



The standardization journey and the path ahead

W. Greg Miller

Department of Pathology, Virginia Commonwealth University, Richmond, Virginia, USA

Correspondence to: W. Greg Miller, Department of Pathology, Virginia Commonwealth University, Richmond, Virginia, 23298-0286, USA.

Email: greg.miller@vcuhealth.org.

Abstract: Standardization of clinical laboratory test results has progressed through several stages. External quality assessment (EQA) identified that results were not equivalent in different laboratories in the 1950s. Approaches to use the same methods were promoted in the 1960s and the challenges recognized during the 1970s as *in vitro* diagnostics manufacturers introduced a variety of technologies to meet the demand for clinical laboratory testing services. Hierarchies of national reference systems were developed in the 1980–1990 period that were hampered by inadequate understanding of the importance of commutable reference materials both for use as calibrators and for EQA. Standards for metrological traceability were developed by the International Organization for Standardization in the 2000s that provided a global approach for establishing robust standardization for approximately 100 measurands. The limitations of non-commutable reference materials became widely appreciated in the 2000s. In 2010 and going forward, laboratory medicine recognized the need for harmonization approaches for the large number of measurands for which higher order certified reference materials and reference measurement procedures are not available. Metrological traceability to certified reference materials and reference measurement procedures remains the first choice when technologically feasible. Metrological traceability to an international harmonization protocol provides an alternative when no other approach is realistically feasible. Achieving equivalent results among different medical laboratory measurement procedures remains an important goal to enable appropriate medical decisions based on laboratory results and decision values included in clinical practice guidelines.

Keywords: Harmonization; metrological traceability; standardization

Received: 10 August 2018; Accepted: 27 September 2018; Published: 17 October 2018.

doi: 10.21037/jlpm.2018.09.14

View this article at: <http://dx.doi.org/10.21037/jlpm.2018.09.14>

Medical laboratories have long recognized the importance of standardizing the steps in the path of workflow for producing a laboratory test result that is suitable for making medical decisions. The critical steps include ordering the correct test, preparing the patient, collecting the specimen, transporting the specimen to the laboratory, performing the test procedure, reporting the test result in the correct units and with the correct interpretive information, and consulting with the clinical provider regarding the results and follow-up testing that may be indicated. Each of these steps influences the effectiveness of the “brain to brain” loop between the patient who seeks medical care, the physician who provides that care and the medical laboratory professional who closes the loop regarding the laboratory medicine results (1). This report reviews progress that has

been made regarding standardizing the results of laboratory measurement procedures and describes the current situation regarding tools and procedures available for this purpose.

External quality assessment (EQA) can assess the need for harmonization of test results and monitor the success of procedures to achieve harmonization of clinical laboratory test results. The first EQA results were published in 1947 and documented poor agreement among 59 hospital laboratories measuring the same analytes (2). For example, in that first EQA assessment, most participants reported values from 40–95 mg/dL (2.2–5.3 mmol/L) for a glucose sample prepared to contain 60 mg/dL (3.3 mmol/L) with 4 laboratories reporting values >333 mg/dL (>18.5 mmol/L). Similarly, most participants reported values from 4.5–9.0 g/dL (45–90 g/L) for a total serum protein sample prepared to be

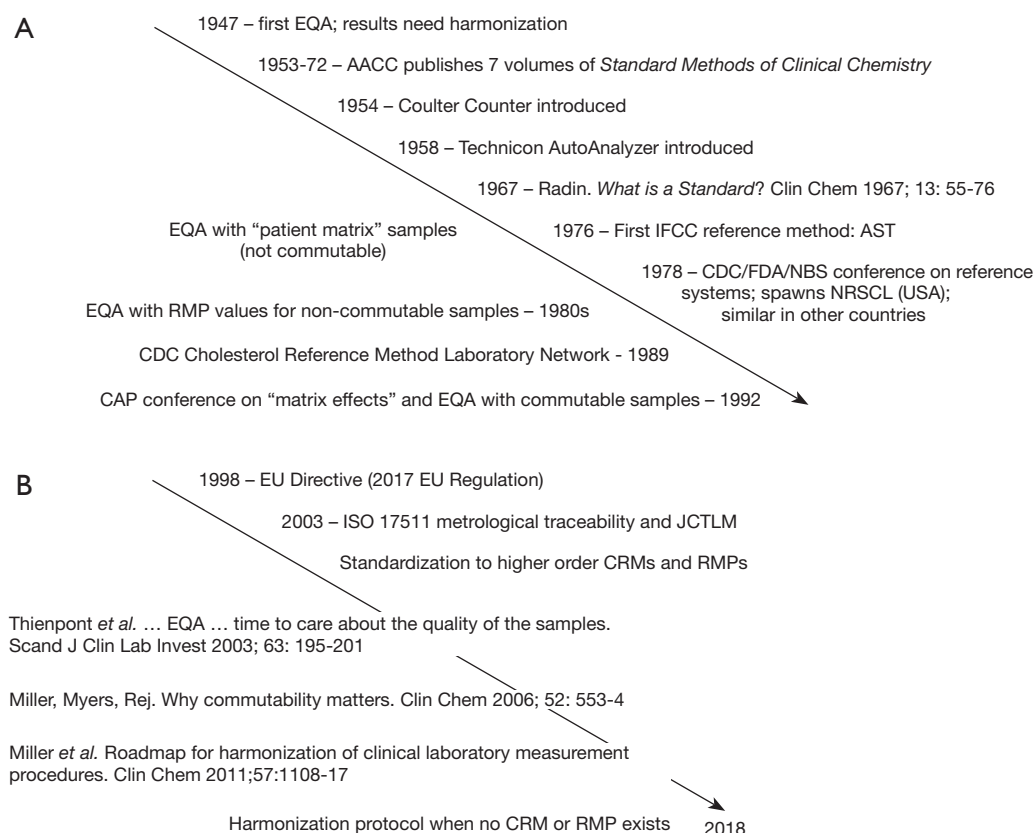


Figure 1 Timeline for key developments in standardization and harmonization of medical laboratory results. AACC, American Association for Clinical Chemistry; CAP, College of American Pathologists; CDC, Centers for Disease Control and Prevention; CRM, Certified Reference Material; EQA, external quality assessment; EU, European Union; FDA, Food and Drug Administration; IFCC, International Federation for Clinical Chemistry and Laboratory Medicine; ISO, International Organization for Standardization; NBS, National Bureau of Standards; NRSCl, National Reference System for the Clinical Laboratory; RMP, reference measurement procedure; USA, United States of America.

6.6 g/dL (66 g/L) with 2 laboratories reporting 12.0 g/dL (120 g/L). Other early EQA programs reported similar discrepancies among results from different laboratories (3,4). At that time, laboratory testing was by manual procedures using reagents and calibrators prepared by a laboratory or purchased from a supplier of reagent kits. Early EQA assessments influenced the laboratory medicine profession to prioritize developing approaches for harmonization of laboratory test results.

Figure 1 shows an overview timeline of key events in developing approaches for standardizing or harmonizing results from different measurement procedures. One early approach was developing standard methods of measurement that could be adopted by all laboratories. In 1953 the American Association for Clinical Chemistry (AACC) published the first volume of a 7 volume series

that extended to 1972 titled “*Standard Methods of Clinical Chemistry*.” The International Federation of Clinical Chemistry and Laboratory Medicine (IFCC) published the first of a series of reference measurement procedures in 1976. Using standard methods was beneficial but did not solve the problem because adoption was voluntary and there was no metrological traceability hierarchy of reference materials and reference measurement procedures to provide an infrastructure for standardization. Automated analyzers were introduced in the 1950s that revolutionized laboratory medicine. The Coulter Counter was introduced in 1954 and was a major advance in standardizing blood cell counting. The Technicon AutoAnalyzer was introduced in 1958 and provided a platform for laboratories to use the same measurement procedures and calibrators for clinical biochemistry testing. Reagents or calibrators from any

supplier, including laboratory prepared, could be easily adopted for the AutoAnalyzer so the potential of consistent measurement procedures was not realized. Beginning in the 1960s, a large number of different *in vitro* diagnostic (IVD) manufacturers began to market automated measurement procedures for clinical biochemistry each of which used different reagent formulations, calibrators and measuring conditions. The combination of different methods and making measurements directly on serum or plasma without any pre-treatment to remove interfering substances lead to an increasing disparity among results from different measurement procedures.

The situation in the mid-1960s was summarized by Radin (5) as a collection of various types of materials marketed as “standards” for calibration few of which were adequately characterized or actually effective for achieving equivalent results among different methods and reagent systems for the same analyte. Radin recognized the need for a system to certify calibration materials and definitive methods as the basis for a calibration hierarchy. A conference in 1978 sponsored by the Centers for Disease Control (CDC), the Food and Drug Administration and the National Bureau of Standards in the United States of America (USA) (6) concluded that a calibration hierarchy of certified reference materials and reference measurement procedures was needed as the basis for standardizing the metrological traceability of results produced by the so called “routine” measurement procedures used in medical laboratories. The National Reference System for the Clinical Laboratory (NRSL) was established in the USA and similar national programs for metrological traceability were established in other countries. The NRSL established criteria for credentialing reference materials and reference measurement procedures and maintained a list of such references that were reviewed and certified as suitable for use. The NRSL and its counterparts in other countries represented a major advance in developing the infrastructure needed to achieve standardized results. An obvious disadvantage of national systems is lack of coordination that lead to different calibration standards in different countries and that became increasingly more challenging as global distribution of IVD medical devices became common. The CDC established the Cholesterol Reference Method Laboratory Network in 1989 to provide the first such network of reference laboratories all providing the same reference measurement procedure as a resource for metrological traceability to standardize cholesterol measurements on a global basis (7). This network became

the model for others that were developed to standardize, for example, hemoglobin A1c and other measurands.

During the 1960s and 1970s the EQA samples were non-commutable with clinical samples and thus not suitable for assessment of accuracy among different measurement procedures (8). However, the limitations of non-commutability were not understood at that time and EQA results were incorrectly used to assess the effectiveness of the calibration hierarchies being deployed to improve standardization among results from different measurement procedures. IVD manufacturers also incorrectly used EQA results as the basis for calibration adjustments so their customers would receive acceptable scores for EQA results without realizing that the EQA results from non-commutable samples were not reflective of the agreement for clinical sample results. During the 1980s, some EQA providers introduced programs that used target values assigned by reference measurement procedures to non-commutable EQA materials in an attempt to improve assessment of standardization among different measurement procedures. When discrepancies between EQA results and results for clinical samples were identified, IVD manufacturers were given a grading exception and the term matrix effect was used to describe a situation when a difference in bias was observed for an EQA sample *vs.* clinical samples.

The College of American Pathologists held a conference in 1992 to discuss matrix effects and accuracy assessment in clinical chemistry (9). This conference clearly identified that commutability was a necessary property of reference materials used in particular for EQA assessment but also as matrix-based certified reference materials in calibration hierarchies when the goal was assessment or establishment of equivalence among results for clinical samples. Following this conference, the College of American Pathologists and other EQA providers began to introduce EQA surveys that used commutable samples for selected analytes for which standardization programs were established such as cholesterol and hemoglobin A1c. Today, the limitations of non-commutable EQA materials are well understood. However, it was not until the 2000s that the impact of non-commutable EQA or matrix-based certified reference materials was fully appreciated (10,11).

The European Union (EU) Directive 98/79/EC was published in 1998 with an effective date of 2003 (12). This legislation was the first requirement that IVD medical devices for medical laboratory testing have metrological traceability to higher order references. This directive

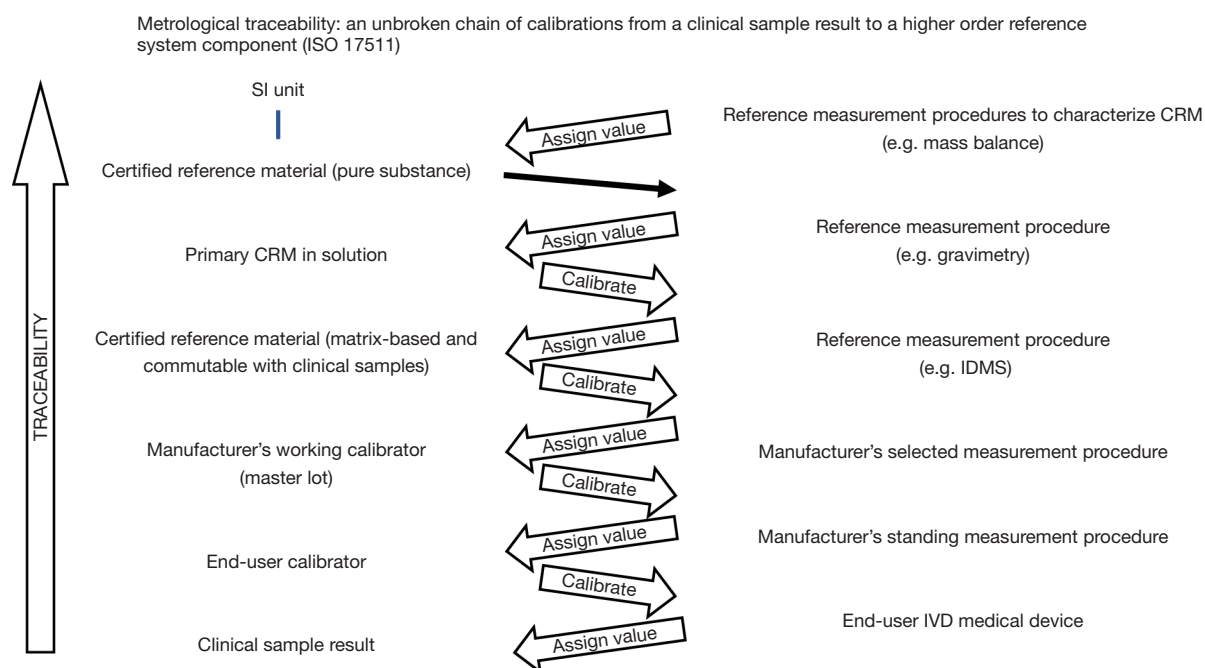


Figure 2 Metrological traceability hierarchy as described in ISO standard 17511 (13). SI, International System of Units (French *Système Internationale d'Unités*); CRM, Certified Reference Material; IDMS, isotope dilution mass spectrometry; IVD, in vitro diagnostic.

created a need for international standards for metrological traceability that were developed by the International Organization for Standardization (ISO) Technical Committee (TC) 212, *Clinical laboratory testing and in vitro diagnostic test systems*, and originally published in 2003 (13-17). The ISO standards have replaced the national programs and are the current system followed by IVD manufacturers to establish calibration hierarchies for their measurement procedures. The Joint Committee for Traceability in Laboratory Medicine (JCTLM) was established in 2003, in response to the EU Directive, to provide a review process and listing of higher order certified reference materials, reference measurement procedures and reference (calibration) laboratory services that conformed to the ISO requirements (18). The EU Directive was replaced with an EU regulation in 2017 (19), with an effective date of 2022, that continues the requirement for metrological traceability with a more stringent review and approval process before IVD medical devices can be sold in Europe.

The ISO standard 17511 describes a metrological traceability hierarchy as shown in *Figure 2*. During the 1980s through 2000s, substantial effort and emphasis was put on developing reference systems that included reference measurement procedures, primary (pure

substance) certified reference materials to calibrate the reference measurement procedures and secondary (matrix-based) certified reference materials based initially on the national and later on the ISO standards. During this period, the limitations of non-commutable reference materials became well documented (20-22). Reports of discrepant clinical sample results among measurement procedures with claimed metrological traceability to the same matrix-based certified reference material highlighted that non-commutable reference materials were not suitable for use in calibration hierarchies (11).

The ISO 17511 standard includes calibration hierarchies applicable when not all of the higher order reference system components are available. For example, if there is no reference measurement procedure, then metrological traceability stops at a secondary matrix-based certified reference material. Such a reference material must be commutable with clinical samples for use with all of the measurement procedures for which it is intended. If an IVD manufacturer claims traceability to a non-commutable reference material, then the bias from non-commutability will be propagated in the calibration hierarchy and non-equivalent results for clinical samples will occur among different measurement procedures. If there is no secondary

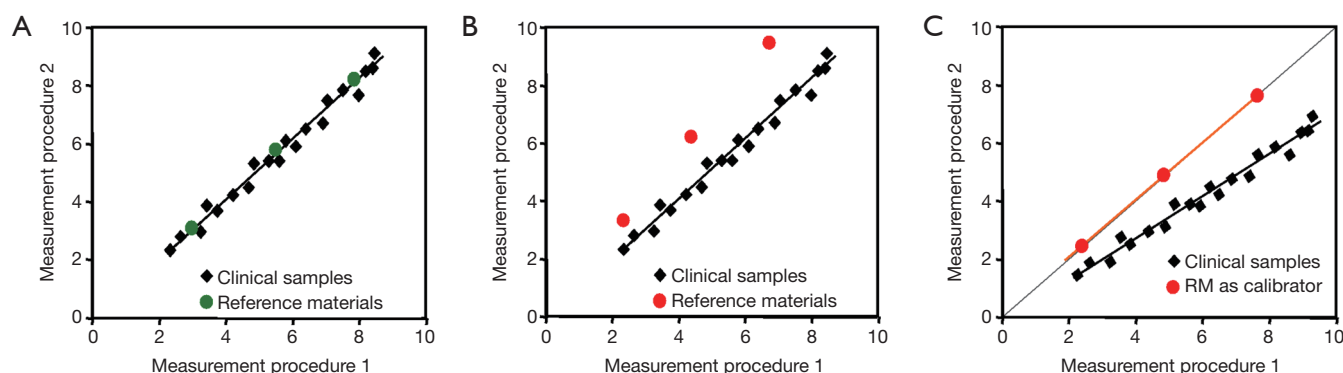


Figure 3 Commutable and non-commutable reference materials. Panel A shows that commutable reference materials have the same relationship between two measurement procedures as do clinical samples. Panel B shows that non-commutable reference materials have a different relationship between two measurement procedures than do clinical samples. Panel C shows that when non-commutable reference materials (RM) are used as calibrators, the calibrators appear to have equivalent values for each measurement procedure but the results for clinical samples are not equivalent between the two measurement procedures.

matrix-based certified reference material, then traceability stops at the manufacturer's working calibrator. Since there is no coordination among different manufacturers, different working calibrators will be used leading to non-equivalent results for clinical samples among different measurement procedures.

Commutability is an essential property of a reference material when used as a matrix-based certified reference material in a calibration hierarchy or as an EQA material for assessment of agreement of results among measurement procedures. Commutability is formally defined as a property of a reference material, demonstrated by the closeness of agreement between the relation among the measurement results for a stated quantity in this material, obtained according to two given measurement procedures, and the relation obtained among the measurement results for other specified materials (23). For medical laboratories, other specified materials are the clinical samples intended to be measured, and the quantity is usually referred to as the measurand. A working definition of commutability can be stated as a property of a reference material whereby the same numeric relationship, within clinically meaningful limits, can be demonstrated between 2 or more measurement procedures for both the reference material and a panel of representative individual patient samples (24). *Figure 3A* shows the commutable condition. When commutable reference materials are used for calibration of each measurement procedure, the results for clinical samples agree. Similarly, when commutable EQA samples are used, the EQA results reflect

the results for clinical samples. *Figure 3B* shows the non-commutable condition where the relationship between the two measurement procedures is different for reference materials *vs.* clinical samples. Results that disagree for non-commutable EQA samples do not mean that results disagree for clinical samples. *Figure 3C* shows that when non-commutable reference materials are used for calibration of each measurement procedure, the relationship established by the non-commutable reference materials will cause results for the clinical samples to disagree. Similarly, when results for non-commutable EQA samples appear to agree between the two measurement procedures, the results for clinical samples will disagree.

JCTLM lists reference materials and reference measurement procedures for approximately 100 measurands. JCTLM did not require commutability data for matrix-based certified reference materials intended for use as calibrators until 2014. Consequently, some of the matrix-based certified reference materials listed are not suitable for use as calibrators. Laboratory medicine needs to develop approaches to achieve equivalent results among different measurement procedures for the hundreds of measurands for which certified reference materials and reference measurement procedures do not exist or for which such higher order references are not likely to be developed for various technical reasons. The AACC organized an international conference in 2010 to discuss improving clinical laboratory testing through harmonization. Harmonization is operationally defined as achieving equivalent results among different measurement

procedures for the same laboratory test and is frequently used to imply there are no certified reference materials or reference measurement procedures available. The term standardization is typically used when equivalent results are achieved by metrological traceability to fit-for-purpose higher order reference system components. Equivalent does not mean identical. Equivalent means within a total allowable error consistent with an acceptable risk of harm from medical decisions based on a laboratory test result.

The 2010 conference recommendations were published as a roadmap for harmonization of clinical laboratory measurement procedures (24). Key recommendations were to establish an organization to prioritize measurands in need of harmonization, to provide an information portal to coordinate globally the work of different organizations to harmonize (or standardize) a measurand, and to develop procedures to achieve harmonization when certified reference materials and reference measurement procedures are not available. The International Consortium for Harmonization of Clinical Laboratory Results (ICHCLR) was formed to fulfil the recommendations (25) and its progress was recently reviewed (26). The ICHCLR web site includes a measurand table with information on priority and harmonization activity, and a resources section with a toolbox of procedures to achieve harmonization for a measurand.

One of the toolbox strategies is called a step-up approach for harmonization that was developed and used by the IFCC Committee for Standardization of Thyroid Function Tests to establish a procedure to harmonize results for thyroid stimulating hormone (TSH) measurements. The harmonization process for TSH has been described (27) and is based on using results from panels of individual clinical samples as harmonization reference materials. Results from the panels are used to progressively step from one experiment to the next to develop and validate correction algorithms for the calibration hierarchies of each IVD manufacturer's TSH measurement procedure to achieve equivalent results for the clinical samples. The technical feasibility of this approach has been validated and is intended to be introduced into clinical laboratory practice with coordination of education of laboratories and clinical providers regarding changes in values and reference intervals, meeting regulatory requirements in different countries, and implementing recalibrated calibration hierarchies by IVD manufacturers at approximately the same time to minimize any disruption of clinical care (28).

The ICHCLR recognized that an ISO standard would

be needed for a harmonization protocol to be recognized as an acceptable approach to achieve metrological traceability. A new work item proposal was introduced and approved by ISO TC212 to develop a new standard designated ISO/CD 21151: *In vitro diagnostic medical devices—Measurement of quantities in samples of biological origin—Requirements for international harmonization protocols intended to establish metrological traceability of values assigned to product (end user) calibrators and patient samples* (29). This new standard has been approved by ISO TC 212 as a draft international standard but has not completed the subsequent ISO voting cycles to be published as an international standard at the time of this report. Consequently, its contents cannot be summarized in this report. The ISO 21151 draft international standard is based on the scientific principles developed for harmonization of TSH by the IFCC Committee for Standardization of Thyroid Function Tests and includes requirements for the various steps in the process. When published, the new ISO standard can be considered by JCTLM as the basis to list harmonization protocols as part of a calibration hierarchy available for use by IVD manufacturers.

In summary, standardization of clinical laboratory test results has progressed through several stages. EQA identified that results were not equivalent in different laboratories in the 1950s. Approaches to use the same methods were promoted in the 1960s and the challenges recognized during the 1970s as IVD manufacturers introduced a variety of technologies to meet the demand for clinical laboratory testing services. Hierarchies of national reference systems were developed in the 1980–1990 period that were hampered by inadequate understanding of the importance of commutable reference materials both for use as calibrators and for EQA. The ISO standards for metrological traceability in the 2000s provided a global approach for establishing robust standardization for approximately 100 measurands. The limitations of non-commutable reference materials became widely appreciated in the 2000s. In 2010 and going forward, laboratory medicine recognized the need for harmonization approaches for the large number of measurands for which higher order certified reference materials and reference measurement procedures are not available. Metrological traceability to certified reference materials and reference measurement procedures remains the first choice when technologically feasible. Metrological traceability to an international harmonization protocol provides an alternative when no other approach is realistically feasible. Achieving equivalent

results among different medical laboratory measurement procedures remains an important goal to enable appropriate medical decisions based on laboratory results and decision values included in clinical practice guidelines.

Acknowledgments

Funding: None.

Footnote

Provenance and Peer Review: This article was commissioned by the Guest Editors (Mario Plebani and Giuseppe Lippi) for the series “Laboratory Medicine-25 Years on” published in *Journal of Laboratory and Precision Medicine*. The article has undergone external peer review.

Conflicts of Interest: The author has completed the ICMJE uniform disclosure form (available at <http://dx.doi.org/10.21037/jlpm.2018.09.14>). The series “Laboratory Medicine-25 Years on” was commissioned by the editorial office without any funding or sponsorship. The author has no other conflicts of interest to declare.

Ethical Statement: The author is accountable for all aspects of the work in ensuring that questions related to the accuracy or integrity of any part of the work are appropriately investigated and resolved.

Open Access Statement: This is an Open Access article distributed in accordance with the Creative Commons Attribution-NonCommercial-NoDerivs 4.0 International License (CC BY-NC-ND 4.0), which permits the non-commercial replication and distribution of the article with the strict proviso that no changes or edits are made and the original work is properly cited (including links to both the formal publication through the relevant DOI and the license). See: <https://creativecommons.org/licenses/by-nc-nd/4.0/>.

References

1. Plebani M, Laposata M, Lundberg GD. The Brain-to-Brain Loop Concept for Laboratory Testing 40 Years After Its Introduction. *Am J Clin Pathol* 2011;136:829-33.
2. Belk WP, Sunderman FW. A survey of the accuracy of chemical analyses in clinical laboratories. *Am J Clin Pathol* 1947;17:853-61.
3. Wootton ID, King EJ. Normal values for blood constituents: inter-hospital differences. *Lancet* 1953;1:470-1.
4. Sunderman FW. Forty-five years of proficiency testing. *Ann Clin Lab Sci* 1991;21:143-4.
5. Radin N. What is a standard. *Clin Chem* 1967;13:55-76.
6. Boutwell JH. Editor. A National Understanding for the Development of Reference Materials and Methods for Clinical Chemistry. The American Association for Clinical Chemistry, Washington DC, 1998.
7. Myers GL, Kimberly MM, Waymack PP, et al. A reference method laboratory network for cholesterol: a model for standardization and improvement of clinical laboratory measurements. *Clin Chem* 2000;46:1762-72.
8. Miller WG, Jones GRD, Horowitz GL, et al. Proficiency Testing/External Quality Assessment: Current Challenges and Future Directions. *Clin Chem* 2011;57:1670-80.
9. College of American Pathologists Conference XXIII: Matrix Effects and Accuracy Assessment in Clinical Chemistry, June 1992. *Arch Pathol Lab Med* 1993;117:343-436.
10. Miller WG, Myers GL, Rej R. Why commutability matters. *Clin Chem* 2006;52:553-4.
11. Miller WG, Myers GL. Commutability still matters. *Clin Chem* 2013;59:1291-3.
12. Directive 98/79/EC of the European Parliament and of the Council of 27 October 1998 on in vitro diagnostic medical devices OJ L 331 of 7 December 1998.
13. ISO 17511:2003 In vitro diagnostic medical devices – Measurement of quantities in biological samples – Metrological traceability of values assigned to calibrators and control materials. International Organization for Standardization, Geneva, Switzerland, 2003.
14. ISO 15193:2009 In vitro diagnostic medical devices – Measurement of quantities in samples of biological origin – Requirements for content and presentation of reference measurement procedures. International Organization for Standardization, Geneva, Switzerland, 2009.
15. ISO 15194:2009 In vitro diagnostic medical devices – Measurement of quantities in samples of biological origin – Requirements for certified reference materials and the content of supporting documentation. International Organization for Standardization, Geneva, Switzerland, 2009.
16. ISO 15195:2003 Laboratory medicine – Requirements for reference measurement laboratories. International Organization for Standardization, Geneva, Switzerland, 2003.
17. ISO 18153:2003 In vitro diagnostic medical devices – Measurement of quantities in samples of biological origin – Metrological traceability of values for catalytic

- concentration of enzymes assigned calibrators and control materials. International Organization for Standardization, Geneva, Switzerland, 2003.
18. Jones GRD, Jackson C. The Joint Committee for Traceability in Laboratory Medicine (JCTLM) – its history and operation. *Clin Chim Acta* 2016;453:86-94.
 19. Regulation (EU) 2017/746 of the European Parliament and of the Council of 5 April 2017 on in vitro diagnostic medical devices and repealing Directive 98/79/EC and Commission Decision 2010/227/EU (Text with EEA relevance). Available online: <http://eur-lex.europa.eu/eli/reg/2017/746/oj>, accessed 4 August 2018.
 20. Miller WG. Specimen materials, target values and commutability for external quality assessment (proficiency testing) schemes. *Clin Chim Acta* 2003;327:25-37.
 21. Thienpont LM, Stockl D, Friedecky B, et al. Trueness verification in European external quality assessment schemes: time to care about the quality of the samples. *Scand J Clin Lab Invest* 2003;63:195-201.
 22. Vesper HW, Miller WG, Myers GL. Reference materials and commutability. *Clin Biochem Rev* 2007;28:139-47.
 23. International vocabulary of metrology—basic and general concepts and associated terms (VIM). 3rd Ed. JCGM; 2012. Sevres, France: International Bureau of Weights and Measures.
 24. Greg Miller W, Myers GL, Gantzer ML, et al. Roadmap for harmonization of clinical laboratory measurement procedures. *Clin Chem* 2011;57:1108-17.
 25. Available online: <http://www.harmonization.net>, accessed 4 August 2018.
 26. Myers GL, Miller WG. The roadmap for harmonization: status of the International Consortium for Harmonization of Clinical Laboratory Results. *Clin Chem Lab Med* 2018;56:1667-72.
 27. Thienpont LM, Van Uytfanghe K, DeGrande LAC, et al. Harmonization of serum thyroid-stimulating hormone measurements paves the way for the adoption of more uniform reference intervals. *Clin Chem* 2017;63:1248-60.
 28. Available online: <http://www.ifcc.org/ifcc-scientific-division/sd-committees/c-stft/c-stft-resources>, accessed 4 August 2018.
 29. ISO/CD 21151 In vitro diagnostic medical devices -- Measurement of quantities in samples of biological origin -- Requirements for international harmonization protocols intended to establish metrological traceability of values assigned to product (end user) calibrators and patient samples (in development). Available online: <https://www.iso.org/standard/69985.html>, accessed 4 August 2018.

doi: 10.21037/jlpm.2018.09.14

Cite this article as: Miller WG. The standardization journey and the path ahead. *J Lab Precis Med* 2018;3:87.