Harmonization of External Quality Assessment Schemes and their role - clinical chemistry and beyond

Ceriotti, F.; Cobbaert, C.

Opinion Paper

Ferruccio Ceriotti* and Christa Cobbaert

Harmonization of External Quality Assessment Schemes and their role – clinical chemistry and beyond

https://doi.org/10.1515/cclm-2018-0265
Received March 12, 2018; accepted April 5, 2018; previously published online May 1, 2018

Abstract: The article tries to reply to the following three questions: *Are* External Quality Assessment Schemes *(EQAS) really fit for purpose? Are all schemes equivalent and sufficiently harmonized? Is the role of EQAS similar and necessary in all branches of laboratory medicine?* Although the reply to the first two questions is, unfortunately, negative for several reasons (lack of commutable material with reference method values, EQAS with different scopes, etc.), the reply to the third one is positive: EQAS are a necessary source of information on trueness and accuracy and must be fully developed for all the branches of the clinical laboratory.

Keywords: accreditation; accuracy; External Quality Assessment Schemes.

Introduction

External Quality Assessment Schemes (EQAS) play a central role in laboratory medicine [1]. In Europe, successful participation in EQA programs is set as a mandatory requisite by country-specific accreditation bodies such as the Italian Accreditation Body (Accredia) and the Board of Accreditation (RvA) in the Netherlands, to have access to the ISO 15189 accreditation according to the flexible scope approach [2]. ISO 15189:2012 dedicates a full paragraph 5.6.3 to interlaboratory comparisons [3]. EQAS are (or should be) the final step to confirm the success of the traceability chain [4] and are also seen as part of the surveillance of the performances of the In Vitro Diagnostic medical tests [5]. Thus, although there is no doubt about their importance, the questions are as follows: Are EQAS really fit for purpose? Are all EQAS equivalent and sufficiently harmonized? Is their role similar and necessary in all branches of laboratory medicine?

***Corresponding author: Ferruccio Ceriotti,** Clinical Laboratory, Fondazione IRCCS Ca' Granda Ospedale Maggiore Policlinico, Via Francesco Sforza 28, Milan 20122, Italy, Phone: +390255032876, Fax: +3902255032219, E-mail: ferruccio.ceriotti@policlinico.mi.it
Christa Cobbaert: Department of Clinical Chemistry and Laboratory Medicine, Leiden University Medical Center, Leiden, The Netherlands

Present situation

Miller already in 2009 clearly indicated advantages and limitations of participation in an EQAS [6] and 2 years later, together with some coworkers designed a table classifying the EQAS in six categories according to their evaluation capabilities [7]. Category 1 schemes, based on commutable materials with target values assigned by reference methods, are the schemes that have full evaluation capabilities, and the authors conclude that, ideally, all EQAS should be category 1 schemes but are rare because of a number of constraints, including costs, technical difficulties and lack of awareness of the importance of these characteristics. Infusino et al. [8] proposed a further subdivision of EQAS categories 1 and 2 in A and B: A, when using the higher order models for performance specifications identified by the EFLM Milan Strategic Conference [9], B if using lower levels.

Unfortunately, with the exception of a few examples [10–14], category 1 schemes remain scarce and so the real benefit of participation in EQAS remains modest. Other areas of laboratory medicine, where the experience of EQAS is more recent and/or deals with different type of measures, seem to benefit more from the participation in EQAS as reported for immunophenotyping [15] or microbiology [16]. In the Netherlands, the Calibration 2.000 initiative is a nationwide program, established in 1998, which strived for either standardization or harmonization of tests in domains beyond clinical chemistry such as hematology, endocrinology, coagulation, immunology, TDM, parasitology, etc. [11]. The output of its research was fed into the EQAS of the Dutch SKML and is reviewed in

this special issue [11]. Notwithstanding the achievements of Calibration 2.000 in the past 20 years, especially in the clinical chemistry domain, EQAS for the other domains are not yet at the category 1 level.

Open questions

Are EQAS really fit for purpose? The concept of traceability and trueness-based grading is typical for clinical chemistry and all other laboratory disciplines producing numerical results on a continuous scale. In the case that a reference measurement system is in place, category 1 EQAS should be implemented for an effective improvement of the analytical quality. On the contrary, participation in Miller's category 5 schemes only allows to make peer group comparisons [6]. As briefly indicated above, for real help in improving clinical laboratories' accuracy, verifying the calibration-related and/or specificity bias as well as imprecision of specific tests or analytical systems, the vast majority of the presently available EQAS are not fit for purpose. Exception to the rule is the EQAS and MUSE scoring system from the Dutch SKML [17], which allows to systematically distinguish bias and imprecision of tests. In other EQAS, the core of the problem is the control materials in the schemes that are not commutable or whose commutability is not verified or demonstrated. The verification of commutability is extremely demanding both in terms of costs and time for realization [18–20] and thus not affordable for small schemes. Moreover, it is quite clear that manipulation of control materials to stabilize them (addition of exogenous substances, lyophilization, etc.) and to obtain predefined measurand concentrations often brings along non-commutability [21, 22]. The production of suitable control materials requires research and investments and their distribution in frozen form is also expensive. Another important cost is the value assignment using reference methods. However, assigning target values with reference methods to non-commutable control materials is not only useless but also dangerous because it may cause biases that do not exist but are only caused by non-commutability, thus inducing undue corrective actions. All these reasons explain why the development of category 1 or 2 EQAS is a significant challenge that therefore progresses slowly, notwithstanding all the literature evidence regarding the absolute necessity of commutability of EQA materials in accuracy-based EQAS. Another relevant drawback that impairs the efficacy of some schemes is the use of inappropriate analytical quality specifications (APS) to evaluate the performances of the laboratories. Only when applying adequate APS, according to the model proposed by the first EFLM strategic conference [9], there is an effective relationship between the test performances and the level of quality needed. EQAS should help laboratorians to make informed decisions about test accuracy by being enabled to verify the adequate implementation of the metrology concept through IVD manufacturers. Unfortunately, most EQAS are not equipped to do this and do not allow a proper verification of the correct implementation of the traceability concept, within allowable measurement uncertainties.

The situation is different when looking at EQAS devoted to assessing performance in pre- or postanalytical phases by circulating questionnaires or clinical cases or when distributing images or slides to assess the competence of the professionals [23–26]. In these cases, usually the improvement obtained is defined, even if it is difficult to quantify it and to demonstrate an objective efficacy.

Are all schemes equivalent and sufficiently harmonized? Also this question has a negative reply. Apart from the different evaluation capabilities depending on the type of EQA materials used, there are several other reasons that create a very large heterogeneity in the EQA programs. Jones and coworkers [27] well identify these reasons that are summarized in six groups: (a) the nature of the EQAS material, including its commutability, which may affect the result interpretation; (b) the procedure used to assign the target value; (c) the data set to which performance specifications are applied (i.e. single result, series of data); (d) the analytical property being assessed (i.e. total error, bias, imprecision); (e) the rationale for the selection of the performance specification; and (f) type(s) of model used to set performance specifications. The problem of the evaluation criteria is very relevant, and their setting depends on the main scope of the program: if it has a regulatory impact, the criteria are looser because the scope is to identify the very poor performer; if it has an educational scope, the criteria are tighter because failing to meet them does not automatically imply sanctions but only a remedial action by the laboratory. This problem exists also outside clinical chemistry as indicated by Olson and coworkers for coagulation [28]. The situation may improve with the implementation on a larger scale of accreditation for EQA providers based on ISO 17043:2010 [29]. Unfortunately, this ISO standard defines no requirements regarding commutability or performance specifications. In this situation of great heterogeneity of EQAS, just the participation in any type of EQAS is not sufficient in terms of verifying precision and metrological traceability of test results as well as compliance with ISO 15189:2012. Therefore, inspectors need to verify also this aspect [30].

Is the role of EQAS similar and necessary in all branches of laboratory medicine? The reply to this last question is

absolutely positive; no diagnostic discipline of the clinical laboratory can operate its test performance without EQAS from independent third parties. EQAS are essential tools for evaluating test result equivalence by comparing test results to value-assigned trueness verifiers in case of commutable EQAS (i.e. category 1 EQAS) or to peer group means in case of category 5 EQAS, and for detecting problematic, non-traceable tests, which can be traced back to specific IVD manufacturers. EQAS are very important tools for standardization and harmonization of laboratory practices. Paradoxically in some laboratory areas, apparently more complex than clinical chemistry, the problems of materials and accuracy evaluation seem less critical, e.g. for microbiology where microorganism identification is the gold standard and both traditional and molecular biology-based techniques can be successfully tested, with very good results [31]. Yet Schuurs et al. [31] demonstrate that harmonization of PCR-based detection of intestinal pathogens is still in its infancy.

We conclude that harmonization of EQAS has still a long way to go, and much technical and organizational work has to be done, but important milestones indicating the way to follow have been defined [1, 7, 9, 11, 18]. Intensive collaborations or alliances between country-specific EQA organizations under the umbrella of the European Organization for External Quality Assurance in Laboratory Medicine are urgently needed, as well as efforts to merge EQAS in countries where different schemes for the same measurands are in use. These efforts should allow to develop jointly affordable and sustainable category 1 EQA schemes by sharing expertise and enlarging market share in medical laboratories.

Author contributions: All the authors have accepted responsibility for the entire content of this submitted manuscript and approved submission.
Research funding: None declared.
Employment or leadership: None declared.
Honorarium: None declared.
Competing interests: The funding organization(s) played no role in the study design; in the collection, analysis, and interpretation of data; in the writing of the report; or in the decision to submit the report for publication.

References

1. Braga F, Pasqualetti S, Panteghini M. The role of external quality assessment in the verification of in vitro medical diagnostics in the traceability era. Clin Biochem 2018 Feb 9. pii: S0009-9120(18)30003-1. doi: 10.1016/j.clinbiochem.2018.02.004. [Epub ahead of print].
2. RT-26 Rev 05 Requirements for accreditation with flexible scope https://www.accredia.it/documento/rt-26-rev-05-prescrizioni-per-laccreditamento-con-campo-di-accreditamento-flessibile/. Accessed: 18 Apr 2018.
3. ISO 15189:2012 Medical laboratories – requirements for quality and competence. Geneva: International Organisation for Standardisation (ISO), 2012.
4. Braga F, Panteghini M. Verification of in vitro medical diagnostics (IVD) metrological traceability: responsibilities and strategies. Clin Chim Acta 2014;432:55–61.
5. EN 14136:2004. Use of external quality assessment schemes in the assessment of the performance of in vitro diagnostic examination procedures. http://esearch.cen.eu/esearch/extended-search.aspx. Accessed: 18 Apr 2018.
6. Miller WG. The role of proficiency testing in achieving standardization and harmonization between laboratories. Clin Biochem 2009;42:232–5.
7. Miller WG, Jones GR, Horowitz GL, Weykamp C. Proficiency testing/external quality assessment: current challenges and future directions. Clin Chem 2011;57:1670–80.
8. Infusino I, Frusciante E, Braga F, Panteghini M. Progress and impact of enzyme measurement standardization. Clin Chem Lab Med 2017;55:334–40.
9. Sandberg S, Fraser C, 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.
10. Cobbaert C, Weykamp C, Franck P, de Jonge R, Kuypers A, Steigstra H, et al. Systematic monitoring of standardization and harmonization status with commutable EQA-samples – five year experience from the Netherlands. Clin Chim Acta 2012;414:234–40.
11. Jansen RT, Cobbaert CM, Weykamp C, Thelen M. The quest for equivalence of test results: the pilgrimage of the Dutch Calibration 2.000 program for metrological traceability. Clin Chem Lab Med 2018;56:1673–84.
12. Burdette CQ, Camara JE, Nalin F, Pritchett J, Sander LC, Carter GD, et al. Establishing an Accuracy Basis for the Vitamin D External Quality Assessment Scheme (DEQAS). J AOAC Int 2017;100:1277–87.
13. Weykamp C, Secchiero S, Plebani M, Thelen M, Cobbaert C, Thomas A, et al. Analytical performance of 17 general chemistry analytes across countries and across manufacturers in the INPUtS project of EQA organizers in Italy, the Netherlands, Portugal, United Kingdom and Spain. Clin Chem Lab Med 2017;55:203–11.
14. González-Lao E, Díaz-Garzón J, Corte Z, Ricós C, Perich C, Álvarez V, et al. Category 1 external quality assessment program for serum creatinine. Ann Transl Med 2017;5:133.
15. Bainbridge J, Rountree W, Louzao R, Wong J, Whitby L, Denny TN, et al. Laboratory accuracy improvement in the UK NEQAS leucocyte immunophenotyping immune monitoring program: an eleven-year review via longitudinal mixed effects modeling. Cytometry B Clin Cytom 2018;94:250–6.
16. Boeras DI, Peeling RW, Onyebujoh P, Yahaya AA, Gumede-Moeletsi HN, Ndihokubwayo JB. The WHO AFRO external quality assessment programme (EQAP): Linking laboratory networks through EQA programmes. Afr J Lab Med 2016;5:560.

17. Thelen MH, Jansen RT, Weykamp CW, Steigstra H, Meijer R, Cobbaert CM. Expressing analytical performance from multi-sample evaluation in laboratory EQA. Clin Chem Lab Med 2017;55: 1509–16.
18. Miller WG, Schimmel H, Rej R, Greenberg N, Ceriotti F, Burns C, et al. IFCC Working Group Recommendations for Assessing Commutability Part 1: General Experimental Design. Clin Chem 2018;64:447–54.
19. Nilsson G, Budd JR, Greenberg N, Delatour V, Rej R, Panteghini M, et al. IFCC Working Group Recommendations for Assessing Commutability Part 2: Using the Difference in Bias Between a Reference Material and Clinical Samples. Clin Chem 2018;64:455–64.
20. Budd JR, Weykamp C, Rej R, MacKenzie F, Ceriotti F, Greenberg N, et al. IFCC Working Group Recommendations for Assessing Commutability Part 3: Using the Calibration Effectiveness of a Reference Material. Clin Chem 2018;64:465–74.
21. Franzini C. Commutability of reference materials in clinical chemistry. J Int Fed Clin Chem 1993;5:169–73.
22. Rej R. Proficiency testing, matrix effects, and method evaluation. Clin Chem 1994;40:345–6.
23. Sølvik UØ, Bjelkarøy WI, Berg KV, Saga AL, Hager HB, Sandberg S. Intensive educational efforts combined with external quality assessment improve the preanalytical phase in general practitioner offices and nursing homes. Clin Chem Lab Med 2017;55:1857–64.
24. Lu Y, Gonzales G, Chen SH, Li H, Cai YC, Chu YH, et al. Urgent needs in fostering neglected tropical diseases (NTDs) laboratory capacity in WHO Western Pacific Region: results from the external quality assessment on NTDs diagnosis in 2012–2015. Infect Dis Poverty 2017;6:106.
25. Bezzegh A, Takács I, Ajzner É. Toward harmonization of interpretive commenting of common laboratory tests. Clin Biochem 2017;50:612–6.
26. Secchiero S, Fogazzi GB, Manoni F, Epifani M, Garigali G, Plebani M. The Italian External Quality Assessment (EQA) program on urinary sediment: results of the period 2012–2015. Clin Chem Lab Med 2015;53 Suppl 2:s1495–502.
27. Jones GR, Albarede S, Kesseler D, MacKenzie F, Mammen J, Pedersen M, et al. Analytical performance specifications for external quality assessment – definitions and descriptions. Clin Chem Lab Med 2017;55:949–55.
28. Olson JD, Jennings I, Meijer P, Bon C, Bonar R, Favaloro EJ, et al. Lack of grading agreement among international hemostasis external quality assessment programs. Blood Coagul Fibrinolysis 2018;29:111–9.
29. ISO 17043:2010. Conformity assessment – general requirements for proficiency testing. Geneva: International Organisation for Standardisation, 2010.
30. Sciacovelli L. External Quality Assessment Programs in the context of ISO 15189 Accreditation. Clin Chem Lab Med 2018.
31. Schuurs TA, Koelewijn R, Brienen EA, Kortbeek T, Mank TG, Mulder B, et al. Harmonization of PCR-based detection of intestinal pathogens: experiences from the Dutch external quality assessment scheme on molecular diagnosis of protozoa in stool samples. Clin Chem Lab Med 2018;56:1722–7.