DOI 10.1515/cclm-2013-0122 — Clin Chem Lab Med 2013; 51(5): 973–979

Opinion Paper

Renze Bais*, Dave Armbruster, Rob T. P. Jansen, George Klee, Mauro Panteghini, Joseph Passarelli and Ken A. Sikaris on behalf of the IFCC Working Group on Allowable Error for Traceable Results (WG-AETR)

Defining acceptable limits for the metrological traceability of specific measurands

Abstract

Although manufacturers are compelled by the European IVD Directive, 98/79/EC, to have traceability of the values assigned to their calibrators if suitable higher order reference materials and/or procedures are available, there is still no equivalence of results for many measurands determined in clinical laboratories. The adoption of assays with metrological traceable results will have a significant impact on laboratory medicine in that results will be equivalent across different laboratories and different analytical platforms. The IFCC WG on Allowable Errors for Traceable Results has been formed to define acceptable limits for metrological traceability chains for specific measurands in order to promote the equivalence of patient results. These limits are being developed based on biological variation for the specific measurands. Preliminary investigations have shown that for some measurands, it is possible for manufacturers to assign values to assay calibrators with a measurement uncertainty that allows the laboratory enough combined uncertainty for their routine measurements. However, for other measurands, e.g., plasma sodium, current assays are too imprecise to fulfil limits based on biological variation. Although an alternative approach based on probability theory is being investigated, the most desirable approach would be for industry to improve measurement methods so that they meet clinical requirements.

Keywords: IFCC; metrological traceability; result equivalence; standardization.

***Corresponding author: Renze Bais**, rbaisconsulting.com, Sydney, Australia, E-mail: renzebais@rbaisconsulting.com
Dave Armbruster: Abbott Diagnostics, 100 Abbott Park Rd, Abbott Park, IL 60064, USA
Rob T. P. Jansen: SKML, Nijmegen, The Netherlands
George Klee: Mayo Clinic, Rochester, 200 First Street, SW, Stabile Building, 3-26, Rochester, MN 55905-0001, USA
Mauro Panteghini: Centre for Metrological Traceability in Laboratory Medicine (CIRME), University of Milan, Milan, Italy
Joseph Passarelli: Roche Diagnostics GmbH, Nonnenwald 2, Penzberg, 82377, Penzberg, Germany
Ken A. Sikaris: Melbourne Pathology, 103 Victoria Parade, Collingwood, Melbourne, VIC 3066, Australia

Introduction

Over recent years, metrological traceability of analytical results has become a major focus in laboratory medicine and there is a concerted effort to standardize results produced by different laboratories and on different platforms [1, 2]. These efforts are being driven by a coalition of professional organizations such as IFCC and AACC, metrological organizations such as BIPM, through Joint Committee for Traceability in Laboratory Medicine (JCTLM), and the in-vitro diagnostic (IVD) manufacturers [3]. The aim of these efforts is to ensure equivalence of patient results for the measurement of the same measurand, that is, laboratory results should be equivalent no matter where and on which platform they are generated. One important reason for adopting a traceable, metrological-based approach is that it will allow the use of common reference intervals and clinical decision limits [4].

A number of documents describing traceability in Laboratory Medicine are available but none of these actually define the analytical requirements for a traceability chain to be clinically acceptable [5, 6]. However, establishing traceability of results for a test measurement should be inseparably linked to the definition of acceptable measurement uncertainty to fit the intended clinical application ("fitness for purpose") [7]. In 2004, Thienpont et al. [8] first pointed out that the absence of clearly defined tolerable deviations derived from clinical needs "might results in a large grey zone with respect to the extent of traceability expected from IVD manufacturers, partially or totally invalidating its theoretical advantages". Consequently, an objective approach needs to be applied to every measurand determined in the clinical laboratory in order to establish if the current status

of the uncertainty budget of its measurement associated with the proposed metrological traceability chain is suitable for the clinical application of the test [9]. As an example, using the measurement of albumin in serum as a model, Infusino et al. [10] recently demonstrated that the reference measurement system (and the associated uncertainty) currently available is probably not enough to guarantee the accuracy needed for the clinical usefulness of this protein.

With these premises, the IFCC WG-AETR has been formed

1. to specifically define clinically acceptable limits for the metrological traceability of specific measurands in order to promote the:
   a. harmonization of patient results for their better clinical application;
2. to cooperate with manufacturers, regulatory bodies and end-users to make patient results traceable to higher order reference materials and methods whenever feasible.

With the adoption of the IVD Directive (Directive of the European Parliament on In Vitro Medical Devices Directive 98/79/EC), manufacturers are compelled to have traceability of the values assigned to their calibrators if a suitable reference measurement system is available. They carry out this process by using a calibration hierarchy to ensure metrological traceability. An example of a traceability chain typically used by manufacturers is shown for glucose in Figure 1. The chain commences with the definition of the measurand and then proceeds through a series of related steps in which a hierarchical series of calibration materials are prepared, each related to the previous step in the chain and ultimately back to the measurand SI definition. Due to cost and limited resources, IVD manufacturers, however, do not perform the full traceable series of steps to value assign every new lot of assay calibrator. They often rely on value transfer from their internally stored ("master") calibrator material. In most cases, this procedure is probably valid, but there have been cases when the stored calibrator has degraded or changed in some way resulting in the final value assignment to laboratory calibrators being incorrect [11].

Although all the major IVD manufacturers have developed traceability profiles for many of their products, there is still no equivalence between results for

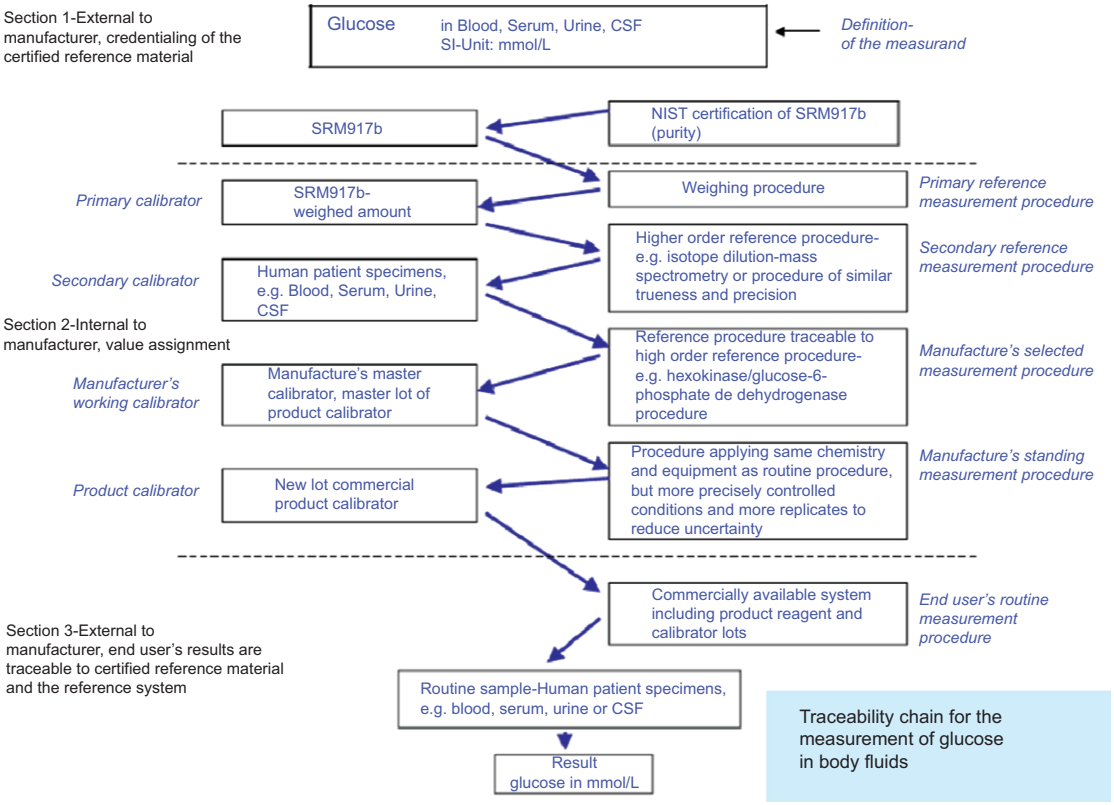

Figure 1 An example of a typical traceability chain used by manufacturers for the measurement of glucose showing the relationship between the definition of the measurand, glucose to the reference method and then to the various calibrators.

many commonly used measurands even though higher order reference material or reference procedures exist. One reason for this is that uncertainty limits for the different steps in the chain and measurement errors have never been appropriately assessed and estimated. In other words, although a particular calibrator from a manufacturer may be traceable along a chain, there are no allowable performance criteria defined to ensure the obtained final measurement meets acceptable specifications, i.e., fit for purpose.

We have illustrated this in Figure 2 in which there are two traceability paths for a measurand, say from two different manufacturers. The assigned values for the end calibrators from the two pathways would have significantly different values resulting in different laboratory results being measured on a patient. The assigned values for the two calibrators are both derived from valid traceability chains, but produce results that would not be equivalent. For these two systems to produce equivalent laboratory results for a measurand, would require the use of a correction factor determined by a correlation study at the steps where there is divergence. Although this would remove the bias between the two chains, it would impact on the measurement uncertainty of the final product. This scenario is commonly seen in immunoassays when manufacturers use proprietary antibodies to measure products. Obviously, this also means that clinical decision limits or reference intervals would be quite different for the two assays unless the bias is removed.

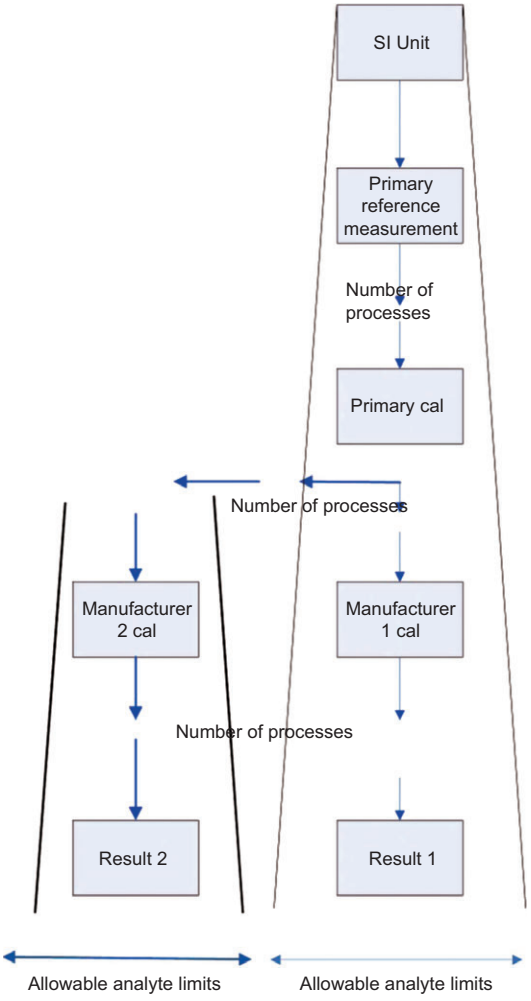

Figure 2 Two schematic traceability processes which are both metrological acceptable.
In both cases, the final product is traceable to the measurand definition but the divergence would result in significantly different patient results. Note that both processes could have the same allowable measurand limit.

Traceability chain and clinically acceptable limits

Within a traceability chain, three main elements contribute to the uncertainty and bias [12, 13]:

1. the components within the chain – reference materials, reference procedures, manufacturer's calibrators;
2. the components of the procedures – values transfer procedures, reagents, instruments, staff;
3. the sample – preanalytical effect, biological variation.

It is generally agreed that the uncertainty that is fit for purpose should be defined for the entire traceability chain, from the reference materials, through the manufacturing processes for calibrator value assignment to the final result reported to the end-users [4, 14, 15]. It is not always possible to determine the individual uncertainties for all the specific components in a traceability chain and thus, an approach can be used in which the chain is divided into parts so that the data generated can be used to calculate the combined uncertainty [6].

There are two important limits that should be defined for the clinical application of the test once a traceability chain has been defined:

1. the allowable limits for the uncertainty of manufacturer's calibrators;
2. the allowable error for measurements done by individual clinical laboratories.

In doing this, one should clearly distinguish between the bias issue and the uncertainty at the level of manufacturer's calibrator. At the calibrator level if a bias is known, the Guide to the Expression of Uncertainty in Measurement [16]

asks manufacturers for its compensation by adjusting the value of the calibrators. After this realignment, only uncertainty of calibrator remains and this will therefore be added to the uncertainty of previous chain levels to obtain the combined uncertainty of the commercial product calibrator [7, 17]. However, also under perfectly traceable conditions, it is impossible for individual laboratories to avoid a bias in their daily activity. This requires the introduction of a goal for tolerable bias when discussing the error of the individual measurements on biological samples. The definition of this error can be made according to total error theory, which requires the definition of a bias goal in addition to the goal for standard uncertainty that, at this lowest measurement level, includes the accumulated uncertainty of the corresponding traceability chain and the uncertainty due to "random" effects in the laboratory.

Approaches for defining allowable limits

The IFCC-IUPAC Stockholm Conference for setting quality specification in Laboratory Medicine [18] agreed on a hierarchical approach to defining analytical performance based on:

1. the evaluation of the effect of analytical performance on clinical outcome in the specific clinical setting;
2. data based on biologic variation information or on clinician survey; and
3. data based on clinical and laboratory experts' opinion and published recommendations.

As the outcome-based information is missing for the majority of measurands, the approach using biological variation data has generated more attention, even though recent publications have questioned the reliability and robustness of available information [19, 20].

Biological variation approach

The concept that for desirable performance, analytical variation (CV_a) should be less than half the within-subject biological variation (CV_i) was developed over 30 years ago, i.e.,

$$CV_a < 0.50\ CV_i. \quad (1)$$

By adopting 0.5 CV_i as the goal for expanded uncertainty of measurement [obtained by multiplying the relative combined standard uncertainty by a coverage factor of 2 (95% level of confidence)], it can be calculated that the true result variability due to biological variation would not be increased by more than 12% due to the analytical uncertainty.

As described above, bias in individual measurements can also have a significant effect on a patient result [7]. For desirable assay performance, it has been accepted that the analytical bias (B_a) should be less than one-quarter of the total biological variation, where total biological variation is made up of within-subject (CV_i) and between-subject biological variation (CV_g), i.e.,

$$B_a < 0.25\ (CV_i^2 + CV_g^2)^{1/2} \quad (2)$$

These two equations can then be used to define general desirable quality specifications for total allowable error (TE_a) as

$$TE_a < 1.65\ (0.5\ CV_i) + 0.25\ (CV_i^2 + CV_g^2)^{1/2} \quad (3)$$

where 1.65 is the Z-score at 95% probability [21].

Different levels of analytical quality can be defined by modulating the described theory [22] so that, minimum analytical performance goals can be defined as

$$\text{expanded uncertainty} < 0.75\ CV_i$$

$$B_a < 0.375\ (CV_I^2 + CV_g^2)^{1/2}$$

$$TE_a < 1.65\ (0.75\ CV_i) + 0.375\ (CV_I^2 + CV_g^2)^{1/2}$$

and optimum desirable performance as:

$$\text{expanded uncertainty} < 0.25\ CV_i$$

$$B_a < 0.125\ (CV_I^2 + CV_g^2)^{1/2}$$

$$TE_a < 1.65\ (0.25\ CV_i) + 0.125\ (CV_I^2 + CV_g^2)^{1/2}$$

If for being acceptable, the degree of uncertainty (expanded) of a measurand in the clinical laboratory (including the accumulated uncertainty of the corresponding traceability chain) using unbiased assays should stay within $\pm 0.25\ CV_i$, $\pm 0.50\ CV_i$ or $\pm 0.75\ CV_i$ (optimum, desirable or minimum quality level, respectively), decisions need to be made on what proportion of this budget can be used up in the traceability chain to ensure enough budget is left for use in routine analysis. In the past, only the contribution of the uncertainty related to the reference procedures has been considered. In particular, Stöckl et al. [23, 24] proposed that this uncertainty should be <0.2 times the maximal tolerated limit, i.e., the clinically allowable uncertainty of measurements.

We have selected three measurands to analyze the current situation and evaluate if traceability is compatible with the application of the biological variation approach.

Example 1

The desirable and minimum expanded uncertainty for serum creatinine are 3.0% and 4.5%, respectively [25]. Currently, NIST provides a reference material, SRM 967a (frozen human serum), in which there are two concentration levels of creatinine [26]. The certified concentration values for each creatinine level assigned by the reference procedure using isotope dilution liquid chromatography/mass spectrometry (LC-IDMS) are as follows: 74.9±1.6 μmol/L for level 1 and 342.7±7.2 μmol/L for level 2. This corresponds to an expanded uncertainty of 2.1% for both levels, which is lower than both desirable (3.0%) and minimum (4.5%) goals of expanded uncertainty derived from biologic variability. If this reference material is used to transfer trueness to manufacturer's calibrators, there is still 30% (if desirable goal is used) or approximately 53% (if minimum goal is used) of the total uncertainty budget available for the remainder of the chain. Only for laboratories working with an uncertainty for serum creatinine (i.e., the random effects of measurement) <2.0%, would a remaining uncertainty budget for the rest of the traceability chain be available to comply with the desirable expanded uncertainty limit of 3.0%. Otherwise, the minimum goal should be employed.

Example 2

Another reference material provided by NIST is the SRM 909c, which contains glucose at a concentration of 5.05±0.059 mmol/L as measured by isotopic dilution gas chromatography-mass spectrometry (GC-IDMS), corresponding to an expanded uncertainty of 1.2%, which is well within the desirable expanded uncertainty for measurement of serum glucose derived from biologic variability (3.1%) [24]. If this reference material is used in the traceability chain for glucose, it would leave more than 60% of the uncertainty budget. In one of our laboratories, the CV for serum glucose is around 1.8%, still leaving more than 30% uncertainty budget for the remaining parts of the traceability chain.

One manufacturer who provided traceability data for glucose to the WG-AETR, used NIST SRM 917b as the higher-order reference material. This is a powder material of D-glucose of purity 99.7±0.2%, which was used to prepare their reference calibration curve by weighing into a volumetric flask. Three final calibrators were prepared that gave the following concentrations: 1.83±0.055 mmol/L, 16.50±0.267 mmol/L, and 33.13±0.587 mmol/L. These equate to expanded uncertainties of 3.0%, 1.6% and 1.8%, respectively, that, at least for the lower glucose concentration, already used all of the desirable uncertainty budget for all the traceability chain of 3.1%.

Example 3

The NIST reference material for sodium in human serum is SRM 956c, which contains sodium at a concentration level of 118.8±1.0 mmol/L, i.e., an expanded uncertainty of 0.84%. The concentration of sodium in this material was assigned by high-performance inductively coupled plasma-optical emission spectrometry (ICP-OES) and ISE potentiometry. As the minimum performance goal for expanded uncertainty of measurement of sodium in serum is 0.50% [24], it is evident that in order to apply the biologic variability approach this reference material is not appropriate because its displays approximately 70% more uncertainty than the allowable total uncertainty.

Alternate statistical approach

As demonstrated by the example of sodium, the approach based on biologic variability may be too rigid to be able to be adopted at present for all measurands. An alternative approach is to use probability theory. The question then becomes *"In a traceability chain, what is the probability that the final sample result has the desired value?"* Obviously, this can never be absolutely certain, but the more times the measurements at the individual steps of a traceability chain are repeated, the more certain the final result. This approach requires a statistical goal to be set for the probability of the final step being correct, e.g., 95% certainty.

In the glucose example shown in Figure 1, there are four measurement steps, weighing of the primary calibrator, measurement of the secondary calibrator using a reference procedure (IDMS) by the manufacturer, measurement of the manufacturer's master calibrator by the manufacturer's selected measurement procedure and finally measurement of the product calibrator by the manufacturer's standing procedure. Statistically independent repeated observations are determined at each step and this chain of repeated measurement procedures would give mean±SD for each step [16]. Assuming the repeated measurements are normally distributed, then the confidence interval (CI) for each step can be calculated and it would be reasonable to have 95% CI as the level of acceptability.

However, for this approach to be valid, there needs to be defined limits for the 95% CI because this interval is dependent on the dispersion of the results so that an imprecise method would have a much wider interval. We suggest the 95% CI be based on desirable TE_a. For serum glucose the TE_a is 6.9% and if the glucose value was 5.05 mmol/L would give a range of 5.07±3.45% (± half the TE_a of 6.9%), that is 4.88–5.22 mmol/L.

In a theoretical experiment in which glucose was measured 10 times so that the final result and uncertainty fitted with the 95% CI of 4.88–5.22 mmol/L, the results were 5, 5.05, 5.1, 4.95, 5.0, 5.1, 4.95, 5.1, 5.15, 4.95 mmol/L. These results give a mean±SD of 5.035±0.0747 mmol/L and 95% CI of 4.889–5.171. These results would be acceptable but it means that the measuring procedure would be required to generate a standard deviation of 0.0747 which is 1.47% or less at the mean concentration of 5.035 mmol/L. In one of our laboratory, the CV for glucose is 1.7% at 4.7 mmol/L so it should be possible for the manufacturer to meet the 1.5% SD requirement using their very precise, unbiased and standardized procedures although it may also require a significant number of repeat determinations at each step.

Conclusions

Patients often have clinical results from different laboratories and without traceability to a reference measurement system, physicians have little or no information about the equivalence of results. To ensure better healthcare for patients, the harmonization of laboratory results is a major initiative involving health professionals, metrological organizations and IVD manufacturers. Metrological traceability means that all measurements made with a calibrated instrument or device are directly traceable back to a stable reference such as the SI. Currently, there are no defined criteria for acceptable limits for the traceability chain of specific measurands and the WG-AETR has been identifying approaches to define these limits.

In particular, two limits should be defined for the clinical application of the test once a traceability chain has been defined: allowable limits for the uncertainty of manufacturer's calibrators and the allowable error for measurements done by individual clinical laboratories. Basically, it is considered appropriate that any defined limits should be based on biological variation in the absence of agreed analytical goals based on clinical outcomes. For the former, desiderable (<0.50 CV_i) or minimum (<0.75 CV_i) goals can be applied for defining the total expanded uncertainty and a concomitant evaluation of the three major sources of uncertainty in the traceability chain (reference materials, manufacturer's calibrators, and value transfer procedures and uncertainty related to the assay) should be performed to understand the contribution of each source and the state-of-the-art of the measurement [9, 10]. The allowable error for measurements done by individual clinical laboratories can actually refer to the classical TE_a concept.

We have used the TE_a for creatinine and glucose to show that it is possible for manufacturers correctly applying the traceability approach to provide commercial calibrators with an uncertainty that leaves enough uncertainty budget to enable clinical laboratories to measure these measurands with an accuracy suitable for clinical application of the test. However, there are other measurands, such as sodium or albumin in serum, in which available reference materials display too high uncertainties and/or current methods are too inaccurate to allow the fulfilment of goals derived from biologic variation. These are measurands for which, depending on their biology and strict homeostatic control, the requested analytical quality is high and the performance of field methods should be extremely good to permit their application in clinical setting.

The next step for the Working Group is to apply this approach to every measurand measured in the clinical laboratory for which a reference material exists in the JCTLM database in order to establish if the current status of the uncertainty budget of its measurement associated with the proposed metrological traceability chain is suitable for clinical application of the test. As reported by others [27, 28], we strongly believe that the specifications of certified reference materials and calibration materials should be defined by the performance needs of the clinical assays; therefore, stakeholders should be prepared to provide new suitable reference materials together with improved assay methods whenever the clinical application of the test is made questionable. The required quality of these materials should be based on data showing the effect of assay performance on clinical outcome.

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.

Received June 13, 2012; accepted January 30, 2013; previously published online March 23, 2013

References

1. Panteghini M. Traceability as a unique tool to improve standardization in laboratory medicine. Clin Biochem 2009;42: 236–40.
2. Vesper HW, Thienpont LM. Traceability in laboratory medicine. Clin Chem 2009;55:1065–75.
3. 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 Rev 2007;28:105–14.
4. Panteghini M, Ceriotti F. Obtaining reference intervals traceable to reference measurement systems: it is possible, who is responsible, what is the strategy? Clin Chem Lab Med 2012;50:813–7.
5. ISO 17511. In vitro diagnostic medical devices – measurement of quantities in biological samples – metrological traceability of values assigned to calibrators and control materials. Geneva, Switzerland: ISO, 2003.
6. IFCC/CLSI X5-R. Metrological traceability and its implementation; a report. Wayne: PA: CLSI, 2006.
7. Panteghini M. Application of traceability concepts to analytical quality control may reconcile total error with uncertainty of measurement. Clin Chem Lab Med 2010;48:7–10.
8. Thienpont LM, Van Uytfanghe K, Rodriguez Cabaleiro D. Metrological traceability of calibration in the estimation and use of common medical decision-making criteria. Clin Chem Lab Med 2004;42:842–50.
9. Panteghini M. Implementation of standardization in clinical practice: not always an easy task. Clin Chem Lab Med 2012;50:1237–41.
10. Infusino I, Panteghini M. Serum albumin: Accuracy and clinical use. Clin Chim Acta 2013;419:15–8.
11. Bais R. What information should manufacturers provide on their procedures? Clin Chem 2006;52:1624–5.
12. European Diagnostic Manufacturers Association. Position Paper. Estimation of uncertainty of measurement in medical laboratories. Brussels, BE, 2006.
13. CLSI document C51-A. Expression of measurement uncertainty in Laboratory Medicine; Approved guideline. Wayne, PA: CLSI, 2012.
14. Armbruster D. Measurement traceability and US IVD manufacturers: the impact of metrology. Accred Qual Assur 2009;14:393–8.
15. Panteghini M. Traceability in laboratory medicine: Copernican revolution or activity for a restricted professional club? 4th International Scientific Meeting " Rethinking quality control in the traceability era ". November 30, 2010, Milano, Italy. Available from: http://users.unimi.it/cirme/public/UploadAttach/Panteghini2010.pdf. Accessed January, 2013.
16. International Organization for Standardization, Joint Committee for Guides in Metrology. Evaluation of measurement data – Guide to the expression of uncertainty in measurement (GUM). Geneva, Switzerland: ISO, 1995.
17. Kallner A. Measurement performance goals: how they can be estimated and a view to managing them. Scand J Clin Lab Invest 2010;70(Suppl 242):34–9.
18. 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.
19. Braga F, Dolci A, Mosca A, Panteghini M. Biological variability of glycated hemoglobin. Clin Chim Acta 2010;411:1606–10.
20. Braga F, Panteghini M. Biologic variability of C-reactive protein: is the available information reliable? Clin Chim Acta 2012;413:1179–83.
21. Fraser CG. Biological variation: from principle to practice. Washington: AACC Press, 2001.
22. 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.
23. Stöckl D, Reinauer H. Development of criteria for the evaluation of reference method values. Scand J Clin Invest 1993;53(Suppl 212):16–8.
24. Stöckl D, Franzini C, Kratochvila J, Middle J, Ricós C, Siekmann L, et al. Analytical specifications of reference methods. Compilation and critical discussion. Eur J Clin Chem Clin Biochem 1996;34:319–37.
25. Westgard J. Desirable Biological Variation Database specifications. Available from: http://www.westgard.com/biodatabase1.htm. Accessed January, 2013.
26. Panteghini M, Myers GL, Miller GW, Greenberg N. The importance of metrological traceability on the validity of creatinine measurement as an index of renal function. Clin Chem Lab Med 2006;44:1187–92.
27. Bunk DM, Welch MJ. Characterization of a new certified reference material for human cardiac troponin I. Clin Chem 2006:52:212–9.
28. CLSI C-53A. Characterization and qualification of commutable reference materials for Laboratory Medicine; Approved guidelines. Wayne, PA: CLSI, 2010.