

Opinion Paper

Hikmet Can Çubukçu*, Marc Thelen, Mario Plebani and Mauro Panteghini

IFCC recommendations for internal quality control practice: a missed opportunity

<https://doi.org/10.1515/cclm-2025-0486>

Received April 22, 2025; accepted April 23, 2025;

published online April 28, 2025

Abstract: The IFCC Task Force on Global Lab Quality (TF-GLQ) has recently released a new guidance for internal quality control (IQC) practice through an approach translating the general principles as stated in the ISO 15189:2022 standard to a series of practical recommendations. The paper contains however important inaccuracies and shortcomings that, in our opinion, make it a missed opportunity for providing a updated guidance for laboratory professionals. In particular, four important issues are discussed: a) how to design IQC strategies in the traceability era, b) how to define IQC acceptance limits, c) how to estimate measurement uncertainty using IQC data, and d) how to manage comparability between the results provided by different analyzers in the same laboratory. Our analysis underscores the necessity for a more systematic, updated, and evidence-based approach to produce an IQC recommendation in line with the IFCC tradition.

Keywords: internal quality control (IQC); acceptance limits; metrological traceability; measurement uncertainty; ISO 15189:2022; result comparability

Introduction

Historically, the internal quality control (IQC) has been one of the main topics of the medical laboratory practice to which

the International Federation of Clinical Chemistry and Laboratory Medicine (IFCC) paid attention. Considering the need of new and updated recommendations reflecting contemporary practices, the IFCC Task Force on Global Lab Quality (TF-GLQ) has recently released a new document for IQC practice through an approach translating the general principles as stated in the ISO 15189:2022 standard to a series of practical recommendations [1, 2]. In our opinion, the paper contains however important inaccuracies and shortcomings that make it a missed opportunity for providing a clear and updated guidance for laboratory professionals. Here we address these flaws to try to overcome their potential misleading effects. In particular, we would like to discuss four important issues: a) how to design IQC strategies in the traceability era, b) how to define IQC acceptance limits, c) how to estimate measurement uncertainty (MU) using IQC data, and d) how to manage comparability between the results provided by different analyzers in the same laboratory.

IQC designing

The TF-GLQ guidance addresses IQC largely from the standpoint of traditional statistical control and pays scant attention to other approaches driven by metrological traceability, patient harm, and measurement system error rate. The fact that also ISO 15189:2022 lacks attention for a potential contribution of IQC to verify metrological traceability and that such application of IQC requires qualities of IQC materials which are often not yet available does not justify to ignore this well documented, important and theoretically superior use of IQC. Laboratory professionals rely to be educated on such concepts by an IFCC guideline.

Classic IQC does not verify that results remain traceable to reference standards, even when metrological traceability is correctly implemented by the IVD manufacturer. In the last 15 years, a number of proposals on how to rethink IQC in the metrological traceability era have been made in order to complement the transition towards more appropriate monitoring of local performance of an *in vitro* diagnostic

*Corresponding author: **Hikmet Can Çubukçu**, Rare Diseases Department, General Directorate of Health Services, Turkish Ministry of Health, Ankara, Türkiye, E-mail: hikmetcancubukcu@gmail.com. <https://orcid.org/0000-0001-5321-9354>

Marc Thelen, SKML, Nijmegen, The Netherlands; and Department of Laboratory Medicine, Radboud University Medical Center, Nijmegen, The Netherlands. <https://orcid.org/0000-0003-1771-669X>

Mario Plebani, University of Padova, Padova, Italy. <https://orcid.org/0000-0002-0270-1711>

Mauro Panteghini, Department of Laboratory Medicine, Ludwik Rydygier Collegium Medicum in Bydgoszcz, Nicolaus Copernicus University in Torun, Torun, Poland. <https://orcid.org/0000-0002-7147-3433>

medical device (IVD-MD) (reviewed in [3, 4]). Unfortunately, none of these proposed approaches are mentioned in the discussed IFCC recommendation. In particular, a proposal was elaborated suggesting that, to obtain information about traceability and its correct implementation (including the suitability of MU on clinical samples), the IQC used to monitor the analytical performance of IVD-MDs in individual laboratories should be properly reorganized into two independent components, one devoted to checking the alignment of the IVD-MD and, indirectly, to verifying the consistency of manufacturer's declared traceability during routine operations (IQC component I), and the latter structured for estimating MU due to random effects (IQC component II) [5, 6]. Avoiding to consider separately the two IQC components may introduce some significant confusion in what is recommended. For example, in the section on selection of IQC materials of the IFCC paper, commutability of these materials is generally recommended. However, while this characteristic is mandatory for IQC materials used to assess IVD-MD variability including drifts and lot-to-lot variations [7], the control materials provided by the IVD manufacturer as an integral part of the IVD-MD does not require this attribute as they are only measuring system dedicated, with the possibility of correction of bias due to matrix-related effect resulting from the interaction of reagents and “matrix-modified” material.

Furthermore, traditional risk-based IQC planning does not directly address the clinical impact of errors related to laboratory tests, although the 2022 revision of ISO 15189 standard has explicitly added to the requirements that IQC planning should involve both the historical performance, representing the probability of failure, and patient harm, representing the severity of the impact of potential failure, regardless its probability [2]. To address this shortcoming, an approach, builds upon former risk-based IQC planning [8], considering the severity of harm category of the measurand in determination of maximum run size (number of patient sample runs between two IQC events) has been proposed [9]. Also, the Clinical and Laboratory Standards Institute guideline EP23, while does not directly assist in determining the maximum run size and selecting proper QC rules, helps laboratory professionals to determine if their IQC strategy has managed risk or not [10].

Finally, the TF-GLQ presented patient result-based real time quality control (PBRTQC) as an alternative in cases where IQC is unavailable. However, PBRTQC can only serve as an extra risk reducing approach alongside IQC and not as a direct replacement for IQC, as even stressed by the propagators of PBRTQC [11, 12]. Furthermore, it is not applicable when the volume of test results for a specific analyte is too low to achieve the sensitivity of error detection at the level of analytical performance specifications (APS), which is

particularly the case for measurands with a high ratio between between-subject variation and APS [12].

Definition of IQC acceptance limits

When introducing an IQC material, the TF-GLQ recommends to calculate the control limits using the means and SD from laboratory results or, alternatively, apply a SD defined by the laboratory based on its experience [1]. After which, to apply IQC validation rules based on statistical principles, e.g., the Westgard approach based on a combination of multiple interpretative statistical rules based on different multiples of SD from the IQC mean. However, the statistical dispersion of data obtained by the laboratory (e.g., ± 3 SD of the mean value) has no relationship with clinically suitable APS. As we are dealing with medical laboratory measurements, simple statistical criteria are not enough [13]. Rather, measurement variability should fall within limits based on medical relevance so that results are reliable for clinical decision-making and patient management. Statistically derived tolerance limits can be to the utmost used after they have been demonstrated to not exceed the set APS during the IVD-MD validation or verification process preceding its introduction in the daily practice. Moreover, in the statistical control rules section, the TF-GLQ appears to advocate questionable criteria as a probability of error detection (PED) higher than 90%. While a $PED \geq 90\%$ is primarily applicable to procedures with high sigma-metric values, for procedures with lower sigma metrics (~ 3), achieving a PED of 90% becomes more challenging, necessitating adjustments in the maximum run size to attain an acceptable maximum expected unreliable final patient results [14]. Statistical quality control concepts free from PED criteria are available in the literature [8–10]. However, there is no international recommendation or consensus on selecting minimum rejection rules, contrary to TF-GLQ claims in the same section.

For the truth, the TF-GLQ also mentions clinical acceptable limits based on “established quality standards”, described as follows: a) the Milan consensus objectives, i.e., the total allowable error (TE_A) derived from desirable biological variation estimates, and b) EQA acceptance limits, designed to reflect the impact of deviations on medical decision-making. Unfortunately, both “standards” lack of evidence. Firstly, the Milan consensus statement in itself did not address the problem of TE_A as it concentrated on models to set APS [15]. On the other hand, biological variation can be a good and usable APS model for many measurands but not for all [16]. The employed reference paper is also wrongly quoted as it was conversely not supporting the TE_A concept by mentioning criticisms to the common model employed to

derive limits for it, using a mathematically incorrect method relying on the sum of mutual exclusive terms [17]. About the definition of EQA acceptance limits, Jones et al. surveyed 10 national EQA organizers about the type of APS used and found that in only 10 % of cases they derived APS from the clinical impact of the results [18].

In the section on evaluation of robustness of the method, the TF-GLQ mentions the six-sigma methodology, which quantifies analytical quality using the formula: $\text{Sigma} = (\text{TE}_A - \text{bias})/\text{CV}$. As already said above, the use of TE_A has been however criticized [19, 20]. A more general and correct definition about the sigma metric is the ratio of tolerance limits to process variability, but this appears not considered in the IFCC recommendations [21]. Even if the sigma metric approach has a potential for becoming one of the parameters for the management of risk-based IQC practice, ambiguities regarding its calculation methods should be eliminated and an international expert consensus is necessary [22].

In the section on key objectives of IQC quality indicators (QI), the TF-GLQ mentions the utility of regular monitoring of Z-score and SD index in evaluating a noticeable bias through “an external comparison of IQC results”. However, how to practically implement this monitoring remains obscure. Again, contrary to TF-GLQ claims, the number of IQC tests performed within a defined time frame is not a QI. Instead, the appropriate frequency should be determined and carried out that can change from one analyte to another. As defined by the IFCC Working Group on Laboratory Errors and Patient Safety, the relevant QI is the number of IQC results outside defined limits/total number of IQC results [23].

MU estimate

How to determine MU has been the subject of many papers for at least the last quarter-century (summarized in [4]). In considering this key performance indicator [24], first the IFCC guidance is unduly brief, consisting of approximately one column of text. More importantly, the TF-GLQ confuses TE_A , which is a type of APS, with MU, which is a parameter characterizing the dispersion of the quantity values being attributed to a measurand [25]. How to deal with bias is also unclearly described and no mention is made on different sources of systematic error in laboratory measurements and their management [26, 27]. Furthermore, the described estimation of MU is incomplete because the MU of the higher-order reference and IVD calibrator, which must be supplied by the manufacturer, are not mentioned as important sources of MU on clinical samples [24, 28, 29]. The statement referring to the possibility that EQA programs may provide an estimate of MU “through an annual review

of results with a comparison of all participating laboratories” is also difficult to accept.

Finally, the last sentence in the MU paragraph reporting that “the choice of performance requirements [for MU] is difficult: the total error is not rigorously statistically comparable and there is little other recent data in the literature” denotes a poor knowledge of the available literature in the field and the lack of systematic research of relevant articles, many of them published in this journal. Overall, the contribution of the EFLM groups is also disregarded. For instance, in the section on new reagent batch validation, the utility of IQC for determining allowable between reagent lot variation as described in an EFLM paper is ignored [30]. In addressing the management of patient results according to analytical deviations/re-analysis, the EFLM contribution on result release process and how to handle erroneous patient results according to ISO 15189 could be also useful [31].

Comparability of between-analyzer results

Following the ISO 15189:2022 structure, one of the last paragraphs of the TF-GLQ guidance discusses the comparability between the results provided by different analyzers used within the same laboratory, which is necessary for equivalent interpretation of results. There the guidance sums situations where comparability is needed with materials and procedures to check comparability, but lacks to give guidance on how to achieve such comparability at an acceptable level and how to act if such comparability is not achieved. With regard to comparability, two main issues need to be addressed: comparability between multiple identical analyzers, and comparability between different measurement procedures for the same measurand.

When the metrologically-based approach to IQC is used, it is implicit that the laboratory uses identical target values and identical tolerances for IQC materials for identical instruments. However, identical target values and tolerances should also be applied if statistically-based target values and tolerances are used. If patient samples are randomly assigned to different analyzers and patient results are combined in one result repository, regardless the instrument used, then the analyzer collective rather than the individual instruments should be considered ‘the analyzer’. This is also known as the ‘virtual analyzer’ concept [32, 33].

With regard to comparability of different IVD-MDs for the same measurand performed on different measuring systems, medical laboratories should verify the same metrological traceability of methods, which is the premise

for obtaining equivalent results. In case this is not verified, ISO 15189:2022 requires laboratories to let know to their requestors the lack of comparability between results of different analyzers [2]. The use of an IQC material with value assignment suitable for the verification of metrological traceability will aid in the identification of such situations [5]. It should not be ignored that even in case when the same metrological traceability is reported by the IVD manufacturers, the performance can substantially differ between IVD-MDs, making the combination of the results of such methods potentially misleading. When different IVD-MDs yield clinically non-equivalent results for the same measurand, those results shall be reported in separate, clearly labelled groups and not merged into a single continuum, so that users can recognise the lack of comparability [2].

In conclusion, although the IFCC TF-GLQ guidance on IQC in the light of the recent revision of ISO 15189 standard uses a helpful risk-based approach, it contains significant inaccuracies, outdated concepts, and oversights across various important aspects of IQC, potentially misleading the laboratory practice. Our analysis underscores the necessity for a more systematic, updated, and evidence-based approach to produce an IQC recommendation in line with the IFCC tradition [34].

Research ethics: Not applicable.

Informed consent: Not applicable.

Author contributions: All authors have accepted responsibility for the entire content of this manuscript and approved its submission.

Use of Large Language Models, AI and Machine Learning Tools: None declared.

Conflict of interest: The authors state no conflict of interest.

Research funding: None declared.

Data availability: Not applicable.

References

- Giannoli JM, Vassault A, Carobene A, Liaudet AP, Blasutig IM, Dabla PK, et al. Ensuring internal quality control practices in medical Laboratories: IFCC recommendations for practical applications based on ISO 15189:2022. *Clin Chim Acta* 2025;571:120240.
- ISO 15189:2022. Medical laboratories – requirements for quality and competence. Geneva: International Organization for Standardization (ISO); 2022.
- Panteghini M. Redesigning the surveillance of in vitro diagnostic medical devices and of medical laboratory performance by quality control in the traceability era. *Clin Chem Lab Med* 2023;61:759–68.
- Panteghini M, Krintus M. Establishing, evaluating and monitoring analytical quality in the traceability era. *Crit Rev Clin Lab Sci* 2025;1:1–34.
- Braga F, Pasqualetti S, Aloisio E, Panteghini M. The internal quality control in the traceability era. *Clin Chem Lab Med* 2021;59:291–300.
- Plebani M, Gillery P, Greaves RF, Lackner KJ, Lippi G, Melichar B, et al. Rethinking internal quality control: the time is now. *Clin Chem Lab Med* 2022;60:1316–7.
- Braga F, Panteghini M. Commutability of reference and control materials: an essential factor for assuring the quality of measurements in laboratory medicine. *Clin Chem Lab Med* 2019;57:967–73.
- Parvin CA. Assessing the impact of the frequency of quality control testing on the quality of reported patient results. *Clin Chem* 2008;54:2049–54.
- Çubukçu HC. QC Constellation: a cutting-edge solution for risk and patient-based quality control in clinical laboratories. *Clin Chem Lab Med* 2024;62:2185–97.
- CLSI. Laboratory quality control based on risk management. CLSI guideline EP23, 2nd ed. PA: Clinical and Laboratory Standards Institute; 2023.
- Panteghini M. Evaluating and monitoring analytical quality by internal quality control. *Clin Biochem* 2023;118:110594.
- van Rossum HH, Bietenbeck A, Cervinski MA, Katayev A, Loh TP, Badrick TC. Benefits, limitations, and controversies on patient-based real-time quality control (PBRTQC) and the evidence behind the practice. *Clin Chem Lab Med* 2021;59:1213–20.
- Panteghini M. Reply to Westgard et al.: ‘keep your eyes wide...as the present now will later be past’. *Clin Chem Lab Med* 2022;60:e202–3.
- Yago M, Alcover S. Selecting statistical procedures for quality control planning based on risk management. *Clin Chem* 2016;62:959–65.
- Panteghini M, Sandberg S. Total error vs. measurement uncertainty: the match continues. *Clin Chem Lab Med* 2016;54:195–6.
- Cerioti F, Fernandez-Calle P, Klee GG, Nordin G, Sandberg S, Streichert T, et al. Criteria for assigning laboratory measurands to models for analytical performance specifications defined in the 1st EFLM Strategic Conference. *Clin Chem Lab Med* 2017;55:189–94.
- Panteghini M. What the Milan conference has taught us about analytical performance specification model definition and measurand allocation. *Clin Chem Lab Med* 2024;62:1455–61.
- Jones GRD, Albaredo S, Kessler D, MacKenzie F, Mammen J, Pedersen M, et al. EFLM Task Finish Group – analytical Performance Specifications for EQAS (TFG-APSEQA). Analytical performance specifications for external quality assessment – definitions and descriptions. *Clin Chem Lab Med* 2017;55:949–55.
- Oosterhuis WP. Gross overestimation of total allowable error based on biological variation. *Clin Chem* 2011;57:1334–6.
- Coskun A, Oosterhuis WP. Six Sigma in laboratory medicine: the unfinished symphony. *Clin Chem Lab Med* 2025;63:e6–8.
- Oosterhuis WP, Coskun A. Sigma metrics in laboratory medicine revisited: we are on the right road with the wrong map. *Biochem Med* 2018;28:020503.
- Badrick T, Theodorsson E. Six Sigma – is it time to re-evaluate its value in laboratory medicine? *Clin Chem Lab Med* 2024;62:2398–400.
- Sciacovelli L, Padoan A, Aita A, Basso D, Plebani M. Quality indicators in laboratory medicine: state-of-the-art, quality specifications and future strategies. *Clin Chem Lab Med* 2023;61:688–95.
- Braga F, Panteghini M. The utility of measurement uncertainty in medical laboratories. *Clin Chem Lab Med* 2020;58:1407–13.
- JCGM 200. International vocabulary of metrology – basic and general concepts and associated terms (VIM), 3rd ed.; 2012. Available at: https://www.bipm.org/utls/common/documents/jcgm/JCGM_200_2012.pdf.
- Panteghini M. Not all biases are created equal: how to deal with bias on laboratory measurements. *Clin Chem Lab Med* 2025;63:916–22.

27. Thelen MHM, van Schrojenstein Lantman M. When bias becomes part of imprecision: how to use analytical performance specifications to determine acceptability of lot-lot variation and other sources of possibly unacceptable bias. *Clin Chem Lab Med* 2024;62: 1505–11.
28. ISO/TS 20914. Medical laboratories – practical guidance for the estimation of measurement uncertainty, 1st ed. Geneva (Switzerland): ISO; 2019.
29. Panteghini M. The simple reproducibility of a measurement result does not equal its overall measurement uncertainty. *Clin Chem Lab Med* 2022;60:e221–2.
30. van Schrojenstein Lantman M, Çubukçu HC, Boursier G, Panteghini M, Bernabeu-Andreu FA, Milinkovic N, et al. An approach for determining allowable between reagent lot variation. *Clin Chem Lab Med* 2022;60: 681–8.
31. Can Çubukçu H, Vanstapel F, Thelen M, Bernabeu-Andreu FA, van Schrojenstein Lantman M, Brugnoni D, et al. Improving the laboratory result release process in the light of ISO 15189:2012 standard. *Clin Chim Acta* 2021;522:167–73.
32. van Rossum HH, van Schrojenstein Lantman M, Severens M, Vermeer HJ, Verboeket-van de Venne WPHG, Oosterhuis W, et al. Quality control in The Netherlands; today's practices and starting points for guidance and future research. *Clin Chem Lab Med* 2024;62:2177–84.
33. Giannoli JM, Bernard M, L'Hirondel J, Heim A, Badrick T. A model for managing quality control for a network of clinical chemistry instruments measuring the same analyte. *Clin Chem Lab Med* 2024;62: 853–60.
34. Büttner J, Borth R, Broughton PM, Bowyer RC. Approved recommendation (1983) on quality control in clinical chemistry. Part 4. Internal quality control. *J Clin Chem Clin Biochem* 1983;21:877–84.