Westgard et al [1]. For example, there are no $R_{1\text{-}2S}$ rule violations in a month. With 180 measurements/month (Romanian laws impose 3 control runs per day), in the case of an assumed normal distribution and a correct SD, the theoretical probability (calculated using normal distribution tables) of such a graph is 0.0224%. The author observed such (and other types) of statistically impossible graphs on all analyzers he practiced: Hitachi Modular, Cobas 6000, Cobas Pro, Cobas Pure, Architect 8000, JEOL, Siemens Advia, and BTS 370.

The former statistically impossible graphs become possible if we design the quality control (QC) graphs with an overestimated SD. For example, assuming an overestimation of 50% of the SD (practically applying instead of the $R_{1\text{-}2S}$ rule, the $R_{1\text{-}3S}$ rule as a warning), the probability of no $R_{1\text{-}2S}$ rule violations in a month becomes 62.58%. The Westgard rules are only correctly applied if we design the QC graphs with the correct SD (the measure of the pure random error component [RE]).

There is no reciprocal relationship between the normal distribution and the SD. We can calculate an SD from any data set, not just from data with a normal distribution. An SD is not proof of a normal distribution. According to Stahl [2]:

> *[The name of] Normal distribution was not the luckiest choice because other distributions are perceived as abnormal.*

Consequently, scientists perceive all distributions as not abnormal and do not verify the Gaussian character. The Gauss equation is only valid if conditions do not change. Westgard rules assume a normal distribution. However, the long-term control data are not normally distributed [3,4]. The significant variation in the monthly biases and SDs also sustains the non-Gaussian distribution (see data published by Kumar and Mohan [5]).

A significant source of error is the definition of the random measurement error in the International Vocabulary of Metrology (VIM) 2.19 [6], which considers random and unpredictable terms equivalent. According to Krystek [7]:

> *We speak of 'random' variations, although we cannot explain what the attribute 'random' actually means.*

There are different types of unpredictable phenomena. Some such examples:

a. A transient phenomenon causing an outlier.
b. An unexpected phenomenon causing a systematic change (shift).
c. A cyclical (eg, sinusoidal) variation can be subjectively perceived as random if checked in more extended time frames than its period.
d. Non-Gaussian (eg, uniformly) distributed random phenomena, like the values generated by the RAND() function in EXCEL.
e. Expected change with unpredictable extent (eg, human interventions), alternating with predictable time frames. It can be named a randomly variable systematic phenomenon.
f. Typical random phenomena caused by the inconstancy of the measuring system (eg, sampling error). Only the last phenomenon is the source of normally distributed data sets.

The confusion between the typical random and the randomly variable systematic phenomena is a severe error source in the QC. The author used the following assumptions:

- Assumption 1: The systematic error component (SE) is concentration-dependent (we perform QC measurements on more levels).
- Assumption 2: The SE is time-dependent (we repeat controls periodically).
- Assumption 3: Calibration is a measurement subject to errors (after calibration, a QC run is compulsory).
- Assumption 4: The instrument is quasi-constant in time. Maintenance does not impose corrective actions (eg, recalibrations), only QC.
- Assumption 5: An instrument failure cannot cause specific systematic variations, and the errors are of aberrant size (eg, a blown lamp).

This study is consistent with the following quintessential principles valid in all sciences:

- Quintessential principle 1: We must determine all parameters under the same conditions under which we use them. For example, if we determine a parameter under specific constant conditions, we cannot use it for predictions in variable conditions. We can extend the use of a parameter obtained within a given time frame to other time frames only if we assume that it is constant.
- Quintessential principle 2: An action (eg, calibration) can efficiently correct neither smaller biases than its average error nor smaller biases than the uncertainty of the bias value.
- Quintessential principle 3: We cannot correct a variable error by adding a constant.

The SE (bias) is dependent on concentration and time (SE≈B(c, t), Assumptions 1 and 2). To apply correctly, we must modify the error model. Westgard et al [8] separated the bias into a constant component (CE) and another proportional to concentration (PE), making it possible to

deal with concentration dependency. If we focus on a single control level, the separation is unnecessary. The corrected error model has a wide range of applicability [9,10].

A similar, generally accepted separation of bias components to deal with time dependency does not exist. Westgard et al [8] started from the assumption of a constant bias. As Badrick [3] observed:

> *[In the Westgard model] One assumption is that the bias is unchanged over time; 'Systematic' implies a specific point in time.*

However, JCGM –6:2020 GUM 10.6 has recommendations in the case of drift effects [11], the JCGM 100:2008 GUM 3.2.4 [12] recommendation "It is assumed that the result of a measurement has been corrected for all recognized significant systematic effects" hides a similar assumption. Neither a correction (GUM B.2.23) nor a correction factor (GUM B.2.24) can eliminate a function (a time-variable bias, quintessential principle 3). The bias is undoubtedly time-variable (Assumption 2). According to Leito [13]:

> *Bias determined within a single day is different from one determined on different days (and averaged).*

If so, the bias measured in external quality assessment (EQA) has a validity term of only 24 hours. When we obtain the result, the value is obsolete. The variable bias is neither eliminated by corrections nor by calibration because it reappears (quintessential principle 3).

When substituting the bias value into an equation, the question arises: Which bias? The bias of today, the value measured in the last EQA, or the long-term mean of the bias values? "The bias" is a definitional uncertainty that imposes a distinction between bias types and their separation into a time-invariable component (CCSE: constant component of systematic error) and a time-variable function (variable component of the systematic error=VCSE[t]). Focusing on a single control level:

$$TE(t) = SE(t) + RE = CCSE + VCSE(t) + RE \quad (1)$$

This study aims to identify, quantify, and characterize these bias components, if possible.

The VIM 2.17 definition [14] by the "or" word indirectly defines 2 SE subcomponents:

> *The systematic measurement error is the component of the measurement error that in replicate measurements remains constant or varies in a predictable manner.*

The CCSE and the VCSE(t) are neither defined nor at least mentioned in VIM. Time variability was known from the beginning [15]. However, the phenomenon has only come into focus in recent years. Due to the lack of standardization, the authors use different names, definitions, and notations [15-23], which cause difficulties in research. The definitions are not (entirely) equivalent. Others only make the difference between short-term bias and long-term bias [9,24] or bias of the moment "t" and mean bias [25], suggesting bias variability.

Several authors built alternative error models to include the VCSE(t) function [15,19,23,25]. A particular case is the graphical model of Theodorsson et al [21], which attempts to prove that: "Variable bias components become random errors over time."

In their model, the variable bias components are included in the SE for short time frames, while in long time frames, they are included in the RE. However, the model is consistent with the VIM 2.17 definition of the SE; its accuracy is debatable because the definition does not distinguish between randomly variable systematic and typical random phenomena (cases e and f of unpredictable).

The transformation of the variable SE components into random ones is only subjective, based on an inaccurate definition. Only the long-term control data are dispersed under the influence of 2 distinct variable phenomena: the RE and bias variation (the VCSE[t]). We can calculate an SD from the VCSE(t) values, as from any variable set of data (cases b-d of unpredictable). Let us note its s_{VCSE} (the SD calculable from the daily [run] mean, bias, or VCSE[t] values). According to more authors (using different names, definitions, and notations), the link between the SD measured in repeatability and reproducibility within laboratory conditions is the s_{VCSE} [19,22,23].

$$S_{RW} = S_r + S_{VCSE} = \sqrt{S_r^2 + S_{VCSE}^2} \quad (2)$$

The VCSE(t) is hidden in the bias of the moment "t" and s_{RW}.

Initially, the bias variations were perceived as unpredictable. Shewhart [15] stated:

> *The causes of this variability are, in general, unknown.*

Similar opinions have been sustained by Westgard et al [8]. Recent studies identified 2 sources of bias variability. According to Marquise [16]:

> *Every new calibration creates a different bias, which appears as a random shift on the chart.*

Magnusson et al [22] referred to the phenomenon as variations in calibration over time. The consequence is an alternation between periods of constant bias with random variations in the SE.

The reagent instability causes a gradually increasing bias (in absolute values) [18,19]. The bias cannot continue to increase indefinitely because we take corrective actions. Consequently, we obtain a sawtooth-like cyclical bias variation. Mackay et al [23] acknowledge both phenomena as sources of bias and variation.

Using computer simulations and real-life QC examples, the author will analyze these phenomena in the Experimental Data section. In the Discussion section, the properties of the VCSE(t) function and the s_{VCSE} will be compared with other bias and SD components.

There are 2 points of view in the clinical laboratory. The accreditation services and clinicians are interested in the limit of credibility of the results: the measurement uncertainty. This point of view is consistent with error parameters measured in reproducibility within laboratory conditions (quintessential principle 1). Unfortunately, this point of view is imposed on all decisions, becoming a source of error.

The laboratory specialist focuses on short-term decisions: May I run patient samples now, or must I make corrective actions before? The decisions are consistent with error parameters measured in repeatability conditions, but not those obtained in long time frames (quintessential principle 1).

There are 2 conflicting approaches in the QC. Gauss [26] introduced the error approach, which was considered valid until the emergence of the measurement uncertainty (MU) approach described by GUM [7]. Usually, there is an expectation to adhere to one of these approaches.

While the theoreticians of the uncertainty of measurement (UM) formulated some pertinent critiques, the UM theory is not perfect. The comparison of the weaknesses and strengths of the error and UM approaches is not the task of this study. Neither the UM approach can challenge the total measurement error (TE) approach-based internal QC decisions, nor can the TE approach substitute the UM in uncertainty calculations [23]. The 2 approaches link to 2 different points of view, and predictably, they will coexist as a state-of-the-art situation. The laboratory specialists must use both, depending on their tasks. Moreover, the 2 approaches share commonalities, using the same (oversimplified) error model. This study challenges the error model, influencing both approaches. The focus of this study is on short-term, internal QC decisions. Therefore, the consequences on UM calculations will only be mentioned.

Methods

This theoretical study uses mathematical statistics. Most statements and observations are present in the literature, but only as mosaic pieces. Critical statements are based on theoretical deductions, computer simulations, and observations made in the author's 40 years of experience in the clinical laboratory. Real-life examples are from the day-to-day IQC of the laboratory of the Brasov County Clinical Hospital for Urgencies (SCJUBv). The author made the exemplified measurements on Cobas 6000 and Cobas Pro analyzers using Roche reagents, but observed similar phenomena on all analyzers he worked with.

A total of 1000 data (expressed with one decimal) with normal distribution, mean 0 (SD 1), were generated to simulate RE. The bias variation was simulated by choosing bias values depending on the task. TE was calculated as the sum of the bias and RE. From the daily RE, B, and TE values, respectively, the s_r (SD measured in constant, repeatability conditions), s_{VCSE}, and s_{RW} were calculated.

To simulate the influence of a single calibration error on the SDs, the bias was maintained at 0 in the first 500 data, and the same chosen value simulating a bias was used for the last 500 in each simulation. Changing the bias from 0 to 2 (0-2 s_r) with increments of 0.25 (0.25s_r), 9 data sets of s_r, s_{VCSE}, and s_{RW} were obtained. The s^2_{RW} was represented in the function of s^2_{VCSE} (Table 1).

Table 1. Computer simulation of a single calibration. In each simulation, "n" takes integer values between 0 and 8 (a total of 9 values). re_i values have a normal distribution with SD=s_r=1.004.

Time (t)	RE[a]	Bias	TE[b]
1	re_1=2.1	0	2.1
2	re_2=–1	0	–1
500	re_{500}=0.1	0	0.1
501	re_{501}=1.7	n × 0.25	1.7 + 0.25 n
502	re_{502}=–0.9	n × 0.25	–0.9 + 0.25 n
1000	re_{1000}=–1.2	n × 0.25	–1.2 + 0.25 n
SD	s_r[c]=1.004	s_{VCSE}[d] = n × 0.125	s_{RW}[e]

[a]RE: random error component.
[b]TE: total measurement error.
[c]s_r: SD measured in constant, repeatability conditions.
[d]s_{VCSE}: the SD calculable from the daily (run) mean, bias, or VCSE(t) values.
[e]s_{RW}: SD measured in variable, reproducibility within laboratory conditions.

To simulate the influence of more calibration errors on the SDs, (3 random changes in the mean) were added 4 × 10 bias values (equal to 1.5, –1, –0.5, and 0) to 2 × 40 normally distributed values (real SD of 1.07), simulating RE on 2 levels. s_{VCSE} from the bias values, s_r from the RE values, and s_{RW} from the TE values were calculated in different time frames.

One thousand and one linearly decreasing bias values were chosen (from 0 to B) to simulate the influence of drift in bias. By changing the slope factor (by changing the value of Bias from 0 to 4 with increments of 0.5), 9 data sets of s_r,

s_{VCSE}, and s_{RW} were obtained. The s^2_{RW} was represented in the function of s^2_{VCSE} (Table 2).

Table 2. Computer simulation of a quasilinear drift caused by reagent degradation. "b"=B/1000. In each simulation, B takes values from 0 to 4 with increments of 0.5 (total 9 values/simulations). re_i values have a normal distribution with SD=s_r≈1.

Time (t)	RE[a]	Bias	TE[b]
0	re_0=0.6	b × 0	0.6 + 0
1	re_1=–2.1	b × 1	–2.1 + b
500	re_{500}	b × 500	re_{500} + b × 500
999	re_{999}=–0.8	b × 999	–0.8 + b × 999
1000	re_{1000}=–1.2	b × 1000	–1.2 + b × 1000
SD	s_r[c]=1.004	s_{VCSE}[d]	s_{RW}[e]

[a]RE: random error component.
[b]TE: total measurement error.
[c]s_r: SD measured in constant, repeatability conditions.
[d]s_{VCSE}: the SD calculable from the daily (run) mean, bias, or VCSE(t) values.
[e]s_{RW}: SD measured in variable, reproducibility within laboratory conditions.

In the real-life data example with drift, the run mean was estimated with the SLOPE and INTERCEPT functions in Excel. A single estimated mean was calculated from the average of the run results expressed as a percentage. The CV_r (CV measured in constant, repeatability conditions) values for each level were calculated from the deviations from the estimated run mean.

The average CV_r for the whole period was calculated as the SD of the half differences of the percent expressed results (an adaptation of a method described in Nordtest 537 TR [22]).

Results

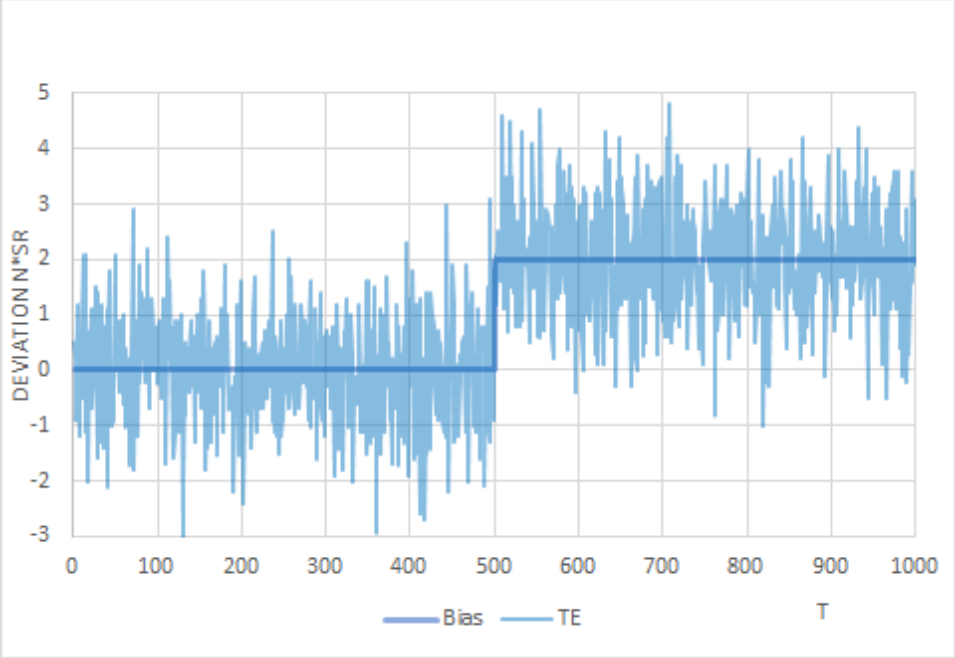
Overview

The computer simulations aimed to demonstrate that the sources of bias variation described in the literature are the true causes of the increased SD in more extended time frames and to confirm the validity of Equation 2. The real-life QC examples demonstrate that computer simulations are grounded in reality.

The Influence of a Single Shift in the Mean Caused by a Calibration

In the computer simulation of a single calibration (a single shift in the mean), the graph of the run mean is a horizontal line with bias=0 before the mean shift (calibration) and a horizontal line with mean=bias after the mean shift (calibration). The results are randomly dispersed around the run mean with SD of 1 (=s_r) (Figure 1). The SDs calculated from 500 data before and from 500 data after the shift are 1 (=s_r), while the SD calculated from all data (s_{RW}=1.43) is significantly bigger according ($F_{0.95,\ 500,500}$=1.43). The SD calculated from runs 480-520 (including the shift) is 1.55, suggesting that the bias variation causes the increase of the SD (s_{RW}). A sudden change of 1 SD (1s_r) in the mean causes an increase of only 12% in the overall SD (s_{RW}), and it is difficult to observe visually such minimal increases.

Figure 1. Computer simulation: a shift in the mean causes an increase in the s_{RW} (SD measured in variable, reproducibility within laboratory conditions). Bias variation=$2s_r$ case. TE: total measurement error.



Representing the s^2_{RW} values as a function of s^2_{VCSE}, a linear graph with slope ≈1, consistent with Equation 2, was obtained, confirming its validity (Figure 2).

An example of magnesium obtained in March 2021 on a Cobas Pro analyzer (2 × 7 runs, 2 levels, and one calibration after 7 runs) is presented in Figure 3 to exemplify real-life data. The results were represented in %, not as absolute values, to reduce the influence of the s_{RW} variability.

Figure 2. Variation of square s_{RW} as a function of square s_{VCSE}. The slope is 1. s_{RW}: SD measured in variable, reproducibility within laboratory conditions; s_{VCSE}: SD calculable from the daily (run) mean, bias, or VCSE(t) values.

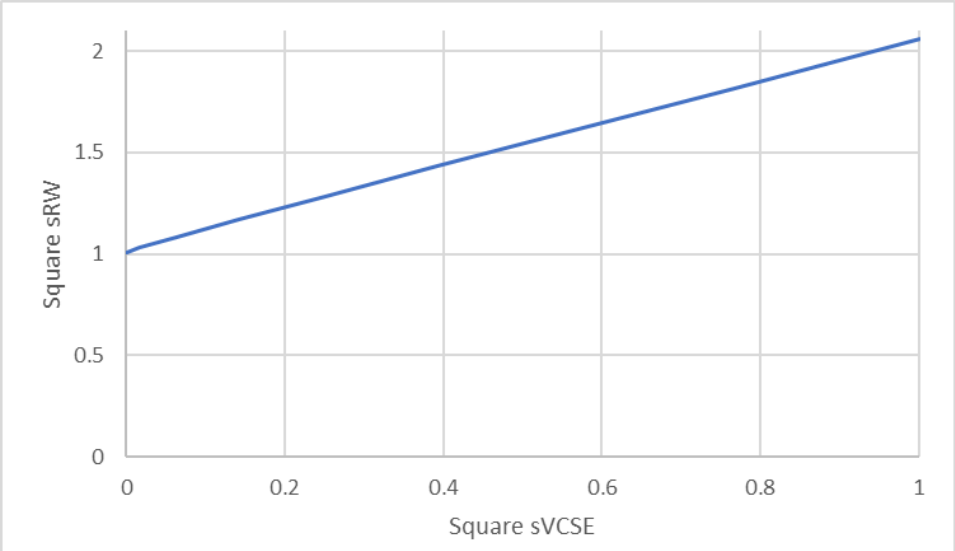
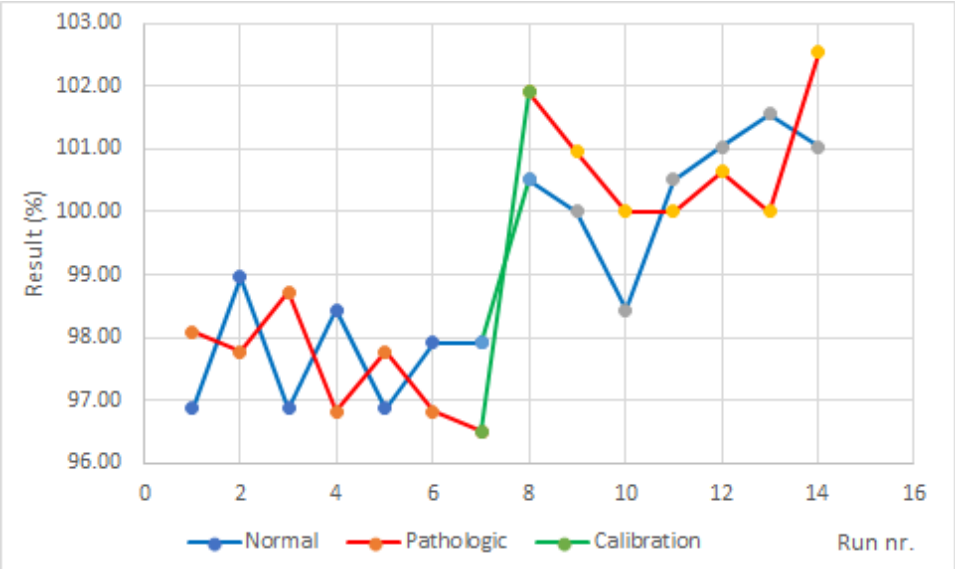


Figure 3. Calibration parameter changes cause bias variations (VCSE(t)). Real-life data. VSCE(t): variable component of the systematic error, a time-variable function.



The graph has an insignificant drift on both levels. Calculations presented in Table 3 show that before and after calibration, the coefficient of variation (CV) is consistent with the CV_r (an *F* test did not reveal significant differences), and the increase in the CV_{RW} (CV measured in variable, reproducibility within laboratory conditions) is due to the shift in the mean. Equation 2 is valid. From the s_{VCSE} calculated from the mean variation and the s_r values, it was possible to predict the value of the CV_{RW}. The *F* test did not find significant differences between the CV of all data, the predicted CV (Equation 2), and the actual CV_{RW}. The actual CV_{RW} (determined from one month's data) is slightly bigger because it includes more calibrations and reagent changes.

Table 3. The increase of the s_{RW}/CV_{RW} (SD measured in variable, reproducibility within laboratory conditions/ coefficient of variation measured in variable, reproducibility within laboratory conditions) caused by a shift in the mean (calibration) can be predicted by Equation 2 (real-life data, magnesium, Cobas PRO).

Analyte and data	Number of data	CV (CV_r)[a], %	CV_r (method validation), %	CV_{VCSE}[b] $= \frac{\Delta B\%}{2}$, %	CV of all data (CV_{RW}), %	Predicted CV_{RW} (Equation 2), %	Actual CV_{RW}, %
Mg level 1							
Before calibration	7	0.86	—[c]	—	—	—	—
After calibration	7	1.01	—	—	—	—	—
All data	14	0.94	1.24	1.39	1.69	1.95	2.14
Mg level 2							
Before calibration	7	0.83	—	—	—	—	—
After calibration	7	1.01	—	—	—	—	—
All data	14	0.92	1.11%	1.69	1.95	2.18	2.54

[a]CV_r: CV measured in constant, repeatability conditions.
[b]CV_{VCSE}: CV of the VCSE(t), s_{VCSE}, expressed as a percent of the target value.
[c]Not applicable.

The Influence of More Random Changes in the Mean (More Calibrations)

Figure 4 shows the simulation graph of more random changes in the mean. Without computer assistance, we can visually detect only the significant mean variation (run 11). As shown in Table 4, the s_r values are quasi-constant. Simultaneously, the s_{RW} values depend on the time frame (variations from 1.10 to 1.94). The bigger the mean change, the bigger the s_{RW}. The validity of Equation 2 is maintained (compare line 4 with line 10).

Figure 4. The influence of multiple mean changes (computer simulation); only significant shifts can be visually observed (run 10-11), and not those that are less significant. s_r: SD measured in constant, repeatability conditions.

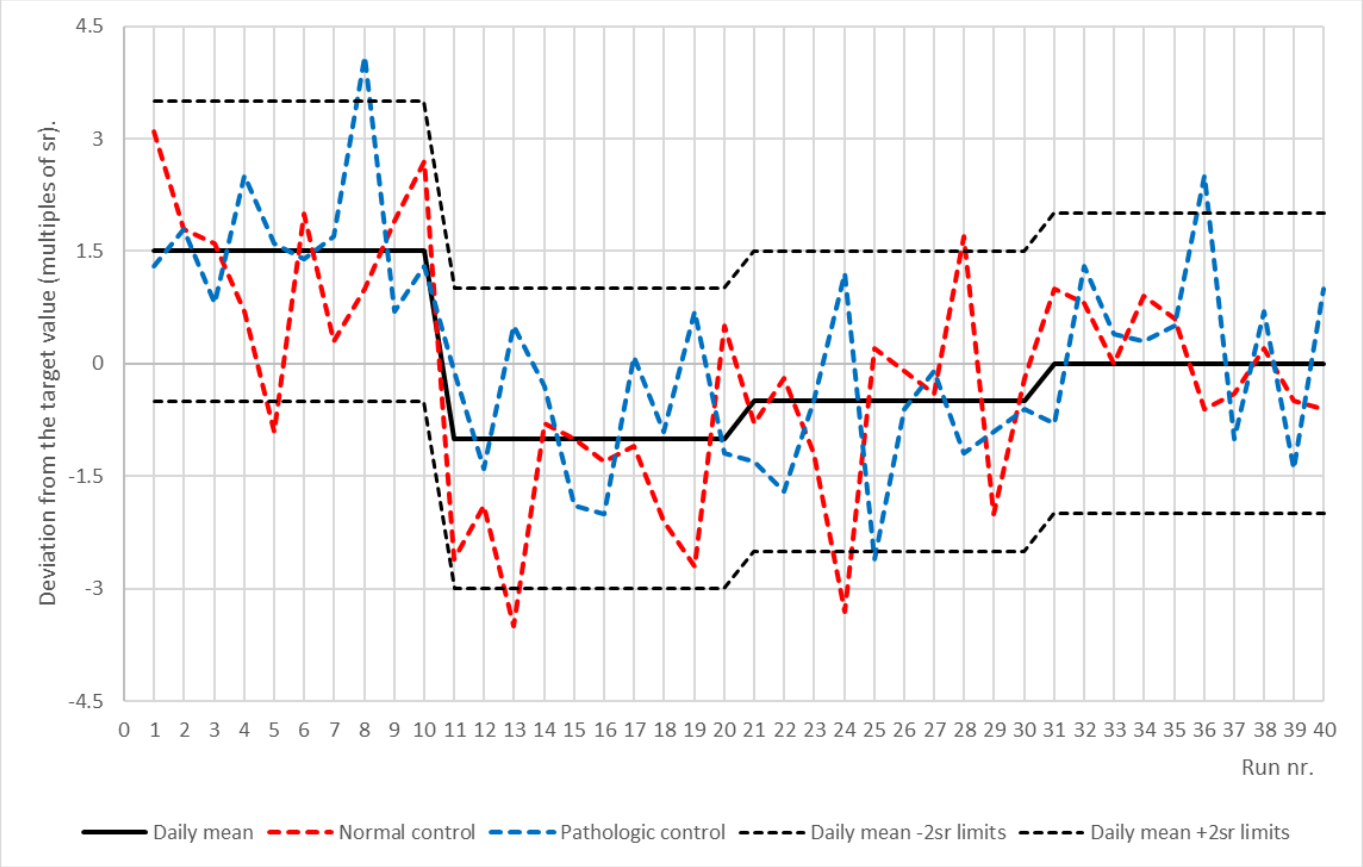


Table 4. There are significant differences in the s_{RW} (SD measured in variable, reproducibility within laboratory conditions) values, depending on the time frame, while s_r (SD measured in constant, repeatability conditions) in the limits of the statistical methods remains constant.

Variable and runs (time frame)	Normal	Pathologic
s_{RW}		
1-20	1.94	1.54
21-40	1.10	1.22
11-30	1.32	0.97
s_{RW} all		
1-40	1.56	1.43
s_r		
1-40	1.10	1.03
1-20	1.17	0.95
11-30	1.15	1.02
21-40	1.03	1.12
S_{VCSE}[a] (SD of bias variation)		
1-40	0.95	0.95
s_{RW} calculated/predicted (Equation 2)		
1-40	1.45	1.40

[a]s_{VCSE}: SD calculable from the daily (run) mean, bias, or VCSE(t) values.

The Influence of Gradual Mean Changes (Drifts) Caused by Reagent Degradation

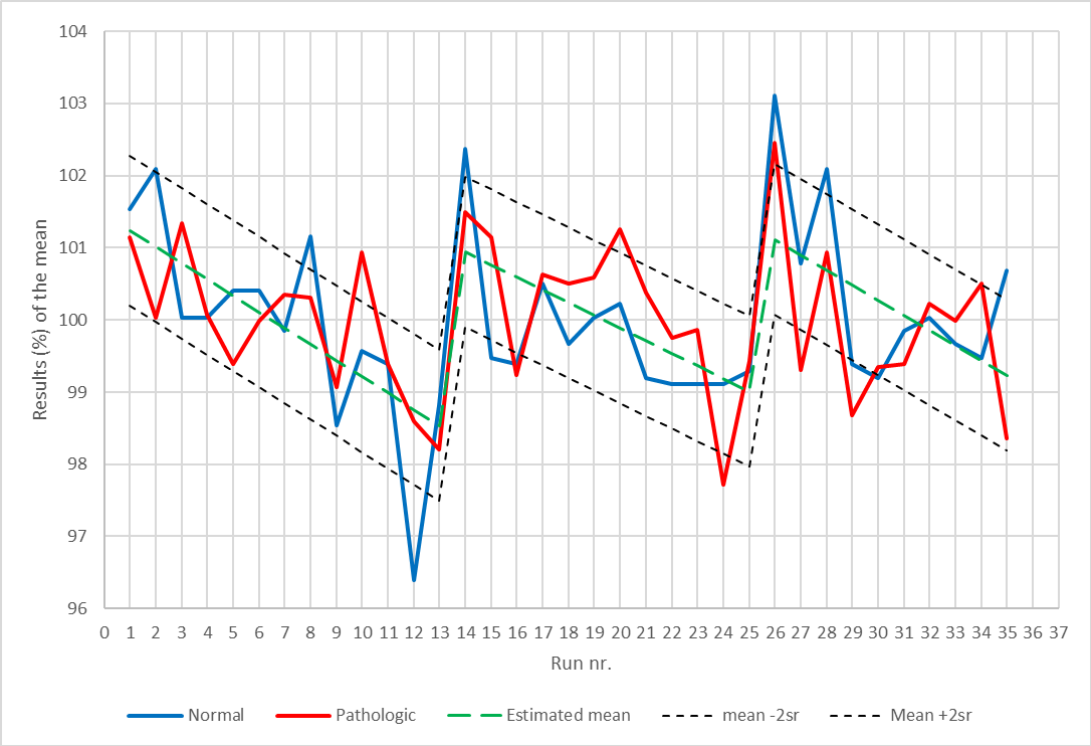
In the computer simulation, the graph of the daily mean was an oblique line with decreasing tendency, with slope=–0.001 × $_{max}$Bias. $_{max}$Bias is the maximum bias in absolute values in each simulation. The SD calculated from the daily means was $s_{VCSE}=\frac{maxBias}{2*\sqrt{3}}$, corresponding to a uniform distribution. The deviation of the results from the daily means had an SD ≈1 (=1s_r) in all simulations. The SD calculated from all 1001 data (s_{RW}) was bigger than 1s_r. A bias variation of 1.5s_r

caused an increase in s_{RW} of only 10%, which was difficult to observe visually.

If we represent the s^2_{RW} values as a function of s^2_{VCSE}, we obtain an identical graph, as shown in Figure 2, consistent with Equation 2 (a linear graph with slope ≈1 and intercept ≈1).

Figure 5 shows a 35-run real-life chart (glucose, Cobas 6000 analyzer, July 2023). The period includes 2 reagent changes (corresponding to the shifts in the mean between runs 14-15 and 25-26). No calibrations were made. In the periods between reagent changes, the means are similar in all time frames (0.22%/run, 0.23%/run, and 0.20%/run), consistent with the degradation tendency of the reagent. Most data are within the estimated mean (SD 2CV_r) limits, suggesting that CV_r (s_r) is the true measure of the RE.

Figure 5. Real-life data sustain the influence of the mean drift on the variable component of the systematic error. s_r: SD measured in constant, repeatability conditions.



The CV_r values calculated from the deviations of the percent expressed results from the estimated mean are similar to the CV_r value calculated from the half differences between the percent expressed results obtained on the 2 control levels (Table 5; a Cochran *F* test for equality of 2 variances did not find significant differences between the CV_r values).

Table 5. The coefficient of variations (CVs) calculated from the deviations from the estimated means are similar to CV_r (CV measured in constant, repeatability conditions; in the limits of the statistical methods; CV_r [half difference, all runs]=0.73%). The CV_{RW} (CV measured in variable, reproducibility within laboratory conditions) is significantly bigger.

	Runs (%)				CV_r (method validation) (%)	CV_{RW} (%)
	1-13	14-25	26-35	All runs		
Normal	0.96	0.72	1.04	0.90	0.81	1.24
Pathologic	0.73	0.80	1.07	0.86	0.80	1.10

A control material handling error (reused control material) in run 26 (false simultaneous increase) caused the slightly bigger s_r in runs 26-35.

Another example with total bilirubin was published by Vandra [27] in a preprint paper.

Discussion

Principal Findings

"The bias" is a definitional uncertainty. The same distinction is necessary between the biases obtained in repeatability and respective reproducibility within laboratory conditions, as in the case of SDs. The need for standardization imposes similar notations. We must highlight the time-variable

function character of the bias as well. The author proposes the following notations:

- $B_r(t)$=Bias measured in repeatability conditions, at the moment t.
- $\overline{B}_{RW}$=Mean bias measured in reproducibility within laboratory conditions. It is the mean of the $B_r(t)$ values in a given time frame. An accent highlights the fact that it is a mean.

We can obtain only a mean bias value in more extended time frames.

A Corrected Error Model

The difference between $B_r(t)$ and $\overline{B}_{RW}$ is VCSE(t), a time-variable function.

$$VCSE(t) = B_r(t) - \overline{B}_{RW} = (B_r(t) - CCSE) \quad (3)$$

Variations in the mean caused by reagent property changes cause drifts. The VCSE(t) cannot increase indefinitely (in absolute values) due to human interventions. It may have only cyclical variations. The cycles depend on external factors (eg, the rhythm of reagent use, frequency of human interventions, and the size of random calibration errors). They have different amplitudes, means, and lengths.

In some cases, a cycle may last even a month. The graphs of the daily means (not of the results) have sawtooth shapes masked by the noise of the RE (easily observed in the case of the unstable reagents, eg, Figure 5). In short or medium time frames, the $\overline{B}_{RW}$ values may have variations. The longer the time frame, the less uncertainty there is for the $\overline{B}_{RW}$ values. Only yearly $\overline{B}_{RW}$ values can be considered quasi-constant [21] and used for accurate corrections. In a chosen time frame, we can identify $\overline{B}_{RW}$ with the CCSE. Consequently, the mean of the VCSE(t) is 0. If we calculate the long-term mean of the $B_r(t)$ values:

$$\frac{\sum_{t=1}^{n} B_r(t)}{n} = \frac{\sum_{t=1}^{n}(\overline{B}_{RW} + VCSE(t))}{n} = \frac{n * \overline{B}_{RW}}{n} + \frac{\sum_{t=1}^{n} VCSE(t)}{n} = \overline{B}_{RW} = CCSE = \overline{TE}_{RW} \quad (4)$$

We obtain the same value for the long-term mean of TE ($\overline{TE}_{RW}$)because $\sum_{t=1}^{n} RE(t)$=0. Similarly, the SD can be calculated from long-term data (s_{RW}):

$$s_{RW} \approx \sqrt{\frac{\sum_{t=1}^{n}(TE(t) - \overline{B}_{RW})^2}{n-1}} = \sqrt{\frac{\sum_{t=1}^{n}(R_E(t) + VCSE(t) + \overline{B}_{RW} - \overline{B}_{RW})^2}{n-1}} = \sqrt{\frac{\sum_{t=1}^{n}(R_E(t))^2}{n-1} + \frac{\sum_{t=1}^{n}(VCSE(t))^2}{n-1}} = \sqrt{s_r^2 + s_{VCSE}^2} \quad (5)$$

Which confirms the validity of Equation 2 (because the long-term mean of RE and VCSE(t) is 0, $\sum_{t=1}^{n}(RE(t) * VCSE(t)) \approx 0$). Regrouping the terms in Equation 5 can be calculated using s_{VCSE}.

Regrouping Equation 3 and adding RE to both parts of the equation yields:

$$TE(t) = CCSE + VCSE(t) + RE(t) = \overline{B}_{RW} + VCSE(t) + RE(t) \quad (6)$$

Equations 3, 5, and 6 define a new error model, which is presented in Figure 6.

Figure 6. A new error model, taking into account the time variability of the bias. $B_r(t)$: bias measured in repeatability conditions at the moment t (a time-variable function); B_{RW}: long-term mean bias, measured in RW conditions, a constant; CCSE: constant component of systematic error; SE: systematic error component; s_r: SD measured in constant, repeatability conditions; s_{RW}: SD measured in variable, reproducibility within laboratory conditions; s_{VCSE}: SD calculable from the daily (run) mean, bias, or VCSE(t) values; TE: total measurement error; VCSE: variable component of the systematic error.

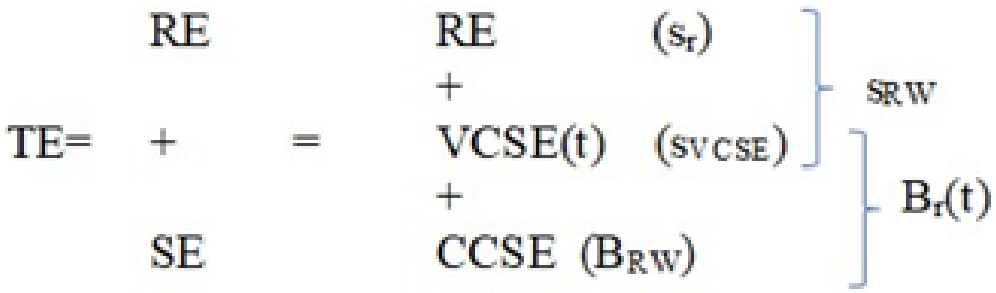


Figure 6 shows that both s_{RW} and $B_r(t)$ include VCSE(t) in a hidden form.

Two Points of View, Two Sets of Error Parameters

We obtain 2 sets of error parameters by separating the bias into a constant and a variable component and distinguishing bias measured in repeatability and reproducibility within laboratory conditions. According to quintessential principle 1, UM calculations must be based on parameters determined in reproducibility within laboratory conditions (s_{RW}, $\overline{B}_{RW}$). In the meantime, the internal QC decisions must be based on parameters determined in repeatability conditions (s_r, $B_r(t)$). The second conclusion contradicts the recommendations of

Westgard et al [1] to design Levey-Jennings charts with an SD calculated from long-term control data (s_{RW}).

Proposed Definitions of CCSE and VCSE(t)

Consistent with the VIM 2.17 definitions [14], we can define the bias components as:

> *The constant component of SE (CCSE) is the component of measurement error that in replicate measurements remains constant.*

Note 1: The CCSE is the long-term mean bias $\overline{B}_{RW}$, depending on the time frame.

> *The variable component of SE (VCSE(t)) is the component of measurement error that in replicate measurements varies predictably.*

Note 2: VCSE(t) is a time-variable function.

Note 3: VCSE(t) is hidden in $B_r(t)$ and s_{RW}.

The simultaneous use of $B_r(t)$ and s_{RW} causes a redundant use of VCSE(t) in equations—for example, ${}_{max}TE=B_{EQA} + z \times s_{RW}$. B_{EQA} is the bias measured in the last EQA round in repeatability conditions, and ${}_{max}TE$ is the TE limit, which includes all TE values with confidence corresponding to z, the confidence factor.

If bias is variable, TE is also variable (contradicting the graphical model of Theodorsson et al [21]). A distinction is necessary between:

a. TE of a given measurement ($TE(t)=B_r(t)+RE$). It has no practical value.
b. The maximum TE value at the moment *t* measured under repeatability conditions with a chosen confidence level.

$$\max TE(t) = B_r(t) + z * s_r \quad (7)$$

Internal QC decisions must be based on ${}_{max}TE(t)$, the maximum value of the TE at the moment t of decisions with a chosen confidence level, where z is the confidence factor.

c. The maximum TE value in long time frames is measured in reproducibility within laboratory conditions with a chosen confidence.

$$\max TE_{RW} = \overline{B}_{RW} + z * s_{RW} \quad (8)$$

Where ${}_{max}TE_{RW}$ is the maximum TE value in long time frames, with a chosen confidence. It must be used when setting limits and is a starting point for uncertainty of measurement (UM) calculations.

TE is also an ambiguous term. It is necessary to specify which TE is mentioned.

TE was the dominant paradigm until the emergence of UM after the publication of GUM in 1993 [11,12,28].

UM mathematically expresses our lack of knowledge about the accuracy of the result. According to VIM 2.26 [14]:

> *Uncertainty of measurement is a non-negative parameter characterizing the dispersion of the quantity values being attributed to a measurand, based on the information used.*

The definition is also mentioned in ISO 15189. According to ISO 15189, 5.6.2 [29]:

> *Sources that contribute to uncertainty may include sampling, sample preparation, sample portion selection, calibrators, reference materials, input quantities, equipment used, environmental conditions, condition of the sample and changes of operator.*

Surprisingly, neither the calibration error nor the reagent instability is mentioned among the uncertainty sources. According to the Hong Kong Association of Medical Laboratories, "the IQC procedure is designed to detect variations in reagents or calibrators" [30].

According to Magnusson and Ellison [28]:

> *The principles laid down by GUM are recognized to apply to all types of quantitative measurements, in all fields of application, and are widely accepted.*

A prerequisite for the application of the GUM [11] is that:

> *The result of a measurement has been corrected for all recognized significant systematic effects.* [GUM 3.2.4]

Using either a correction (GUM B.2.23) or a correction factor (GUM B.2.24) [11]. Then, the uncertainty of the correction is included in the uncertainty budget. Unfortunately, according to Magnusson and Ellison [28]:

> *…instances in which bias is known or suspected, but in which a specific correction cannot be justified, are comparatively common. The ISO Guide to the Expression of Uncertainty in Measurement does not provide well for this situation.*

The uncorrected bias must be included in the uncertainty budget, and due to the VCSE, it is not negligible. There is a debate in the literature about incorporating the uncorrected bias in the expression of total uncertainty (eg, Magnusson and Ellison [28], Westgard [31]). A review of this debate is not the task of this study.

UM equations start from the same error model, and TE equations (TE=SE+RE). They substitute the error parameters with the uncertainties caused in patient results.

$$UM = U_{tot_{SE}} + U_{tot_{RE}} \quad (9)$$

Where $Utot_{SE}$ is the total uncertainty of the patient's result caused by the SE, and $Utot_{RE}$ is the total uncertainty caused by the RE.

There are 2 types of uncertainty in the case of both parameters: the uncertainty of the result because the error parameters exist, and our uncertainty about the value of the parameters. For example, the uncertainty of a patient's result, caused by the RE, is as follows:

$$U_{RE} = z * SD \quad (10)$$

Unfortunately, the SD value is not accurate. Therefore:

$$U_{\text{totRE}} = U_{RE} + U_{SD} = z * SD_{\max} \quad (11)$$

where U_{SD} is the uncertainty of the SD value, and SD_{max} is the maximum value of the SD. However, the adepts of the UM critique TE theory because TE equations do not include the uncertainty of the error parameters, nor do UM equations include the uncertainty of the SD. However, s_{RW} has big monthly variations [5].

The uncertainty caused by the SE (bias) equals the bias value. U_{SE}=B. Because the bias value is uncertain, it must be added to the U_B term.

$$U_{\text{totSE}} = U_{SE} + U_B = B + U_B \quad (12)$$

According to the GUM recommendations, all discovered bias sources must be corrected. The bias becomes insignificant, and the B term can be neglected.

Applying the UM, the first step is to correct for bias (if possible and recommended). Having 2 sets of error parameters, according to the presented error model and 2 TE equations, 2 different UM equations can be obtained.

The first, calculated in repeatability conditions, starts from Equation 7. The bias, which must be corrected in the first step, is the average of the bias measurements in the same EQA round ($\overline{B}_{EQA}$). The bias value after correction can be considered negligible. The uncertainty of the correction can be determined in 2 ways (bottom-up and top-down methods). In the bottom-up approach, the bias uncertainty is calculated as the sum of uncertainties of the reference value and the uncertainty of the measurement in repeatability conditions (s_r); in the top-down method, as the sum of the uncertainty of the reference value and the root mean square of the corrected bias values ((B_{EQA_i} -$\overline{B}_{EQA}$) (Where B_{EQA_i} are the individual bias results.) The 2 methods give similar results (within the statistical methods' limits) because RMS(B_{EQA_i} -$\overline{B}_{EQA}$)≈s_r.

$$\text{UM}(t) = U_{\text{tot}} = U_{SE} + U_{RE} = B_r(t) + z * \left(\sqrt{u_{C_{ref}}^2 + u_{rec}^2 + \frac{s_r^2}{n}} + s_{r_{max}}\right) \approx$$
$$\approx B_r(t) + z * \left(\sqrt{u_{C_{\text{ref}}}^2 + u_{\text{rec}}^2 + \frac{RMS^2_{(B_{\text{EQA}_i} - \bar{B}_{\text{EQA}})}}{n}} + S_{r_{\max}}\right) \quad (13)$$

Where n is the number of measurements, $s_{r_{max}}$ is the estimated maximum value of the s_r, $u_{C_{ref}}$ is the uncertainty of the nominal value of the reference material, and u_{rec} is the uncertainty of its reconstitution, equal to the uncertainty of 2 volume measurements (≈$\sqrt{2}$ × 0.5% — the accuracy of the actual pipettes is 0.5%-0.6%). However, u_{rec} is not a negligible value; the recommended uncertainty equations do not include it. The division of the uncertainty of the bias with $\sqrt{n}$ was necessary because the bias value is a mean. As the number of measurements increases, the uncertainty of a mean value decreases $\sqrt{n}$ times). An equivalent equation was published by White [32] and in Nordtest TR 537 [22], except for the neglected u_{rec} value.

In repeatability conditions, the bottom-up and top-down methods, within the limits of the statistical measurements, give similar results for the uncertainty because $\frac{RMS^2_{(B_{EQA_i} - \overline{B}_{EQA})}}{n}$ ≈ $\frac{s_r^2}{n}$. The SD is an RMS of the deviations from the mean, with a correction: n is substituted with n-1. If a calibration is made between measurements, the top-down uncertainty will be bigger due to the bias variability. This is similar to the case of Mg: Table 3 and Figure 3.

Unfortunately, Equation 13 has no practical value in the clinical laboratory. There is a significant delay between the measurement and the moment when the results are obtained. In the meantime, reagent changes and calibrations are done, and the bias is changed. A constant cannot correct a variable. In addition, there is insufficient information to determine whether the bias is constant or proportional. Due to bias variability, the calculated uncertainty value cannot be used for extended time frames. UM is a long-term parameter.

The situation changes over time. The $\overline{B}_{RW}$=CCSE is a constant, which can be corrected without contradicting quintessential principle 3.

Each EQA round measures a different bias using different reference materials with different u_{Cref} and with varying errors of reconstitution. The average of the measured bias values in different rounds is $\overline{B}_{RW}$ (absolute mean bias). Starting from Equation 8, with bottom-up and top-down approaches, we obtain:

$$UM = U_{\text{tot}} = U_{\text{SE}} + U_{\text{RE}} = \overline{B}_{RW} + z * \left(\sqrt{\frac{RMS^2_{u_{C_{\text{ref}}}} + u_{\text{rec}}^2 + S_{RW}^2}{n}} + S_{RW_{\max}}\right) \approx$$
$$\approx \overline{B}_{RW} + z * \left(\sqrt{\frac{RMS^2_{(B_{\text{EQA}_i} - \overline{B}_{\text{EQA}})}}{n}} + S_{RW_{\max}}\right) \quad (14)$$

Actual recommendations suggest calculating the uncertainty of the bias correction as the root mean square (RMS) of the bias values [22], but this equation assumes "…a variance of bias based on assumed mean of zero" [28].

The assumption is only valid, and the equation is correct if the bias is corrected efficiently. If not, RMS_{bias} is not only u_B but includes the mean bias in its expression.

$$RMS_B = \sqrt{\bar{B}_{RW}^2 + u_B^2} = \left(\sqrt{\bar{B}_{RW}^2 + RMS^2\left(B_{EQA_i} - \bar{B}_{EQA}\right)}\right) \quad (15)$$

Which is only correct if we accept the quadratic addition law between bias and its uncertainty (questioned by the debates in the literature).

If n=1 and the $\bar{B}_{RW}$ term is added quadratically to the other terms under the square root, the top-down term of Equation 14 is equivalent to the equation proposed by Nordtest TR 537, except for the missing $RMS^2_{u_{C_{ref}}}$term. The equation in Nordtest TR 537 expresses the uncertainty of a single value, not the uncertainty of a mean (n=1) [22].

$$U_{\text{bias}}\ (\text{literature}) = \sqrt{RMS_B^2 + RMS^2_{u_{C_{\text{ref}}}}} = \sqrt{\bar{B}_{RW}^2 + RMS^2_{(B_{\text{EQA}_i} - \bar{B}_{\text{EQA}})} + RMS^2_{u_{C_{\text{ref}}}}} \quad (16)$$

In repeatability conditions, the u_{Cref} and u_{rec} caused an unknown bias in the bias value, and these terms expressed our uncertainty about this value. Making more measurements decreases the influence of random errors; however, our uncertainty about the reference value remains unchanged. In the case of different EQA rounds, these biases of the bias values are variable and contribute to the bias variability. Therefore, to avoid redundancy, the $RMS^2_{u_{C_{ref}}}$ term (included in RMS_B) must be eliminated from the top-down equation. While $u_{C_{ref}} or RMS_{u_{C_{ref}}}$) are bottom-up parameters, whereas the RMS_B is a top-down parameter, considering the consequences of the individual sources. Their mix causes redundancy in equations.

While the uncertainty caused by the bias variability (the s_{VCSE}) term) is included in both expressions in the $s_{RW_{max}}$, the top-down values are significantly bigger than the bottom-up ones. In the meantime, in the case of calculations based on internal QC data, there are no significant differences (as in the case of EQA in a single round).

Table 6 presents the differences between the bias uncertainty results obtained with top-down and bottom-up methods. Similar calculations based on internal QC data and those from a single EQA round are provided for comparison. The number of measurements is considered n=1 in all calculations for the sake of better comparison.

Table 6. Differences between the uncertainty results on 2 analyzers and 5 analytes obtained in different conditions (real-life data). All values are in percentages.

Conditions/analyte	Cobas 501 13 EQA[a] rounds top-down	Biomajesty 13 EQA rounds top-down	Cobas 501 Bottom-up	Biomajesty Bottom-up	Cobas 501 Internal QC[b] bottom-up	Cobas 501 1 EQA round repeatability
ALT[c]	4.1	5.4	2.38	4.2	2.27	1.80
AST[d]	3.07	4.3	1.91	2.04	1.73	1.62
Glucose	2.1	4.02	1.94	1.9	1.71	1.26
Urea	2.8	5.59	2.5	2.04	2.39	1.38
Potassium	1.66	1.37	1.66	1.36	1.34	1.12

[a]EQA: external quality assessment.
[b]QC: quality control.
[c]ALT: alanine aminotransferase.
[d]AST: aspartate aminotransferase.

In long time frames (more EQA rounds), the uncertainty is more significant than in a single round because variable bias values are measured. The differences between internal QC and the bottom-up method are not significant and are caused by the u_{Cref} and u_{rec} included in the bottom-up uncertainty. Except for potassium, in almost all cases, the top-down method gives a bigger value due to the difference between the declared and true u_{Cref} values.

In the bottom-up equation, the declared $u_{C_{ref}}$value is substituted. In the meantime, the top-down equation includes the real one in the RMS_B term, causing the differences. There are 2 conditions for a correct EQA. The sample must be commutable and must have predetermined values [33]. Neither of these conditions is fulfilled in EQA with surrogate reference values (the mean of participants). The equation used to evaluate the uncertainty of the reference value may only be correct if the peer groups are homogeneous, and they are not [34]. This error causes an additional and significant uncertainty.

The uncertainty equations can be corrected by eliminating the confusion hidden in the bias definitional uncertainty. A key conclusion: only the long-term mean biases can be corrected efficiently. Correcting individual values is risky due to the variability of bias and delay. The actual EQA bias determinations conceal a significant source of uncertainty: the uncertainty of the surrogate reference values. Not even the bias variability can explain the differences between the uncertainties calculated in single and multiple EQA rounds, as well as between bottom-up and top-down methods. Following studies are necessary to sustain the former theoretical conclusions; the proofs and discussion do not fit within the limits of this study.

The existence of the VCSE suggests a change in the point of view. Even after correction, the bias reappears due to its variable properties. The confusion between the bias and the mean of the variable bias is a source of error.

The (immediately) incorrigible biases bring to attention the debates about including uncorrected biases in uncertainty

equations. If they are not corrected immediately, the mean bias must be included in the uncertainty budget.

Sources of Bias Variations

We cannot quantify the preanalytical and postanalytical errors in the QC, nor can we measure the method and matrix errors only in EQA. The analytical errors detectable in IQC are:

- Environmental errors
- Laboratory errors
- Human (operator) errors
- Noninstrumental errors
- Instrumental errors [21,24]
- Rounding errors

In the case of a laboratory with air conditioning, using liquid phase reactions in thermostated conditions, the influence of the environment is quasi-negligible. The laboratory and human errors are redundant in the list. Neither specific laboratory nor specific human errors exist. Laboratory and human errors are a sum of preanalytical, noninstrumental, and instrumental errors.

We can include rounding errors in the instrumental error category. They have similar properties (both are nonspecific and time-invariable).

The instrumental errors are linked to the construction and functionality of the analyzer. They are always constant and nonspecific (assumptions 4 and 5). An instrumental failure will influence all measurements in an aberrant manner. Instrumental errors may be the sources of the constant error components, but never of the variable ones.

There are only 2 noninstrumental error sources: the reagent stability and the calibration graph (see quote from HKALM recommendations [30]). Both are specific and variable. Each measurement has its specific reagents with variable properties. Producers only guarantee that we can successfully recalibrate the reagents in the validity term, not that the properties remain constant. Random changes in the reagent properties contradict the laws of chemistry. The changes are always unidirectional and gradual. The variation is not perfectly linear; however, linearity is an acceptable approximation in short intervals. The phenomenon is consistent with the linear bias variation model of J. Krouwer ($B=B_0+b_1t$) [19]. It applies only to time frames that do not include human interventions (such as calibrations, reagent changes, or control bottle changes).

The noise of the RE usually covers the drift. We can observe only significant drifts (if the mean change is $>1.5s_r$); however, all contribute to the increase of the s_{RW}. The significant drifts cause R_{7T}, $R_{2\text{-}2S}$, and $R_{1\text{-}3S}$ violations.

Many authors consider the calibration a quasi-perfect process [35]. Raúl Girardi, on an IFCC webinar (Metrology and uncertainty, August 21, 2021), even presented an alternative equation that reduced the bias uncertainty to the nominal value uncertainty of the reference material. Other authors share similar opinions [29]. Such an attitude neglects the most significant causes of the calibration graph error. On one hand, the measured reference material does not have the same composition as the material analyzed by the producer. It undergoes a lengthy process before being measured. Even if we neglect human errors (stability, homogenization, temperature errors), the reconstitution includes 2 volume measurements: one at the producer and another at the user. Badrick [36], referring to the Tietz Textbook of Clinical Chemistry [37], underlines:

> *The act of reconstitution can introduce an error far greater than the inherent error of the rest of the analytical process.*

Each reconstituted reference material bottle has a different concentration. We generate similar systematic errors until we use the same reconstituted calibrator bottle.

On the other hand, calibration is a measurement subject to systematic and random errors. In a linear calibration, we make 2 × 2 measurements and calculate the slope factor as a difference. We make calibrations in repeatability conditions. The average calibration random error is $\approx \sqrt{2} * \frac{1}{\sqrt{2}} CV_r = 1CV_r$ (the error of the null-point absorption A_0 was neglected in the former estimation).

The calibration error introduces a systematic error in measurements, which remains constant within the time frame between calibrations, but each calibration induces an unpredictable variation in the systematic error. The result is a randomly variable systematic error. The phenomenon is consistent with the models presented by Marquise [16] and Magnusson et al [22]. We can observe only the significant shifts in the mean (>1 sr).

Because we can observe only the significant drifts and shifts, we tend to consider the bias variations unpredictable (case b of unpredictable), contradicting the bias definition (predictable).

The mostly predictable character of the bias suggests that a focus change in the internal QC is necessary. The QC system must also have a strategy to predict bias variations and detect unpredictable changes.

Properties of the VCSE(t) Function

The VCSE(t) is a time-variable function that describes the bias variations around the CCSE. It is a variable error component but different from RE. The RE changes unpredictably from measurement to measurement; meanwhile, VCSE(t) remains quasi-constant on a given day. The bias variations have unequal cycles, while the long-term mean of VCSE(t) is 0. Its values are not normally distributed.

VCSE(t) has 2 primary sources. Both are noninstrumental and specific. The variation in reagent quality follows a predictable pattern, and this variation is also predictable. After the calibration, we can predict the mean and bias variation from the old and new calibration parameters.

$$\frac{c_{before\ calibration}}{c_{after\ calibration}} \frac{F_{cal_{before\ calibration}}}{F_{cal_{after\ calibration}}}$$

VCSE(t) is a mostly predictable phenomenon. We can correct it for a moment, but not definitively eliminate it. In repeatability conditions, the VCSE(t) is nonsense. The differences between CCSE (B_{RW}) (B_{RW}: long-term mean bias, measured in RW conditions, a constant), VCSE(t), and RE are presented in Table 7.

Table 7. Differences and similarities between the error components.

Criterion	RE[a]	VCSE(t)[b]	CCSE[c]
Predictability	Unpredictable	Yes, from the preceding data	Quasi-constant
Variability	Yes	Yes	No
Distribution caused	Normal	Non-Gaussian	Quasi-constant
Influence on the mean in reproducibility within laboratory conditions	Negligible (≈0)	Only after several complete cycles, it becomes negligible≈0	Yes
Calibration influence	Insignificant	It can be corrected, but not eliminated	Not significant
Corrections or correction factors, according to GUM	No effect	In the short term, yes, on long-term reappears	Yes
Measurable under repeatability conditions	s_r[d]	$B_r(t)$ includes VCSE(t)	No
Measurable under reproducibility within laboratory conditions	s_{RW}[e] includes s_r	s_{RW} includes s_{VCSE}[f]	$\overline{B}_{RW}$

[a]RE: random error component.
[b]VCSE: variable component of the systematic error.
[c]CCSE: constant component of systematic error.
[d]s_r: SD measured in constant, repeatability conditions.
[e]s_{RW}: SD measured in variable, reproducibility within laboratory conditions.
[f]s_{VCSE}: SD calculable from the daily (run) mean, bias, or VCSE(t) values.

We cannot ignore the differences between VCSE(t), RE, and CCSE. If, and only if, we are conscious that both $B_r(t)$ and s_{RW} contain VCSE(t), it is not an erroneous practice to measure RE and VCSE(t) together and to include VCSE(t) in s_{RW}. The origins of the equations must be known, as well as the risk of redundant use.

Determination of the CCSE and the VCSE(t)

The determination of CCSE$\equiv\overline{B}_{RW}$ is possible using the control results and Equation 4. Such CCSE values only show the difference between the mean of control measurements and the target specified by the producer. We can obtain an absolute value of CCSE from the percent expressed EQA results. Due to the low number of measurements, the value has significant uncertainties.

The comparison between the 2 types of CCSE values is not the task of this study. A single mention: the difference between the 2 CCSE-s is predictably constant until we use the same control material. Another study is necessary to verify this prediction.

As a consequence of the constant difference, the VCSE(t) measured in internal QC and EQA predictably is the same; however, the statement needs confirmation. The accurate determination of the VCSE(t) function has a high cost-effectiveness ratio and negligible practical importance, mainly due to its short validity term. The computer-assisted estimation of the run means (Figure 5) is a promising solution, but needs a separate study to confirm its efficiency.

The same statement applies to the s_{VCSE}. To estimate s_{VCSE} using Equation 2 is more practical than calculating it from daily VCSE(t) values [27].

The increased s_{VCSE}/s_r ratio indicates wrong internal QC decisions (delayed calibrations); however, we can also use the $s_{RW}/s_{r\ ratio}$ without calculating s_{VCSE} [23].

The paramount importance of VCSE(t) and s_{VCSE} lies in the distinction between SD and bias types, not their absolute value. We do not need their accurate values; we do not make decisions based on them. These 2 parameters are always included in $B_r(t)$ or s_{RW}. However, we must be aware of where they are hidden. Highlighting VCSE(t) and s_{VCSE} in equations helps us avoid redundant use.

The Proposed Error Model and the Westgard Rules–Based Internal QC System

The original aim of this study was to draw attention to the neglected VCSE(t) and s_{VCSE}. The proposed new error model (Figure 6, Equations 2, 3, 5) also uncovers the weaknesses of the actual Westgard rules–based internal QC system. By distinguishing the biases measured in repeatability and reproducibility within conditions ($B_r[t]$ and B_{RW}), 2 sets of error parameters are obtained ($B_r[t]$ and s_r, respectively, for B_{RW} and s_{RW}). The link between them is VCSE(t) and s_{VCSE} (which are usually hidden in the $B_r(t)$ and s_{RW}). Avoiding redundant use by highlighting VCSE(t) and s_{VCSE} in equations is not the only advantage of the proposed error model. The non-Gaussian distribution of the VCSE(t) values explains the non-Gaussian distribution of the long-term QC data [3,4] and the significant monthly variability of s_{RW} [5], which contradicts the laws of the normal distribution. The Gauss-Laplace equation is valid only under constant repeatability conditions (if the mean remains constant). Therefore, s_r is the correct estimator of the Gaussian σ parameter and the mean RE. While the sources of specific bias variability (reagent property and calibration

parameter changes) are known [16,18,19,22,23], the sources of specific RE variability cannot be identified. All identifiable RE sources are linked to the inconsistent functionality of the instrument and, therefore, are constant (nonvariable) and nonspecific [38]. In contrast to s_{RW}, s_r is invariant within the limits of accuracy of the statistical methods (Vandra's unpublished data [38]).

The constant RE (s_r) questions the efforts of Westgard et al [1,8,39] to detect variations in RE. The primary objective of internal QC is to detect risky variations in bias, and, by definition, the bias between human interventions is predictable [14]. Anyway, according to Westgard JO [40], the QC rules cannot be applied across corrective actions. The objective change changes the way of thinking in QC. The focus is not on the immediate detection of unpredictable changes, but rather on following tendencies in bias to predict the moment when the run bias will reach a critical value.

There are 4 different mechanisms to reach a critical bias, imposing different decision strategies, because the QC rules (especially the cross-run rules: $R_{4\text{-}1S}$ and R_{10X}) have different efficiencies in each case.

1. Immediately after a calibration (Was the calibration successful?)
2. Constant bias in the case of a stable reagent (Is the new mean acceptable?)
3. Gradually increasing bias (in absolute values) in the case of an unstable reagent (When will the bias reach critical values?)
4. Unexpected shift in bias.

The immediate error detection is compulsory only in cases 1 and 4. In cases 2 and 3, bias is predictable. However, the QC system must be able to detect changes in the tendencies.

GRD Jones was the first to notice the difference between cases 1 and 4 [41], highlighting that in case 1, the cross-run rules ($R_{4\text{-}1S}$ and R_{10X}) cannot be applied due to a lack of data. However, he did not observe the hidden assumption in Westgard's calculations, which falsely assumes a constant bias in all runs. While focusing on immediate error detection in case 4, the calculations are based on case 2 (constant bias). If the cross-run rules detect a constantly critical bias, it indicates delayed, rather than immediate, error detection. In cases 3 and 4, the previous bias value is lower than in the last run, and the efficiency of the cross-run rules was overestimated.

In cases 2 and 3, the QC rules are applied repeatedly, increasing the efficiency of error detection. Instead of applying the $R_{1\text{-}3S}$ rule, the $R_{1\text{ of n-}3S}$ rule is used de facto. All runs are only accepted if neither of them violates the 3 SD decision limit.

The former observations impose the reevaluation of the efficiency of the Westgard rules in a subsequent study.

The Westgard rules are only correctly applied if the QC graphs are designed with σ or the correct estimator. As previously concluded, the correct estimator of the σ parameter and the mean RE is s_r, and Westgard's assumption that $s_{RW} \approx \sigma$ is false. Not else, but Westgard and Groth [39] acknowledged that:

> *The calculations based on computer simulations behind the power function graphs are made assuming within-run SD, while the graphs are designed with total SD.*

Considering the s_{RW}/s_r ratio, this results in an overestimation of the decision limits 1.5-2 times. Respecting Westgard's recommendations, intending to apply the $R_{1\text{-}3S}$ rule, de facto, we use the $R_{1\text{-}4.5S}$ or the $R_{1\text{-}6S}$ rule ($3s_{RW} \approx 4.5\text{-}6s_r$.) This contradiction and overestimation explain the existence of the statistically impossible graphs observed in practice (mentioned in the Introduction).

Correcting the estimator of σ (from s_{RW} to s_r) requires recalculating all parameters that include SD in their equations: TE, MU, sigma metrics, the critical SE, not just a change in the design of the QC graphs. This means an entirely new QC system, using different rules and strategies.

Sounds bizarre, but according to calculations based on normal distribution tables, a correctly applied Westgard rules–based QC system (designing the graphs with σ) would be dysfunctional due to several false alarms. Despite the efforts to correct them, half of the monthly biases measured in the internal QC are around 1 s_{RW} or bigger, and two-thirds of them are bigger than $1s_r$. According to quintessential principle 2, it cannot be corrected by calibration for smaller biases than the average calibration error, questioning another assumption of Westgard et al [39]: the assumption of error-free calibrations. According to Vandra [38], the average calibration error is ≈1-2 s_r (consistent with the observed monthly biases). If such biases are incorrigible, the QC rules must avoid alarms in these cases. The correctly applied Westgard rules alarm in the first run only by exception if the bias is 0. The statement is not valid if $B>1s_r$ and the rules are applied in several runs.

Conclusions

This study is a theoretical one. It aims to draw the attention of the scientific community to the fact that the VCSE is a neglected phenomenon and a source of several errors. Because it is hidden in the inaccurately defined bias and the s_{RW}, there is the risk of its redundant use in equations. This study also aimed to uncover the primary sources of bias variations (both present in the literature in mosaic pieces), propose corrected equations, and describe the properties of the VCSE. Because several problems were uncovered, the proofs, based on computer simulations and real-life data for each issue, neither fit within the limits of a single study nor are consistent with the declared aims. To analyze them, subsequent studies will be necessary in the future. This study intends to be a starting point for building a new QC system based on a different error model, a different strategy, and a rule system. The theoretical foundations, description, proofs with computer simulation, and real-life data do not fit within the limits of this study.

The time variability of bias is a well-known but neglected phenomenon. A variable bias does not fit into the

classical error model. If bias has variations, a question arises: Which bias is being referred to? A new error model was obtained by (1) separating the bias into a constant and a variable subcomponent and (2) distinguishing between bias measured in repeatability and reproducibility within laboratory conditions. The error model is consistent with similar attempts found in the literature; however, it questions the theory of transformation of variable biases into random errors (based on an inaccurate definition of 'random' in VIM), which forces the VCSE into the Procrustes' bed of the old error model. The author proposed definitions consistent with the VIM 2.17 definition of the SE and abbreviations consistent with those used for SD ($\overline{B}_{RW}$, $B_r(t)$).

The bias variability has 2 sources. Both are noninstrumental and specific to each measurement, and neither causes normally distributed biases. One is reagent instability, and the other is human intervention, including reagent changes and calibrations. Reagent instability causes gradually increasing, quasilinear biases, whereas calibrations result in alternation between constant periods with random shifts in the calibration parameters. Computer simulations and real-life QC data presented in this study support that these are real sources of bias variability.

The 2 phenomena occur simultaneously, resulting in sawtooth-like variations in bias. In the time frames between human interventions, the biases are predictable. However, they are hidden behind the noise of the RE. Without computer assistance, we can observe only significant shifts and drifts. For this reason, the increase of the SD in longer time frames was erroneously considered unpredictable, with an unknown cause (type b of unpredictable). An unpredictable bias contradicts its definition in VIM.

We must change our way of thinking in the QC by focusing on predictive actions instead of corrective ones.

The properties of the CCSE, the VCSE(t) function, and the RE differ, justifying the distinction between them. Accurately determining the SE subcomponents theoretically is possible; however, it has a high cost/effectiveness ratio. The significance of their separation is that they help us avoid the redundant use of the VCSE(t) classically hidden in $B_r(t)$ and s_{RW}.

Two sets of error parameters are obtained by separating biases measured in repeatability and reproducibility within laboratory conditions. We must determine the parameters under the same conditions we use them. UM calculations must be based on parameters determined under reproducibility within laboratory conditions, whereas internal QC decisions must be based on parameters determined under repeatability conditions. This conclusion is thought-provoking because it contradicts the recommendations for designing the Levey-Jennings graphs based on the SD calculated from long-term control data. In the meantime, the calculations behind the Westgard rules assume pure RE.

The actual Westgard rules–based internal QC system is not consistent with two quintessential principles valid in all sciences:

1. We must determine the parameters under the same conditions we use them.
2. A calibration cannot efficiently correct smaller biases than the mean calibration error.

The proposed error model uncovered several false assumptions behind the actual Westgard rules–based QC system.

1. The internal QC aims to detect variations in RE and SE. (RE is not variable.)
2. Bias variations are unpredictable. (Correct: between human interventions are predictable.)
3. The same rules are efficient in all cases. (Correct: there are 4 different situations of decisions, imposing different rules and strategies.)
4. Cross-run rules can be applied in immediate error detection. (Correct: they can be applied only with a delay.)
5. The estimator of the σ parameter and the measure of the mean RE is s_{RW} (Correct: it is s_r.)
6. QC graphs must be designed with s_{RW}. (Correct: with s_r, highlighting the incorrigible biases.)
7. Calibrations are error-free, and all biases are correctable by calibration. (Correct: smaller biases than $1\text{-}2s_r$ are incorrigible.)

The false assumptions 6 and 7 cause 2 compensating errors. The compensation explains the long-term success of the Westgard rules. If we use s_{RW} in the design of the Levey-Jennings graphs, we use larger, increased decision limits, de facto applying different rules (eg, the $R_{1\text{-}5S}$ rule instead of the intended $R_{1\text{-}3S}$). As a consequence, the alarms for incorrigible biases become less frequent. However, this compensation is not accurate. The observed statistically impossible QC graphs sustain the overestimation of the RE by the s_{RW}.

Based on the proposed error model, correcting the former false assumptions, and considering the 4 different decision situations, the Westgard rules–based QC system must be mathematically reevaluated. It can be predicted that patching it is not a solution, and a new QC system is necessary, based on the s_r, and the avoidance of alarms in the case of incorrigible biases.

The proposed error model also suggests corrections to the MU equations. MU is a long-term parameter, and therefore, its equation must be based on long-term parameters. The uncertainty of the inaccurately defined bias (Which one?) must be substituted with the uncertainty of the long-term mean bias, measured in reproducibility within laboratory conditions ($U(\overline{B}_{RW})$), and must be considered the uncertainty caused by the variability of s_{RW}, substituting it with its maximal value in the MU equation.

Furthermore, the proposed error model, together with quintessential principle 1 (that all parameters must be determined under the same conditions under which they are used), explains why the more correct MU theory cannot substitute for TE in internal QC decisions. MU is a long-term parameter, while internal QC decisions are made under repeatability conditions.

Acknowledgments

The author thanks Dr Prof Marius Mărușteri for the initial reading, valuable advice, and constructive critiques that helped improve the study, as well as the reviewers' critical opinions. To this study, no other persons contributed except the author. The author created all images and tables. The author attests that this manuscript did not use generative artificial intelligence (AI) technology to generate figures, ideas, data, or other informational content. AI was used only for grammar correction and for unintentional plagiarism detection. To assist with the language correction, the author used the following Grammarly AI prompts: “Improve it” and “Find synonyms.”

Data Availability

All computer simulation files were uploaded as Multimedia Appendices 1 and 2 (in Excel format). The data, which constituted the basis of the real-life data graphs, were also uploaded as Multimedia Appendices 3 and 4 (Excel files). The latter data source is the quality control results obtained in the Brasov County Clinical Hospital for Urgencies (Romanian abbreviation: SCJUBv), part of a protected database; therefore, these cannot be made available. The author did not use patient data in this study. In the real-life examples, reference materials produced by Roche were used.

Conflicts of Interest

None declared.

Multimedia Appendix 1

Random values generation and shift and drift computer simulation.
[XLSX File (Microsoft Excel File), 926 KB-Multimedia Appendix 1]

Multimedia Appendix 2

Differences in calibrations and shift and drift computer simulations (protected file).
[XLSX File (Microsoft Excel File), 408 KB-Multimedia Appendix 2]

Multimedia Appendix 3

Real-life data for glucose. The influence of reagent degradation on the bias variation.
[XLSX File (Microsoft Excel File), 278 KB-Multimedia Appendix 3]

Multimedia Appendix 4

Real-life data for magnesium. Shift caused by calibration.
[XLSX File (Microsoft Excel File), 49 KB-Multimedia Appendix 4]

References

1. Westgard JO, Barry PL, Hunt MR, Groth T. A multi-rule Shewhart chart for quality control in clinical chemistry. Clin Chem. Mar 1981;27(3):493-501. [doi: 10.1093/clinchem/27.3.493] [Medline: 7471403]
2. Stahl S. The evolution of the normal distribution. Math Mag. Apr 2006;79(2):96-113. [doi: 10.1080/0025570X.2006.11953386]
3. Badrick T. Biological variation: understanding why it is so important? Pract Lab Med. Jan 2021;23:e00199. [doi: 10.1016/j.plabm.2020.e00199] [Medline: 33490349]
4. Katayev A, Fleming JK. Past, present, and future of laboratory quality control: patient- based real-time quality control or when getting more quality at less cost is not wishful thinking. J Lab Precis Med. 2020;5:28-28. [doi: 10.21037/jlpm-2019-qc-03]
5. Kumar BV, Mohan T. Sigma metrics as a tool for evaluating the performance of internal quality control in a clinical chemistry laboratory. J Lab Physicians. 2018;10(2):194-199. [doi: 10.4103/JLP.JLP_102_17] [Medline: 29692587]
6. [VIM3] 2.19 random measurement error. Joint Committee for Guides in Metrology. 2025. URL: https://jcgm.bipm.org/vim/en/2.19.html [Accessed 2025-06-30]
7. Krystek M. Calculating Measurement Uncertainties. Beuth Verlag GmbH; 2016. ISBN: 978-3-410-26091-2
8. Westgard JO, Carey RN, Wold S. Criteria for judging precision and accuracy in method development and evaluation. Clin Chem. Jul 1974;20(7):825-833. [doi: 10.1093/clinchem/20.7.825] [Medline: 4835236]
9. How to calculate your long term bias for your uncertainty calculation?. Weqas. Jul 13, 2020. URL: https://www.weqas.com/download/how-to-calculate-your-long-term-bias-for-your-uncertainty-calculation/ [Accessed 2022-05-29]
10. Tholen DW. Evaluation of the linearity of quantitative measurement procedures: a statistical approach; approved guideline. Vol 23. NCCLS; 2003. URL: https://mdcpp.com/doc/standard/NCCLSEP6-A-2003.pdf [Accessed 2026-01-28]

11. JCGM GUM-6:2020. Bureau International des Poids et Mesures (BIPM). URL: https://www.bipm.org/documents/20126/2071204/JCGM_GUM_6_2020.pdf/d4e77d99-3870-0908-ff37-c1b6a230a337 [Accessed 2025-05-06]
12. JCGM GUM 100:2008 evaluation of measurement data — guide to the expression of uncertainty in measurement. Bureau International des Poids et Mesures (BIPM). Jan 1, 2008. URL: http://www.bipm.org/en/publications/guides/gum.html [Accessed 2025-06-29]
13. Leito I. Validation of liquid chromatography mass spectrometry (LC-MS) methods. University of Tartu. URL: https://sisu.ut.ee/lcms_method_validation/51-Bias-and-its-constituents [Accessed 2024-05-29]
14. [VIM3] 2.17 systematic measurement error. Joint Committee for Guides in Metrology. 2025. URL: https://jcgm.bipm.org/vim/en/2.17.html [Accessed 2025-06-30]
15. Shewhart WA. Economic quality control of manufactured product. Bell Syst Tech J. 1930;9(2):364-389. [doi: 10.1002/j.1538-7305.1930.tb00373.x]
16. Marquis P. Common misconceptions in medical laboratory quality control. Service de Biochimie, Centre hospitalier régional Metz – France; URL: https://files.secure.website/wscfus/4091441/31621936/misconceptions.pdf [Accessed 2026-01-28]
17. Eisenhart C. Realistic evaluation of the precision and accuracy of instrument calibration systems. J Res Natl Bur Stan Sect C Eng Instr. Apr 1963;67C(2):161. [doi: 10.6028/jres.067C.015]
18. Haeckel R, Schneider B. Detection of drift effects before calculating the standard deviation as a measure of analytical imprecision. Clin Chem Lab Med. 1983;21(8):491-498. [doi: 10.1515/cclm.1983.21.8.491]
19. Krouwer JS. Setting performance goals and evaluating total analytical error for diagnostic assays. Clin Chem. Jun 2002;48(6 Pt 1):919-927. [doi: 10.1093/clinchem/48.6.919] [Medline: 12029009]
20. Kadis R. Evaluation of measurement uncertainty in analytical chemistry: related concepts and some points of misinterpretation. ResearchGate. 2008. URL: https://www.researchgate.net/publication/277054223_Evaluation_of_measurement_uncertainty_in_analytical_chemistry_related_co ncepts_and_some_points_of_misinterpretation [Accessed 2025-03-07]
21. Theodorsson E, Magnusson B, Leito I. Bias in clinical chemistry. Bioanalysis. 2014;6(21):2855-2875. [doi: 10.4155/bio.14.249] [Medline: 25486232]
22. Magnusson B, Näykki T, Hovind H, Krysell M, Sahlin E. Handbook for calculation of measurement uncertainty in environmental laboratories (NT TR 537 - edition 4). NORDTEST. Nov 29, 2017. URL: http://www.nordtest.info/wp/2017/11/29/handbook-for-calculation-of-measurement-uncertainty-in-environmental-laboratories-nt-tr-537-edition-4/ [Accessed 2025-03-07]
23. Mackay M, Hegedus G, Badrick T. Assay stability, the missing component of the error budget. Clin Biochem. Dec 2017;50(18):1136-1144. [doi: 10.1016/j.clinbiochem.2017.07.004] [Medline: 28733188]
24. Oosterhuis WP, Bayat H, Armbruster D, et al. The use of error and uncertainty methods in the medical laboratory. Clin Chem Lab Med. Jan 26, 2018;56(2):209-219. [doi: 10.1515/cclm-2017-0341] [Medline: 28796637]
25. Vandra AB. Incertitudini… în lumea incertitudinii. Deplasarea [Uncertainties... in the world of uncertainties. The bias]. Revista română de laborator medical. Sep 2014.
26. Gauss CF. Bestimmung der Genauigkeit der Beobachtungen [determining the accuracy of the observations]. Z Astron Verw Wiss [J Astron Relat Sci]. 1816;1:187-197.
27. Vandra AB. Reevaluation of the variable component of the systematic error calls for paradigm change in clinical laboratory quality control. Health Systems and Quality Improvement. Preprint posted online on May 28, 2023. [doi: 10.1101/2023.05.24.23290382]
28. Magnusson B, Ellison SLR. Treatment of uncorrected measurement bias in uncertainty estimation for chemical measurements. Anal Bioanal Chem. Jan 2008;390(1):201-213. [doi: 10.1007/s00216-007-1693-1] [Medline: 18026721]
29. White GH, Farrance I, AACB Uncertainty of Measurement Working Group. Uncertainty of measurement in quantitative medical testing: a laboratory implementation guide. Clin Biochem Rev. 2004;25(4):S1-24. [Medline: 18650962]
30. Pang R. A guide on how to implement internal quality control (IQC) HKAML 2024. ResearchGate. Mar 27, 2024. URL: https://www.researchgate.net/publication/379333346 [Accessed 2025-07-01]
31. Westgard S. Uncertainty in how to calculate measurement uncertainty: different approaches for incorporating effects of clinically significant bias. WestgardQC. 2023. URL: https://westgard.com/essays/iso/uncertainty-in-uncertainty.html [Accessed 2025-05-05]
32. White GH. Basics of estimating measurement uncertainty. Clin Biochem Rev. Aug 2008;29 Suppl 1(Suppl 1):S53-S60. [Medline: 18852859]
33. Kristensen GBB, Meijer P. Interpretation of EQA results and EQA-based trouble shooting. Biochem Med (Zagreb). Feb 15, 2017;27(1):49-62. [doi: 10.11613/BM.2017.007] [Medline: 28392726]

34. Vlašić Tanasković J, Coucke W, Leniček Krleža J, Vuković Rodriguez J. Peer groups splitting in Croatian EQA scheme: a trade-off between homogeneity and sample size number. Clin Chem Lab Med. Mar 1, 2017;55(4):539-545. [doi: 10.1515/cclm-2016-0284] [Medline: 27658147]
35. Toacşe G, Toacşe AM. Controlul de Calitate Şi Validarea Metodelor Analitice Cantitative: Pentru Laboratoarelor Medicale Bucureşti [Quality Control and Validation of Quantitative Analytical Methods: For the Use of Clinical Laboratories Bucharest]. Editura Tehnică; 2010. ISBN: 978-973-31-2368-2
36. Badrick T. The quality control system. Clin Biochem Rev. Aug 2008;29 Suppl 1(Suppl 1):S67-S70. [Medline: 18852861]
37. Burtis CA, Ashwood ER, editors. Tietz Textbook of Clinical Chemistry 2. W.B. Saunders; 1994. ISBN: 9780721644721
38. Vandra AB. Calibration error, a neglected error source in the clinical laboratory quality control. EJIFCC. Dec 2025;36(4):443-451. [Medline: 41459182]
39. Westgard JO, Groth T. Power functions for statistical control rules. Clin Chem. Jun 1979;25(6):863-869. [doi: 10.1093/clinchem/25.6.863] [Medline: 445821]
40. QC - the multirule interpretation. WestgardQC. URL: https://www.westgard.com/lessons/basic-qc-practices-l/31-lesson18.html [Accessed 2025-07-01]
41. Jones GRD. Reevaluation of the power of error detection of Westgard multirules. Clin Chem. Apr 2004;50(4):762-764. [doi: 10.1373/clinchem.2003.025585] [Medline: 15044336]

Abbreviations

B_{RW}: long-term mean bias, measured in RW conditions, a constant
CCSE: constant component of systematic error
CV: coefficient of variation, the SD expressed as a percent of the mean of measurements
CV_r: CV measured in constant, repeatability conditions
CV_{RW}: CV measured in variable, reproducibility within laboratory conditions
EQA: external quality assessment
IQC: internal quality control
QC: quality control
RE: random error component
SE: systematic error component
s_r: SD measured in constant, repeatability conditions
s_{RW}: SD measured in variable, reproducibility within laboratory conditions
s_{VCSE}: the SD calculable from the daily (run) mean, bias, or VCSE(t) values
TE: total measurement error
UM: uncertainty of measurement
VCSE: variable component of the systematic error
VIM: International Vocabulary of Metrology

Edited by Tiffany Leung; peer-reviewed by Elvar Theodorsson, Anonymous; submitted 05.Jun.2023; final revised version received 09.Jul.2025; accepted 30.Nov.2025; published 27.Feb.2026

Please cite as:
Vandra AB
Investigating the Variable Component of the Systematic Error, a Neglected Error Parameter: Theoretical Reevaluation Study
JMIRx Med 2026;7:e49657
URL: https://med.jmirx.org/2026/1/e49657
doi: 10.2196/49657