Special issue: External Quality Assessment in Laboratory Medicine

Review

Interpretation of EQA results and EQA-based trouble shooting

Gunn Berit Berge Kristensen*[1], Piet Meijer[2]

[1]The Norwegian EQA Program (NKK), Bergen Norway
[2]ECAT Foundation (External quality Control for Assays and Tests), Netherlands

*Corresponding author: gunn.kristensen@noklus.no

Abstract

Important objectives of External Quality Assessment (EQA) is to detect analytical errors and make corrective actions. The aim of this paper is to describe knowledge required to interpret EQA results and present a structured approach on how to handle deviating EQA results. The value of EQA and how the EQA result should be interpreted depends on five key points: the control material, the target value, the number of replicates, the acceptance limits and between lot variations in reagents used in measurement procedures. This will also affect the process of finding the sources of errors when they appear. The ideal EQA sample has two important properties: it behaves as a native patient sample in all methods (is commutable) and has a target value established with a reference method. If either of these two criteria is not entirely fulfilled, results not related to the performance of the laboratory may arise. To help and guide the laboratories in handling a deviating EQA result, the Norwegian Clinical Chemistry EQA Program (NKK) has developed a flowchart with additional comments that could be used by the laboratories e.g. in their quality system, to document action against deviations in EQA. This EQA-based trouble-shooting tool has been developed further in cooperation with the External quality Control for Assays and Tests (ECAT) Foundation. This flowchart will become available in a public domain, i.e. the website of the European organisation for External Quality Assurance Providers in Laboratory Medicine (EQALM).

Key words: external quality assessment; commutability, reference method; EQA error; between lot variation

Received: September 30, 2016 *Accepted: December 24, 2016*

Introduction

The scope of External Quality Assessment (EQA) in laboratory medicine has evolved considerably since Belk and Sunderman performed the first EQA scheme in the late 1940's (1). Today, EQA schemes are an essential component of a laboratory's quality management system, and in many countries, EQA is a component of laboratory accreditation requirements (2,3). EQA should verify on a recurring basis that laboratory results conform to expectations for the quality required for patient care.

A typical EQA scheme (EQAS) consists of the following events: A set of samples is received by the laboratory from an external EQA organization for measurements of one or more components present in the samples. The laboratories do not know the concentration of the components in the samples and perform measurements in the same manner as for patient samples. The results are returned to the EQA organizer for evaluation and after some time the laboratory receive a report stating the deviation of their results relative to a "true" value (assigned value). Reports may also include evaluation of whether the individual laboratory's results met the analytical performance specifications and an evaluation of the performance of the various measurement procedures used by the participants.

Important objectives of EQA are, beside monitoring and documenting the analytical quality, to identify poor performance, detect analytical errors, and make corrective actions. Participation in EQA gives an evaluation of the performance of the individual laboratory and of the different methods and instruments (3,4). Therefore, proper and timely evaluation

of EQA survey reports are essential and even a must for accreditation (see ISO 15189, paragraph 5.6.3.4). In this opinion paper, we focus on the knowledge required to interpret an EQA result and present a structured approach on how to handle an EQA error. The paper is limited to EQA for evaluation of quantitative measurement procedures.

Knowledge required to interpret EQA results

The value of participating in EQAS for the laboratory depends on proper evaluation and interpretation of the EQA result. Key factors for interpreting EQA results are knowledge of the *EQA material* used, the process used for *target value assignment*, the number of *replicate measurement* of the EQA sample, the range chosen for acceptable values around the target (*acceptance limits*), and the impact of *between lot variations* in reagents used in measurement procedures (4-6).

EQA material

The most important property of the EQA sample is commutability (7-9). The significance of this is something that one has become more and more aware of in recent years. A commutable EQA sample behaves as a native patient sample and has the same numeric relationship between measurements procedures as is observed for a panel of patient samples. A non-commutable EQA sample includes matrix related bias that occurs only in the EQA sample but not in authentic clinical patient samples and therefore, does not give meaningful information about method differences. Matrix related bias is due to an unwanted distortion of the test result attributed to physical and chemical differences in the samples, compared to the patient material the measurement procedures are directed towards. In a recently published article concerning method differences for immunoassay's, non-commutability for EQA materials was observed on 13 out of 50 occasions (5 components, 5 methods and 2 EQA samples) (9). The bias demonstrated by the EQA samples was five times found to be in an opposite direction compared with the native serum samples. Therefore, EQA materials should be tested for commutability and if evaluation of method differences is intended, it is mandatory. Additionally, the sample should be stable during the survey period, homogeneous, available in sufficient volume and have clinical relevant concentrations (10,11). Higher concentrations of components can be achieved by adding components (spiking) to pooled unaltered samples but this may induce non-commutability (12,13). In practice, the EQA sample very often is a compromise between ideal behaviour in accordance with native samples and stability of the material and therefore, may not be commutable, which limited the opportunities in EQA result evaluation (4).

Assignment of target values

If the EQA sample is commutable, target value assignment could be made by using a reference measurement procedure or a high-specificity comparative method that is traceable to a reference measurement procedure (14,15). In this case, all participants are compared to the same assigned value and trueness can be assessed. Target assignment by value transfer based on results from certified reference materials is possible if the commutability of the reference materials has been verified (16-18). An example is Labquality's EQAS 2050 Serum B and C (2-level) that use transferred values from NFKK Reference Serum X (Ref. NORIP home site (http://nyenga.net/norip/index.htm) – Traceability), as assigned values for 16 components. Serum X has certified values from IMEP 17 Material or Reference Serum CAL (19). For many components, a reference method or certified reference material is not available. In that case, an overall mean or median can be used as the assigned value, after removal of outliers or by the use of robust statistical methods (20). All measurement procedures are expected to give the same results for a commutable sample. That gives the possibility to compare the result with other methods. However, the measurement procedure with most participants will have greatest influence on the overall mean or median, and you do not know what the true value is. An alternative is to use the mean (or median) of the peer-group (see below) means (or medians) in order to give the same weight to each peer-group (21). A common reference assigned value should not be used if the commutability of the EQA sam-

ple is unknown because it is not possible to determine if a deviation from the assigned value is due to matrix-related bias, calibration bias or that the laboratory did not confirm to the manufacturer's recommended operation procedure.

The most common procedure used to assign a target value if the commutability of the EQA sample is unknown is to categorize participant methods into peer-groups that represent similar technology and calculate mean or median of the peer-group, after removal of outlier values, and use this as the assigned value. A peer-group consist of methods expected to have the same matrix-related bias for the EQA sample and it is possible to assess quality, i.e. verifying that a laboratory is using a measurement procedure in conformance to the manufacturer's specification and to other laboratories using the same method. A limitation is the number of participants in each group. The uncertainty of the calculated assigned value would be larger in a peer group with few participants compared to a group with many participants. The variability of results in the group will also influence the uncertainty of the assigned value. A high variability combined with few participants will give the greatest uncertainty of the assigned value.

Acceptance limits

To assess if the EQA result is acceptable, acceptance limits (i.e. analytical performance specifications) around the target value must be established (22-24). The acceptance limits can be considered *regulatory*, *statistical* or *clinical*.

Regulatory limits have the intention to identify laboratories with sufficiently poor performance that they should not be able to continue to practice. These limits tend to be wide and are often based on "fixed state-of-the-art". The German RiliBÄK and the USA Clinical Laboratory Improvement Amendments (CLIA) have defined such regulatory limits (25,26).

Statistical limits are based on "state-of-the-art" and the assumption that the measurement procedures is acceptable if it is in concordance with other using the same method. The assessment of the individual laboratory is given as z-scores, which is the number of standard deviations (SD) from the assigned value the EQA result. Assessment of z-scores is based on the following criteria: $-2.0 \leq z \leq 2.0$ is regarded as satisfactory; $-3.0 < z < -2.0$ or $2.0 < z < 3.0$ is regarded as questionable ('warning signal'); $z \leq -3.0$ or $z \geq 3.0$ is regarded as unsatisfactory ('action signal'). These criteria is stated in ISO/IEC standard 17043:2010 (27). The performance of the individual laboratory is compared against the dispersion of results obtained by the participants in the peer-group in each survey. A disadvantage is that these limits are variable and may change with time as methods and instruments evolve. Another disadvantage with statistical based criteria is that the limits may vary between peer-groups measuring the same component. Imprecise-method peer groups will have a large acceptance interval whereas precise-method peer groups will have a small interval for acceptable results, independent of what is required for clinical needs. Several EQA organizations use z-scores in the feedback reports to the participants.

Clinical limits can be based on a difference that might affect clinical decisions in a specific clinical situation (28). These limits are desirable but may be difficult to implement because very few clinical decisions are based solemnly on one particular test. More common are clinically established limits derived from biological variation in general (29,30). A challenge is the fact that the existing database on biological variation is based on few studies or studies with rather poor quality. However, in the strategic conference to arrive at a consensus on how to define analytical performance goals that took place in Milan 2014, a working group for revising the current biological variation database was established (31-33).

Both regulatory and clinical limits are fixed limits and the uncertainty of the assigned value will be a fraction of the acceptance interval. To account for the uncertainty of the definitive value, Norwegian Quality Improvement of Laboratory Examinations (Noklus) have added a fixed interval around the target value in their acceptance limits (34). When the acceptance interval is expressed as a percent, it might also be necessary to include a fixed unit interval below a concentration at which a percent is not reasonably achievable because the concen-

tration-independent variability of a measurement procedure becomes a larger fraction of the acceptance interval.

Replicate measurements

EQA results are meant to reflect results of patient samples and in most of the schemes, the participant is asked to perform a single measurement of the EQA sample. The acceptance limits are often given in %, and are established according to a Total Error allowable (TEa) concept (35,36). Total error is assessed because bias, imprecision, and analytical non-specificity can contribute to variation in a single result. If replicate measurements of the samples are included, it may be appropriate to have different limits to separately assess bias and imprecision.

Between lot variation

Between lot variation in the reagents used in measurement procedures may influence participant assessment in EQA (5,37). The percentage of participants with a "poor" quality assessment declined from 38% if using a common target value to 10 and 4% when using a method specific target value and a lot specific target value, respectively (5). Between lot variation has been described in several publications for glucose strips (38-41). Ideally, the use of lot-specific target values in EQAS would allow assessment of the individual participant's performance, but such assessments are not feasible in routine EQAS due to the larger number of lots on the market. EQA organizers should, however, register lot numbers when relevant and in some instances comment on lot variation in feedback reports (37). Additionally, between lot variation found when using control materials may not mirror results when using native blood (5,37). To evaluate the clinical importance of between lot variation discovered in routine EQAS, the actual lot should therefore be examined using native blood.

A structured approach for handling unacceptable EQA results

An unacceptable EQA result should be investigated by the participant (the person in charge of EQA in the laboratory) to find the cause of the deviation and make corrective actions. According to ISO 15189, an accredited laboratory shall participate in EQAS, monitor and document EQA results, and implement corrective actions when predetermined performance criteria are not fulfilled (3). In spite of the extensive use of EQAS in evaluating the quality of the analytical work done in medical laboratories, it is remarkable that there is little aid in the process of finding the sources of errors when they appear. Therefore, the Norwegian Clinical Chemistry EQA Program (NKK) has developed a tool for handling deviating EQA results.

All the mentioned key factors that must be taken into consideration when interpreting an EQA result also apply for handling an EQA error. The ideal EQA sample has two important properties; it behaves as a native patient sample toward all methods (is commutable) and has an assigned value established with a reference method with small uncertainty. If either of these two criteria are not entirely fulfilled, results with errors NOT related to the quality of the laboratory may arise. Therefore, the EQA provider should take steps in the scheme design to avoid or ameliorate adverse consequences. This could be done for example, by using peer-group assigned values for a non-commutable material. It is important to distinguish between different types of error (external, generating cost without benefit) and those important ones that are caused by the laboratory itself (internal). For the laboratory, errors caused by themselves are most important and of their primary interest. However, errors made by either manufacturers and/or EQA organizers (external) may also affect the quality of laboratory performance and therefore could have a major impact.

A simple relation has to be fulfilled if a deviation is to be further investigated:

$$|R - AV| > L$$

where R is the laboratory result, AV is the assigned value and L is the maximum acceptable deviation, i.e. acceptance limits. Many EQA organizers have suggested acceptance limits for their EQAS. The laboratories should be aware of these limits, and in countries where participating in EQAS is not

mandatory/regulatory, it is the laboratories responsibility to define which limits is relevant for their use. In reports from EQA organizers, the laboratory's performance history is often shown graphically together with the EQA organizer's acceptance limits. Of the three variables in the above equation, only one, R, is the immediate responsibility of the laboratory. Errors in AV has an external source while an error in L is fundamentally internal as commented above even if most laboratories tend to adopt the limits proposed by the EQA organizer. To understand the complexity of finding the cause of an EQA error all sources of deviation in an EQA result are included in a flowchart and have to be considered. In those EQAS using the z-scores as an individual performance index R should be within the range - 2 ≤ z-score ≤ 2. This indicate that the laboratory result is within the 95% range of the distribution of all results. Results with a z-score < - 3 or > 3 can be identified as unsatisfactory, while results with a z-score between -3 and -2 or 2 and 3 are questionable (a warning signal). This means the laboratory should investigate whether there is a reason why the results tend to become an outlier.

The history of developing a flowchart

In 2008 and 2009, the topic for group works at NKK's annual meetings was "How to handle a deviating EQA result". The result of this work was further processed by the NKK expert group and resulted in a flow chart with additional comments that could be used by the laboratories, e.g. in their quality system, to document actions against deviations in EQA.

In 2012-2013 NKK carried out a follow-up and an evaluation of the flowchart. Deviating EQA results from Labquality's EQA scheme 2050 Serum B and C (2-level), survey 4 and 6, 2012) were selected and the laboratories were asked to use the flowchart to assess the EQA error and state the cause of the error. They were also asked if they use the flowchart regularly, and if not, why they do not use it. Finally, they were asked if they have any suggestions for improvement of the flowchart. Fifty-six percent of the invited laboratories replied (39/69). The results showed that most errors (81%) were the laboratory's responsibility (internal causes), 15% the EQA provider's responsibility (external causes), whereas 4% were a mix (internal/external causes). The most common errors were transcription errors (72%) both with respect to internal and external causes. For 4% of the deviating EQA results the participants did not reach any conclusion. Fifty-eight percent of the laboratories that responded used the flowchart regularly. Of these, 37% commented that they found the flowchart comprehensive and a bit complicated, but very useful in training/educational situations. They suggest changing the order of the items in the flowchart and start with transcription errors, the most common cause to a deviating EQA result (unpublished data).

The recommendations from the evaluation has been taken into account and a new version of the flowchart has been developed in cooperation with the External quality Control for Assays and Tests (ECAT) Foundation in the Netherlands (Figure 1). The content of the original flowchart is kept and where necessary expanded and re-structured.

Description of the flowchart

The flowchart starts with the most frequently errors followed by the logical steps in the flow of an EQA survey (from pre-survey issues to report and interpretation – see Figure 1). Four different aspects elucidate each item in the flowchart: **Observation** – what is the potential error, **Responsibility** – who is responsible for the error, **Comment** – a short comment on action to undertake, **Note** – a more detailed description of actions (see Figure 2). The responsible could be the participant, the EQA-provider (EQAP), and/or the manufacturer, each marked with different colour.

Before starting evaluating the potential cause for a deviating result, the report and/or comment letter should be read carefully for a possible explanation for deviating results. If no explanation is given, the flowchart should be used (Figure 1 and Figure 2) to reveal the potential cause(s).

The flowchart starts with the most probable causes of error; "*Transcription errors*" (item 1-6). The EQA provider may wrongly enter the data or the laboratory may record or report a wrong result. In the evaluation of the first version of the flowchart,

Introduction

Before you start evaluating the potential cause for a deviating result, please read carefully the report and/or comment letter for a possible explanation for deviating results (see pre-note). If no explanation is given please use the flowchart below to reveal the potential cause(s).

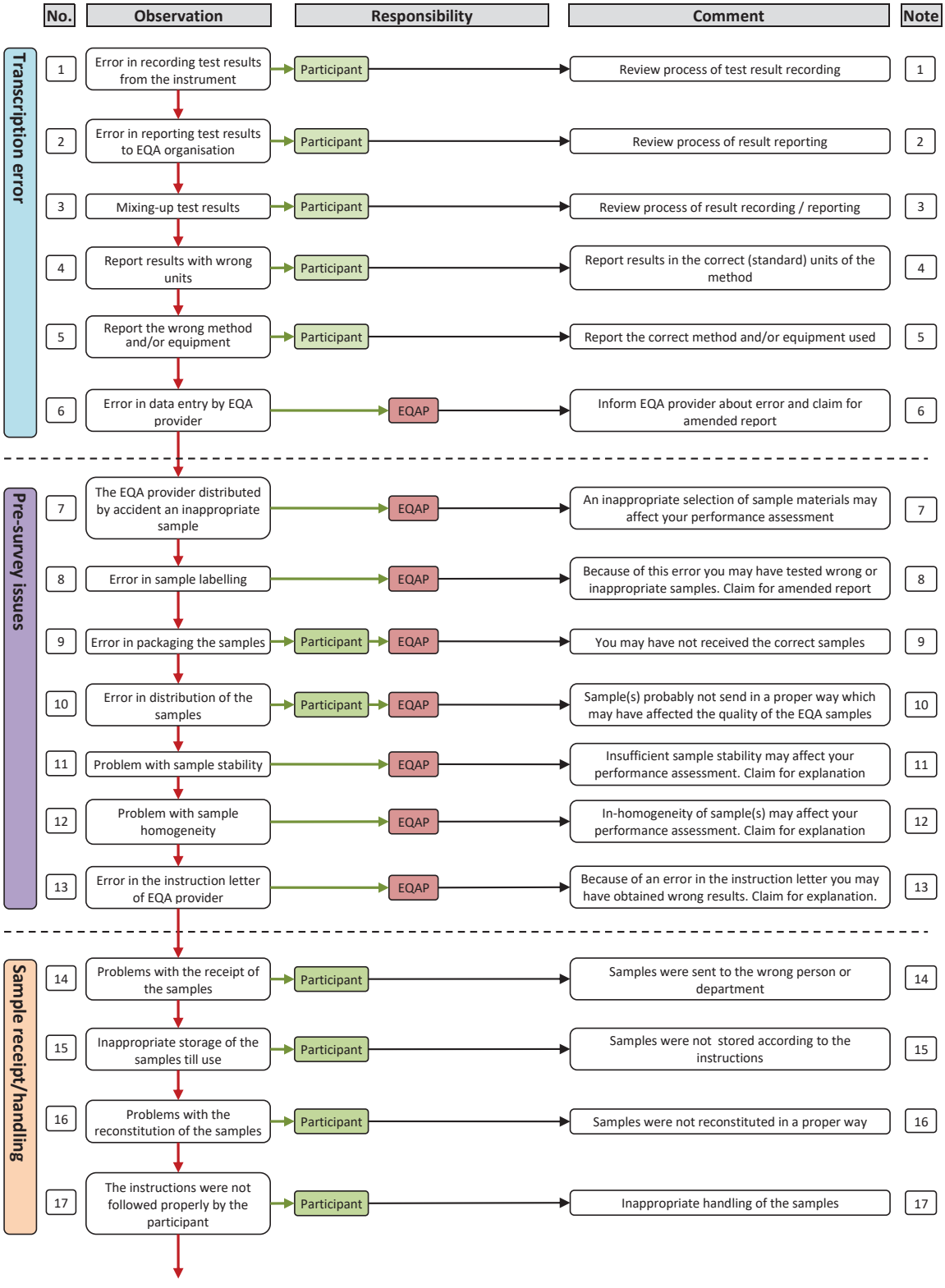

Figure 1. Flowchart for handling deviating EQA results.

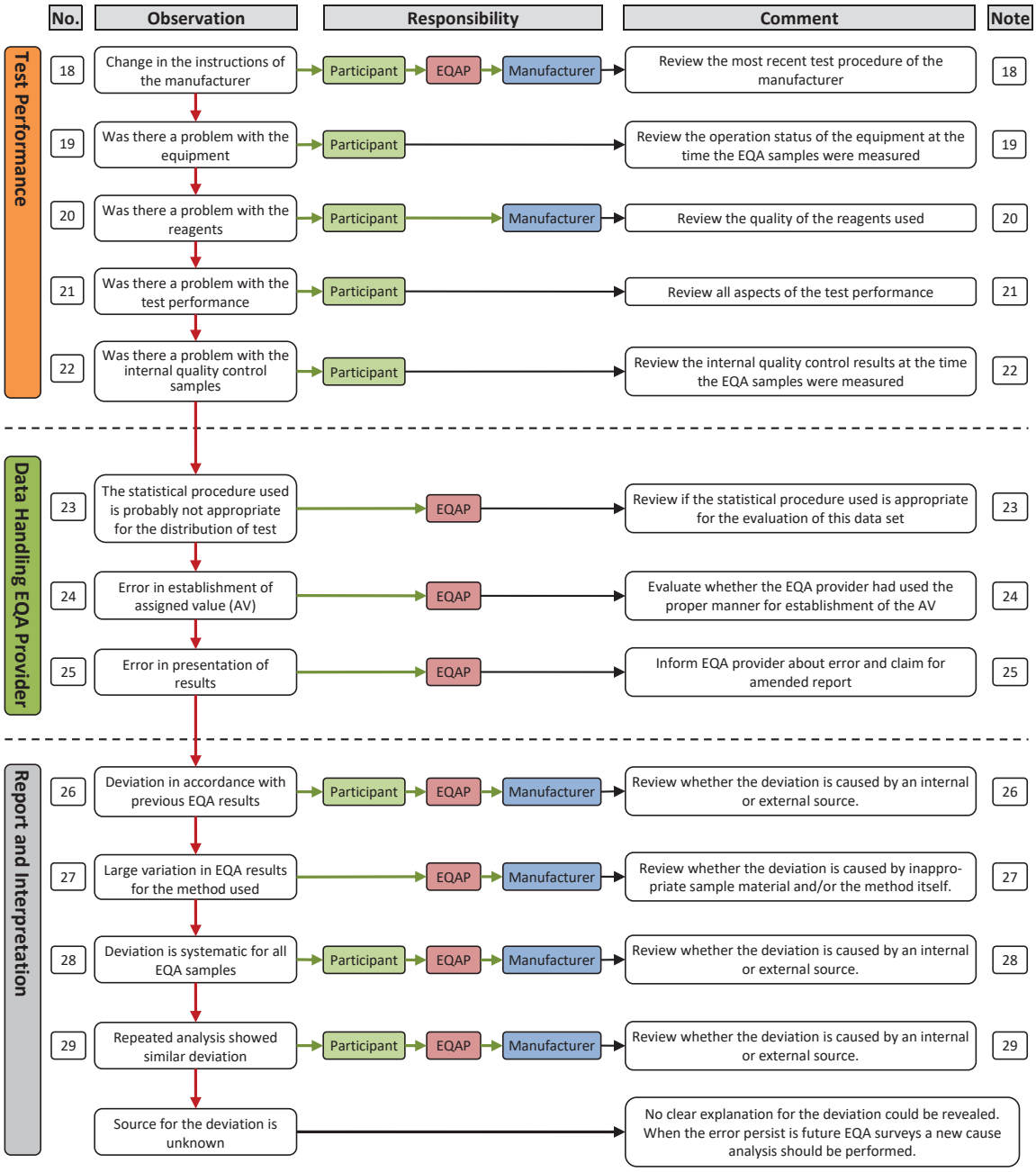

Figure 1. Flowchart for handling deviating EQA results (continued).

transcriptional errors were the most common cause for a deviating result.

The next is "*Pre-survey issues*" (item 7-13). Obviously, a lot may go wrong before the sample reaches the laboratory like sample selection, inappropriate stability or homogeneity, a mistake in labelling or an error in packaging. These errors are the EQA organizer's responsibility and should have been commented on in the comment letter (see above). Unfortunately, this is often not the case even though these errors are hard, and often impossible, to detect for the laboratory. Examples of more subtle origin are related to the stability of the samples. A good procedure is always to store the EQA

Note	Remark
Pre-note	The comment letter or comment section in the report may include comments regarding remarkable observations (e.g. a relative large deviation for a specific method mean with respect to the overall assigned value). Evaluate carefully the remarks and consider whether this could be explaining your deviating result.
1	The participant had made an error in recording the test result of the EQA sample(s) and as such reported a wrong result(s). This is an internal cause for the error made. The participant should carefully review the process of result recording and take appropriate action to avoid this problem in future surveys.
2	The participant had made an error in reporting the test result of the EQA sample(s) to the EQA organisation. This is an internal cause for the error made. The participant should carefully review the process of result reporting and take appropriate action to avoid this problem in future surveys.
3	The participant had mixed-up the test results either at the level of recording the test result from the instrument or when the test results were reported to the EQA organisation. The participant should carefully review the process of result recording and/or reporting and take appropriate action to avoid this problem in future surveys.
4	The participant had reported the result with the wrong unit (e.g. report the result in U/dL instead of U/mL). This may lead to incorrect treatment of the data in the evaluation software of the EQA organiser. In this case the result could be assigned as an outlier. The participant should select the correct unit when reporting a result to avoid this problem in future surveys.
5	The participant had reported the wrong method and/or equipment used. This may lead to inclusion of the result in the wrong method/equipment group. This may affect both the total evaluation of that group(s) as well as the evaluation of your own result and performance assessment. The participant should select the correct method/equipment when reporting a result to avoid this problem in future surveys.
6	The EQA provider had wrongly entered your data (e.g. result, unit, method, instrument etc.) into the database which leads to a wrong evaluation of your performance. Please inform the EQA provider and ask for an amended report.
7	The EQA provider had distributed inappropriate sample material (e.g. the sample material is not commutable for your specific method. I.e. the sample material used by the EQA provider behave for your specific method not identical as a real patient sample). This may affect your performance assessment. It is the responsibility of the EQA provider to write a note about this issues in the report or comment letter. If this is not done, please inform the EQA provider and ask for an amended report or comment letter.
8	The EQA provider had made a mistake in the labelling of the sample(s). As such the evaluation is not in correspondence with the description of the sample information and/or expected target values. It is the responsibility of the EQA provider to write a note about this issue in the report or comment letter. If this is not done, please inform the EQA provider and ask for an amended report or comment letter. In addition, the EQA provider should provide replacement of the sample to give you the opportunity to rerun the sample or the EQA provider has to rerun the whole survey.
9	You have received wrong samples because of an error in the packaging by the EQA provider. Because you were not aware of this fact you had measured the wrong samples which may had affect your performance assessment. It is the responsibility of the EQA provider to write a note about this issues in the report or comment letter. If this is not done, please inform the EQA provider and ask for new samples and an amended report. It is also the responsibility of the participant to check whether they have received the correct samples as soon as possible after receipt. If this is not done by the participant it is also the responsibility of the participant if incorrectly received samples are still measured.
10	The EQA provider selected an inappropriate manner to distribute the sample (e.g. inappropriate packaging material, distribution at an inappropriate time period, etc.). This may lead to delay in the receipt, damage of the package etc. Measuring such samples may have affected your performance assessment. Please inform the EQA provider in time about such issues in future surveys. In addition, the EQA provider should provide replacement of the sample(s) to give you the opportunity to rerun the sample and provide you with an amended report. In general, inform the EQA provider immediately after the delayed receipt of samples or when the package is damaged and sample are probably affected and ask for replacement of the sample(s).
11	The EQA provider had distributed a sample with insufficient stability. This may affect your performance assessment. It is the responsibility of the EQA provider to write a note about this issue in the report or comment letter and rerun the survey. If this is not done, please inform the EQA provider and ask to rerun the survey with better samples and provide the participants with an amended report.
12	The EQA provider had distributed a sample with insufficient homogeneity. This may affect your performance assessment. It is the responsibility of the EQA provider to write a note about this issue in the report or comment

Figure 2. Notes to flowchart for handling deviating EQA results

	letter. If this is not done, please inform the EQA provider and ask for an amended report or comment letter. In addition, the EQA provider should provide replacement of the sample(s) to give you the opportunity to rerun the sample and provide you with an amended report.
13	The EQA provider had provided inappropriate information in the instruction letter. Because of this you may have not handled the sample in an appropriate way. This may have affected your performance assessment. It is the responsibility of the EQA provider to write a note about this issues in the report or comment letter. If this is not done, please inform the EQA provider and ask for an amended report or comment letter. In addition, the EQA provider should provide replacement of the sample(s) to give you the opportunity to rerun the sample and provide you with an amended report.
14	The participant is informed by the EQA provider about the dispatch date of the samples (e.g. by an annual survey schedule). There is a delay in the delivery of the samples in your laboratory. This may be caused for instance by wrong information available by the EQA provider about the address details or wrong distribution within the hospital. This may lead to bad test results. Please evaluate carefully the delivery and/or distribution of the samples and take appropriate action to avoid this problem in future surveys. If there is a systematic delay in the delivery of the sample (e.g. because of the post services in for in your particular country), the EQA provider should be informed and another way of delivery of the samples should be used (e.g. courier service).
15	Samples were not stored in a proper way after receipt. For instance, they were not put into the refrigerator. This may lead to bad test results. Please evaluate carefully the procedure for storage of the samples and take appropriate action to avoid this problem in future surveys.
16	The samples were not reconstituted in a proper way because, for instance, you did not use calibrated pipettes. Therefore, you pipette the wrong volume. Another possibility is that you did not mix the sample properly. This may lead to bad test results. Please evaluate carefully the procedure for reconstitution of the samples and take appropriate action to avoid this problem in future surveys.
17	You have not read the instruction letter carefully. Therefore, you may have made a mistake in the handling of the sample (for example: the sample used in this survey was stable for a shorter time period as usual and you did not notice this. Therefore, you did not measure the sample within the stable time period). This may lead to bad test results. Please read carefully the instruction letter every time.
18	If the manufacturer has made changes in the test constitution and/or procedure they should have made the user aware of this. If this is not done the manufacturer should be informed and asked for improved communication in future cases. The laboratory should evaluate/validate the revised test procedure and adapt the test procedure when necessary. If the laboratory has not changed the test procedure based on information given by the manufacturer, internal actions have to be carried out.
19	If there was any problem related to the equipment at the time the EQA samples were measured, (e.g. problems due to maintenance issues, calibration, test settings) internal actions have to be carried out to prevent repeated problems.
20	If there was any problem related to the reagents used at the time the EQA samples were measured, (e.g. problems with a specific lot. no., reconstitution, storage) actions have to be carried out to prevent repeated problems. If the problem is caused by external factors the distributor or manufacturer of the reagents should be contacted.
21	If there was any problem related to the performance of the test at the time the EQA samples were measured, (e.g. test settings, local modifications, calibration) internal actions have to be carried out to prevent repeated problems.
22	If the results of the internal quality control samples could explain the deviation of the EQA results, internal actions have to be carried out to prevent repeated problems. Make sure that the patient results were correct during the period the EQA samples were measured. Look for trends in internal quality control results.
23	It is the responsibility of the EQA provider to provide the participant with information about the statistical procedure used (e.g. information included in each report or provided on an annual basis in a survey manual). Review whether there was maybe a problem with the statistical method used for this particular data set, e.g. due to a non-normal distribution of the results or the size of the data set. Contact the EQA provider and ask for further explanation by the EQA provider.
24	If the assigned value (AV) was established by a reference method the deviation might be caused by either the assignment of the AV or your method. Check the deviation of other methods used. If the deviation is similar to your method the AV might be incorrect. Inform the EQA provider. If the deviation is not similar to other methods, analyse reference material and some patient samples in parallel with another reliable method. Make the manufacturer of your method aware of the deviation. Ask for appropriate action. If the AV is a consensus value (calculated from all the results in a method group consisting of different methods) and the deviation is representative for your method, realise that the AV of the total group is not representative for your method. The laboratory should evaluate their EQA results to the method specific consensus value and not to the AV of the total group. If the method specific AV is not given in the report, ask the EQA provider for possibilities to provide

Figure 2. Notes to flowchart for handling deviating EQA results (continued).

	this information. If the AV is calculated from a small number of results this might result in a less reliable AV. Interpret the result carefully. If the AV is not calculated from original results but from a mix of original and modified results (e.g. by the use of a local correction factor), this might contribute to a large variation and an incorrect AV. When reporting results, it should be possible to mark if the results are not original and as such should be excluded from the assignment of the AV or treated as a separate method group (in the case sufficient participants use the same modification of results).
25	The EQA provider had probably wrongly presented the results in the report (e.g. results for your specific method are linked to another method.). Please inform the EQA provider and ask for an amended report.
26	A similar deviation has been observed earlier. If the deviation is typical for the method the cause is external [see notes 18 (e.g. change in method) and 20 (e.g. problem with specific lot of reagents)]. The cause can also be due to the EQA samples used (e.g. non-commutable for a particular method). If not, the cause is internal. Make sure that the internal quality control and patient results were correct at the time the EQA samples were measured. Undertake appropriate corrective actions.
27	If relevant, complain the EQA organiser that the sample material used in the survey was probably inappropriate for your method. Some materials show especially large variation for one specific method and in that case the deviation might be large compared to your acceptable limits without being large compared to the variation for your method. Inform EQA organiser and ask for the use better commutable material (when possible).
28	When multiple samples with different concentration were used, investigate thoroughly if a systematic error is present. A systematic error may have different sources, e.g. problems with the calibrator (external), pipetting error (internal), problems with a certain lot no of the reagents used (external), problem with the value assignment (external). The potential cause should be investigated thoroughly and appropriate actions should be undertaken.
29	Repeat analysis on stored EQA material. If the repeat analysis shows no deviation anymore the method seems to be OK. Make sure that the internal quality control and patient results were correct at the time the EQA samples were measured. When necessary undertake appropriate corrective actions. If repeat analysis shows the same deviation, ask the EQA organiser for a repeat sample(s) for re-analysis. If the repeat analysis shows no deviation anymore the method seems to be OK. There was most likely something wrong with the EQA sample (e.g. wrong sample sent by EQA organiser, pipetting error during reconstitution). If repeat analysis show the same deviation there seems to be, for instance, something wrong with the method used (external cause) or with the sample sent by the EQA organiser (external cause). Investigate thoroughly potential causes and undertake appropriate actions.

Figure 2. Notes to flowchart for handling deviating EQA results (continued).

sample at stable conditions at least until the report is received – a reanalysis of the sample may eliminate many sources of error. If you do not have any sample material left, you should ask the EQA organizer for a new sample.

The next section, "*Sample receipt/handling*" (item 14-17), is solely the laboratory's responsibility. The laboratory should carefully check that the EQA provider has the correct address details and that all instructions for handling the sample from the EQA organizer, has been followed. The visual appearance of the specimen should be checked by reception for an immediate check of sample quality and physical integrity and also that the sample identifiers match the documentation.

"*Test Performance*" (item 18-22) is next. The laboratory or the instrument or kit manufacturer is responsible for errors in this section. Local documentation of measurement is important: who/when/how. You may locally have changed the procedure of measurement (e.g. factorized results: internal source) or the producer of the method may have changed the calibrator/reagents/procedure (external) without informing you (example creatinine, ALP). The problem may be related to the equipment, the reagents or the test performance. Is the problem new to your laboratory or is it an old problem? Have the error occurred before? It is important to evaluate results in relation to previous surveys. In other words, evaluation of the results of a single survey may be insufficient to reveal the cause of the problem. If it is new, look at your internal quality control data (IQC). First, look for systematic deviations (bias/trends) that may

explain the EQA result. If this is the cause of the error, are your IQC rules not stringent enough or is L too narrow? In any case, is there a need for reanalysis of any of the patient samples in the relevant analytical run? If no hints can be found in IQC, you should proceed in the flowchart.

Errors in "*Data handling*" (item 23-25) are external and usually, not the responsibility of the laboratory. The problem could be related to the statistical procedure used in handling the data, e.g. parametric methods used when the data are not normally distributed, the consensus value is based on few participants causing a large deviation, or it may stem from uncertainty caused by a mix of factorized and original results. The establishment of the assigned value (AV) is a challenge. All participants, regardless of instrument or method, should be evaluated against the AV established by a reference method when this is available and commutable material is used. A deviation that is representative for one particular instrument or method is caused either by the EQA provider (non-commutable material) or the instrument or method used (e.g. a problem with a certain lot of reagents). An evaluation based on a reference value for a non-commutable EQA sample is a mistake by the EQA provider. Another example applies to a deviation between a particular instrument or method and the peer-group AV, based on results from a large number of instruments or methods. The deviation is similar for all participants with the particular instrument or method, and in that case, the instrument or method is linked to the wrong peer-group. It is important to check that the grouping of the instrument or method is correct, by both the EQA provider and the participant. This is a frequent cause of error unless the method is stated and adjusted at each survey. One should also be aware of that in a peer-group consisting of several instruments or methods the instrument or method with most participants will have a greater influence on the assigned value. Errors in this section may be difficult for the participant to detect and should have been commented on in the feedback report.

The last section is "*Report and Interpretation*" (item 26-29). Is the deviation clinically important? If not, the acceptable limits should be reconsidered, and may be expanded. Limits expressed in percent are probably not suitable for the lowest concentrations of the component because the measurement uncertainty may be larger than the acceptance limits if the concentration is low. Especially high concentrations are often less interesting and therefore also the deviation. However, from an analytical point of view it might still be worth reducing the error. This does not apply to limits based on state-of-the-art. A deviation in accordance with previous results has probably been handled earlier. The error may be the responsibility of the participants, the EQA provider or the manufacturer. The mean of all results for one particular method, however, may always be used to distinguish between errors general for the method (external) and errors in the laboratory (internal), even if you do not know the commutability of the sample. It may be that the error is already recognized as a general problem or specific for your method. An unusually large variation for a particular method may be caused by poor EQA material (external/EQA provider) or between lot variation in reagents for that specific method and several lots present (external/manufacturer). It could also be due to change in the method by the manufacturer. A suspected internal error requires review of the internal quality control (IQC) and the patient results in the period where the EQA sample was analysed. A similar deviation observed in several samples with different concentrations, may suggest that a systematic error is present. In that case, it may be wise to check previously EQA results to look for a trend. For more details, look closer to Figure 1 and Figure 2.

Sometimes there is no explanation to the EQA error. It may have been a transient error in the system at the time of measurement. The error should be followed up in later EQA surveys.

It should be realised that an error made by the EQA provider or manufacturer may cause a deviating result for a participant in an EQAS. The participant should therefore also consider this possibility when evaluating deviating EQA results. Errors caused by the EQA provider's should have been commented in the comment letter. These errors

are often hard and sometimes impossible for the participant to discover and handle. In order to improve their schemes, the EQA provider should create a checklist based on this flowchart as a tool in their work to make ongoing EQA schemes more useful for the participant.

Limitations

The flowchart presented in this paper is limited to cover mistakes that occur in the analytical phase of the total testing process. Transcription errors, which counted for about three quarters of the mistakes or errors, could be classified as post-analytical errors, i.e. not part of the analytical process, and therefore may "falsely" affect the evaluation of the analytical performance. Today, writing down the patient results are not part of the daily routine when laboratories are highly automatized. The fact that a laboratory professional does not check written results, might reflect lack of attention to deliver correct results and hence, indicate a lack of quality. Another limitation is the limited use of the flowchart so far.

Future directions

The flowchart itemizes the steps taken by many EQA providers when working with participants to understand and correct adverse performance and is used in the format of Corrective and Preventative Action (CAPA) documentation or Root Cause Analysis (RCA) tools. The flowchart is a useful addition to these as it summarizes these processes for participants. To our knowledge, this is the first time such a structured approach on how to handle deviating EQA results, have been published. So far, the flowchart has had a very limited use. However, the flowchart will soon become available in the public domain, i.e. the website of the European organisation for External Quality Assurance Providers in Laboratory Medicine – EQALM. This flowchart can be the basis for modified versions for specific EQA areas and be further improved based on the experience of users.

Potential conflict of interest

None declared.

References

1. *Belk WP, Sunderman FW. A survey of the accuracy of chemical analyses in clinical laboratories. Am J Clin Pathol 1947;17:853-96. https://doi.org/10.1093/ajcp/17.11.853.*
2. *Libeer JC, Baadenhuijsen H, Fraser CG, Petersen PH, Ricos C, Stockl D, et al. Characterization and classification of external quality assessment schemes (EQA) according to objectives such as evaluation of method and participant bias and standard deviation. External Quality Assessment (EQA) Working Group A on Analytical Goals in Laboratory Medicine. Eur J Clin Chem Clin Biochem 1996;34:665-78.*
3. *International Organization for Standardization. ISO 15189: medical laboratories: particular requirements for quality and competence. Geneva, Switzerland: International Organization for Standardization; 2012.*
4. *Miller WG, Jones GR, Horowitz GL, Weykamp C. Proficiency testing/external quality assessment: current challenges and future directions. Clin Chem 2011;57:1670-80. https://doi.org/10.1373/clinchem.2011.168641.*
5. *Kristensen GB, Christensen NG, Thue G, Sandberg S. Between-lot variation in external quality assessment of glucose: clinical importance and effect on participant performance evaluation. Clin Chem 2005;51:1632-6. https://doi.org/10.1373/clinchem.2005.049080.*
6. *Miller WG, Erek A, Cunningham TD, Oladipo O, Scott MG, Johnson RE. Commutability limitations influence quality control results with different reagent lots. Clin Chem 2011;57:76-83. https://doi.org/10.1373/clinchem.2010.148106.*
7. *Miller WG, Myers GL, Rej R. Why commutability matters. Clin Chem 2006;52:553-4. https://doi.org/10.1373/clinchem.2005.063511.*
8. *Vesper HW, Miller WG, Myers GL. Reference materials and commutability. Clin Biochem Rev 2007;28:139-47.*
9. *Kristensen GB, Rustad P, Berg JP, Aakre KM. Analytical Bias Exceeding Desirable Quality Goal in 4 out of 5 Common Immunoassays: Results of a Native Single Serum Sample External Quality Assessment Program for Cobalamin, Folate, Ferritin, Thyroid-Stimulating Hormone, and Free T4 Analyses. Clin Chem 2016;62:1255-63. https://doi.org/10.1373/clinchem.2016.258962.*
10. *Middle JG, Libeer JC, Malakhov V, Penttila I. Characterisation and evaluation of external quality assessment scheme serum. Discussion paper from the European External Quality Assessment (EQA) Organisers Working Group C. Clin Chem Lab Med 1998;36:119-30. https://doi.org/10.1515/CCLM.1998.023.*

11. Schreiber WE, Endres DB, McDowell GA, Palomaki GE, Elin RJ, Klee GG, et al. Comparison of fresh frozen serum to proficiency testing material in College of American Pathologists surveys: alpha-fetoprotein, carcinoembryonic antigen, human chorionic gonadotropin, and prostate-specific antigen. Arch Pathol Lab Med 2005;129:331-7.
12. Miller WG, Myers GL, Ashwood ER, Killeen AA, Wang E, Thienpont LM, et al. Creatinine measurement: state of the art in accuracy and interlaboratory harmonization. Arch Pathol Lab Med 2005;129:297-304.
13. Ferrero CA, Carobene A, Ceriotti F, Modenese A, Arcelloni C. Behavior of frozen serum pools and lyophilized sera in an external quality-assessment scheme. Clin Chem 1995;41:575-80.
14. Stockl D, Franzini C, Kratochvila J, Middle J, Ricos C, Siekmann L, et al. Analytical specifications of reference methods compilation and critical discussion (from the members of the European EQA-Organizers Working Group B). Eur J Clin Chem Clin Biochem 1996;34:319-37.
15. Thienpont L, Franzini C, Kratochvila J, Middle J, Ricos C, Siekmann L, et al. Analytical quality specifications for reference methods and operating specifications for networks of reference laboratories. discussion paper from the members of the external quality assessment (EQA) Working Group B1) on target values in EQAS. Eur J Clin Chem Clin Biochem 1995;33:949-57.
16. Broughton PM, Eldjarn L. Methods of assigning accurate values to reference serum. Part 1. The use of reference laboratories and consensus values, with an evaluation of a procedure for transferring values from one reference serum to another. Ann Clin Biochem 1985;22:625-34. https://doi.org/10.1177/000456328502200613.
17. Eldjarn L, Broughton PM. Methods of assigning accurate values to reference serum. Part 2. The use of definitive methods, reference laboratories, transferred values and consensus values. Ann Clin Biochem 1985;22:635-49. https://doi.org/10.1177/000456328502200614.
18. Blirup-Jensen S, Johnson AM, Larsen M, Proteins ICoP. Protein standardization V: value transfer. A practical protocol for the assignment of serum protein values from a Reference Material to a Target Material. Clin Chem Lab Med 2008;46:1470-9. https://doi.org/10.1515/CCLM.2008.289.
19. Rustad P, Felding P, Franzson L, Kairisto V, Lahti A, Martensson A, et al. The Nordic Reference Interval Project 2000: recommended reference intervals for 25 common biochemical properties. Scand J Clin Lab Invest 2004;64:271-84. https://doi.org/10.1080/00365510410006324.
20. International Organization for Standardization. Stastistical methods for use in proficiency testing by interlaboratory comparisons. ISO 13528. Geneva, Switzerland: International Organization for Standardization; 2005.
21. Van Houcke SK, Rustad P, Stepman HC, Kristensen GB, Stockl D, Roraas TH, et al. Calcium, magnesium, albumin, and total protein measurement in serum as assessed with 20 fresh-frozen single-donation sera. Clin Chem 2012;58:1597-9. https://doi.org/10.1373/clinchem.2012.189670.
22. Fraser CG. General strategies to set quality specifications for reliability performance characteristics. Scand J Clin Lab Invest 1999;59:487-90. https://doi.org/10.1080/00365519950185210.
23. Fraser CG, Kallner A, Kenny D, Petersen PH. Introduction: strategies to set global quality specifications in laboratory medicine. Scand J Clin Lab Invest 1999;59:477-8. https://doi.org/10.1080/00365519950185184.
24. Fraser CG, Petersen PH. Analytical performance characteristics should be judged against objective quality specifications. Clin Chem 1999;45:321-3.
25. Richtlinie der Bundesärztekammer zur Qualitätssicherung laboratoriumsmedizinischer Untersuchungen. Available at: http://www.bundesaerztekammer.de/fileadmin/user_upload/downloads/Rili-BAEK Laboratoriumsmedizin.pdf. Assessed September 3rd 2016. [In German.]
26. Department of Health and Human Services. Centres for Disease Control and Prevention. Current CLIA regulations. Available at: www.cdc.gov/ophss/csels/dls/clia.html. Assessed September 3rd 2016.
27. International Organization for Standardization/International Electrotechnical Commission. Conformity assessment - General requirements for proficiency testing. ISO 17043. Geneva: ISO/IEC; 2010.
28. Sandberg S, Thue G. Quality specifications derived from objective analyses based upon clinical needs. Scand J Clin Lab Invest 1999;59:531-4. https://doi.org/10.1080/00365519950185292.
29. Stockl D, Baadenhuijsen H, Fraser CG, Libeer JC, Petersen PH, Ricos C. Desirable routine analytical goals for quantities assayed in serum. Discussion paper from the members of the external quality assessment (EQA) Working Group A on analytical goals in laboratory medicine. Eur J Clin Chem Clin Biochem 1995;33:157-69.
30. Ricos C, Baadenhuijsen H, Libeer CJ, Petersen PH, Stockl D, Thienpont L, et al. External quality assessment: currently used criteria for evaluating performance in European countries, and criteria for future harmonization. Eur J Clin Chem Clin Biochem 1996;34:159-65.
31. Roraas T, Stove B, Petersen PH, Sandberg S. Biological Variation: The Effect of Different Distributions on Estimated Within-Person Variation and Reference Change Values. Clin Chem 2016;62:725-36. https://doi.org/10.1373/clinchem.2015.252296.
32. Panteghini M, Sandberg S. Defining analytical performance specifications 15 years after the Stockholm conference. Clin Chem Lab Med 2015;53:829-32. https://doi.org/10.1515/cclm-2015-0303.
33. Sandberg S, Fraser CG, Horvath AR, Jansen R, Jones G, Oosterhuis W, et al. Defining analytical performance specifications: Consensus Statement from the 1st Strategic Conference of the European Federation of Clinical Chemistry and Laboratory Medicine. Clin Chem Lab Med 2015;53:833-5. https://doi.org/10.1515/cclm-2015-0067.
34. Bukve T, Stavelin A, Sandberg S. Effect of Participating in a Quality Improvement System over Time for Point-of-Care C-Reactive Protein, Glucose, and Hemoglobin Testing. Clin Chem 2016;62:1474-1481. https://doi.org/10.1373/clinchem.2016.259093.

35. Hens K, Berth M, Armbruster D, Westgard S. Sigma metrics used to assess analytical quality of clinical chemistry assays: importance of the allowable total error (TEa) target. Clin Chem Lab Med 2014;52:973-80. https://doi.org/10.1515/cclm-2013-1090.
36. Westgard JO. Useful measures and models for analytical quality management in medical laboratories. Clin Chem Lab Med 2016;54:223-33. https://doi.org/ 10.1515/cclm-2015-0710.
37. Stavelin A, Riksheim BO, Christensen NG, Sandberg S. The Importance of Reagent Lot Registration in External Quality Assurance/Proficiency Testing Schemes. Clin Chem 2016;62:708-15. https://doi.org/10.1373/clinchem.2015.247585.
38. Harrison B, Markes R, Bradley P, Ismail IA. A comparison of statistical techniques to evaluate the performance of the Glucometer Elite blood glucose meter. Clin Biochem 1996;29:521-7. https://doi.org/10.1016/S0009-9120(96) 00092-6.
39. Skeie S, Thue G, Nerhus K, Sandberg S. Instruments for self-monitoring of blood glucose: comparisons of testing quality achieved by patients and a technician. Clin Chem 2002;48:994-1003.
40. Skeie S, Thue G, Sandberg S. Patient-derived quality specifications for instruments used in self-monitoring of blood glucose. Clin Chem 2001;47:67-73.
41. Voss EM, Cembrowski GS. Performance characteristics of the HemoCue B-Glucose analyzer using whole-blood samples. Arch Pathol Lab Med 1993;117:711-3.