

External quality assessment schemes in Latin America

Gabriel Alejandro Migliarino

Gmigliarino Consultants, Buenos Aires, Argentina

ARTICLE INFO

Corresponding author:

Gabriel Alejandro Migliarino, Biochemist
Director at Gmigliarino Consultores
Carlos Tejedor 1323 1 A
1706 Haedo, Buenos Aires, Argentina
Phone: +54 11 4460 2527
E-mail: gmigliarino@gmigliarino.com

Key words:

PT, EQA, assigned value, standard deviation,
assigned value standard uncertainty,
comparison group, performance, intended use

ABSTRACT

As professionals of the clinical laboratory we must generate clinically useful results, products and services for the patients' health care. Laboratories must participate in one or more proficiency testing (PT) or external quality assessment (EQA) programs as part of routine quality assurance. Nevertheless participating per se is not enough. There are critical factors to take into consideration when selecting a PT or EQA providers. In most cases the survey's providers offer assigned values obtained from consensus of results provided by the participants for comparison, it is critical to evaluate consistency of the comparison group before interpretation and decision-making.

As far as possible, one must participate in schemes accredited under the ISO 17043 [4] regulations or those that substantially comply with their guidelines. It assures a correct statistical treatment of data through robust statistical techniques like those proposed by ISO 13528 [5], these provide necessary information for evaluation of consistency of the group dedicated for comparison. In Latin America, the External Quality Assessment programs participation rate is not high. Providers of local schemes face difficulties putting together comparison groups due to multiple reagents and instruments from different commercial

brands. It results in non-consistent comparison groups due to the small number of participants. Besides, in general the attitude of laboratories is reactive instead of proactive. They tend to pay more attention to rejected results than to accepted ones. At the same time, it is not common practice that laboratories evaluate the latest survey with previous ones. In conclusion, the information offered by these schemes is underused and there is a lot of work to carry out.

INTRODUCTION

The main objective of the clinical laboratory is to yield results, products and services clinically useful health care. Laboratories must participate in one or more Proficiency testing (PT) PT or External Quality Assessment (EQA) programs as part of routine quality assurance [1].

These schemes evaluate the laboratory's analytical performance compared to its peers (other laboratories using the same method and instrument), reference standards and/or reference laboratories [2].

These schemes are designed to monitor the laboratory performance in a retrospective manner using "blind" samples analyzed as if they were patient samples. The results are sent to the scheme organizer on a timely manner for statistical analysis; each laboratory will then receive a report comparing its performance with those of other participants of the same program. These schemes provide an external validation of the laboratory results and also constitute a valuable tool for the internal monitoring of the laboratory performance. This usually benefits the laboratory, its clients and furthermore the accreditation and/or regulation agencies.

Using these schemes as a tool for internal monitoring is not limited to the investigation of unacceptable results. Monitoring of all results (accepted and rejected), always evaluating the latest survey against previous ones, enables the

laboratory to identify deviations and tendencies which generate a disqualification, as such exposing potential problems related to the presence of a significant random error, a significant systematic error or even a human error.

The evaluation of the comprehensive report (summary) allows one to get insight into measurement techniques that may well be favored over others that may not have reported satisfactory analytical performance. Exhaustive reporting enables identification of differences among different measurement procedures for the same analyte, significant systematic error or issues related to poor reproducibility.

Traditional schemes (PT/EQA) tend to address only the analytical procedure (examination procedures), but some innovative schemes have recently been introduced to evaluate the pre-analytical activities as well as the post-analytical activities of the clinical laboratory [3].

It is also important to bear in mind that each scheme (PT/EQA) has limitations and that it is no appropriate using these schemes (PT/EQA) alone to evaluate laboratory quality. Therefore, it is necessary to underline that internal quality control (IQC), PT/ EQA and other methods must be utilized to supervise and improve clinical laboratory performance.

WHAT IS THE DIFFERENCE BETWEEN EQA AND PT?

Nowadays, definitions of external quality assurance programs (EQA) and proficiency test programs (PT) are indiscriminately used as valuable tools for the clinical laboratories services quality improvement [2]. Anyways, the main objectives of EQA are educational and can be supported by complementary elements like specific plans designed to extend the evaluation along all the steps of the test cycle, including the results interpretation. ISO / IEC 17043 Standard [4] refers on its development to the proficiency tests

but mentions the particularities of the external quality assessment programs applied on the clinical area: *“Some providers of Proficiency Testing (PT) in the medical area use the terms external quality assessment (EQA) for their proficiency tests programs, for wider programs, or for both”.*

At the same time ANNEX A (A.4) of the same standard states: *“EQA programs (such as those provided for clinical laboratory testing) offer a variety of programs for interlaboratory comparisons based on this traditional method of proficiency testing but with a wider application of the programs. Many EQA programs are designed to evaluate the volume of work of the laboratory and not only the testing process. Most EQA programs are on-going programs including a long term follow up of a laboratory performance. A typical characteristic of the EQA programs is that they instruct the participants and foster quality improvement. Comments of advisory an educational status is part of the report delivered to the participants aiming at this end”.*

According to a widely accepted definition, PT's are programs to evaluate the participants' performance concerning previously established criteria through interlaboratory comparisons [4]. A PT is a program by which multiple samples are regularly sent to the members of a group of participant laboratories for their analysis and/or identification. Each of the results informed by the laboratory is compared with those from other laboratories belonging to the same group or with an assigned value for a valid procedure.

STATISTICAL TREATMENT OF DATA

The results of the PT/EQA programs may be of many different formats, covering a wide range of data and following different statistical distributions. It is important that the design used by the provider of the PT/EQA is appropriate for the type and purpose of the PT/EQA scheme it is

organizing. Additionally, the design the PT/EQA provider uses must be thoroughly described to the participants. The preferred statistical techniques are those described on ISO 13528 Standard [5], although other valid approaches may be used [13].

The main statistical approaches used on PT/EQA programs are based on the normal distribution of data. However, it is common that, even though the set of results of the participants essentially reflect a normal distribution, at both ends of the distribution they are contaminated by a small proportion of extreme values. The original approach used by the PT/EQA providers (and still used by some PT/EQA programs) was using statistical testing to identify the presence of extreme values in the original data set. However, the most common approach currently used by PT/EQA providers, recommended by the ISO 13528 Standard [5], is using robust statistics [14, 15]. Robust statistics have the advantage of reducing the contribution of extreme values to calculated statistical parameters like the median and the standard deviation. A series of models apply robust statistics. Some of which are described by the ISO 13528 Standard [5].

One of the basic elements of any PT/EQA program is the performance evaluation of each participant. To this end, the PT/EQA program provider must basically establish two values, used to evaluate performance:

- Assigned Value
- Standard Deviation of the comparison group in the scheme

Additionally, the PT/EQA program provider is expected to offer a measurement uncertainty estimation associated to the assigned value and a statement of its metrology traceability when possible. This concept has been included on the ISO / IEC 17043 Standard [4]. The relevance, necessity and feasibility of this estimation will be determined by the PT/EQA scheme design.

Different methods can be used to establish these values [5, 16, 35]. To date there is no standardization or agreement as to which protocol to apply. The statistical design must be documented by the PT/EQA program provider and must be taken into account at the time of choosing a PT/EQA scheme. Knowing the measurement uncertainty associated with the estimation of the assigned value will be relevant at the time of evaluating how reliable the comparison group offered by the PT/EQA program is.

Assigned value

Essentially, there are, as described by the ISO 13528 Standard [5], five methods available to obtain the assigned value and its typical associated uncertainty:

1. Formulation
2. Certified reference materials
3. Reference values
4. Consensus values provided by a group of expert laboratories
5. Consensus value obtained from the participants

Using a consensus value, produced in each round of the PT/EQA program, based on the results obtained from the participants is widely spread along the clinical laboratory area (option 5). The consensus value is often estimated using robust statistical techniques. The consensus approach is clearly the most simple and, sometimes, for example, as it uses natural matrix samples, it is frequently the only way to establish an estimation of the actual value [31, 32].

As expected, this model has its limitations:

- a. A real consensus among the participants may not exist;
- b. The consensus may be biased by the general use of a flawed methodology and this bias will not be reflected on the standard uncertainty of the assigned value.

Standard deviation of the comparison group in the scheme

As described by the ISO 13528 Standard [5], there are essentially five approaches to determine the standard deviation of the comparison group in the PT/EQA that will be used to establish the acceptable range of the participants' results:

1. Prescribed value.
2. Perception.
3. Based on a general model.
4. Based on the results of a precision experiment.
5. Based on the data obtained from a round of a PT/EQA scheme.

As expected, at the clinical laboratory environment most of the providers of PT/EQA schemes resort to this last option (option 5). With this approach, the standard deviation to evaluate the PT/EQA program used in a round of a scheme derives from the results reported by the participants in the same round. It will be the robust standard deviation of the results reported by all the participants.

A disadvantage of this approach is that the value may substantially vary from one round to the other. This variation makes difficult the follow up by the laboratory of the standard deviation index (Z score) along different rounds in search of deviations and tendencies [13].

Measurement uncertainty associated to the assigned value estimation

As already mentioned, the PT/EQA programs for Clinical Laboratories providers generally resort to the participants consensus to obtain the assigned value and the standard deviation of the comparison group. In other cases, providers of PT/EQA programs resort to simpler statistic models, not based on robust statistics techniques [33].

We will work on an assigned value and a standard deviation established from the participants' consensus, using robust statistics techniques [40] like those described by ISO 13528 [5].

The uncertainty associated to the assigned value is calculated as (equation 1):

$$u(x_{pt}) = 1,25 \times S^* / \sqrt{p}$$

Equation 1

Where:

$u(x_{pt})$: Uncertainty associated to the estimation of the assigned value in the round

S^* : Robust Standard Deviation of the group of participants in the round

p : Amount of participants of the comparison group in the round

CRITICAL ASPECTS AT THE TIME OF SELECTING AN EXTERNAL QUALITY ASSESSMENT PROGRAM (EQA) OR A PROFICIENCY TESTING PROGRAM (PT)

We will evaluate critical aspects to take into account at the time of selecting an external quality assessment program. Not all the external quality assessment schemes offer the same. ISO 15189 [1] Standard states: *"It is recommended that the laboratory participates in external quality assessments programs which substantially comply with the relevant requirements of the ISO/IEC 17043 Standard [4]."* It is worth mentioning that the statistical treatment of data proposed by this standard complies with what ISO 13528 Standard establishes [5].

The critical aspects we are going to take under consideration (figure 1) are dealt with by these two standards [4,5].

Let's remember that a PT/EQA survey is carried out by sending a sample or set of samples by an organizer entity to a participant laboratory. The laboratory must process the samples in the

same manner it will do with routine samples, namely, as far as possible as if they were patients' samples. (Fig. 1)

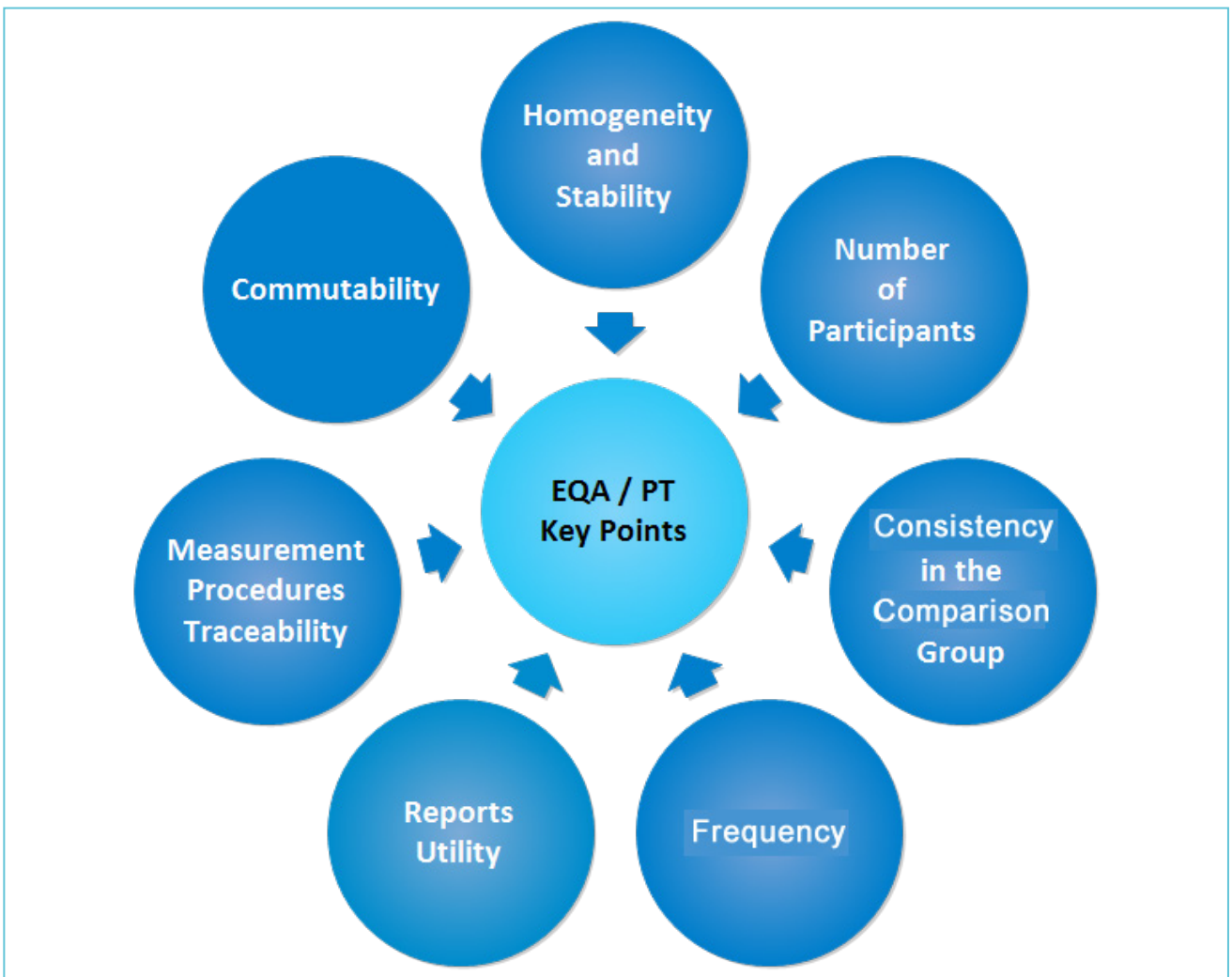
Traceability of the measurement procedures

A primary objective is to achieve that the obtained results over the same sample at comparable times in different laboratories using different measurement procedures are equivalent, within clinically significant limits, to enable optimal use of clinical practice guidelines to diagnose illnesses and an optimal patient's health care. When the results are obtained from non-standardized or harmonized measurement procedures, in different laboratories, it is expectable that their numerical value will be different or even, it is expectable that their clinical interpretation will be different.

Currently, standardization and harmonization of measurement procedures (traceability of calibration) are based on traceability criteria described in ISO 17511 [9], including five categories of reference systems. Categories 1, 2, and 3 stand for standardized measurement procedures. Category 4 stands for harmonized measurement procedures. In category 5 we find the measurement procedures where the manufacturer is solely responsible for the calibration traceability chain [6, 7, 8,17].

It is essential that clinical laboratories know the metrology traceability of their measurement procedures to group accurately in the PT/EQA programs. We must remember that most of PT/EQA program providers estimate the assigned value by the participant's results consensus of the comparison group. If we do not group correctly, it is very likely that the assigned value will not correspond to the best estimation of the true value of our measurement procedure taking into account its metrology traceability (Calibration Traceability).

Figure 1 PT/EQA key points



Commutability

The objective of the PT/EQA programs is to verify on a recurring basis that laboratory results comply with the quality requirements established according to the intended use of measurement procedures for the optimal patients' health care.

A key factor for the correct interpretation of the results of PT/EQA programs is the knowledge of the samples commutability and the procedure used to assign the true value for them. Commutable PT/EQA samples demonstrate the same numerical relationship among

different measurement procedures as the one expected for patients' samples. Non commutable PT/EQA samples present a bias, caused by matrix effects of unknown magnitude which limits the clinical interpretation of results [10, 11, 37, 41].

Concerning the commutability of the PT/EQA samples, ISO 17043 [4] states: *"It is convenient that the items proficiency testing coincide, in terms of matrix, measurand and concentrations, as much as is feasible, with the type of items or materials of the routine or calibration testing."*

Homogeneity and stability

Appropriate homogeneity and stability criteria must be established, based on the effect the absence of these characteristics would have on results and the evaluation of the participants' performance.

Sometimes it is not possible to submit the PT/EQA testing items to homogeneity and stability testing. An example would be when having limited material available to prepare the items for a PT/EQA testing. Sometimes the best available option is with materials which are not homogeneous or stable enough. In these cases, they still can be useful as PT/EQA testing items, as long as the uncertainty of the assigned values is taken into account during the results evaluation. In cases when the determination of homogeneity and stability is not feasible, the proficiency testing provider must demonstrate that the procedures used to get together, produce, pack and distribute the PT/EQA testing items are enough for the purpose of the proficiency testing.

The procedures to evaluate the homogeneity and stability must be documented and implemented, when corresponding, according to appropriate statistical designs. It must be demonstrated that the PT/EQA testing items are stable enough to assure that they will not suffer significant changes along the performance of the PT/EQA testing, including the storage and transport conditions [4].

Number of participants

The amount of participants, which integrate a comparison group, is a limiting factor of the scheme usefulness. When ISO 13528 [5] sets the equation to estimate the uncertainty associated to assigning a consensus value (Equation 1), it refers to $p > 10$, where p is the amount of participants in the comparison group.

It is interesting the information presented by the IUPAC guidelines [18]. It was designed to

establish guidelines for PT/EQA schemes with few participants. It establishes 30 as a minimum p (on the IUPAC guidelines $P=N$), considering that groups with $20 \leq p < 30$ must be evaluated in a critical manner to judge its usefulness.

At the PT/EQA program providers' in clinical environment we frequently find groups with less than 10 participants ($5 \leq p < 10$). It is very likely that an appropriate statistical analysis in these cases would indicate that the comparison group is not consistent.

Consistency of the pair comparison group

The assigned value "x" has a standard uncertainty u_x which depends on the method used for its estimation.

In general, at the clinical laboratory environment the assigned value (x) is estimated by the participants' consensus using robust statistics [5] according to the following equation:

$$u(x_{pt}) = 1,25 \times S^* / \sqrt{p}$$

Equation 1

The standard deviation for PT/EQA testing " σ_{pt} " is used to evaluate the size of the laboratory bias estimations found in a scheme round of a PT/EQA scheme. At the clinical laboratory environment, the standard deviation for the evaluation of the PT/EQA program, used in a round of one scheme, derives from the results reported by the participants of the same round based on the use of robust statistical techniques [5].

Therefore, in our case:

$$\sigma_{pt} = S^*$$

Equation 2

If the standard uncertainty for the assigned value " $u(x_{pt})$ " is too big as compared to the standard deviation for PT/EQA scheme round (σ_{pt}), then, there is a risk that some laboratories

would receive signals of action and warning due only to the inaccuracy of the assigned value determination, not to an issue of the measurement procedure performance at the laboratory itself. This is the reason why PT/EQA schemes providers must report the uncertainty associated to the true value assignment [19].

We must consider if the following relationship occurs (equation 3):

$$u(x_{pt}) < 0,3 \times \sigma_{pt}$$

Equation 3

If we replace equation 2 in equation 3 we obtain (equation 4):

$$u(x_{pt}) < 0,3 \times S^*$$

Equation 4

If this relationship occurs, the uncertainty associated to the estimation of the assigned value is negligible and the comparison group can be considered acceptable [5].

Frequency of submissions

At the clinical laboratory environment, the frequency of submissions varies and depends on the specific area of the laboratory and the scheme provider. For example, submissions at the clinical chemistry and hematology areas are usually monthly or fortnightly. At other areas like hemostasis or serology it is frequent that submissions are quarterly or bimonthly. The greater the amount of submissions per year, the greater usefulness it offers to the scheme.

RESULTS

Reports utility

PT/EQA testing reports must be clear and exhaustive and include information on the results of all the participants, together with

an indication of the individual participants' performance.

As we have already mentioned, the statistical treatment of data may be different among providers. However, we must remember that it is advisable that laboratories are able to participate in PT/EQA schemes accredited by ISO 17043 [4] or substantially complying with its guidelines. This standard establishes a series of requirements about the information that must be included on the reports.

Besides the information considered at the reports, their time of delivery is important. Reports must be available for participants within the established time [4].

THE SITUATION IN LATIN AMERICA

In the area, we will find laboratories of different sizes and varied complexity.

In general, in the Capital and/or big cities we will find a limited number of high complexity laboratories (considering the total amount of laboratories). These laboratories work under international standards (ISO, CAP, etc.) with a strong pressure on quality.

However the critical mass of laboratories in the area lives a very different situation. We frequently find very small laboratories and sometimes one-person laboratories. At these laboratories, due to different causes, there are flaws at quality level [20].

Regional regulations at quality level present inconveniences in many countries of the area and at the same time there are problems at the time of controlling its effective compliance.

If we consider the higher complexity laboratories we may see that they participate in PT/EQA programs voluntarily or complying with regulations according to the country considered. These laboratories participate in international and national schemes.

If we consider the many small laboratories of the area we will notice that voluntary participation in PT/EQA schemes is very low (remarkably low). When they participate they generally do in national or regional schemes.

At the same time, we have identified several flaws in interpreting results offered by the reports. We will deal with some of the critical aspects identified in the area concerning the participation in external quality assessment schemes.

Problems with the metrology traceability of the measurement procedures (calibration traceability)

There is a particularly critical situation in the area of clinical chemistry. Because of financial considerations, smaller laboratories tend to use reagents and calibrators of a commercial brand in instruments of a different commercial brand. Furthermore, not infrequently, they use calibrators of a third commercial brand. At this time of grouping in a PT/EQA scheme, inconveniences arise. Scheme providers need to open multiple comparison groups in order to cover all the possible combinations. As it is expectable, these groups count on very few participants, they are inconsistent groups and therefore the information provided by the reports has a limited value.

In other cases, the situation is so heterogeneous that scheme providers do not open individual groups and group these laboratories by method. Even though it is a single method, different commercial brands offer different traceability for the same method and comparison groups are again inconsistent.

These types of situations are a true inconvenience in the area.

Commutability

Considering the total amount of PT/EQA program providers in the area, only a few are accredited for the ISO 17043 Standard [4]. The

commutability of the samples is not always assured by the providers and therefore inconveniences related to matrix effects for specific measurement procedures may arise.

Consistency of the comparison group

Few local PT/EQA providers, considering the total amount of providers, report the uncertainty associated to the assigned value estimation, and besides, even if the information is available, laboratories do not frequently evaluate the consistency of the comparison group before making decisions. If we add that due to grouping issues (mixture of reagents and instruments) comparison groups with a very small number of participants ($p < 10$) are open or comparison groups with a very big p , but with an enormous S^* due to metrology traceability themes (calibration traceability) in the grouping by method or are massive (all the participants together in a single group), we come to the conclusion that it is crucially important to evaluate the consistency of the comparison groups before making decisions and this is not done in a routine way by laboratories.

Reactive behavior against proactive behavior

Laboratories' behavior faced to PT/EQA schemes is essentially reactive, namely, they react in front of exclusion. It means that laboratories specifically pay attention to rejected results, particularly to the ones from the last survey.

This as such generates non-conform products, i.e., patient results bear so huge an error that their clinical usefulness is invalidated.

Everyone working on analytical quality in the area intends that laboratories evaluate all the results, rejected and accepted, always reviewing the latest survey against previous ones. As such, through a correct interpretation of the laboratories reports, they could detect latent deviations and tendencies that still have not invalidated the

clinical usefulness of the routine results. If we achieve this change, laboratories will be able to anticipate potentially dangerous situations.

Acceptance and rejection criteria

In general, laboratories in the area trust the acceptance and rejection criteria established by the schemes providers. On numerous occasions, due to the reasons already stated, comparison groups are not consistent in several local schemes, with very big standard deviations that end up generating a big room for errors. The recommendation is that laboratories use the quality requirements that they must individually select for each measurement procedure to evaluate the measurement error of each individual survey. As such, to establish an acceptance criterion considering what is necessary for the measurement procedure considering its intended use.

Based on a set of surveys, using a valid statistical model [21], laboratory can estimate the bias of the measurement procedure (generally from 6 surveys). Once again the laboratory can evaluate the bias obtained in front of an established percentage (for example 50%) of the quality requirement to know if there is a clinically significant systematic error.

Underutilization of the information provided by the reports

Let's remember that the attitude of the laboratories towards the PT/EQA schemes is reactive, not proactive. If laboratories could evaluate the last survey against previous ones they could obtain valuable information to:

- Estimate biases on the measurement procedures [21].
- Integrate the information of these schemes with the information of the internal quality control to estimate the uncertainty of the

measurement procedures [21, 22, 23, 24, 28, 29, 30, 36].

- Estimate quality requirements according to the state of the art (metrology considerations) [25, 26].
- Carry out a follow up of the Z score considering several rounds along the time to detect deviations and tendencies [2].

Rejected results

It is less frequent in the area, except for accredited or certified laboratories according to different schemes, to record rejected results. Besides, it is less frequent that laboratories evaluate the impact of these non-conformities on the already liberated results and less frequent still is that take steps to recall results of concern [2, 27].

CONCLUSIONS

At the level of PT/EQA, there is much to be done by the participating laboratories and the scheme providers.

It is fundamental to increase the level of participation and at the same time to work on training to achieve that laboratories can use these tools for on-going improvement. To achieve this we must generate simple and useful information; work on training in a planned way to eradicate incorrect practices.

It can be stated that laboratories participate in PT/EQA schemes voluntarily or due to pressure by regulating authorities. They spend money and time but do not recover the investment because they underuse the information.

Bachelor degree careers at university level do not update their academic programs and it is very little what is learnt at quality level and less still concerning analytical quality. It is frequent that a professional is formed without having seen anything of this matter during his/her university career.

The economic-financial situation of the region is complicated and in some countries, critical.

Many professionals of the clinical laboratory environment or blood bank are formed and trained to be able to transmit the acquired knowledge at their work places to improve the analytical quality. At the time of trying to implement these improvements they frequently crash with the laboratory heads due to resource issues.

This hostility sometimes is due to lack of information, lack of understanding of the matter by the laboratory heads; and at other times is caused by the lack of appropriate resources.

The health system of the region has not incorporated the concept of quality and this is also true for analytical quality as a requirement for reliable laboratory results. They are not willing to pay for quality.

This concept of not paying for quality is attributed to the lack of resources, although I do want to mention that many times it is due to lack of training and knowledge on the subject.

There is a lot of available information to improve different aspects of the management of PT/EQA schemes [39].

Local PT/EQA schemes providers are very different. There are compliant schemes offering understandable schemes, with ISO 17043 [4] accreditation and others that have a lot to improve, for example, at the level of commutability of their samples, the statistical management of data and timing and particularly laboratory grouping [34, 38].

REFERENCES

1. ISO 15189 (2012) Medical laboratories—Requirements for quality and competence. ISO, Geneva.
2. CLSI GP-27 A2 (2007) Using proficiency testing to improve the clinical laboratory. CLSI, Wayne, PA
3. Sciacovelli L, Secchiero S, Zardo L, Plebani M The role of the External Quality Assessment. *Biochimica Medica* 2010; 20(2):160-4.
4. ISO/IEC 17043 (2010) Conformity assessment—general requirements for proficiency testing. ISO, Geneva.
5. ISO 13528 (2015) Statistical methods for use in proficiency testing by interlaboratory comparisons. ISO, Geneva.
6. Siekmanna L. "Metrological traceability – a concept for standardization in laboratory medicine". *Clin Chem Lab Med* 2013; 51(5): 953–957
7. Vesper HW, Thienpont LM. "Traceability in laboratory medicine". *Clin Chem* 2009; 55: 1067–1075.
8. Plebani M "Harmonization in laboratory medicine: the complete picture". *Clin Chem Lab Med* 2013; 51(4): 741–751.
9. ISO 17511 (2003) In vitro diagnostic medical devices -- Measurement of quantities in biological samples -- Metrological traceability of values assigned to calibrators and control materials. ISO, Geneva.
10. Miller GW, Jones GRD, Horowitz GL, and Weykamp C. "Proficiency Testing/External Quality Assessment: Current Challenges and Future Directions". *Clinical Chemistry* 57:12 (2011).
11. Miller GW, Myers GL, Rej R. "Why Commutability Matters". *Clinical Chemistry* 52, No. 4, 2006.
12. Armbruster D, Miller RR. "The Joint Committee for Traceability in Laboratory Medicine (JCTLM): A Global Approach to Promote the Standardisation of Clinical Laboratory Test Results. *Clin Biochem* 2007; 105- 114.
13. EURACHEM Selection, Use and Interpretation of Proficiency Testing (PT) Schemes. Second edition (2011). EEE-PT WG.
14. Analytical Methods Committee - Robust Statistic Part I & II. *Analyst* 1989; 114, 1693-1702.
15. Thompson, M. and Ellison, S.L.R., "Fitness for purpose – the integrating theme of the revised Harmonised Protocol for Proficiency Testing in Analytical Chemistry Laboratories". *Accred.Qual. Assur.* 2006;11, 373-378,
16. Tholen, D.W., "Statistical treatment of proficiency testing data". *Accred. Qual. Assur.*, 3 (1998),362-366.
17. Miller WG, Myers GL, Gantzer ML, Kahn SE, Schönbrunner ER, et al. "Roadmap for Harmonization of Clinical Laboratory Measurement Procedures". *Clinical Chemistry* 2011;57:8 1108-1117.
18. IUPAC/CITAC Guide: Selection and use of proficiency testing schemes for a limited number of participants—chemical analytical laboratories (IUPAC Technical Report). *Pure Appl. Chem.*, Vol. 82, No. 5, pp. 1099–1135, 2010.
19. ISO/IEC Guide 43-1: 1997, Selection and use of proficiency testing scheme by laboratory accreditation bodies.

20. Westgard JO. "Basic QC Practices. Spanish". Translation by GA Migliarino Third Edition. Madison WI, 2010. IISBN 978-1-59425-098-9.
21. Nordtest Report TR 537, Version 3.1 (May 2012), Handbook for Calculation of Measurement Uncertainty in Environmental Laboratories. Nordic Innovation.
22. Ellison SLR and Williams A (Eds). Eurachem/CITAC guide: Quantifying Uncertainty in Analytical Measurement, Third edition, (2012) ISBN 978-0-948926-30-3. Available from www.eurachem.org.
23. EUROLAB Technical Report 1/2006. "Guide to the Evaluation of Measurement Uncertainty for Quantitative Test Results". Paris, France.
24. EUROLAB Technical Report 1/2002 –"Measurement Uncertainty in testing". Paris, France.
25. Hyltoft Petersen P, Fraser CG. Strategies to set global analytical quality specifications in laboratory medicine: 10 years on from the Stockholm consensus conference. *Accred Qual Assur* 2010; 15:323–330.
26. Klee GG. Establishment of outcome-related analytical performance goals. *Clin Chem* 2010;56:714–722.
27. CLSI. "Nonconforming Event Management"-Second Edition. CLSI QMS11-A2. Wayne; PA: Clinical and Laboratory Standards Institute; 2015.
28. EUROLAB Technical Report 1/2007. "Measurement uncertainty revisited: Alternative approaches to uncertainty evaluation". Paris, France
29. White GH. Basics of estimating measurement uncertainty. *Clin Biochem Rev* August 2008; 29(Suppl. (i)):S53.
30. BIPM, IEC, IFCC, et al. Evaluation of measurement data—guide to the expression of uncertainty in measurement GUM. JCGM 1002nd Ed. 2008 [<http://www.bipm.org/>].
31. Koch M and Baumeister F. "On the use of consensus means as assigned values". *Accred Qual Assur* 2012;17:395–398.
32. Heydorn K. "The quality of consensus values". *Accred Qual Assur* 2013; 18:243–245.
33. Ya L. Xiao, Chuan B. Zhang, Hai J. Zhao, Feng F. Kang, Wei Wang, Kun Zhong, Shuai Yuan, Zhi G. Wang. "Application of ISO 13528 robust statistical method for external quality assessment of blood glucose measurements in China". *Accred Qual Assur* 2014; 19:397–401.
34. Gun-Munro J. "The challenges and benefits of implementing the requirements of ISO/IEC 17043 by PT/EQA providers". *Accred Qual Assur* 2012; 17:363–370.
35. Srnková J and Zbiral J. "Comparison of different approaches to the statistical evaluation of proficiency tests". *Accred Qual Assur* 2009; 14:467–471.
36. Patriarca M, Chiodo F, Castelli M, Menditto A. "Estimates of uncertainty of measurement from proficiency testing data: a case study". *Accred Qual Assur* 2006; 11: 474–480.
37. Unsal I, Coskun A, Serteser M, Inal TC, Ozpinar A. "Toward standardization of quality assessment in laboratory medicine by using the same matrix samples for both internal and external quality assessments". *Accred Qual Assur* 2010;15:621–627.
38. Lehmann C. "Accrediting PT/EQA providers to ISO/IEC 17043". *Accred Qual Assur* 2012; 17:371–374.
39. de Albano FM, ten Caten CS. "Proficiency tests for laboratories: a systematic review". *Accred Qual Assur* 2014; 19:245–257.
40. Kang F, Wang W, Zhang CB, Wang ZG. "Establishment of an assigned value and its uncertainty for tumour markers in proficiency testing in China". *Accred Qual Assur* 2013; 18:435–439.
41. Kim SY, Chun S, Lee W and Min WK. "Commutability of proficiency testing (PT): status of the matrix-related bias in general clinical Chemistry". *Clin Chem Lab Med* 2013; 51(8): e169–e173.