

Behind the scenes of EQA - Characteristics, Capabilities, Benefits and Assets of External Quality Assessment (EQA)

Part II - EQA cycles

Christoph Buchta^{1*}, Rachel Marrington², Barbara De la Salle³, Stéphanie Albarède⁴, Tony Badrick⁵, Heidi Berghäll⁶, David Bullock⁷, Wim Coucke⁸, Vincent Delatour⁹, Wolf-Jochen Geilenkeuser¹⁰, Andrea Griesmacher^{1,11}, Gitte M. Henriksen¹², Jim F. Huggett^{13,14}, Jonna Pelanti⁶, Peter B. Lupp^{15,16}, Paola Pezzati¹⁷, Sverre Sandberg¹⁸, Michael Spannagel¹⁶, Marc Thelen^{19,20}, Veronica Restelli²¹, Lucy A. Perrone²¹

- 1) Austrian Association for Quality Assurance and Standardization of Medical and Diagnostic Tests (ÖQUASTA), Vienna, Austria;
- 2) Birmingham Quality (UK NEQAS), part of University Hospitals Birmingham NHS Foundation Trust, Birmingham, United Kingdom
- 3) UK National External Quality Assessment Scheme (UK NEQAS) for Haematology, Watford, United Kingdom
- 4) Centre Toulousain pour le Contrôle de qualité en Biologie clinique (CTCB), Toulouse, France
- 5) Royal College of Pathologists of Australasia Quality Assurance Program, St Leonards, Sydney, Australia
- 6) Labquality, Helsinki, Finland
- 7) former Director, Birmingham Quality, former President, UK NEQAS
- 8) Sciensano, Brussels, Belgium
- 9) Laboratoire National de Métrologie et d'Essais (LNE), Paris, France

- 10) former Reference Institute for Bioanalytics, Bonn, Germany
- 11) Central Institute of Clinical and Chemical Laboratory Diagnostics, University Hospital of Innsbruck, Innsbruck, Austria.
- 12) DEKS, Rigshospitalet-Glostrup, Denmark
- 13) National Measurement Laboratory, LGC, Teddington, United Kingdom
- 14) School of Biosciences and Medicine, Faculty of Health & Medical Sciences, University of Surrey, Guildford, United Kingdom
- 15) Institut für Klinische Chemie und Pathobiochemie, Klinikum rechts der Isar, Technische Universität München, Munich, Germany
- 16) INSTAND e.V. Society for Promoting Quality Assurance in Medical Laboratories, Düsseldorf, Germany
- 17) SOD Sicurezza e Qualità AOU Careggi Firenze, Italy
- 18) The Norwegian Organization for Quality Improvement of Laboratory Examinations (Noklus), Bergen, Norway
- 19) Foundation for quality assessment in laboratory medicine (SKML), Nijmegen, the Netherlands
- 20) Department of Laboratory Medicine, Radboud Medical Center, Nijmegen, the Netherlands
- 21) Canadian Microbiology Proficiency Testing Program (CMPT), Department of Pathology and Laboratory Medicine, University of British Columbia, Vancouver, Canada

*Corresponding author: Christoph Buchta, Austrian Association for Quality Assurance and Standardization of Medical and Diagnostic Tests (ÖQUASTA), Hörlgasse 18/5, 1090 Vienna, Austria; christoph.buchta@oequasta.at

Running head: Behind the scenes of EQA Part III

Word count: 4780

Tables: 1

Figures: 2

Key words: EQA; external quality assessment; PT; proficiency testing; interlaboratory comparison;

Statement

Points of view in this document are those of the authors and do not necessarily represent the official position of their affiliated organisations. Where certain commercial software, instruments, and materials are identified in order to specify experimental procedures as completely as possible, such identification does not imply recommendation or endorsement by the authors, nor does it imply that any of the materials, instruments, or equipment identified are necessarily the best available for the purpose.

Abstract

External quality assessment (EQA) cycles are the smallest complete units within EQA programs that laboratories can use to obtain external assessments of their performance. In each cycle, several samples are distributed to the laboratories registered for participation, and ideally,

EQA programs not only cover the examination procedures but also the pre- and post-examination procedures.

The properties and concentration range of measurands in individual samples are selected with regard to the intended challenge for the participants so that each sample fulfils its purpose. This aims to ensure the most significant possible information gain in every cycle using the lowest possible number of EQA samples and thus, under economically optimal conditions. Participants examine samples and the results are reported to the EQA provider, who compares them with the target values for individual measurands in every sample. The EQA provider assesses the laboratory performance, and finally communicates the assessment results to the participant. After completion, each cycle is evaluated by the provider so that limitations and weaknesses of the EQA program can be identified and appropriate measures taken, or to confirm its continued suitability and appropriateness.

Introduction

This is Part II of a five-part series of articles describing the principles, practices and benefits of External Quality Assessment (EQA) of the clinical laboratory. Part I describes the historical, legal and ethical backgrounds of EQA and properties of individual programs (ref). Part II deals with key properties of EQA cycles. Part III is focused on the characteristics of EQA samples (ref). Part IV summarises the benefits for participant laboratories (ref), and Part V addresses the broad benefits of EQA for stakeholders other than participants (ref).

EQA providers are important quality partners in laboratory medicine because they operate from a position of neutrality and provide an objective evaluation of laboratory performance during the total testing process (TTP) (Figure 1). EQA providers collect and store laboratory

performance data, which enables their longitudinal assessment. All laboratories enrolled in an EQA program receive material distributed at a similar time period and thus have the comparable initial conditions for analysis. Participants measure levels/concentrations of measurands or determine the properties of samples and accordingly submit quantitative, ordinal and/or nominal results to the EQA provider. Individual results are evaluated by comparison to an assigned value/ target; for details see Part II, chapter “*Determination of the target value*” (ref). Grading criteria and the statistics employed are commonly specific to each EQA provider and are informed by the subdiscipline in laboratory medicine (e.g., chemistry versus microbiology) and the specific program design. Participants receive feedback on their examination performance that they can then use to make informed decisions. EQA performance data from a large enough sample size of participants can also be used as an evaluation of a particular examination procedure’s performance in the field. EQA programs usually consist of several individual cycles per year, and the number of samples in cycles and their composition varies depending on the provider and the program’s objectives.

The EQA cycle

EQA usually runs in recurring cycles (also referred to as “rounds”, “challenges”, or “distributions”). After participants have been administered and the EQA samples have been shipped to them, they examine the measurands contained in samples or determine their properties. The EQA provider evaluates the results submitted by the participants, assesses each result, and reports the results of the evaluations back to the participants. The steps in EQA cycles and their embedding in EQA programs are shown in Figure 2.

Participant management

When a participant is enrolled in a service, there is a certain amount of 'participant maintenance', in terms of programs / measurands / examination procedures, and ensuring that contact details are correct. Enrolment of participants for the correct measurands, in-vitro diagnostic medical devices (IVD-MDs) and units is essential for successful result interpretation. Units are an important factor that must be considered when results are evaluated. Some EQA providers allow participants to return results in the participant's units, and the EQA provider will convert to a single unit for all result evaluation or the EQA provider may ask the participant to convert the units to a 'program unit' before the result submission. Different EQA providers will have various processes for handling enrolment, both in terms of time points for enrolment - specific dates versus any time - and whether participants manage all enrolment processes, online, or whether the EQA provider does it centrally.

Sample preparation

Though each EQA provider will have their own program design for their EQA service, all EQA cycles require considerable planning. This starts with the selection of suitable EQA sample material. Samples are selected to present the intended challenges to participant laboratories or examination procedures. The samples should correspond as closely as possible to patient samples and cover both physiological reference ranges and pathological ranges above or below them. The material selected needs to be in sufficient volume and stable enough to reach the participants without deterioration, and it must also be in a form that ensures homogeneity has been preserved. This is to ensure that all participants have an equal opportunity to analyse the same material to allow the possibility of comparing results with each other. Part III of this series provides a more in-depth discussion on preparation of EQA samples (ref).

Sample dispatch

Samples are dispatched to registered participants under such conditions that the retention of physical, chemical and biological properties is maintained within the specified period for analysis. Participants should be informed when the cycle starts. The time from dispatch from the EQA provider to receipt and analysis by the laboratory can be critical to the EQA sample. A reliable postal or courier service should be used to avoid delays in the distribution process since deterioration of samples in transit may be an area where inappropriate sample handling may occur. The use of temperature-regulated transport, monitoring of conditions in transit and a record of the sample analysis date can provide some assurance.

The EQA provider may choose to dispatch materials from multiple cycles simultaneously for the participant to store these until the required time of analysis or the EQA provider may dispatch samples from a single cycle at a time. The advantage of the latter is that the EQA provider has more control over storage conditions for the samples before analysis and can quickly adapt if there is a change in participation or if changes in sample characteristics or program design is required, as was often necessary during the COVID-19 pandemic due to the changing pathogen variants. It is more cost effective to send a single dispatch of samples in cases where sample stability allows for such an approach, but this does transfer the responsibility of long-term storage to the participant doing this correctly and then analysing the samples at the correct time points. This can work very successfully, but the EQA providers are responsible for informing participants about storage conditions and when the samples should be analysed. Shipments must always include clear information on the sample stability period and storage conditions, handling and preparation, analysis, mechanism for submission of results to the EQA provider, safety requirements and instructions for disposal of the material. All

transportation should comply with local and national safety requirements and may necessitate the inclusion of permits to avoid delay in receipt. The EQA provider must take transportation into consideration when making the EQA program available to certain regions or territories.

Examination and reporting

The laboratory is responsible for handling EQA samples as if they were from a patient (though pre-treatment may be required, for example, reconstitution of lyophilised material, or thawing and allowing it to reach room temperature). Some EQA providers may ask for EQA samples to be measured through replicate analysis, e.g. to estimate measurement precision. Samples will usually be provided with a 'request form' and/or 'instruction leaflet'. This must be read before analysis as it could contain important information and advice regarding EQA sample handling before analysis. EQA samples should be processed by personnel routinely performing pre-examination, examination, and post-examination procedures. In most cases, this is achievable and practicable by booking the samples into a laboratory's laboratory information management system (LIMS) and sending the samples for analysis; however, many laboratories return patients' results electronically to the requester, and this is an area which is still manual and labour intensive concerning EQA samples. Some EQA providers are offering a direct integration service which allows the electronic transfer of EQA requests and EQA results between the LIMS and the electronic platform of the EQA provider. This follows the same process as laboratory-to-laboratory communication and supports the handling of EQA samples in the same manner as patient specimens. Other mechanisms for participants to submit results include web interface, email and mail. The results must be reported within the instructed time period to be included in the assessment. Late results cannot always be included in the evaluation and certainly not after the target or assigned value has been published.

Evaluation and feedback to participants

Upon closing a cycle, the EQA provider will undertake statistical analysis and produce a report. The level of detail in a report will vary from provider to provider, but as a minimum, the laboratory should be able to easily see their performance against an assigned value or target, and compare it to acceptable limits of performance. There is no 'regulation' of what Analytical Performance Specification (APS) an EQA provider should use, but several publications may have recommendatory character as they report on APS for EQA [1]. Several steps in the EQA cycle can be outsourced, but not the statistical evaluation.

The choice of assigned value and performance specifications varies from provider to provider. The EQA provider is responsible for informing their participants how the target value was assigned and the laboratory needs to assess whether the assigned value and performance specifications they are being assessed against are beneficial to their patients' requirements [1]. See the *EQA Review* section for more information on the determination and selection of assigned values.

In most cases, an EQA provider will notify the participant when a report is ready for review. This may be an interim report or the final report. Participants will receive an individual report where their performance is assessed. In addition, there may be, where appropriate, expert comments discussing the general performance, possible sources of errors, suggestions on corrective actions and other scientific content. Depending on the respective national legislation, results of the assessments are communicated only to the participant, or to supervisory authorities.

Laboratory professionals and point-of-care test (POCT) users usually have different criteria for the required information in the reports. A clinical laboratory should have the option of detailed statistical information regarding their individual performance as compared to peers, whereas POCT sites, which are not always supervised by scientists, may benefit most from a very straightforward report where the main focus is on their individual results. More in-depth statistical information/reports should be available for a POCT coordinator. It would be beneficial for all participants if the EQA provider could produce different types of reports that reflect the requirements of the end user.

Finally, participants need to be aware of how they can give a response, or appeal, to the EQA provider about the evaluation of their EQA results if required. All information related to a participant's results and performance evaluation is confidential unless supervisory authorities require it differently according to national legislation. Participants can also waive their anonymity to obtain advice, seek help from other experts or in-vitro diagnostics (IVD) manufacturers, or contribute to regional or national comparisons for rare measurands.

Participant's review of EQA reports

The role of the laboratory is not merely to participate and analyse EQA samples, they also need to actively review and act upon their results and any feedback reported by the EQA provider.

It is the participant's responsibility to review each report, and the participant's local procedures will determine how they record within their Quality Management System (QMS) and what action is taken for poor performance. A participant may record all EQA returns within their QMS, or they may only record poor or borderline performance. The EQA provider will have performance limits/criteria for their own EQA programs and participants should adhere to these

as they are all part of the EQA program design. However, they may also have their own internal assessment triggers.

One mechanism of recording borderline/poor performance is using the Corrective Action/Preventive Action procedure, which should be established in all ISO-accredited laboratories. The depth of investigation and recording of borderline / poor performance will depend on the individual laboratory. Examples of root causes of an out-of-consensus EQA return are shown in Table 1.

Suppose the EQA issue is isolated to a single user or a small subset of users. In that case, it is more likely that there are problems with the handling of the EQA material, operator or reporting of results. Individuals may also see calibration or imprecision issues may also be seen by individuals; but it would be hoped that these would have been identified by internal quality control (IQC) procedures if robust IQC performance criteria were in place and suitable IQC material was used. In these cases, a review of what actually happened with this cycle of EQA, and a review of EQA policies and procedures, review of IQC, calibration and reagent records may elucidate the cause of an error or identify an opportunity for improvement (a reminder that the participant is in the best position to know what is happening in their laboratory).

Analytical issues that impact several users are more difficult to rectify. This could be due to calibration errors, lot to lot variation in reagent or quite more generally that the IVD-MD is not suitable for its intended use, be it in terms of selectivity, specificity or imprecision or how the laboratory uses the IVD-MD. It can also be that EQA materials have inadequate commutability for some IVD-MDs. Regulations may differ between countries on actions that should be taken (whether a laboratory continues reporting clinical results), escalation, etc. It is of utmost importance that laboratories work in a triangular relationship with EQA providers and manufacturers to address analytical concerns. The ISO 15189:2022 standard emphasises the

laboratory's responsibility for risk assessment of service provision; therefore, laboratories have a crucial role in this process [2].

Evaluation of cycles by the EQA provider

EQA is more than just an assessment of a laboratory's performance. It also collects data on the performance of examination procedures and IVD-MDs in routine use, thus supporting manufacturers to fulfil their obligation of post-market surveillance. At the close of a cycle, the program organiser / EQA provider will review the overall performance of all examination methods. Changes in assay performance and/or changes in clinical requirements may lead to the development of EQA services by the EQA provider or discussion/cooperation between the EQA provider and manufacturer. Depending on the nature and extent of the issue and local / national regulations, the EQA provider may be required to take further action. Significant changes to EQA program design may not be covered under the scope of an EQA provider's ISO 17043:2023 accreditation; their local accreditation body may require further assessment. EQA program review and development are all part of ongoing quality improvement.

Failure in EQA - What should the laboratory do?

Beyond what regulatory bodies may require after an EQA failure, the laboratory director needs to assess the risk of continuing testing, whether testing should be stopped until a root cause can be identified, or whether the failure is unlikely to affect patient health and testing can continue.

The laboratory director must also assess the need to control all patients tested since the last EQA conform result.

The first step to resolution after an EQA challenge failure is a Root Cause Analysis (RCA). A root cause analysis – as the name implies – is an investigation to uncover the ultimate cause of a failure. Once the root cause has been identified, it is essential to determine whether and how many patient test results were potentially affected by the gap in the process. These changes in results need to be listed in corrective reports [3]. Patient impact is usually more likely when a problem is systemic. Some regulatory bodies guide this root cause analysis and reporting process using an investigation response form that a laboratory has to complete for all unacceptable proficiency testing results [4]. The form is intended to guide laboratories through investigating a failure as much as it is to inform regulatory bodies of the laboratory's steps to identify the root cause, corrective actions, patient impact, etc. Accreditation bodies will review the file, communicate with the laboratory about any concerns, and store the report in the laboratory's file for the next accreditation assessment so the assessors can follow up on corrective actions [4]. Particularly challenging situations, accreditation bodies can also contact subject matter experts for advice, which can help inform the laboratory's response to the problem [4].

The laboratory may question whether the EQA results reflect its current performance on patients samples. The first check to perform when EQA results are outside acceptable performance limits is whether the observed performance relates to a clerical or data entry error that would not occur with a patient's sample, e.g., where patients' data is automatically transferred from the analyser to the laboratory information system. However, lapses in transcription may reflect pre or post-analytical errors and should be addressed by the laboratory. If the performance reflects the results of a single EQA sample, verification by repetition on the same or a repeat may be required before action is taken. Suppose apparently if unsatisfactory performance is concluded from an EQA report considering multiple samples at different levels, measured at different time points, i.e., providing a long term retrospective

assessment of performance. In that case, the laboratory should evaluate the significance of the performance before assessing the need for corrective action and the possible impact on the validity of patients' results. Firstly, the laboratory should consider if the metrological traceability of the assigned value is identical to that of the examination method that the laboratory uses, in which case the deviation of such target value is immediately meaningful for the laboratory. For that reason, an EQA with method-independent target values assigned in a reference laboratory using a reference method traceable to a reference material has a different applicability from an EQA with target values based on peer group consensus of participants. The first can be used for trueness verification, whereas the second is used to check conformity to peer group performance. Target value assignment by a reference method alone is not enough to allow trueness verification. It also requires samples to be commutable because only then between-method differences will reflect differences that would also be seen in patient samples at such concentrations. To take corrective action, the laboratory also needs to know to what extent the inaccuracy is caused by imprecision around the single estimate of the true result of an EQA sample and to what extent it is caused by bias. To gain such insight, the EQA report should ideally take into account the results of several samples and/or be based on the analysis of replicate samples. An EQA program that reports on multiple or replicates commutable samples with value assignment and metrological traceability to a reference method is considered category 1 in the EQA categorisation proposed by Miller et al. [5]. EQA programs without value assignment traceable to reference methods are category 3, and those without commutable samples are category 5. (See Part III)

When evaluating an EQA report, a laboratory also needs to consider the performance specifications used by the EQA provider and how much bias from the assigned value requires corrective action, with its uncertainty of the bias correction [6]. According to ISO 15189:2022 laboratories should define their criteria for successful EQA participation [2]. Whether a

laboratory sets different satisfactory performance criteria than those used by the EQA provider will depend on the type and rationale of the tolerance limits provided by the EQA organiser. When selecting a suitable EQA, laboratories are advised by ISO 15189:2022 to consider sample commutability, assigned value and tolerance limits in relation to the intended use of the particular EQA program. The ISO/IEC 17043:2023 [7] accreditation of an EQA program or individual measurand(s) demonstrates the competency of the EQA provider in terms of statistical analysis, information to participants, homogeneity and stability of EQA samples, which is necessary for the EQA results to reflect the laboratory's performance with patients' specimens. However, crucially, ISO/IEC 17043:2023 does not guide on EQA program design.

Once a laboratory has decided that corrective action is justified, ISO 15189:2022 requires the laboratory to investigate the potential impact of unsatisfactory performance on patients' results. The time elapsed since the last satisfactory EQA participation determines when the patients' results should be reconsidered. The laboratory should also have a process to monitor whether the corrective action has been successful from future EQA results or repeated measurement on replicate EQA samples, where these are available and remain viable after the survey is closed. This requires the EQA provider to provide an EQA cycle frequency that reflects the needs of its participants.

It is important to remember that although a failure in EQA may have some immediate consequences for the laboratory, the process to correct this failure strengthens and improves the laboratory's quality and subsequently, patient safety [6].

EQA review

Determination of the assigned value

The basis for evaluation of results is the target value (also known as assigned value) which can be determined via a number of mechanisms, each with its own merits and pitfalls. Ideally, the assigned value should be the ‘best estimate of the truth’.

There are different ways to set or determine the assigned value [8]. They can be classified into two groups depending on the source data used. One group uses external data that are not reported by the participants in the EQA cycle, for example, when known amounts of a substance of a defined purity are added to a matrix not containing this measurand, the assigned value can be calculated based on the quantity of pure substance added and the volume of the matrix. The other group uses individual results reported by the participants, either from all participants or from a subgroup, e.g. reference laboratories.

The assigned value may also be obtained using certified reference materials (CRMs), obtaining a reference value by application of the reference method, or as a consensus value when analysed by a small group of experts or reference laboratories. Reference methods are high-accuracy examination procedures capable of delivering SI-traceable results with high specificity and low measurement uncertainties. Reference methods are available for some, but by far not for all measurands. There is no point in using reference methods if the material is not commutable or the IVD-MD is not specific for the measurand. Given that reference methods are usually labour-intensive, not automated and require the involvement of very specialised staff, they cannot be used routinely in medical laboratories and hospitals due to high turnaround time. This also impacts EQA programs as the use of reference method assigned values may impact the number and frequency of samples that can be dispatched. Therefore, reference methods are almost exclusively operated in National Metrology Institutes and calibration laboratories accredited

according to ISO 17025:2018 and/or ISO 15195:2018 [9,10]. The measurement uncertainty on the assigned value plays a crucial role in evaluating laboratories. The more uncertainty of an assigned value, the higher the risk of a wrong evaluation of an EQA result. This uncertainty should be available for participants, or, as an alternative, the evaluation limits should be extended by including this uncertainty. Taking into account the uncertainty of the assigned value is advised to be done for all types of assigned values, independently from how they are obtained. A special note should be added for EQA results that are evaluated for the variability of the reported EQA results, as is the case for Z-scores. The uncertainty of the estimate of variability should be taken into account as well.

Specific requirements exist for determining the assigned value based on the reported EQA results. Quantitative EQA data are expected to be normally distributed (log-normally in the case of logarithmically scaled measurement results, e.g. virus load), with possible contamination of outliers [11]. A mean value of participant results can be used for all results, and a measurement procedure-specific or an individual IVD-MD-specific mean can be determined, giving three further options of assigned value. Reference method analysis can be used to validate these three assigned values. Once again, there are advantages and disadvantages of these methods for determining assigned values. Comparing a participant to their IVD-MD group peers will ensure that they are performing as well as other users of that IVD-MD. Although EQA materials commutability is not a prerequisite, there are limitations on the information that can be obtained from a post-market surveillance perspective on overall performance compared to other IVD-MDs on the market. To base the assigned value on all results from one manufacturer is more popular but may have significant limitations, as results of the comparator method may be biased, which could result in a flawed estimate of accuracy.

Since the proportion of outliers in EQA data is low, the breakdown point of robust estimators is a less critical parameter to distinguish between robust estimators for determining assigned

values. The most widely known robust estimator for the assigned value is the median [12]. It is the middle value when EQA data are sorted from smallest to largest. Medians have, in comparison with other estimators, a low efficiency. Other estimators are the H1.5 algorithm [13], which is equal to the estimator of algorithm A from ISO 13528:2022 [8], L1.5 estimator [13] and the MM-estimator [14]. Another class of statistical estimators for determining the assigned value consists of using the classical, not robust, mean after the exclusion of outliers [15]. Although less widely used, the principle of excluding outliers before performing statistical analysis is an approach that allows the use of more flexible statistical methods [15]. Unimodality (data clustered around *one* value) of data is another important condition for having a reliable assigned value. Before calculating any assigned value, the EQA provider should check the unimodality of the data [16].

In all cases, the EQA provider should be able to justify the use of any specific assigned value within their EQA program. When the commutability of the sample is questionable, the EQA provider is limited to the use of so-called peer group comparisons [5]. When a sample is commutable, laboratories can be evaluated without dividing the results into peer groups.

EQA also covers qualitative assays; a consensus or mode value is often used to assign the value. This could be based on a participants' consensus for that examination, expert consensus or an overall analytical consensus considering results from more than one examination. In all cases the EQA provider needs to define the minimum number of participants to meet statistical design objectives. Also, the procedures on handling outliers and result exclusion need to be described and communicated to participants.

Participant peer groups

Examination methods cannot always be compared with each other and the results need to be divided into groups that have the same chemical principles, same examination procedures or use the same IVD-MDs - participant peer groups. The grouping is always done based on the program and the performance of the examination methods. Participants can use data from peer group analysis to one the one hand be informed on their own performance using a method, and on the other hand, if commutable materials are used, be informed about performance of the IVD-MD used. Participant peer group analysis is, therefore, a useful tool in EQA reports, either with or without method-independent assigned values.

Although ISO 15189:2022 requires interlaboratory comparison, a minimum number of participants within peer groups is required to allow a statistical evaluation. A comparison of two laboratories can be useful if they are confirmatory of each other, but in case of discrepancy, at least a third participant is needed to form a majority. Depending on statistical procedures, EQA providers will have to validate and document their choice of a minimal number of results contributing to an assigned value. This could be as low as five, but in practice is usually higher. However, the fact that the results of the majority do not necessarily have to be the correct ones applies to peer groups of all sizes. Ultimately, only a comparison with a measurement result obtained with a reference method in a commutable sample material is sufficient to verify the accuracy. For more details see Part III, chapter “*Commutability*” (ref).

Qualitative analysis

Results of qualitative analyses, like immunohaematology examinations (e.g., ABO- and D-typing), some infection diagnostic examinations (e.g., positive or negative (“not detected”) for

a pathogen), and cell morphology EQA (e.g., blood smear, cytology) do not require any statistical evaluation. To assess the results in such EQA programs, they are matched with the target, which may be determined by one or several expert laboratories or by the most reported results (consensus value). For both procedures, assigned determination and evaluation criteria, including consequences of results for individual samples not matching the assigned, must be agreed upon before the start of the cycle (e.g., for “core samples”, correct results must be reported, but participants can still pass an EQA cycle if they have not reported correct results for “educative samples”).

Conclusion

EQA programs are performed in cycles where participants receive challenges (Figure 2). This article presented characteristics of EQA cycles, criteria for evaluating results and suggestions for actions when a participant was unsuccessful in a challenge.

Author contributions:

Christoph Buchta: Conceptualized this review, wrote the manuscript draft, edited and critically reviewed the manuscript; Rachel Marrington: Provided scientific advice, wrote, edited and critically reviewed the manuscript; Barbara De la Salle: Provided scientific advice, wrote, edited and critically reviewed the manuscript; Stéphanie Albarède: Provided scientific advice, wrote, edited and critically reviewed the manuscript; Tony Badrick: Provided scientific advice, wrote, edited and critically reviewed the manuscript; Heidi Berghäll: Provided scientific advice,

wrote, edited and critically reviewed the manuscript; David Bullock: Provided scientific advice, wrote, edited and critically reviewed the manuscript; Wim Coucke: Provided scientific advice, wrote, edited and critically reviewed the manuscript; Wolf-Jochen Geilenkeuser: Provided scientific advice, wrote, edited and critically reviewed the manuscript; Andrea Griesmacher: Provided scientific advice, wrote, edited and critically reviewed the manuscript; Gitte M. Henriksen: Provided scientific advice, wrote, edited and critically reviewed the manuscript; Jim F. Huggett: Provided scientific advice, wrote, edited and critically reviewed the manuscript; Peter B. Lippa: Provided scientific advice, wrote, edited and critically reviewed the manuscript; Jonna Pelanti: Provided scientific advice, wrote, edited and critically reviewed the manuscript; Paola Pezzati: Provided scientific advice, wrote, edited and critically reviewed the manuscript; Sverre Sandberg: Provided scientific advice, wrote, edited and critically reviewed the manuscript; Michael Spannagl: Provided scientific advice, wrote, edited and critically reviewed the manuscript; Marc Thelen: Provided scientific advice, wrote, edited and critically reviewed the manuscript; Annette Thomas: Provided scientific advice, wrote, edited and critically reviewed the manuscript; Veronica Restelli: Provided scientific advice, wrote, edited and critically reviewed the manuscript; Lucy A. Perrone: Provided scientific advice, wrote, edited and critically reviewed the manuscript. The authors have accepted responsibility for the entire content of this manuscript and approved its submission.

Funding sources:

None declared.

Acknowledgements

The authors wish to express their gratitude to Anna Malikovskaia for her support with the management of the extensive bibliography of this five part paper series.

Preprint not peer reviewed

References:

1. Jones GRD, Albarede S, Kessler D, MacKenzie F, Mammen J, Pedersen M, et al. Analytical performance specifications for external quality assessment – definitions and descriptions. *Clin Chem Lab Med* 2017;55:949-55
2. Medical laboratories - Requirements for quality and competence. Geneva: International Organization for Standardization (ISO); 2022.
3. Review of patient results in response to a PT failure. College of American Pathologists; Available at: https://documents-cloud.cap.org/appdocs/learning/LAP/FFoC/InspectingPT_2016/presentation_content/external_files/PT_review_guidance.pdf
4. College of Physicians and Surgeons of British Columbia. Diagnostic Accreditation Program. Laboratory Medicine Proficiency Testing. Surgeons of British Columbia; 2023.
5. Miller WG, Jones GR, Horowitz GL, Weykamp C. Proficiency Testing/External Quality Assessment: Current Challenges and Future Directions. *Clin Chem* 2011;57:1670–80.
6. Kristensen GBB, Meijer P. Interpretation of EQA results and EQA-based trouble shooting. *Biochem Med (Zagreb)* 2017;27:49–62.
7. ISO 17043:2023 Conformity assessment - General requirements for the competence of proficiency testing providers. Geneva: International Organization for Standardization (ISO); 2023.
8. ISO 13528:2022 Statistical methods for use in proficiency testing by interlaboratory comparison. Geneva: International Organization for Standardization (ISO); 2022.
9. ISO/IEC 17025:2017 General requirements for the competence of testing and calibration laboratories. Geneva: International Organization for Standardization (ISO); 2017.
10. ISO 15195:2018 Laboratory medicine - Requirements for the competence of calibration laboratories using reference measurement procedures. Geneva; 2018.
11. Duewer DL. A comparison of location estimators for interlaboratory data contaminated with value and uncertainty outliers. *Accred Qual Assur* 2008;13:193–216.
12. Sciacovelli L, Secchiero S, Zardo L, Plebani M. External Quality Assessment Schemes: need for recognised requirements. *Clin Chim Acta.* 2001;309:183–99.

13. Huber PJ. Robust Statistics: Huber/Robust Statistics [Internet]. Hoboken, NJ, USA: John Wiley & Sons, Inc.; 1981. Available from: <http://doi.wiley.com/10.1002/0471725250>
14. Ellison SLR. Performance of MM-estimators on multi-modal data shows potential for improvements in consensus value estimation. *Accred Qual Assur* 2009;14:411–9.
15. Coucke W, China B, Delattre I, Lenga Y, Van Blerk M, Van Campenhout C, et al. Comparison of different approaches to evaluate External Quality Assessment Data. *Clin Chim Acta* 2012;413:582–6.
16. Lowthian PJ, Thompson M. Bump-hunting for the proficiency tester—searching for multimodality. *Analyst* 2002;127:1359–64.

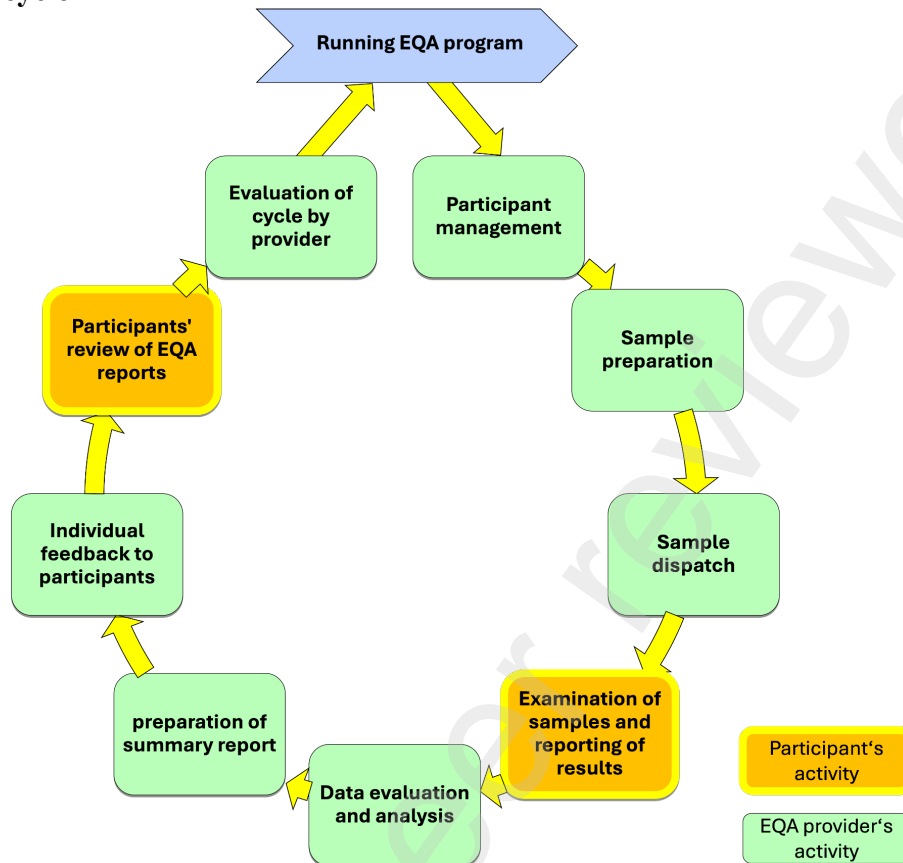
Abbreviations

APS	Analytical performance specification
CRM	Certified reference material
COVID-19	Coronavirus disease-19
EQA	External Quality Assessment
IQC	Internal quality control
ISO	International Organization for Standardization
IVD	In-vitro diagnostic
IVD-MD	In-vitro diagnostic medical device
LIMS	Laboratory information (management) system
POCT	Point-of-care testing
QMS	Quality management system
RCA	Root cause analysis
TTP	Total testing process

Table 1: Examples of root causes of an out-of-consensus EQA return

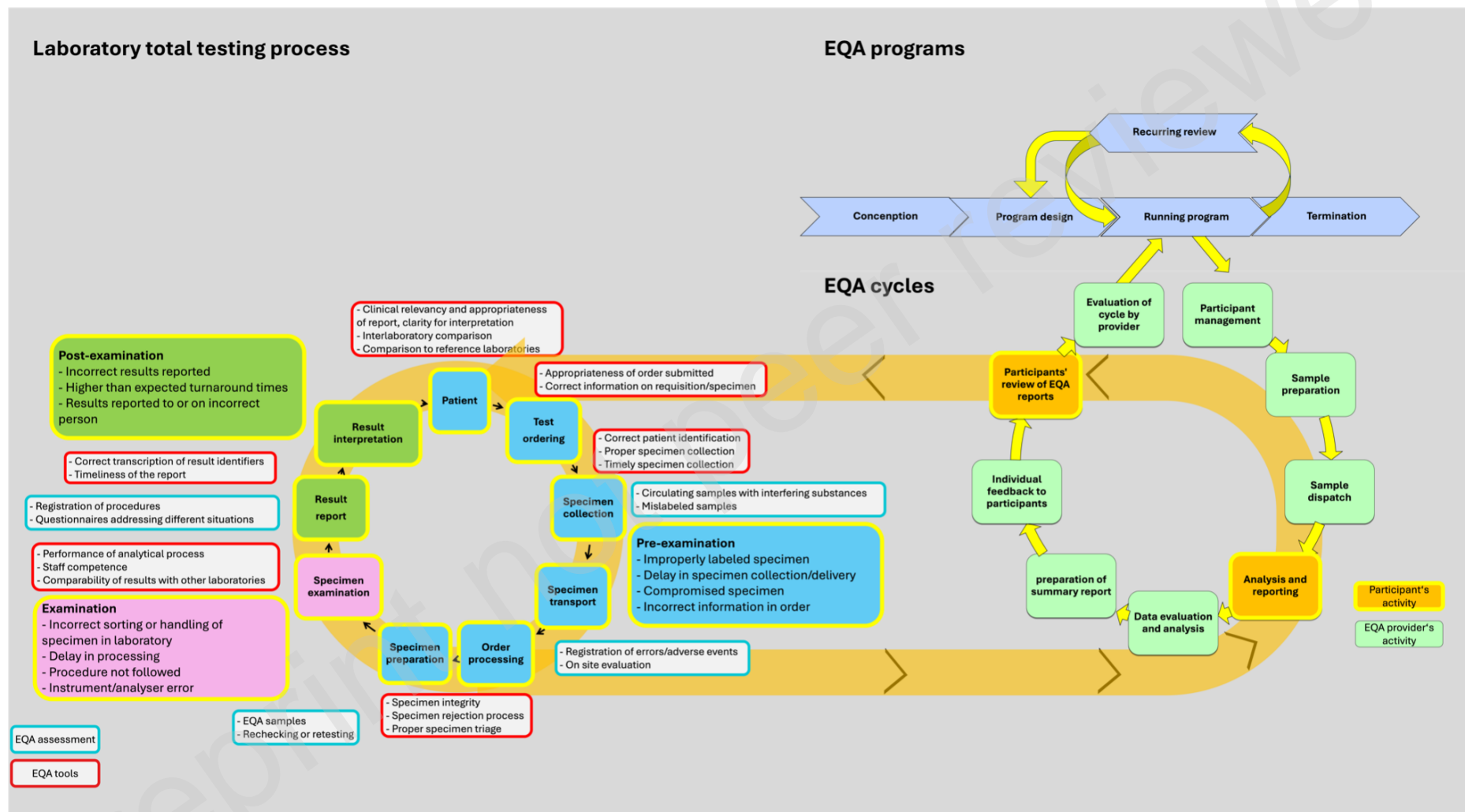
- Analysis of the incorrect sample(s)
- Incorrect storage before analysis
- Incorrect pre-treatment of the EQA sample(s)
- Analytical issues - calibration, reagent, specificity, imprecision
- Operator issues - training, competency
- Incorrect reporting of result / unit or examination procedure / transcription error and/or incorrect interpretation of result.
- Method related issues, either the examination is performing as intended and the laboratory is not considering this when reviewing their EQA, or there is genuinely a method related issue.

Figure 1: EQA cycle



Legend: EQA cycles are those essential components of ongoing EQA schemes through which participants regularly receive challenges in the form of samples for examination. The responsibility for each step in the process lies with the EQA provider, with the exception of *examination of samples and reporting of results to the provider*, and *review of EQA reports*, which are to be performed by the participants.

Figure 2: EQA programs and cycles



Legend: Relationship of the laboratory total testing process to EQA cycles and EQA programs