Clin Chem Lab Med 2012;50(9):1547–1558 © 2012 by Walter de Gruyter • Berlin • Boston. DOI 10.1515/cclm-2012-0003

Opinion Paper

Collective opinion paper on findings of the 2011 convocation of experts on laboratory quality

Ortwin Adams[1], Greg Cooper[2], Callum Fraser[3], Michael Hubmann[4], Graham Jones[5], Mario Plebani[6], Oswald Sonntag[7,*] and Michel Vaubourdolle[8]

[1] Universitätsklinikum Düsseldorf, Institut für Virologie, Düsseldorf, Germany
[2] W. Gregory Cooper LLC, Denton, TX, USA
[3] Scottish Bowel Screening Centre, Dundee, Scotland, UK
[4] MZ Labor, Feldkirch, Austria
[5] St Vincent's Hospital Sydney Ltd, Sydney, Australia
[6] Department of Laboratory Medicine, University Hospital of Padua, Padua, Italy
[7] Bio-Rad Laboratories GmbH, Munich, Germany
[8] Saint-Antoine Hospital, Paris, France

Abstract

In April of 2011, Bio-Rad Laboratories Quality System Division (Irvine, CA, USA) hosted its third annual convocation of experts on laboratory quality in the city of Salzburg, Austria. As in the past 2 years, over 60 experts from across Europe, Israel, USA and South Africa convened to discuss contemporary issues and topics of importance to the clinical laboratory. This year's conference had EN/ISO 15189 and accreditation as the common thread for most discussions, with topics ranging from how to meet requirements like uncertainty to knowledge gained from those already accredited. The participants were divided into five discussion working groups (WG) with assigned topics. The outcome of these discussions is the subject of this summary.

Keywords: accreditation; ISO 15189; proficiency testing; quality control; quality specifications; traceability; uncertainty of measurement; uncertainty.

Introduction

ISO 15189:2007 (ISO 15189) (1) was first introduced to the medical laboratory community in 2003. Nearly 60 countries participated at one point or another in its development over a period of years. The standard was created because those laboratories that wished to have and maintain certification/accreditation to some recognized standard, chose ISO 17025 which was specific to calibration and testing laboratories. While many of the quality management system (QMS) requirements are transferrable to medical laboratories, the technical requirements were not a perfect fit. So ISO Technical Committee 212 (ISO/TC 212) undertook development of a practice standard that would be specific for medical laboratory practice. Acceptance was initially slow until various countries and professional societies became aware of the standard, understood its implications and either promoted or required its adoption. Acceptance was fairly quick in Australasia followed by Canada and Europe. In some isolated cases, countries adopted "their version" of ISO 15189 and so the laboratories in these countries would not necessarily comply with the official ISO version or be recognized by ISO or those who obtain accreditation under the official version.

One question that has persisted throughout recent years surrounding the discussion of laboratory quality in particular and ISO 15189 in general continues to be what level of quality is truly appropriate for medical testing. Various governments, regulators, professional societies, standards organizations, such as ISO and medical laboratory experts have all, at some point in time, advised or directed laboratories to define quality expectations for analytical testing. Competing sources of total allowable error specifications abound. The US (and others) uses CLIA proficiency testing limits (2), but these limits are not intended for use as overall quality specifications. In fact, some of the limits are too broad when compared to contemporary method capabilities. In contrast, use of biological variation for some analytes can result in total allowable error specifications that are difficult to attain. Moreover, the requirement in ISO 15189 for laboratories to estimate and make available the measurement uncertainty for each measurand assayed has been confounding for many laboratories because there is no international agreement on how this measure is to be reasonably calculated by medical laboratories and for biological measurands.

Since its introduction in ISO 15189 in 2003, Uncertainty of Measurement for medical laboratories has been controversial. Anecdotally, many laboratories questioned the need for uncertainty and some felt that metrologists and the framers of ISO 17025 conspired to force medical laboratories to estimate uncertainty for biological samples. In some sectors there have been outcries that this calculation is burdensome, provides no perceived value and the cost/benefit ratio is much

*Corresponding author: Oswald Sonntag, Bio-Rad Laboratories GmbH, Heidemannstrasse 164, 80939 Munich, Germany
E-mail: Oswald_Sonntag@Bio-Rad.com
Received January 5, 2012; accepted February 6, 2012; previously published online March 16, 2012

too high. Some believe uncertainty unfairly creates an image with physicians and care givers, unfamiliar with the concept, that laboratories are not accurate.

ISO 15189 also does not give specific direction for qualitative tests but the requirements set in Section 5.6 *Assuring quality of examination procedures* still apply to accredited laboratories. Performance evaluations of such tests are difficult to evaluate because interpretation is often subjective and results are expressed as attributes (e.g., positive, negative, 3+). These tests may involve color development on a strip or porous cartridge/slide while others may be performed on an automatic analyzer consisting of some reaction between the measurand and reagents creating positive or negative outputs based on some cut-off value that can be recorded and characterized statistically. Regardless, there is not much in the literature that guides laboratories about how to evaluate the performance of these tests.

All these aspects have been discussed in the convocation in individual WGs to create some proposals to deal with these issues. Finally, the proposals were presented and discussed in the plenary session. The summarized results are reported below.

Results

Using ISO 15189 as the key to quality: practical experience and peer guidance for success

The WG examined the ISO 15189 accreditation process to identify what processes helped with accreditation and potential pitfalls. It was reported that nearly 50% of the participants, representing seven European countries, have already been accredited to ISO 15189. The others had undertaken quality initiatives with hopes of accrediting in the future.

There are many reasons for laboratories to seek accreditation including: a) improving quality; b) improving management processes; c) to meet local or national regulations; and d) gaining international recognition.

However, there is little or no published evidence of measurable improvements achieved in laboratories that have accredited to ISO 15189 in the medical area. Some participants have anecdotally claimed that laboratories do not want to publically admit to peers and users that they may have not been as good as they should have been or, more importantly, may have had analytical failures that resulted in harm to the patient. Others do not want to admit to wasting valuable resources under their prior system.

Members of the WG identified four steps to obtaining accreditation. These are: a) laboratory (management) recognition of the standard and acceptance; b) implementing a unique quality management system (QMS); c) achieving accreditation for a part of the laboratory scope of work; and d) obtaining final accreditation of the total laboratory scope of work.

The WG suggest that a successful accreditation event can be traced directly to management involvement and commitment combined with a positive attitude and leadership, which is most important. The laboratory director must be intimately involved. It also requires the laboratory staff to work as a team and to avoid turf wars. Moreover, it is also important to lead people to embrace change. Management should recognize that involving all staff in change eases later acceptance.

Getting started is always a big question mark for those unfamiliar with ISO accreditation. Members of the discussion WG advised that looking to colleagues with ISO experience or other accredited laboratories for advice and direction can be of big help. Also, private consultants are available to help guide the process. Regardless, the discussion WG felt that first steps should include: a) appointment of a quality manager; b) competency assessment for all tasks and workstations (gap analysis); c) standardization of all procedures in large laboratories; d) implementation of management programs for internal QC, PT and EQA; and e) ensure traceability of calibration.

But even with good preparations, challenges can be encountered including: a) achieving and maintaining competence; b) moving to a quality culture and making sure all staff understand the process and the final outcome; c) managing work time to include time for quality at all levels of the organization; d) allowing the quality system and document control system to grow and evolve as conditions in the laboratory change; e) managing a QMS without a QMS software; f) qualifying external vendors when a department outside the laboratory makes those decisions; and g) keeping up with changes in product inserts when they occur.

Laboratories new to ISO should expect to take time to achieve accreditation. Participants felt that 3–4 years would be typical but if a laboratory has good systems in place and once obtained then laboratory needs to focus on maintaining quality and evolving their QMS, keeping in mind that a QMS must be flexible and able to grow, change, evolve as laboratory needs and conditions change. Some of the maintenance challenges include: a) handling of complaints (complaints and non-conformities need to be addressed quickly and responsively); and b) internal audit plans (laboratories need to have a robust internal audit program that helps identify underlying non-conformities that could lead to failure in the system or processes. Internal audits lead to corrective or preventive actions).

System and process failures can be mitigated by: a) using independent controls – when possible; b) having EQA results that are traceable to internal QC results (e.g., matrix, concentration etc.) – when possible; c) calculating means and standard deviations and CV for the laboratory reflecting actual performance; and d) documenting actions taken to remediate failures or potential failures (corrective and preventive action).

While a majority of the focus on quality is typically on QC plans, ISO also sets other requirements for assuring the quality of examination procedures by requiring laboratories to participate in external quality assessment. The WG identified several important observations about EQA including its purpose and important characteristics to look for in an EQA program.

Purpose of EQA is: a) evaluating method performance; b) means to monitor diagnostic devices as a part of vigilance;

c) resource for continuing education for participating laboratories; and d) means to help with test interpretation and advice to clinicians.

Most important EQA characteristics are: a) target values with metrological traceability; b) commutability; c) values near decision limits; d) high number of participants; e) homogeneity of methods within the group; f) statistical treatment and comments; and g) frequency of samples.

Defining appropriate analytical objectives in the laboratory

Many different requirements are identified in ISO 15189 (1). Particularly germane to the discussion held by the WG charged with debating aspects of setting appropriate analytical objectives in the laboratory is the following: *Clause 5.5.4 – Performance specifications for each procedure used in an examination shall relate to the intended use of that procedure*.

Thus, ISO 15189 mandates that laboratories set analytical quality specifications (AQS), sometimes termed analytical goals or analytical performance goals, for each examination and that these be delineated with respect to clinical needs. This is undoubtedly a difficult task since the results of examinations are used in many clinical settings including in diagnosis, monitoring, case-finding and screening. They are also used for other purposes including in education, teaching and training, and in research, development and clinical trials. Moreover, in order to assist in the interpretation of test results, laboratories provide tools including population-based reference values, reference change values and clinical decision limits, and other data, such as analytical detection limit. It is vital to recognize that the analytical reliability performance characteristics affect all of these, as outlined recently (3) and thus the analytical quality required to facilitate optimal patient care must be specified.

AQS are required for a number of purposes in the establishment of a new method (4), are necessary pre-requisites for quality planning, that is, based upon the AQS and the actual performance achieved in practice, deciding the number of QC samples to be analyzed and the rules that should be applied for acceptance or rejection of analytical runs (5) and are needed for use in proficiency testing (PT) or external quality assessment schemes (EQAS) so as to provide information to laboratories on whether their analytical quality is acceptable (6).

All members of the WG did use objectively set AQS in a variety of aspects of quality management in laboratory medicine. The hierarchy agreed at the consensus conference held in 1999 (7) under the auspices of the IUPAC, IFCC and WHO, based in large part upon a model published shortly before the conference (8), was used by all, namely: Level 1: Assessment of the effect of analytical performance on specific clinical decision-making in specific clinical situations. Level 2: Assessment of the effect of analytical performance on general clinical decision-making. 2A: General quality specifications based on biological variation. 2B: General quality specifications based on medical opinions. Level 3: Professional recommendations. 3A: Guidelines from national or international expert groups. 3B: Guidelines from expert individuals or institutional groups. Level 4: Quality specifications laid down by regulation or by EQAS organizers. 4A: Quality specifications laid down by regulation. 4B: Quality specifications laid down by EQAS organizers. Level 5: Published data on the state of the art. 5A: Published data from EQAS and PT programs. 5B: Publications on individual methodology.

It was recognized that there were some interesting newer approaches to derivation of AQS based upon analysis of effect of performance on requesting patterns (9) and reference intervals (10) but it was considered that, since these strategies had advantages and disadvantages (11), it would be of great interest to see if these approaches become widely used and worthy of incorporation into the hierarchy.

Members of the WG assessed, inter alia, the most important reliability performance characteristics of imprecision and bias using well-established guidelines and recommendations when new analyzers, methods and systems were established. In general, total allowable error [TEa] based on biological variation data [Level 2a] was used as the basis for setting AQS to undertake quality planning after the quality attained had been determined. However, it was widely recognized that there were problems with the ubiquitous use of the existing data on biological variation as has been described in detail at previous convocations (12, 13). As a consequence, the EFCC should continue with the remit to assess and document stringent criteria regarding how to determine components of biological variation in the style of the internationally accepted IFCC and CLSI guidelines on the proper derivation of population-based reference intervals. It is the opinion that progress on this was needed now, since works on the determination of biological variation of varying quality do keep being published in the literature of laboratory medicine.

Although, in general, AQS used were based on biological variation data, it was recognized that examinations done on samples from special groups, in particular those in which quantity concentrations were rather different to most adults, e.g., in pediatric practice or for groups with immunodeficiency, less stringent AQS were probably required. The WG recommended the use of Level 3B AQS in such situations, that is, locally agreed guidelines developed by specialists in laboratory medicine and clinicians together and then supported by all.

Professional recommendations on AQS from guidelines from national or international expert groups [Level 3a] were used by all members of the WG when appropriate, e.g., for examinations for glucose, cholesterol and troponin. However, it was recognized that there were many such guidelines in both the general and specialty literature of clinical medicine as well as in that of laboratory medicine. The quality of some of the guidelines was based on rather less than good evidence. It is necessary to create a database of all guidelines that contained numerical guidelines for acceptability of reliability performance characteristics in laboratory medicine. It was thought that the guidelines should be assessed using the well-established techniques of evidence-based medicine. Moreover, just as the database on components of biological variation (14), it would be appropriate to update this database

regularly. It was uncertain whether a professional body, such as EFCC or IFCC or an individual or group should undertake this task.

ISO I5198 states: *Clause 5.6.2-The laboratory shall determine the uncertainty of results, where relevant and possible.* Uncertainty (of a result of measurement) is defined as: a parameter associated with the result of measurement, which characterizes the dispersion of values that can be reasonably attributed to the measurand. The intent of ISO is to make measurements transferable, or comparable, on a global basis. This requires eliminating or correcting biases or systematic errors between measurement systems and then reporting any remaining variance of a test result (uncertainty) to inform the user of its quality (15). Thus, a germane question is whether AQS should be set and used for uncertainty of measurement or TEa? Both concepts have advantages and disadvantages and there are protagonists and antagonists for and against uncertainty of measurement that have detailed the arguments (15). The WG considered whether both should now be used by the laboratory or only one, and also debated whether uncertainty of measurement, and the associated concepts of verification of trueness so as to ensure traceability, would eventually replace use of total error. The consensus was that use of total error concepts was the best means to use objectively set AQS for imprecision and bias in daily laboratory quality management as well as monitoring ongoing performance. If estimates of uncertainty of measurement were to be reported to users in the future, significant education would be required. Moreover, the WG thought that it was important to recognize that clinicians were probably more interested in uncertainty of results rather than the metrological concepts of uncertainty of measurement. In addition, it was noted that there were several methods for calculation of uncertainty of measurement and the broader acceptance and use of the concepts would be aided by a unified approach to calculation.

As others (6), the WG recognized that the Stockholm consensus conference did not explore every situation in quality management in which AQS would be useful. As a consequence, the WG considered the use of AQS in EQAS. Consensus was that on the use of biological variation data [Level 2A] was the best approach and that such should be used as criteria for acceptability of performance. However, it was recognized that this approach might need some modification and it was accepted that state of the art goals [Level 5] were attainable. The WG also supported use of commutable materials that are traceable to reference systems as well as material challenges that were at medical decision points. It was considered that future effort should be directed to investigation of means to harmonize criteria for acceptance and rejection in EQAS across Europe. EQAS organizers should consider now how to deal with submissions that reported not only the numerical results found but an attached uncertainty of measurement, exactly as results of samples from patients might be reported in the future.

Moreover, setting AQS for qualitative tests was seen as difficult and AQS be best based on analysis of the effect of performance on clinical outcomes, i.e., Level 1 approaches. The WG also wondered why laboratories persisted in using qualitative examinations, such as dipsticks when quantitative methods were now available and therefore quantitative procedures should replace qualitative approaches: as one example, it has been advocated (16) that fecal immunochemical tests (FIT) should replace guaiac-based fecal occult blood tests (gFOBT). Setting AQS for examinations done as point-of care-tests (POCT) was also seen as difficult and AQS should be best based on local discussions between laboratory medicine experts and clinicians, i.e., Level 3B approaches.

TEa is typically estimated using "desirable" bias and "desirable" imprecision. The WG considered whether TEa should be estimated for an examination using the well-established three level strategies (17) by using combinations of "optimum", "desirable" and "minimum" bias and imprecision. The AQS derived from the three levels were used as appropriate by members of the WG for imprecision and bias separately and then combined to set AQS for TEa. However, in view of the performance achieved for some examinations that was superior to even current optimum specifications, it was recommended that a new fourth level, termed "ideal", be considered and, at least on an interim basis, calculated as 0.1 $CV_{within\text{-}subject}$ for imprecision and 0.1 $CV_{population}$ for bias. However, the WG is not supporting the semantics used to describe the three (or four) levels and it was advised that a new nomenclature based on sigma-metrics should be developed.

Effects of lot to lot variation in reagents, particularly calibrators, were seen as the major cause of changes in bias over time. In consequence, it was considered that more information is required from manufacturers. Not only were valid estimates of uncertainty of measurement of the calibrator values assigned required, but also the AQS for allowable lot to lot variation should be documented along with detailing of the objective rationale for such AQS. It was noted that "acceptance quality checks" could be adopted for specific reagents, such as those for gFOBT, FIT, POCT glucose and blood-bank reagents, and this strategy was a relatively novel way to improve quality. When representative material from candidate reagent lots is obtained and tested for compliance with AQS prior to acceptance for delivery and use (18): rejection of candidate lots would ensue if agreed AQS were not met.

The WG debated whether TEa and/or uncertainty of measurement could be defined in the context of multiple instruments examining the same quantity. It was considered that AQS adopted for daily rejection or acceptance of daily QC or should be used also for long-term monitoring of quality. Different laboratory organizational approaches were clearly feasible, including different instruments in one laboratory, different modules in one analytical system and different instruments in different locations. Such multiple analytical systems should be treated as a single "virtual" system with the same objectively set AQS for each analytical component of the system. The WG considered the AQS for the allowable difference between multiple systems examining the same analyte in one laboratory organization. There were few publications in this area of quality management in laboratory medicine: however some time ago, it had been advocated (19) that the allowable delta should be $<1/3\ CV_{within\text{-}subject}$. It was recognized that attainment of this was very difficult for

certain quantities and, analogously to the three level models for setting AQS for imprecision and bias, the WG recommended wide adoption of a new approach. The AQS for the allowable difference between analytical systems in a single laboratory organization should be set at levels of: $1/3CV_{within-subject}$, termed "optimum", $2/3CV_{within-subject}$ termed "desirable" and $3/3\ CV_{within-subject}$ termed "minimum".

Implementation of metrological traceability and estimation of uncertainty of measurements in the context of ISO 15189 requirements

Uncertainty of measurement (UM) is a concept developed in the field of metrology to enable assessment and communication of the uncertainty in results produced by laboratories. The key document describing UM is the Guide to the Expression of Uncertainty in Measurement (GUM) (20), which is produced by the Bureau International des Poids et Mesures (BIPM). This document clearly argues the importance of providing an estimate of the uncertainty with every result produced by a laboratory. Typically a result is expressed in the form $X\pm\mu$ (where X is the result of the measurement and μ is the uncertainty of the measurement). By contrast the vast majority of the millions of results produced in laboratory medicine are released for clinical use with no uncertainty estimate attached. Recently, the advent of ISO 15189 as a widely used standard for accreditation of clinical laboratories has necessitated the estimation of UM for assays "where relevant and possible". The aim of this WG was to assess the current and possible future processes for UM in clinical laboratories.

1) What is UM in the context of laboratory medicine?

Despite definitions of UM being available in the International Vocabulary of Metrology (VIM) (21) and in other resource documents, the concept of UM was not easily described amongst this group of interested laboratory scientists. Descriptions of UM included the following: imprecision+bias; imprecision with bias removed; uncertainty+traceability; total error; all variability in the result (including biological and pre-analytical). The official definition, *"A non-negative parameter characterising the dispersion of values being attributed to a measurand"* (21) did not seem to provide the WG with assistance in putting the concept into words which were useful to clinical scientists. The two major issues that were raised regarding the concept of UM were which factors to include in the uncertainty estimate (e.g., variation in measurement only; or measurement variation combined with within-subject biological variation, pre-analytical variation, and any post-analytical variation), and the approach required for dealing with bias.

Given these different understandings it was recognized that clear and practical definitions are required within an accreditation jurisdiction to ensure that both laboratory scientists and accreditors are aware of what is required. With regard to bias, there was a tension recognized between the GUM approach, where bias should be measured and corrected, and a concept of total error, where combination of bias and imprecision of a method should meet specified performance criteria (15). It was agreed that adjusting bias within individual laboratories based on single experiments, such as a comparison with a reference method or measurement of a certified reference material may add to the total uncertainty compared with acceptance of manufacturer assigned calibrator values.

In summary UM was considered to be an estimate of the imprecision and bias achieved in an assay in routine use. Also it was considered that UM should be limited to the measurement procedure only, however recognizing that for clinical interpretation this must be combined with pre-analytical and within-subject biological variation.

The processes of estimating UM for an assay may vary, but reasonable approaches would be to determine precision from internal QC and bias from external quality assurance (EQA) or the use of Certified Reference Materials or reference methods. This bias and precision can be combined for total error if required. These processes are of course very different from the "bottom up" processes described in GUM and have been described as "top down".

The process of satisfying regulatory requirements has generally been the preparation of a report on UM for accreditation purposes only and it was the feeling of many that this was the most onerous component of the ISO 15189 accreditation process.

2) Introduction of "UM" as part of ISO 15189

Two differing views on the introduction of UM determination and reporting were identified. Broadly these can be seen as positive or negative. A supportive view would be that UM is good laboratory practice by another name. Laboratories should know the imprecision (uncertainty) of their assays, should routinely review these as part of assay performance to ensure fitness for purpose, and be able to communicate the effect of imprecision on clinical interpretation. The contrary view is that the time spent on UM for accreditation purposes is not commensurate with the benefit, the reports prepared are for accreditation purposes only, and that this has led to waste of senior scientist's time, one of our most precious resources. The imposition of UM as it has been done may, to some extent, have damaged the relationship between laboratories and accreditors and possibly with metrologists and this damage may be in both directions. The clinical culture is one of evidence-based medicine where the introduction of a new, time consuming (and therefore costly) activity should be assessed against the benefit obtained. The feelings of the laboratory scientists were generally that this activity is required for accreditation only, and not for clinical care. By contrast, for the metrological tradition from where this concept has arisen, a result without a stated uncertainty would seem incomplete and inadequate. In this regard the process can be viewed as a clash of cultures where the cultures of laboratory medicine and of metrology have clashed, with accreditors as adjudicators, but siding with the ISO standard which is based on metrological traditions.

The recommendations of the WG are that before implementation of new practices all parties should be involved in consultation, that clear guidelines on definitions and protocols are required (although these may vary by jurisdiction) and that compromise may be required for an optimal outcome. In this case discussion could involve laboratories, professional bodies, accreditors and metrologists.

It was also recognized that local documents describing the processes in terms suitable for laboratory scientists are vital. An example of this is the Australian document by White and Farrance (22). While the contents were not discussed, the concept of an agreed local standard was supported, for example the NPAAC guidelines in Australia (23).

3) Why should laboratories estimate UM?

The WG produced the following reasons for laboratories to be involved in the UM process: to assist with interpreting results; to assess assay quality; to ensure maintenance of assay quality over time; to assist with method selection; to assess ability to share results in a common IT system; and "because we have to" for ISO 15189 accreditation.

In order to use UM to assess assay quality it was noted that there must be a quality standard for assessment against. Possible standards include the various components of the Stockholm consensus hierarchy (7), ideally with a standard reaching "six sigma" to allow for variation on assay performance. It was, however, also noted that a more realistic criterion for a commercial assay may be the manufacturer's claims, as performance better than these cannot be expected. It was also noted that performance should particularly be considered near important clinical decision points.

4) Clinical use of UM

As background to the discussion it was noted that all interpretation of laboratory results is by comparison. This may be with a previous result from the patient (monitoring), with a population reference interval (diagnosis) or with a clinical decision point (diagnosis). These different decisions indicate the need for focus on different aspects of the uncertainty of the result. For the above settings the key factors are within-laboratory precision for monitoring, bias from reference interval method or bias from method used to determine the clinical decision point for the two diagnosis examples.

Even within these broad categories the knowledge required for use of UM in clinical decisions will vary by the geography of the included testing. For example, if a patient is measured only at one laboratory, the precision within that laboratory is the required knowledge for monitoring; if the patient moves within a city, the UM should reflect the variation in the city; movements within a region, knowledge of the UM of the system comprising all regional laboratories may be required.

It was recognized that in discussing the uncertainty of a network or a city, an expanded terminology may be required to ensure clear transmission of information. While terminology, such as the CV for repeatability (CVr), between-run (CVbr), total analyzer precision (CVtot) within-individual variation (CVi) are commonly used, other concepts may be required. Examples of other precision estimates may include the following: between analyzers of the same method (CVba), between methods (CVbm), pre-analytical (CVpa), some of which may be combined to give the uncertainty of the result issued (CVres) and the CV of estimate of patient set point (CVpat).

The uncertainty of a result can also be considered with regard to the clinical question. For comparison with a fixed decision point, the uncertainty of a single result is required, for monitoring purposes the uncertainty of the difference (comprising two measurements) is required. Clarity is needed to decide which information is communicated to clinicians.

5) Communicating UM for clinical decision-making

One of the reasons for estimating UM is to support better clinical decision-making. In this context the UM itself, i.e., the uncertainty of the estimate of the concentration of the analyte in the sample, is likely to be less useful than an estimate of the uncertainty of the estimate of the homeostatic set point in the patient (CVpat). Obviously the latter requires combination of the appropriate analytical uncertainty with the expected variation in the patient (CVi) and, if known, pre-analytical variation (CVpa). Indeed provision of UM alone may be misleading when it is small compared with the biological variation. In the laboratory we tend to use CV or SD as terms to express imprecision/uncertainty, whereas UM is typically expressed as twice this value and described as having a "coverage factor of 2". This seems reasonable as our experience is that the concepts of CV and SD are often poorly understood by clinicians and a 2×CV range provides a 95% confidence interval for the result which may be more easily comprehended. If both concepts are in use we need to ensure that they are clearly defined when used.

The WG considered possible ways of communicating the uncertainty of a result to our clinical colleagues (Table 1).

Several general issues were raised concerning the reporting process, which were considered important. These include the need for a balance between more information against the benefit of that information; the need for clarity of the results on the report and for highlighting important results. It was also noted that it may be easier to respond to a specific question,

Table 1 Methods for communicating uncertainty of a result.

1. Choice of significant figures (reporting interval)
2. Supporting information (e.g., website, lab handbook)
3. Give a range for the result (e.g., 2.2–2.4 mmol/L)
4. "± ×" (e.g., 2.3±0.1 mmol/L)
5. Graphical displays with error bars
6. On-line calculator to calculate statistical significance of differences
7. On request, i.e., as part of a consultation, when specifically requested or by pre-arrangement for specific clients or for certain tests (e.g., serum creatinine or PSA)
8. As a text comment attached to the result
9. Placing symbols alongside results (e.g., asterisks as a marker of changes exceeding the critical difference)

e.g., has this PSA changed since last measurement, rather than responding to a set of results for which the clinical background is uncertain.

The sought-after advantages of reporting the uncertainty of a result would be improved clinical decision-making, e.g., the recognition of true changes in a patient and avoiding responding to insignificant changes. The flow-on effects may be better test requesting, and better use of medicines. Reporting uncertainty may also highlight method differences, e.g., point of care vs. main laboratory analyzer, and highlighting specific limitations (e.g., low concentration precision near a CV of 20%).

Discussion, however, did raise a number of potential disadvantages to reporting uncertainties with patient results. These include the following: a) confusion in understanding the meaning of the expressed uncertainty; b) confusion in reading a report which is more complex; c) the uncertainty estimate may be misleading if not all factors are taken into account (e.g., pre-analytical factors); d) medicolegal implications of the report – "you told me the change was not significant"; e) differences in formats from different laboratories adding complexity; f) consumption of precious "Real estate" on reports; and g) reduced confidence in a laboratory ("the other laboratory is not uncertain").

It was also noted that the concept of a formally expressed uncertainty was not generally used for any numerical values in medicine and also that laboratory results are combined with a wide range of other data from history, examination and other tests and a single "best estimate" may be all that can be handled in the decision-making process. From an evidence-based approach we are unaware of research demonstrating that the reporting of uncertainty improved patient outcomes – the highest standard that should guide our actions.

Thus, our recommendations are that uncertainty of a result or of an estimate of a patient's homeostatic set point should only be done after consultation, education and agreement of all involved parties. It is necessary to avoid a "Clash of Cultures" between laboratories, where recognition and handling of imprecision/uncertainty is routine and clinical services, where formal statements of uncertainty are rarely used. It is also recommended that adoption of similar practices in a city, region or country would be beneficial, with different methods of reporting uncertainty in reports from different laboratories likely to add to clinician confusion.

In order to further consider the benefits, risks and methods of reporting uncertainty to clinicians, it is necessary that there is widespread involvement of laboratory scientists, clinicians, relevant researcher, information technologists and the professional bodies representing these parties.

6) Critical difference theory

The development and popularization of critical difference theory by Fraser (4) and others has been a major step forward in the area of clinical decision-making based on laboratory results. This is however a theory which may benefit from active research to improve the quality of the interpretation of laboratory data. Some examples of additions to the basic theory are the effect of the number of significant figures (24) and the effect of varying SD with concentration (25).

Transforming advanced concepts in analytical quality to everyday laboratory practice: supporting and promoting best practices of internal QC

The use of internal QCs and the interpretation of the results are regulated by various locally agreed upon or national mandatory guidelines and rules (e.g., Westgard Rules or German RiliBÄK). The correct implementation and execution of these guidelines are continuously monitored by national authorities and by accreditation or certification bodies under ISO 15189 and ISO 9001. However, many different factors influence both the handling of QC samples and the corrective actions which need to be taken when non-conformities arise. The discussion focused on determining these factors and on defining best practice solutions in terms of effectiveness and efficiency.

Before deciding what best practice was with regard to the handling of internal QC, one first had to recognize, that the concept of best practice will depend on the different stakeholders and their different interests. For example, for patients, medical and laboratory staff, implementation of best practice will prevent the reporting of incorrect results thereby saving the patient from incorrect treatment. For the laboratory owner, best practice will result in a cost effective QC system which is legally compliant. For manufacturers, best practice can be achieved by receiving and listening to valid suggestions and complaints from their customers.

The handling of internal QC material can be affected by a number of factors, such as the type of QC material, the number of levels and the analyte concentrations, whether the QC material is assayed or unassayed, the form in which the QC material is supplied (i.e., lyophilized vs. liquid-stable), the time and frequency with which the QC material is tested, the statistical methods used to interpret the QC data and the criteria for acceptance or rejection.

1) Types of QC material (assay manufacturer, third party or patient samples)

The quality of QC material plays a major role in this context. The non-commutability of QC materials with patient samples is a recognized phenomenon, indeed the occurrence of non-commutable results for QC materials has been found to occur frequently enough that QC results cannot always be used to verify consistency of results for patient samples when changing lots of reagents (26).

Although the use of third party QC materials is recommended in different regulatory standards (1, 27), we found that using a mixture of assay manufacturer's QC and third party QC was most convenient for the laboratory. The advantage of using assay manufacturer's QC material is that, when troubleshooting, the laboratory is only dealing with a single supplier who then becomes responsible for helping to resolve the problem. Conversely, it was recognized that the acceptable

ranges quoted by assay manufacturers for their QC materials, are usually quite broad.

The use of unassayed QC material is often necessary, because assayed material is not always available. It was felt strongly that the manufacturers of QC material should offer suitable QC materials for all measurands being determined in medical laboratories (e.g., serum indices and electrophoresis). When validating unassayed QC material, a stepwise procedure for establishing target values is recommended (e.g., an initial target value is established by analyzing a small number of samples, but is reviewed regularly following its introduction into routine use as the number of measurements increases).

If the shelf life and the stability of a QC material are of reasonable duration, liquid QC material has several advantages over lyophilized material. The major advantage is the removal of any variability of results caused by pipetting errors during reconstitution of the lyophilized material. In the case of stable measurands, aliquots of the QC material may be stored frozen. Patient samples may also be used as QC material, although legal or ethical constraints may prevent this being possible for some laboratories.

2) Number of QC levels and analyte concentrations

The number of QC levels chosen to monitor assay performance will generally depend on both the type of calibration (i.e., linear or non-linear) and the Westgard Rules being used to interpret the QC data. It is recommended that at least two levels, one in the normal and one in the abnormal range, or levels close to clinical decision limits, are appropriate (1, 27). Sometimes it might be necessary to dilute QC material to reach the concentration of interest (e.g., female testosterone or sodium in sweat). In such cases, the influence of matrix effects should be taken into consideration and the use of the diluted material must be validated.

3) Frequency and time of testing

When deciding on the frequency and time of testing of QC material, the requirements of the IVD directive, and of national laws have to be considered as well as the recommendations of the manufacturers (28) or other references (29). In addition, the use of internal QC may be indicated before and after calibration, after a reagent change, at the beginning of an assay run or after a certain number of patient samples have been analyzed, at fixed times during the day and after maintenance or repair of analyzers or other relevant technical equipment (e.g., water preparation units).

4) Criteria of acceptance (rejection) and rules

When available, use of the manufacturer's QC target values and limits are considered to be the minimum acceptance criteria. Also, use of Westgard Rules 1–3s, 2–2s or R–4s were regarded as the basic criteria for acceptance or rejection (nine out of 12 participants in the WG used Westgard Rules). Some manufacturers supply QC samples with fixed target values (i.e., target values which do not change from lot to lot). Use of such material significantly reduces the efforts of administration.

If the target means and acceptable limits (standard deviation, coefficient of variation or total allowable error) are to be established by the laboratory, the methods used to determine these values should be validated (e.g., to ensure adequate sample size and a robust method for the elimination of outliers) and deemed to be suitable. The use of recommended statistical methods (e.g., those published by regulatory authorities) is highly recommended. Any statistical method used should be kept constant over time to enable long-term comparison of data. It was recognized that the process of determining target means and acceptable limits might be time consuming in case of analytes that were measured infrequently. It was agreed that if the same methods are used (e.g., total protein in serum or plasma), inter-analyzer limits should be the same as intra-analyzer limits.

5) Corrective actions

The best practice procedure to follow when resolving QC rejections was discussed. The WG suggested that QC rejections could be handled as follows, with the initial trouble shooting being performed by technicians: a) eliminate human error (e.g., analysis of wrong QC, incorrect QC preparation, incorrect storage conditions or use of the incorrect target value following a change in QC lot no); b) decide whether the error observed is a random or systematic error by considering all QC values measured and comparing them to previous values and/or other analytes; c) re-measure the QC, use new QC material or new reagent, recalibrate as appropriate; d) review the effectiveness of your corrective action (rerun some relevant patient samples); and e) document the QC rejection and comment on the action taken.

6) Best practice for handling complaints

It was recognized that problems with the quality of the laboratory service might be highlighted by service users from outside the laboratory. The WG suggested that such complaints can be processed as follows: a) document the complaint and inform the person in charge (e.g., head of department); b) determine whether or not the laboratory is responsible for the complaint; c) inform the person who complained that the complaint is being dealt with; d) take action (the resolution of the complaint will depend on the investigation, the type of problem and the impact of the problem); e) review the effectiveness of the corrective action; f) document and close the complaint; and g) inform external bodies (if relevant and necessary).

Quality control for qualitative assays: special challenges and best practices in virology and serology laboratories (30–33)

Over the last decades there have been many changes in virological and serological techniques. Former dilution-based tests

like complement fixation or hemagglutination inhibition tests were replaced in most laboratories by fully automated ELISA techniques. Many of these applications lead to quantitative results which are partially calculated in international units. A complete paradigm change has taken place in the detection of viral pathogens: The formerly less sensitive and not quantitative technique of virus isolation has been replaced by real-time PCR-assays leading to measurements of "viral loads" in different body fluids. Unfortunately for both serologic and PCR-testing, international standards are only partially available for the assessment of the accuracy of measurement.

Modern virology and serology is now similar to clinical chemistry in terms of measurement methods and technical platforms. This leads to the possibility of the implementation of QC procedures in virology laboratories where daily QC data are archived and statistically evaluated. Even if most methods are only qualitative, a sample/cut-off ratio can be calculated and this number can be tracked by a Levey-Jenning chart for assuring that method performances are stable especially when a change in reagents occurs. Nevertheless, there are some specific aspects in testing for serologic immune responses as well for direct measurement of pathogens that are challenging, e.g., the need for negative or below cut-off controls, the non-linear but logarithmic scale in quantitative real-time-PCRs and the uncertainty of what serological titer or viral load is considered a pathological result.

The harmonization of QC in virology and serology laboratories is additionally complicated by the fact that there are generally two main working fields, namely blood bank testing and other infectious diagnostics (microbiology, virology) with different aims. The organization of blood banking differs widely among European countries varying from complete centralization in France accompanied by nearly the same spectrum of tests and technical platforms to a greater variety of forms of organization and techniques in other countries.

1) NAT-testing

The WG identified some special challenges that exist in PCR-techniques in contrast to clinical chemistry: a) existence of negative results requires testing of negative controls; b) existence of more standards will lead to more harmonization between different technical platforms. At present primary standards (WHO standards) are only available for HIV, HBV, HCV, Parvovirus B19. There are no standards for HAV (mainly for the plasma industry), the family of herpesviruses (EBV, HHV-6, HSV1 and -2), *Chlamydia trachomatis, Mycoplasma genitalium*, noroviruses, polyomaviruses (JCV and BKV). Finding the right donors for less frequent infectious diseases is a problem in terms of preparing standards; c) the suitability of natural analytical standards versus synthetic standards (plasmids) needs to be established; d) suitable QC materials are needed for the different steps in PCR techniques with inherent problems, such as specimen processing, nucleic acid extraction, and amplification. Rules for run-on-controls are needed; e) specificity of the PCR must be verified; and f) QC for "drifting" targets (e.g., influenza virus-PCRs).

In recent years there has been a clear trend to quantitative monitoring in IU/mL wherever possible. This leads to special problems with reporting and interpreting of PCR-results: a) What is considered a significant increase/decrease in viral loads: 1.0 log10 (10-fold) or 0.5 log10 (3-fold) or what?; b) Should results be reported as integers (e.g., 1000 copies/mL), in scientific notation (1.0×10^3 copies/mL) or as a log10 (3.0)?; c) Should the limits of tolerance for the controls be calculated as integer values or transformed data?; and d) Matrix effects: results from dedicated specimen may not accurately reflect the real site of virus replication (e.g., CMV in cells or in plasma).

2) Serology

There are some special problems in serology which should be addressed in the future: a) more international standards are needed for serologic infectious parameters leading to the reporting of results in U/mL (as in rubella and anti-HBs, etc.). Working with arbitrary units in some cases is usual, but should be avoided in the long-term; b) there exist only old WHO standards with measles and mumps virus. There is the need for renewal; and c) primary standards are not available for some antigens for technical and financial reasons and for some antigens standards are not available for reasons of regional epidemiological variation of pathogens. Examples of standards that are not available are CMV, EBV, *Chlamydophila pneumoniae*, VZV and toxoplasmosis (only old standards exist), borrelia, mycoplasma. Finding the right donors for standards for less frequent infectious diseases is a problem (see above for NAT-testing). Calibrators should be included in all IgG-assays to better quantify the specific immune responses. Standards should include diagnostically relevant isotypes, e.g., IgM, IgA. Clinical/diagnostic sensitivity and specificity studies are needed to address what clinical titers or U/mL are relevant for seroprotection and acute infection. The crude calculation of indices or sample/cut-off values leads only to qualitative results.

Generally there is no reason why infectious parameters (both, serologic and molecular) should not be included in the quality assurance programs of a laboratory. Calculation and documentation of QC markers (precision, accuracy, repeatability, random/systematic errors, etc.) is possible. There are some special conditions in quality assurance programs of infectious diseases that need to be considered and some problems need to be addressed, such as primary standards, calculation and interpretation of PCR-results.

Discussion

It should be noted that metrological traceability is directly linked to the traceability of the calibrator on the device used to quantify the measurand. If one assumes the patient sample is traceable, then any material treated like a patient sample, such as a control or PT material should likewise be traceable.

Commutability is a different issue. While commutability of EQA materials is an admirable goal, laboratories need to

understand that to provide materials that are stable, challenge the diagnostic range, and contain multiple analytes must sacrifice commutability in some cases. Granted, fresh frozen plasma samples can be used that guarantee commutability but here, all the convenience of a material with long-term stability, that covers the analytical range and large peer groups is likely sacrificed. Laboratories should also require from suppliers of fresh frozen plasma samples, evidence of testing for infectious diseases, and compliance with ISO 17043 which sets requirements for EQA providers.

It was recognized that laboratories will evolve differently from country to country but that ISO 15189 should be considered as a good basis to achieve consistent quality among laboratories in the European community.

The communication of quality to users of the laboratory, particularly when objectively set AQS, were attained or surpassed. Traditional population-based reference values (stratified when required), reference change values, clinical decision limits and other information were all considered appropriate. The key communication of quality needed and achieved was seen as constructive discussions between laboratory medicine experts and health care users, whenever appropriate and possible.

It has shown that the concept of UM can be viewed in many ways. There have been many concerns about the ISO 15189 requirements for estimation of UM in clinical laboratories. These have been due to uncertainty about how to "do" UM and the time required for an activity primarily aimed at meeting accreditation requirements. The WG recommends that where possible the requirements are clarified and simplified. It was noted that communication of uncertainty to clinicians has the potential for both benefit and for harm. It was recognized that both in the introduction of UM into clinical laboratories, and in communication of UM to clinicians, there is the potential for a clash of cultures, which can only be addressed by meaningful communication and respect for the science and traditions present in different fields of endeavor.

Various guidelines and rules are available to be used for internal QC and its interpretation. The implementation and execution of the guidelines are monitored by national authorities and by accreditation or certified bodies under ISO 15189 and ISO 9001. Many factors may affect the handling of the internal QC material, such as the type of QC material, if it is lyophilized or liquid, the timing and frequency, the statistical methods to be used for interpretation and finally the criteria for the acceptance or rejection of the results.

The QC in virology and serology seems to be somewhat complicated. There are different working fields (blood bank testing, microbiology, virology) with different aims. As the blood banks are completely centralized in France but not in other European countries it is challenging to create a common strategy of QC. There is no reason why infectious parameters should not be include in QC programs as calculation of precision, accuracy, etc. is possible. Nevertheless there are new fields like PCR testing, which may need special attention.

Acknowledgments

1. Using 15189 as the key to quality: practical experience and peer guidance for success:
Michel Vaubourdolle (CHAIR), Saint Antoine Hospital, Paris, France; Philippe Aubert, Unilabs, Perrières, Switzerland; Nathalie Colard, Laboratoire Colard, Fresnes Sur Escaut, France; Roberto Guerranti, Azienda Ospedaliera Universitaria Senese, Siena, Italy; Jacques Izopet, Toulouse University Hospital, Toulouse, France; Matthias Orth, Marienhospital, Stuttgart, Germany; Paul-Louis Courrier, Centre de Transfusion Croix Rouge Luxembourg, Luxembourg, Luxembourg; Katri Puukka, Oulu University Hospital, Oys, Finland; Francisco Ramon Bausa, University Materno-Paediatric Hospital, Barcelona, Spain; Ángel Salas, Consorci Sanitari Terrassa, Terrassa, Spain; Florian Scherrer, LBM de l'Avenue, Roussillon, France; Marc Thelen, Amphia Hospital Breda, Breda, The Netherlands.

2. Defining appropriate analytical objectives in the laboratory:
Callum Fraser (CHAIR), Scottish Bowel Screening Centre, Dundee, Scotland; Anna Carobene, Ospedale San Raffaele, Milano, Italy; Jean Pierre Dehorne, LBM Dehorne Reseau Bio, Chapelle-sur-Erdre, France; Nathalie Durgniat, Unilabs, Neuchâtel, Switzerland; Pilar Fernández Calle, La Paz University Hospital, Madrid, Spain; Hazel Finney, St Bartholomew's Hospital, London, UK; Davide Giavarina, San Bortolo Hospital, Vicenza, Italy; Maria Adelina Gomes, Instituto Nacional, Dr. Ricardo Jorge, Lisboa, Portugal; Rob Jansen, SKML, Nijmegen, The Netherlands; Michele Mussap, University-Hospital, Genova, Italy; Jocelyn Naicker, NHLS/University of Witwatersrand, Johannesburg, South Africa; Carmen Perich, Laboratorio Clínico Bon Pastor Calidad, Barcelona, Spain; Viviane van Hoof, University Hospital Antwerp, Edegem, Belgium; Douwe van Loon, Sint Antonius Ziekenhuis, Nieuwege, The Netherlands.

3. Implementation of metrological traceability and estimation of uncertainty of measurements in the context of ISO 15189 requirements:
Graham Jones (CHAIR), St Vincent's Hospital Sydney Ltd, Sydney, Australia; Francisco A. Bernabeu, Hospital Universitario Príncipe de Asturias, Alcalá de Henares, Spain; Maguy Bernard, GH Pitié-Salpêtrière, Paris, France; Carinke Buiting, Euprax, Tilburg, The Netherlands; Niels de Jonge, Leids Universitair Medish Centrum, Leiden, The Netherlands; Brendan Fitzpatrick, Bon Secours Health System, Dublin, Ireland; Carsten Schriver Højskov, Aarhus University Hospital, Aarhus, Denmark; Nuthar Jassam, Harrogate District Hospital, Harrogate, UK; Solveig Linko, USLAB, Helsinki, Finland; James Nichols; Baystate Medical Center, Springfield, USA; David Rambau, Lancet Laboratories, Johannesburg, South Africa; Mzougui Taoufik, Hôpital de Jolimont, Haine-Saint-Paul, Belgium; Gerda Ziervogel-Lukas, Klinikum Klagenfurt am Wörthersee, Klagenfurt, Austria.

4. Transforming advanced concepts in analytical quality to everyday laboratory practice: supporting and promoting best practices of internal QC:
Michael Hubmann (CHAIR), MZ Labor, Feldkirch, Austria; Frances Flores, The Royal London Hospital, London, UK; Jean-Marc Giannoli, Laboratoire des Charpennes, Villeurbanne, France; Jeremie Gras, Clinique Saint-Luc Bouge, Bouge, Belgium; Adrian Holt, Manchester Royal Infirmary, Manchester, UK; Veronica Lee, Barts and the London NHS Trust, London, UK; Guillaume Lefevre, Hopital Tenon, Paris, France; Simon Lykkeboe, Aalborg University Hospital, Aalborg, Denmark; Edward Kearney, East Kent Hospital University Foundation NHS Trust, Margate, UK; Rui Pimentel, LAC Joaquim Chaves, Miraflores Algés, Portugal; Santiago Prieto, Hospital Universitario de Fuenlabrada, Madrid, Spain; Birgitte

Reinholdt, Lillebaelt Hospital, Vejle, Denmark; Hanna Maria Schumacher, Universitätsmedizin, Mainz, Germany; Jean Michel Vialle, Laboratoire Vialle, Bastia, France.

5. Quality control for qualitative assays: special challenges and best practices in virology and serology laboratories:
Ortwin Adams (CHAIR), Universitätsklinikum Düsseldorf, Düsseldorf, Germany; Mira Barak, Clalit Health Fund, Nesher Haifa, Israel; Anton Buiting, St Elisabeth Hospital, Tilburg, The Netherlands; Denise Annie De Smet, Rode Kruis Vlaanderen, Leuven, Belgium; Peter Declercq, Jessa Ziekenhuis, Hasselt, Belgium; Françoise Le Vacon, EFS Laboratory of Blood Donor Qualification, Rennes, France; Elisabeth Legenstein, ÖQUASTA, Vienna, Austria; Mihai Pecec, National Institute of Blood Transfusion, Bucharest, Romania; Marion Riffelmann, Helios Klinikum, Krefeld, Germany; Volkmar Schottstedt, DRK-Blutspendedienst West, Hagen, Germany; Jukka Suni, Helsinki University Hospital, Helsinki, Finland; Paulo Tavares, Unidade Local Saude da Guarda, Guarda, Portugal; Karl-Heinz Wirsing von König, Helios Klinikum Zentrallabor, Krefeld, Germany.

Conflict of interest statement

Authors' conflict of interest disclosure: The authors stated that there are no conflicts of interest regarding the publication of this article.
Research funding: None declared.
Employment or leadership: None declared.
Honorarium: None declared.

References

1. EN/ISO 15189:2007 Medical laboratories – Particular requirements for quality and competence.
2. Clinical Laboratory Improvement Amendments. Centers for Medicare & Medicaid Services, HHS, 2003. 42 CFR Ch. IV (10–1–03 ed.).
3. Fraser CG, Scott MG. Analytical performance requirements for point-of-care testing: glucose as an example. In: Price CP, StJohn A, Kricka LJ, editors. Point-of-care testing: needs, opportunity, and innovation, 3rd ed. Washington, DC: AACC Press, 2010.
4. Fraser CG. Biological variation: from principles to practice. Washington DC: AACC Press, 2001.
5. Westgard JO. Basic QC practices. 3rd ed, Madison, WI: Westgard QC, 2010.
6. 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;15:323–30.
7. Hyltoft Petersen P, Fraser CG, Kallner A, Kenny D. Strategies to set global analytical quality specifications in laboratory medicine. Scand J Clin Lab Invest 1999;59:475–585.
8. Fraser CG, Hyltoft Petersen P. Analytical performance characteristics should be judged against objective quality specifications. Clin Chem 1999;45:321–3.
9. Klee GG. Establishment of outcome-related analytic performance goals. Clin Chem 2010;56:714–22.
10. Haeckel R, Wosniok W. A new concept to derive permissible limits for analytical imprecision and bias considering diagnostic requirements and technical state-of-the-art. Clin Chem Lab Med 2011;49:623–36.
11. Hyltoft Petersen P, Sandberg S, Fraser CG. Do new concepts for deriving permissible limits for analytical imprecision and bias have any advantages over existing consensus? Clin Chem Lab Med 2011;49:637–40.
12. Burnett D, Ceriotti F, Cooper G, Parvin C, Plebani M, Westgard JO. Collective opinion paper on findings of the 2009 convocation of experts in quality control. Clin Chem Lab Med 2010;48:41–52.
13. Cooper G, DeJonge N, Ehrmeyer S, Yundt-Pacheco J, Jansen R, Ricos C, et al. Collective opinion paper on findings of the 2010 convocation of experts on laboratory quality. Clin Chem Lab Med 2011;49:793–802.
14. Ricos C, Garcia-Lario JV, Alvarez V, Cava F, Domenech M, Hernandez A. Biological variation database, and quality specifications for imprecision, bias and total error (desirable and minimum). The 2010 update. Accessed May 3, 2011. Available from: http://www.westgard.com/biodatabase1.htm13.
15. Westgard JO. Managing quality vs. measuring uncertainty in the medical laboratory. Clin Chem Lab Med 2010;48:31–40.
16. Fraser CG. Fecal occult blood tests: life savers or outdated colorectal screening tools? Clin Lab News 2011;37:8–10.
17. Fraser CG, Hyltoft Petersen P, Libeer JC, Ricos C. Proposals for setting generally applicable quality goals solely based on biology. Ann Clin Biochem 1997;34:8–12.
18. McDonald PJ, Anderson CM, Fraser CG. Acceptance quality checks for qualitative fecal immunochemical tests ensure screening program consistency. Int J Cancer 2011;128: 247–8.
19. Hyltoft Petersen P, Fraser CG, Westgard JO, Larsen ML. Analytical goal-setting for patient monitoring when two analytical methods are used. Clin Chem 1992;38:2256–60.
20. Evaluation of measurement data — Guide to the expression of uncertainty in measurement. JCGM 2008. Accessed May 3 2011. Available from: http://www.bipm.org/en/publications/guides/vim.html.
21. International Vocabulary of Metrology – Basic and general concepts and associated terms. VIM, 3rd ed. JCGM 200:2008. Accessed May 3 2011. Available from: http://www.bipm.org/en/publications/guides/vim.html.
22. White GH, Farrance I. Uncertainty of measurement in quantitative medical testing – a laboratory implementation guide. Clin Biochem 2004;25(Suppl ii):S1–24.
23. Council, National Pathology Accreditation Advisory. Requirements for the estimation of measurement uncertainty (2007 ed.). Accessed May 3 2011. Available from: http://www.health.gov.au/internet/main/publishing.nsf/Content/health-npaac-emu.htm.
24. Jones GR. Effect of the reporting-interval size on critical difference estimation: beyond "2.77". Clin Chem 2006;52:880–5.
25. Jones GR. Critical difference calculations revised: inclusion of variation in standard deviation with analyte concentration. Ann Clin Biochem 2009;46:517–9.
26. 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.
27. Statistical quality control for quantitative measurement procedures: principles and definitions. CLSI 2006, C24-A3, Vol 26, No 25. Wayne, PA: Clinical and Laboratory Standards Institute.
28. In vitro diagnostic medical devices, Directive 98/79/EC of the European Parliament and of the Council. 1998.
29. Parvin CA. Assessing the impact of the frequency of quality control testing on the quality of reported patient results. Clin Chem 2008;54:2049–54.

30. Banoo S, Bell D, Bossuyt P, Herring A, Mabey D, Poole F. Evaluation of diagnostic tests for infectious diseases: general principles. Nat Rev Microbiol 2006;4(Suppl 12): S20–32.
31. Peeling RW, Smith PG, Bossuyt PM. A guide for diagnostic evaluations. Nat Rev Microbiol 2006;4(Suppl 12): S2–6.
32. Domingo C, Niedrig M, Teichmann A, Kaiser M, Rumer L, Jarman RG, et al. 2nd International external quality control assessment for the molecular diagnosis of dengue infections. PLoS Negl Trop Dis 2010;4:e833.
33. Cruciani M, Mengoli C. An overview of meta-analyses of diagnostic tests in infectious diseases. Infect Dis Clin North Am 2009;23:225–67.