Research Article

An optimized method for setting Internal Quality Control targets (mean and limits) for multi-instrument Internal Quality Control strategies in hematology?

Tony Badrick[1*], Mathieu Bernard[2], Lionel Tabard[3], Jean-Marc Giannoli[4]

[1]*Royal College of Pathologists of Australasia Quality Assurance Programs, 8 Herbert Street, St Leonards, NSW, Sydney, Australia*
[2]*Bioesterel-Biogroup - Technical Platform, Sanary sur Mer, France*
[3]*Biogroup AURA, Lyon, France*
[4]*Cerballiance Laboratory, CerbaHealthCare, AURA, Lyon, France*

Article Info

[*]***Corresponding Author:***
Tony Badrick
Royal College of Pathologists of Australasia Quality Assurance Programs
8 Herbert Street, St Leonards, NSW, Sydney, Australia
E-mail: tony.badrick@rcpaqap.com.au

Keywords

components of random error in multi-instrument sites, peer group analyzers statistics, IQC of multiple analyzers

Abstract

Objectives: Internal Quality Control (IQC) procedures must evolve as laboratory practices change to meet the requirements of higher patient volumes and the greater need for comparability of results across laboratories. This study developed a strategy to detect significant errors in routine hematology analytical processes using a novel approach to setting IQC material targets (mean and limits) across multiple instruments. This has not been described for hematology laboratories before.

Content: The approach described used a common IQC sample mean and limits determined from a manufacturer-based peer group using the same IQC material. The limits involved a new parameter, the random error of the peer group analyzers around their own mean ($SD_{intra}90$), to detect bias or imprecision. The model was assessed over five months using 17 analyzers that measured hemoglobin and red cell count. Using a common mean requires that the different analyzers exhibit no significant bias.

Summary: The model effectively controls analytical error in a network of laboratories using the same measurement system. This model aligns with the best theoretical principles for patient risk reduction and harmonizes practices for acceptance and rejection across the network.

Outlook: The model presents an approach to IQC that meets the demands of real-world practice.

Introduction

Clinical hematology laboratories employ statistical methods and tools, primarily control charts, to monitor and control analytical processes, ensuring they operate efficiently and produce reliable results for diagnosing and monitoring patient diseases [1–3].

The nature of the cellular measurands in hematology presents special challenges in providing suitable internal quality control (IQC) samples, leading to the use of preserved human or animal blood materials to ensure pre-analytical stability [4]. Consequently, commercial controls from manufacturers and independent providers are widely used. These samples are stabilised to extend their shelf life but are often very sensitive to temperature and storage conditions [5]. Using the manufacturer-provided limits of acceptability in the insert is common [6], but not considered best practice as they are usually broader than the laboratory's self-determined limits [6,7].

In 2024, the International Council for Standardization in Haematology (ICSH) published the results of an international survey into IQC practices.[6]. They found diversity in both guidance and practice worldwide. There were two primary approaches to setting IQC targets: either use the IQC limits supplied by the manufacturer in the insert or calculate the targets using the IQC values. For hemoglobin (Hb), the limits used ranged from mean ± 1.5 to 6 SD (or 11.5S to 16S as Westgard rules [3]), with a median of 4 SD. These limits were significantly broader for manufacturer-provided limits. Responding to this variation, the ICSH produced guidelines to promote best practices [8].

However, they neglect the issue of the IQC situation, where multiple instruments measure the same analyte, a very real problem in modern practice, as all instruments will have a slight difference in bias and imprecision when compared to each other. There is little guidance in the literature on this situation, but what is suggested is individual instrument targets, which we will show is not theoretically optimal [9].

Suppose two or more different analyzers are measuring the same analyte. In that case, two alternative IQC strategies can be employed [10]; the first is to manage each instrument separately, that is, set an instrument-specific mean and SD. The other option is to have a common mean and SD (and IQC batch) for all instruments. Either option has advantages and disadvantages depending on what the laboratory aims to achieve. If the goal is to reduce unreliable patient results during an out-of-control situation, then Parvin et al [10] showed a need to centre the IQC rejection limits on a common mean. It has been demonstrated in simulation studies that entering the control rule on the common mean is an optimal strategy for reducing patient risk with analyzers that have significant and stable uncorrected bias [10,11]. Not surprisingly, Yago and Pla found that IQC procedures are generally not optimal in the presence of significant bias [11]. This must be assessed and monitored.

Figure 1 illustrates an example of an analytical site with a clustering of four analyzers where the laboratory uses individual instrument targets to develop its control charts. Each analyser has its own targets (mean/SD), and an identical IQC value can result in rejecting one analyzer run and accepting another.

Figure 1: The data from multiple instruments measuring the same IQC sample. The coloured plots represent separate analyzers' mean(m) and error distributions.

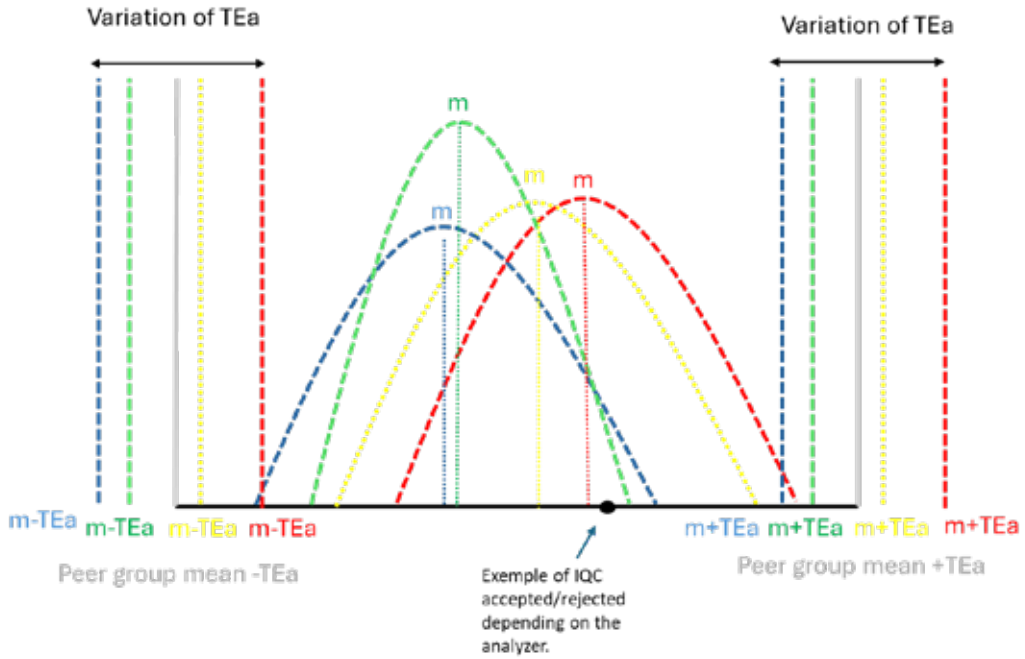


Figure 2: An example of the data from multiple instruments measuring the same IQC sample. The coloured plots show the error distribution of separate analyzers.

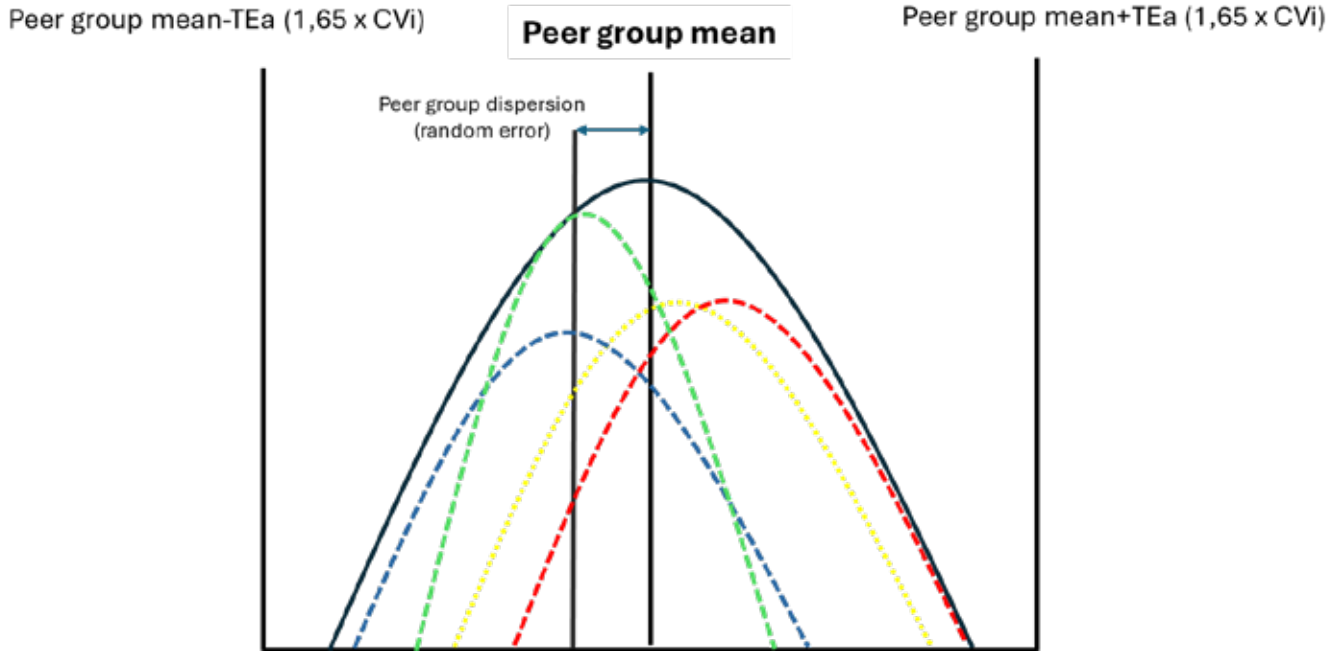


Figure 2 shows the effect of using a common CV and mean for error detection.

All IQC strategies rely on accurate IQC targets of mean and standard deviation (SD) to set rules for the acceptance/rejection of an IQC sample result. The ICSH Guideline recommended that "Laboratories should choose whether to apply their own calculated action limits as prescribed by CLSI H26-A2 [12], or to tighten the limits provided by the manufacturer. For example, suppose a cell counter manufacturer does not recommend that the target and limit values supplied with the assay sheet be used routinely, or whether they should be verified before use by the diagnostic laboratory. This should be clearly stated in the product insert" [6].

Deciding to use common IQC targets raises the question of how to derive them. The obvious and most optimal option for a laboratory is to use its own IQC data. However, this approach has disadvantages, including small sample sizes and the risk that small biases will accumulate over time, potentially leading to a significant bias that may not be detected [13]. External Quality Assurance schemes are designed to detect these trends but may be slow. Some manufacturers calculate peer-group means and SDs for the provided IQC material, based on all instruments supported by the supplier. A peer group is defined as a group of analyzers from the same manufacturer's analyzer series using the same lot number of QC material and measurement method. These are provided regularly and automatically calculated from results downloaded from each analyser. The advantage of this approach is the number of data points used to determine the IQC targets.

Different ways exist to set control rules (reject/accept) based on the mean and SD. Usually, the rule is to reject the batch of results if the IQC sample value lies outside the mean ± 2 or ± 3 SD. Based on the Gaussian distribution, these rules will only flag if the IQC result is outside the 95th or 99th percent of expected values. However, there are other options for the percentile to choose for rejection in the IQC rule.

This paper describes an IQC model for managing multiple hematology analyzers that measure the same analytes at the same site or within a network. The approach will utilize common, manufacturer-provided IQC targets and a novel method, the peer group rejection limit, specifically the 90th percentile ($CV_{intra}90$), as the limit of acceptability. The red cell count (RBC, measured using the impedance principle of measurement) and Hb (measured using the spectrophotometric principle of measurement) were used as examples in this study.

Methods

There were two arms to this study. A retrospective arm that collected long-term data to analyze long-term data to determine the IQC parameters that would be used for the construction of the IQC control charts, specifically the target mean and dispersion limit.

The second arm involved a prospective study using these targets and assessing their reliability.

Sources of retrospective peer group data

Sysmex provides data by instrument peer group, all instruments of the same model. These Sysmex instrument peer groups consisted of between 1,269 to 1,786 analyzers distributed

throughout Europe. The first step of the study consists of retrospectively analyzing Sysmex peer group data from previous IQC batches (years 2021/2022) to determine the parameters necessary for the construction of common control charts.

Target

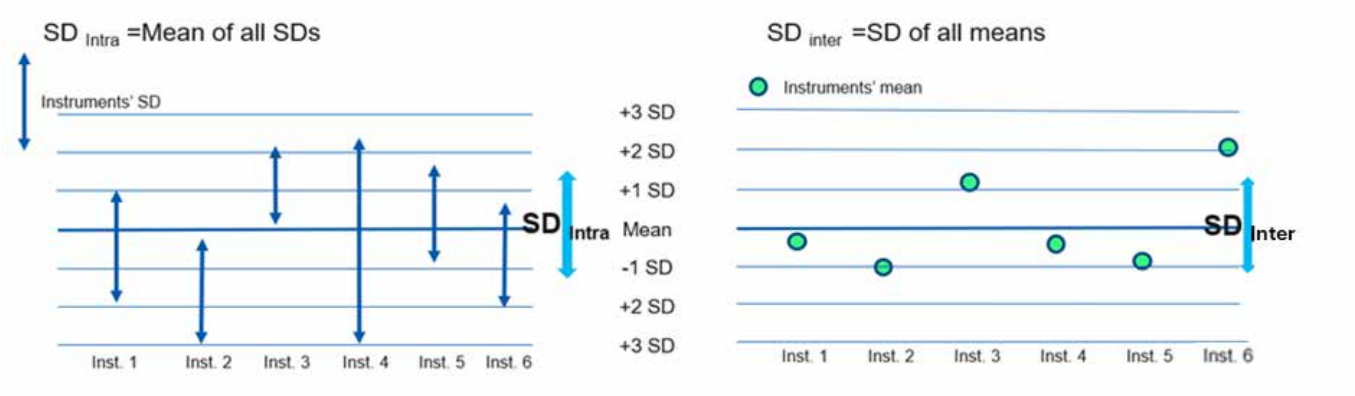
The robustness of the peer group averages, which will serve as target averages for each batch, is checked by recording the number of participants (N). The reliability of this target average is evaluated by assessing the distribution of these averages. The stability of the group average is crucial.

Determining the dispersion target

Sysmex outsourced IQC reports offer two peer group result variation points, the SD_{inter} and the SD_{intra}. The SD_{intra} is the random error of the peer group analyzers around their own mean (the mean of all the SDs), whereas the SD_{inter} is the dispersion of the analyzer means about the consensus mean (the SD of all the means). The total SD, SDtotal, is the sum of these two components and is a form of measurement uncertainty of the peer group (see Figure 3).

$$SDtotal = \sqrt{\{(SD_{intra})2 + (SD_{inter})2\}}$$

Figure 3: Components of the imprecision of separate analyzers.



Sysmex provided the authors with SD_{intra} percentile data from fifteen batches of IQC material (30 months) that had been used in Europe between 2021 and 2022. These SD_{intra} percentiles were converted to CV_{intra} with the mean of the peer group corresponding to each batch to obtain $CV_{intra}50$ / $CV_{intra}90$ / $CV_{intra}95$. After checking the normality of distribution of the percentile CVs of the fifteen batches (Shapiro-Wilk test), an average of the CVs weighted with the number of analyzers included in each batch was calculated.

The CV_{intra} chosen to set the limits is the average CV_{intra} 90, the state-of-the-art imprecision for specific analyzers and IQC level. The selection of the $CV_{intra}90$ over the $CV_{intra}95$ or $CV_{intra}50$ was empirical. The relevance of this choice was verified by comparing the $CV_{intra}50/90/95$ with the pooled CVs data from the analyzers used in the following part. The pooled CV was calculated as the total CV for an IQC sample batch from the seventeen analyzers and compared to the percentile data from the peer group provided by Sysmex.

The prospective study - using the IQC targets on real data

The prospective study was conducted over five months, monitoring the analytical process from December 1, 2023, to April 30, 2024. This period is covered by batch numbers 3339 and 4023. It was carried out with seventeen Sysmex XN analyzers (Sysmex Europe, Hamburg, Germany) distributed across four different sites covering the Lyon region (France). Five analyzers on a site that we will call "D", six analyzers on a site named "S", three analyzers on a site named "R" and three analyzers on a site named "F". These seventeen analyzers share the same reference intervals on the reports. Whether in a diagnostic context or during monitoring or screening, one of these seventeen instruments can analyze patient samples indifferently.

The total acceptable error is defined with the biological variations of intra-individual CV data obtained from the EFLM database [14,15]: 1.65 x CVi (intraindividual biological variation) for the IQC Hb and RBC level 2 (normal level). This approach comes from the notion of MAU EFLM (allowable measurement uncertainty) [16].

IQC rules

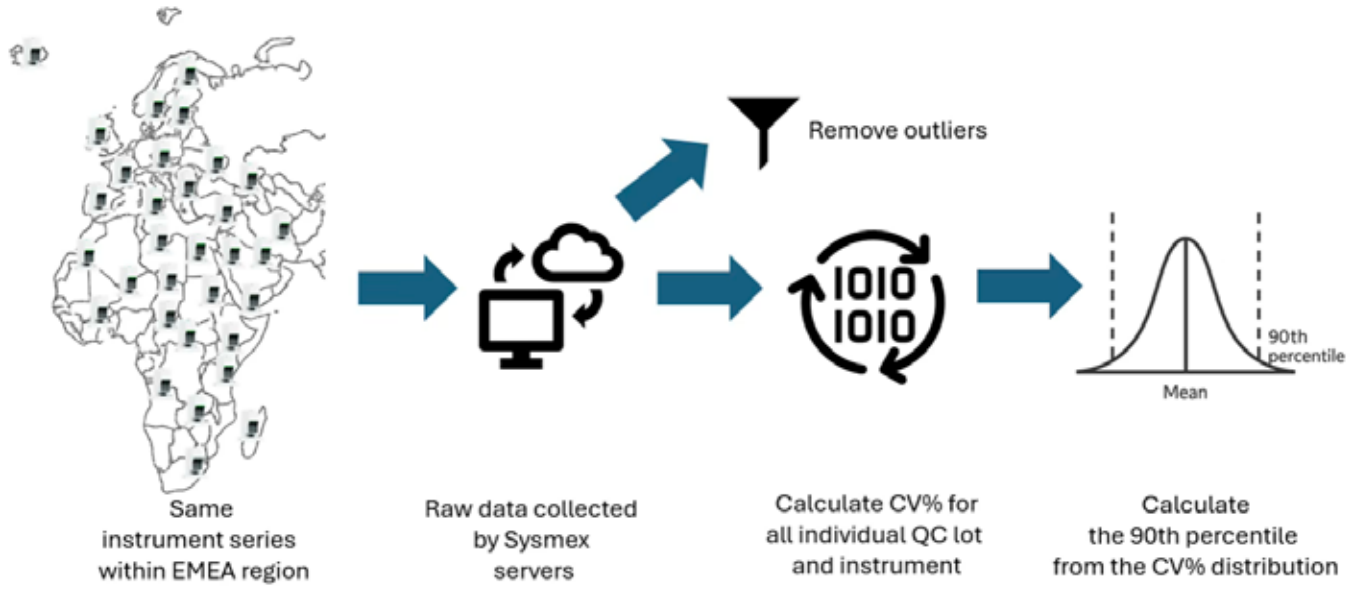
Depending on the number of samples, which routinely range between 300 and 500 per instrument per day, two or three levels of IQC are used at the start of the analysis series: one level in the middle of the series, and a third IQC event of one or two levels at the end. At the start or middle of the series, the Westgard 22s rule was used as an alarm rule, while the 13s and 2/32s rules (two levels of out-of-bounds controls in the same direction) were used to signal a rejection.

Use of patient-based moving medians.
Patient-based quality control is well established in hematology with Bull's algorithm [17] however, other models are available with better detection rates [18]. The ISLH suggested that "haematology analyzers may be able to provide regular moving averages for other parameters and programmable for adjusting outlier limits and the number of patient results averaged" [6]. Patient measurement data for Hb and RBC were extracted retrospectively. The moving median of block 50 for each analyzer was calculated after Winzorization from the usual values of each parameter. From the 15th day, the median ± 2 SD of each parameter was calculated and fixed for the remainder of the study to detect any retrospective drift or shift during the analyzed period. This analysis, presented in Figure 7, enables us to superimpose the variations observed in the IQC with those observed in the patient data.

Results

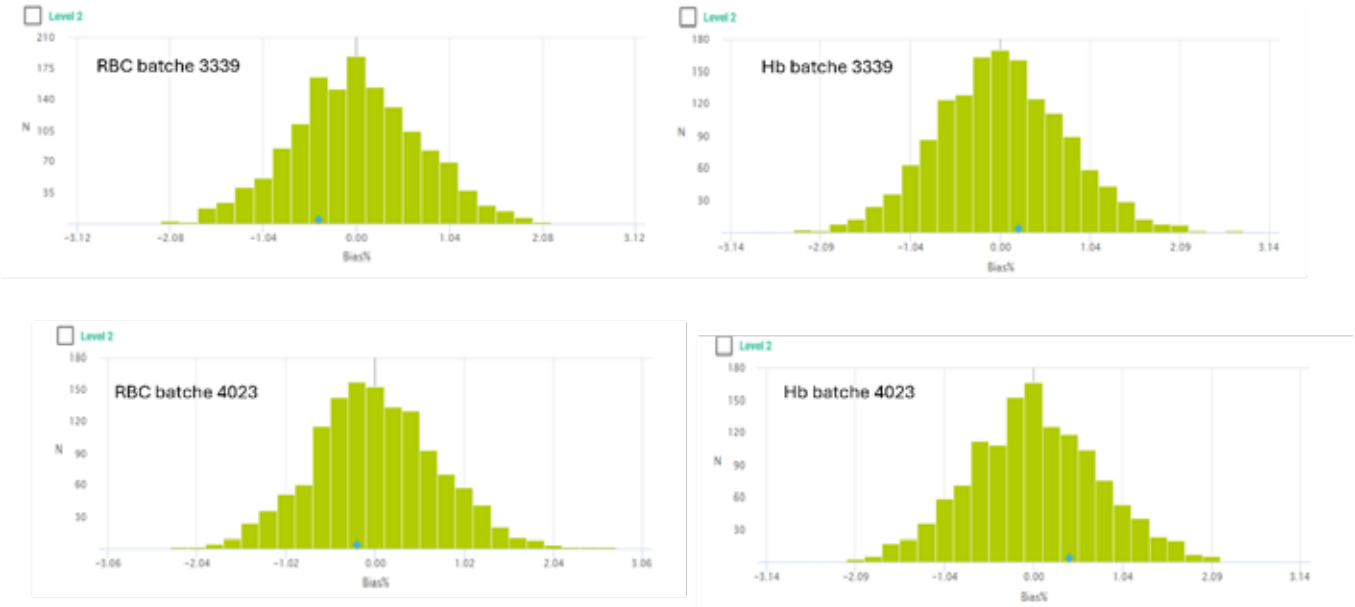
Prospective study IQC and patient data
Each new batch of manufacturer IQC samples (Sysmex Europe, Hamburg, Germany) was received the week before implementation. During the overlap period, samples from the latest batch of material are measured four times on the seventeen instruments. Then, two to three days before the routine implementation of this IQC batch, the control chart target is determined using the Caresphere data hub (https://www.sysmex-europe.com/services/qc-online-services/caresphere-xqc/). This is calculated after all values outside the group mean ± 5 SD_{inter} are excluded using the European reference group. This group average is available to all users in real-time and is updated daily. The group averages nevertheless remained stable throughout the validity period of the batches studied here. The concentration levels between IQC batches are always in the same part of the method accuracy profile. Thus, the SD used as limits for the seventeen analyzers is the $CV_{intra}90$ described above, multiplied by the peer group average corresponding to each of the two batches studied. Caresphere XQC is an EQA provider (ISO/IEC 17043) accredited by the Japanese Accreditation Body (JAB) (https://www.sysmex-europe.com/services/qc-online-services/caresphere-xqc/) (Figure 4).

Figure 4: The Caresphere process for determining IQC material targets (mean and CV) is provided to laboratories daily.



Target
The peer groups of batches 3339 and 4023 comprised 1,249 and 1,294 analyzers, respectively, on the first startup day and 1,474 and 1,313 on the last day. The distribution of the analyzer averages around the consensus average was normal (Figure 5).

Figure 5: The peer group data of IQC batches 3339 and 4023.



The group average on day one for lot 3339 was 4.310 T/L for RBC and 11.29 g/dl for Hb. The end-of-use group average for this lot was 4.306 T/L for RBC and 11.28 g/dl for Hb. Regarding batch 4023, the first and last day averages were 4,314 T/L and 4,310 T/L for RBC and 11.37 g/dl and 11.38 g/dl for Hb.

Verification of assumptions

It is essential to verify that all seventeen instruments have minimal differences in bias or imprecision. The absence of significant bias was assessed by constructing a figure showing the daily biases (%) of each analyzer relative to the peer-group average (Figure 6). The total error (TE) is defined as the bias plus 45% of the dispersion due to random error: bias + 1.65 CV. The system is in control if TE < TEa or bias + 1.65 x CV < TEa. The upper limit before significance of the positive bias at the 5% risk level can then be calculated as follows: upper limit of bias = TEa - 1.65 x actual CV. And the lower limit before significance of a negative bias is calculated as follows: lower limit of bias = -TEa + 1.65 x CV. The CV used corresponds to the actual mean CV of the analyzers over the studied period. In other words, if the bias of an analyzer plus the dispersion due to random error does not exceed the total error, we cannot conclude that it is significant. The significance of the bias between analyzers was estimated using the same principle, but this time considering the analyzer with the lowest average (MINa) and the one with the highest average (MAXa) for each batch of IQC analyzed. If the delta% between the MINa and MAXa plus the random error coefficient (CV) of the analyzer with the greatest dispersion is less than the TEa, we cannot conclude that the difference between the two most deviant analyzers is biologically significant.

The SD_{intra}90 used as the dispersion limit for the seventeen analyzers must represent their overall performance. The comparison of the overall actual dispersion of the method with the theoretical dispersions (SD_{intra}50/90/95) is carried out by calculating the pooled variance (S2pooled) converted to pooled CV from the SD, means and number of points of each IQC batch for the seventeen analyzers [19].

$$s^2_{pooled} = \frac{1}{N-1}\left(\sum_{k=1}^{m} s_k^2(n_k-1) + \sum_{k=1}^{m} n_k\bar{x}_k^2 - \frac{1}{N}\left[\sum_{k=1}^{m} n_k\bar{x}_k\right]^2\right)$$

Bias and comparability

The acceptable bias between analyzers and the group average over five months was verified through a daily visual review, which confirmed that the bias of the analyzers did not exceed the defined bias objective. The difference of the average MAX and average MIN of hemoglobin and RBC IQC values for each batch was compared.

For batch 3339, the MIN and MAX averages corresponded to analyzers D4 and D2 for Hb and D4 and to D1 for RBC. The deltas plus 1.65 x CV were 3.29% for Hb and 4.43% for RBC, with a TEa of 4.5% for Hb, and 4.6% for RBC. For batch 4023, the MIN and MAX averages corresponded to analyzers R3 and D2 for Hb, and to analyzers R1 and D1 for RBC. The deltas plus 1.65 x CV were 3.92% for Hb, and 4.43% for RBC, with a TEa of 4.5% for Hb and 4.6% for RBC. It cannot be concluded that the analyzers are not comparable.

Figure 6: Daily IQC bias for all analyzers over the assessment period.

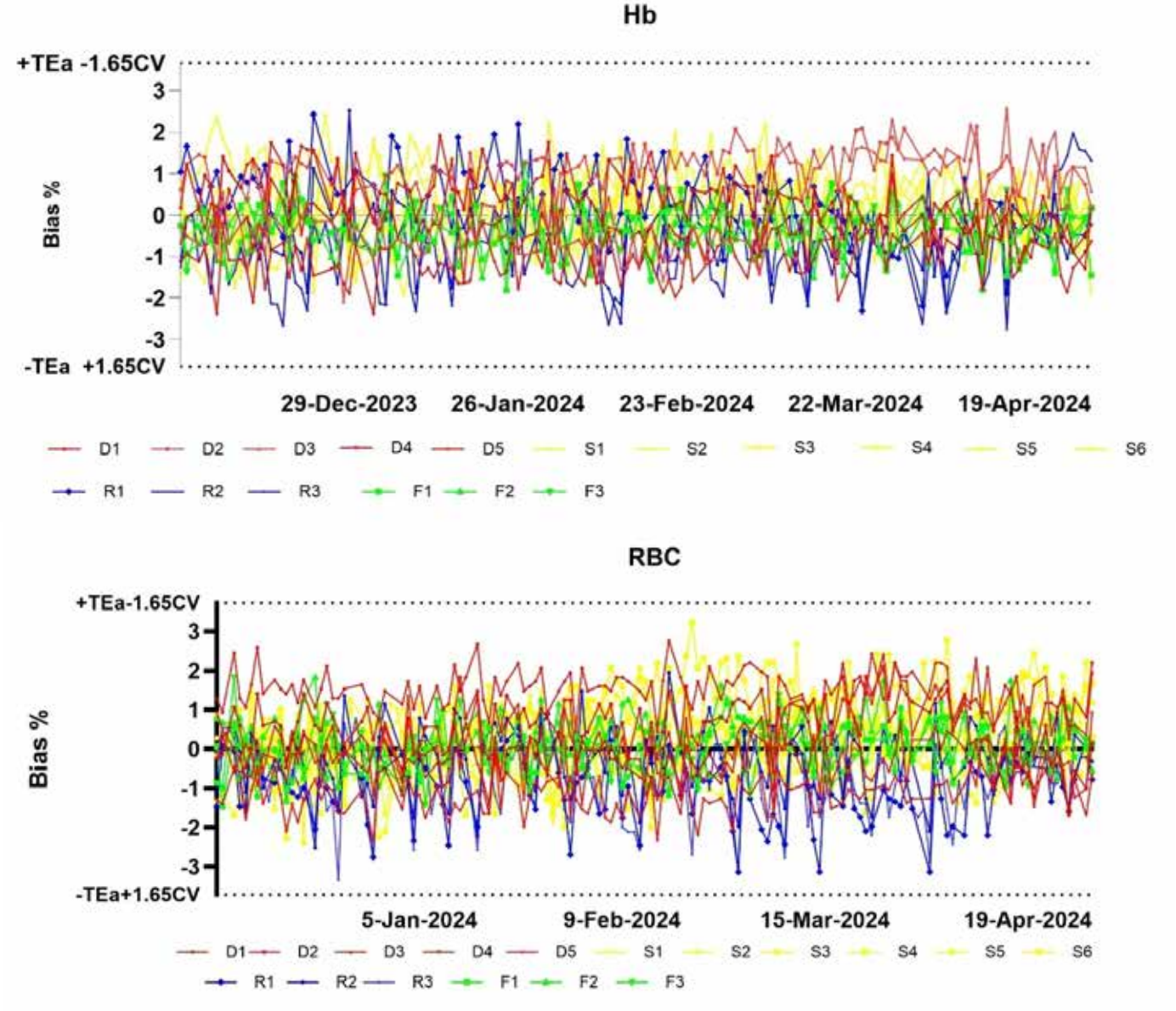


Figure 6 shows that there is no daily bias over the period studied.

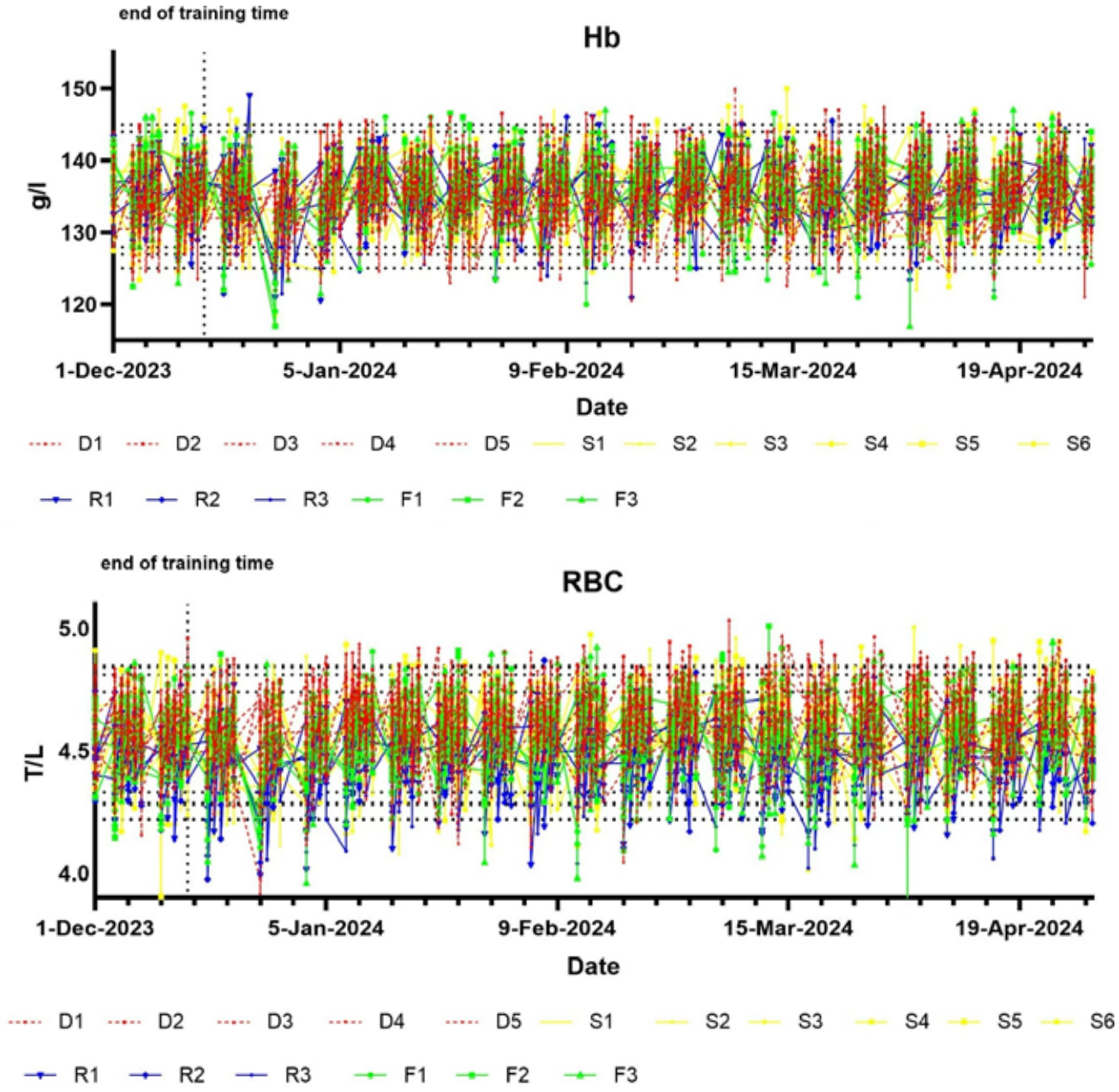
Imprecision

The percentile CV closest to the pooled CVs is consistently $CV_{intra}90$. These two CVs are not significantly different [19]. This is not the case, however, when considering $CV_{intra}50$ and $CV_{intra}95$ (see Supplementary Table 1). Therefore, it was set for the entire lifetime of the analyzer, unless the manufacturer informed us of a change in IQC manufacturing or packaging, or a global shift in the analyzer bias. In which case, these CVs should be revised.

Retrospective analysis of patient outcomes from moving median

We utilised the moving median of patient data to demonstrate that the model is effective in practice (Figure 7).

Figure 7: Evolution of patient medians for all analyzers with a block size of 50 and Winsorization from usual values (11.5 to 16.7 g/dl for Hb and 3.8 to 6 T/L for RBC).



The four analytical sites are identified by color. From December 1st to 15th, there was a training period. The black dotted lines represent the mean ± 2 SDs for each analysis site at the end of this test period.

There does not appear to be any seasonal effect over the analyzed period. All analyzers displayed a negative spike during the Christmas vacation, when most patients come from emergency departments or present serious pathologies that require frequent monitoring and explanation of lower hemoglobin levels. The revision of the IQC curves that day confirmed this hypothesis, making it possible to exclude a technical problem.
Supplementary Figure 1 demonstrates a simulation of the impact of + 1 SD bias on the seventeen analyzers and the error detection rate. For analyzers perfectly centered on the group mean before the introduced bias, the error will be detected at the first IQC point. The error will be immediately apparent for analyzers slightly above the group mean. However, not surprisingly, for analyzers running below the group mean (e.g., D4 and F2), the added bias cancels the negative bias, and there is no IQC alarm.
Supplementary data Figure 2 shows Violin plots for the patient medians of all analyzers with a block size of 50 and Winsorization outside reference intervals (11.5 to 16.7 g/dl for Hb and 3.8 to 6 T/L for RBC). The four analytical sites are identified by color. From December 1st to 15th, there was a training period. The colored dotted lines represent the median ± 2 SD.

Error detection
Once the $CV_{intra}90$ was determined, the alarm rates ±2 SD and rejection rates ±3 SD were calculated from the $SD_{intra}90$ for each analyzer. The rates of ± 2 $SD_{intra}90$ for hemoglobin are 4.2% (121/2894 QC points), and those of 3 $SD_{intra}90$ are 0.3% (8/2894 QC points). For RBC, the rates of ± 2$SD_{intra}90$ are 5.5% (178/2951 QC points), and those of 3 $SD_{intra}90$ are 0.5% (15/2951 QC points). These fall within the expected error detection range. Recall that the analyzers are centered on the mean of the peer group and not on their own mean. Thus, an analyzer exhibiting, for example, a slight downward shift, will trigger more -2s or -3s alarms than a perfectly centered analyzer with the same random error. It is this characteristic that makes our strategy effective for more quickly detecting the introduction of bias on one of the analyzers our instrument compared to the true value but also compared to other mirror analyzers (peer group average). See Supplementary Data Figure 3 for an example on an individual analyzer IQC.

Discussion
This paper aims to revisit IQC target selection in hematology, where a network of instruments is involved. An IQC strategy CLSI C24-Ed4 defines an SQC strategy as the "number of IQC materials to measure, the number of IQC results and the IQC rule to use at each IQC event, and the frequency of IQC events" [20]. Critical to the success of any IQC rule is the accuracy of the IQC target values of mean and acceptable range (IQC limits) [21].

Traditionally, many laboratories have determined these IQC targets using manufacturer-provided values on the IQC insert or by modifying these values [8]. This approach overlooks the fact that many laboratories operate multiple instruments to measure the same test. Each instrument has some inherent variation, and using different IQCs for every instrument is less optimal for managing patient risk. The optimal strategy in this situation is to use a single reference mean across all instruments, as shown in Figures 1 and 2.

As mentioned, setting accurate IQC targets is crucial for running an effective IQC strategy. The IQC targets (mean and CV) for the model were obtained from the manufacturer's website. These are dynamic and robust targets because of the contemporary instrument peer group sample size. The novel aspects of this investigation include using common IQC targets for all analyzers and utilising the $SD_{intra}90$ as the dispersion limit (Figure 5). This model has been applied to routine clinical chemistry networks [22].

The significant benefits of centering the analyzers on a consensus average from a robust peer group and common SD are:

- to detect and immediately correct a systematic error as soon as a new batch of IQC is set up (without waiting for a possible poor EQA result or a monthly retrospective outsourced IQC report with a poor SDI).
- to guarantee the comparability of intra- and inter-site analytical results via IQC data (without requiring exchanges of patient samples between technical sites.
- harmonize practices in terms of acceptance and rejection across the network.

In fact, given that the group average is the reference and the group CV is known, this is almost equivalent to performing a form of internal group quality assessment three times a day. This is why the daily bias was compared against the TEa, the dispersion due to random error, as a limit. However, some Westgard rules cannot be used in these circumstances (41S and 10x).

The $SD_{intra}90$ ($CV_{intra}90$) is a practical control limit. The total SD, through these two components, represents the overall dispersion of a method (mean dispersion and variance dispersion). For the network analyzers, it accurately reflects the overall performance of these instruments. The $SD_{intra}90$ is the maximum all-instrument imprecision achieved when all instruments have long-term stable performance, that is, no detected errors.

For any IQC strategy, it is essential not to neglect the critical importance of the run size, the number of IQC samples measured at each IQC event, and the placement of the IQC

samples [23–26]. The decrease observed in MaxE(NUF), or the Maximum number of patients affected in an undetected out-of-control condition, for a given control rule, becomes more significant as the number of control samples analyzed in each IQC event increases. This enables reaching the same target MaxE(NUF) (i.e., the same risk level for the patient) while analyzing a larger number of patient samples between IQC events, as the value of MaxE(NUF) obtained during an out-of-control condition is directly proportional to the run size [8].

The IQC rules used were standard. An assessment of intra-instrument bias was conducted before the model was implemented, and the $CV_{intra}90$ was used as the acceptable variation about the mean as this was closest to the pooled CV.

Validating the success of an IQC rule is difficult, but this was undertaken by assessing any significant statistical differences in daily IQC means and CVs for the different instruments. Patient moving medians were also evaluated for bias between instruments, and none was detected. The laboratories are enrolled in the national EQA scheme, and during the investigation period, no nonconforming results were found in the seven EQA cycles.

The model presented in this investigation provides a mechanism to assign the most statistically appropriate targets for the mean and allowable IQC value ranges to achieve the goal stated above. The model was validated prospectively.

Conclusion

Quality Control strategies must accommodate technological changes, even though the fundamental principle of reducing introduced variation remains unchanged. The increased number of samples processed in routine hematology laboratories has led to the use of multiple instruments to handle the workload, which needs a common IQC strategy to control for intra- and inter-instrument variation. It has added to the problem of ensuring comparability between measurement systems and managing bias between individual instruments. The goal of any IQC strategy needs to change to reduce patient risk, ensuring that patients obtain the same results regardless of the measuring instrument used in a laboratory or a network of laboratories [10]. The model presented focuses on identifying the bias between different instruments which can reduce variability between patient results.

Quality Control strategies continue to evolve in response to changing laboratory requirements. Most notably, consistent results are required when a patient sample is tested on different analyzers over time.

Acknowledgement

The authors wish to thank Mattheus Hofman for his help in obtaining the data.

Figure 4 was used with permission from Sysmex.

CRediT Author Statement

Conceptualization, Methodology: JMG, MB, TB.
Software: LT, MB
Validation, Analysis: LT, MB
Investigation: JMG, MB
Writing (Original draft, review and editing): TB, MB, JMG
Visualization: MB
Supervision, Project Admin: JMG, TB.

Declaration of Conflict of Interest

No conflicts of interest to declare.

Authors Disclosures

No disclosures to declare.

Funding

No funding was obtained for this work.

Ethical Approval for Research

Not applicable for this research.

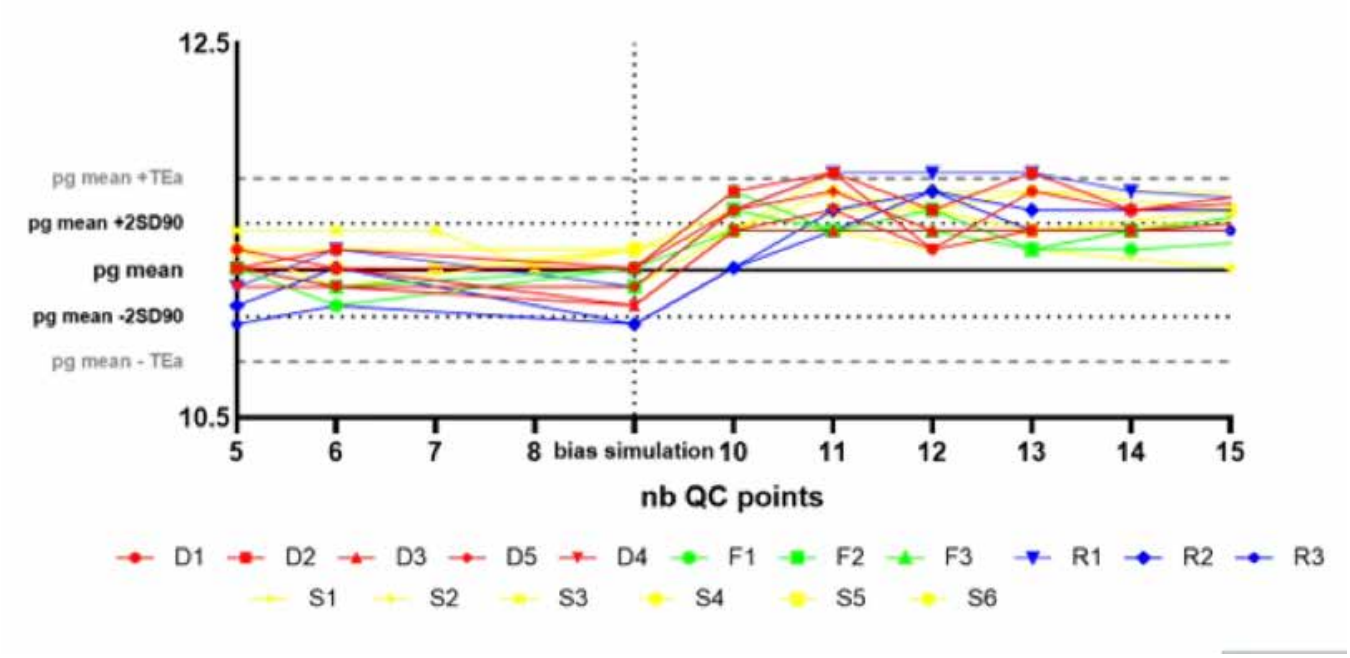
References

1. G.S. Cembrowski, Thoughts on quality-control systems: a laboratorian's perspective, Clin. Chem. 43 (1997) 886–892.
2. H. Kinns, S. Pitkin, D. Housley, D.B. Freedman, Internal quality control: Best practice, J. Clin. Pathol. 66 (2013) 1027–1032. https://doi.org/10.1136/jclinpath-2013-201661.
3. J. Westgard, T. Groth, Power Functions for Statistical Control Rules, Clin. Chem. 25 (1979) 863–869. https://doi.org/10.1093/clinchem/25.6.863.
4. S.M. Lewis, Quality control in haematology, Proc. R. Soc. Med. 68 (1975) 615–619.
5. H.T. Michael, M.B. Nabity, C.G. Couto, A. Moritz, J.W. Harvey, D.B. DeNicola, J.M. Hammond, Improving quality control for in-clinic hematology analyzers: Common myths and opportunities, Vet. Clin. Pathol. 51 (2022) 302–310. https://doi.org/10.1111/vcp.13154.
6. R. McCafferty, G. Cembrowski, B. de la Salle, M. Peng, E. Urrechaga, ICSH review of internal quality control policy for blood cell counters, Int. J. Lab. Hematol. 46 (2024) 216–226. https://doi.org/10.1111/ijlh.14220.
7. G.S. Cembrowski, G. Clarke, Quality Control of Automated Cell Counters, Clin. Lab. Med. 35 (2015) 59–71. https://doi.org/10.1016/j.cll.2014.10.006.
8. R. McCafferty, G. Cembrowski, B. de la Salle, M. Peng, E. Urrechaga, ICSH guidance for internal quality control policy for blood cell counters, Int. J. Lab. Hematol. 46 (2024) 227–233. https://doi.org/10.1111/ijlh.14212.
9. R. Pang, ed., A Practical Guide to Internal Quality Control (IQC) for Quantitative Tests in Medical Laboratories

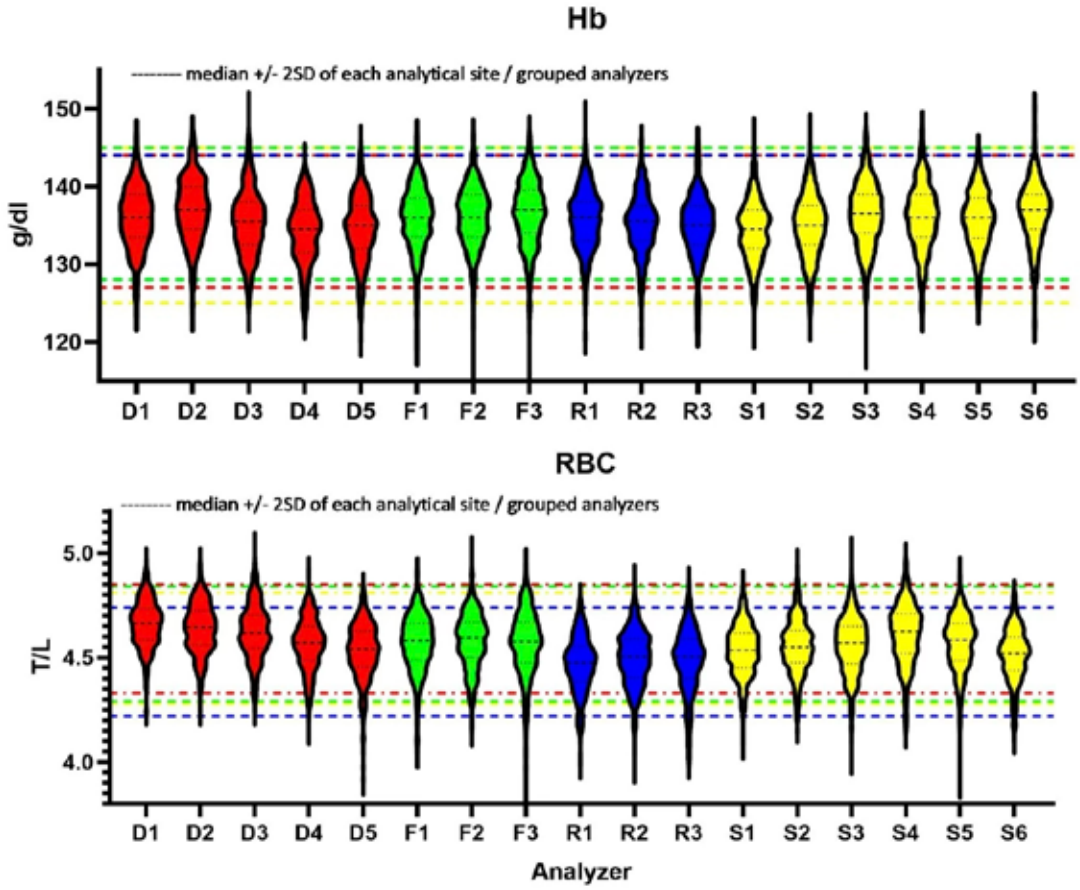
(Proposed Guidelines), Hong Kong Association of Medical Laboratories Ltd. , Hong Kong, 2010.
10. C. Parvin, L. Kuchipudi, J. Yundt-Pacheco, Designing QC rules in the presence of laboratory bias: should a QC rule be centered on the instrument's mean or the reference mean?, (2019). https://www.researchgate.net/publication/233854769_Designing_QC_Rules_in_the_Presence_of_Laboratory_Bias_Should_a_QC_Rule_be_Centered_on_the_Instrument's_Mean_or_the_Reference_Mean (accessed September 1, 2025).
11. M. Yago, C. Pla, Reference-mean-centered statistical quality control, Clin. Chem. Lab. Med. 58 (2020) 1517–1523. https://doi.org/10.1515/cclm-2019-1034.
12. H26-A2, Validation, Verification, and Quality Assurance of Automated Hematology Analyzers; Approved Standard, 2nd ed., Clinical Laboratory Standards Institute, Wayne, PA, 2010.
13. A. Algeciras-Schimnich, D.E. Bruns, J.C. Boyd, S.C. Bryant, K.A. La Fortune, S.K.G. Grebe, Failure of current laboratory protocols to detect lot-to-lot reagent differences: Findings and possible solutions, Clin. Chem. 59 (2013) 1187–1194. https://doi.org/10.1373/clinchem.2013.205070.
14. EFLM, EFLM Biological Variation Database, (n.d.). https://biologicalvariation.eu/ (accessed September 1, 2025).
15. S. Sandberg, A. Carobene, A.K. Aarsand, Biological variation - eight years after the 1st Strategic Conference of EFLM, Clin. Chem. Lab. Med. 60 (2022) 465–468. https://doi.org/10.1515/cclm-2022-0086.
16. T. Badrick, Biological variation: Understanding why it is so important?, Pract. Lab. Med. 23 (2021) e00199.
17. G.S. Cembrowski, J.O. Westgard, Quality control of multichannel hematology analyzers: Evaluation of Bull's algorithm, Am. J. Clin. Pathol. 83 (1985) 337–345. https://doi.org/10.1093/ajcp/83.3.337.
18. T. Badrick, A. Bietenbeck, A. Katayev, H.H. Van Rossum, M.A. Cervinski, T.P. Loh, Patient-Based Real Time QC, Clin. Chem. 66 (2020) 1140–1145. https://doi.org/10.1093/clinchem/hvaa149/5880074.
19. J.M. Giannoli, S. Albarede, T. Avellan, J.P. Bouilloux, R. Cartier, R. Cohen, N. Colard, L. Essemilaire, J.L. Galinier, M. Kuentz, M. Paris, H. Portugal, F. Scherrer, J.P. Siest, A. Vassault, J.M. Vialle, Recommendations for the application and follow-up of quality controls in medical laboratories, Biochem. Med. (Zagreb). 31 (2021) 020501. https://doi.org/10.11613/BM.2021.020501.
20. CLSI, Statistical quality control for quantitative measurement procedures: principles and definitions. Vol. CLSI Guideline C24. 4th ed, Wayne, 2016.
21. T. Badrick, C. Parvin, Letter to the Editor on article Dimech W, Karakaltsas M, Vincini G. Comparison of four methods of establishing control limits for monitoring quality controls in infectious disease serology testing, Clin. Chem. Lab. Med. 56 (2018) 1970–1978. https://doi.org/10.1515/cclm-2018-1276.
22. J.M. Giannoli, M. Bernard, J. L'Hirondel, A. Heim, T. Badrick, A model for managing quality control for a network of clinical chemistry instruments measuring the same analyte, Clin. Chem. Lab. Med. 62 (2023) 853–860. https://doi.org/10.1515/cclm-2023-0965.
23. C.A. Parvin, Assessing the impact of the frequency of quality control testing on the quality of reported patient results, Clin. Chem. 54 (2008) 2049–2054. https://doi.org/10.1373/clinchem.2008.113639.
24. C.A. Parvin, A.M. Gronowski, Effect of analytical run length on quality-control (QC) performance and the QC planning process, Clin. Chem. 43 (1997) 2149–2154.
25. C.A. Parvin, J. Yundt-Pacheco, Six QC recommendations to consider today, n.d. www.mlo-online.com.
26. M. Yago, Selecting statistical quality control procedures for limiting the impact of increases in analytical random error on patient safety, Clin. Chem. 63 (2017) 1022–1030. https://doi.org/10.1373/clinchem.2016.267682.

Supplementary Data

Supplementary Figure 1: Simulation of the impact of +1CVi bias on the seventeen analyzers and the error detection rate.

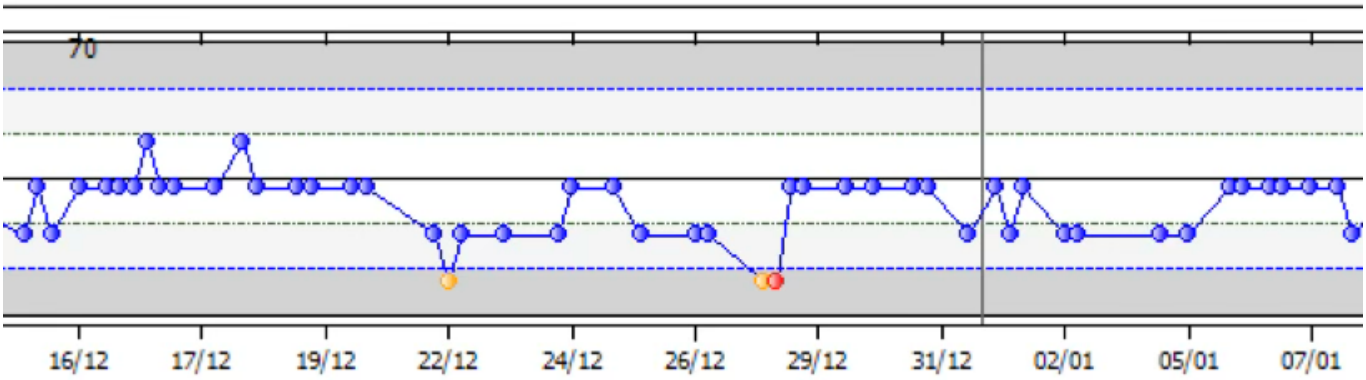


Supplementary Figure 2: Violin plots for the patient medians of all analyzers with a block size of 50 and Winsorization outside the reference interval (11.5 to 16.7 g/dl for Hb and 3.8 to 6 T/L for RBC).



The four analytical sites are identified by color. From December 1 to 15th, there was a training period. The colored dotted lines represent the median +/- 2 SD for each analysis site at the end of this test period.

Supplementary Figure 3. Example of IQC chart for one representative analyzer showing the impact of centering the IQC targets on the peer group data.



The analyzer exhibits a subtle systematic negative error that cannot be corrected with calibration, generating more -2S alarms.

Supplementary Table 1: Comparison of Pooled CV, CV_{50}, CV_{90} and CV_{95}.

Comparison of percentiled CV_{intra} with pooled CVs					
Parameter/level/Batch	**CV_{intra}90**	**pooled CV**	**CV_{intra}50 / pooled CV**	**CV_{intra}90/ pooled CV**	**CV_{intra}95/ pooled CV**
HB L2 3339	1,1	1,07	0,71	1,03	1,26
RBC L2 3339	1,2	1,13	0,78	1,06	1,22
HB L2 4023	1,1	1,03	0,74	1,07	1,31
RBC L2 4023	1,2	1,16	0,76	1,03	1,19