Current practices and challenges in the standardization and harmonization of clinical laboratory tests[1–3]

Hubert W Vesper,[4] Gary L Myers,[5] and W Greg Miller[6]*

[4]Division of Laboratory Sciences, National Center for Environmental Health, CDC, Atlanta, GA; [5]American Association for Clinical Chemistry, Washington, DC; and [6]Virginia Commonwealth University, Richmond, VA

ABSTRACT

Effective patient care, clinical research, and public health efforts require comparability of laboratory results independent of time, place, and measurement procedure. Comparability is achieved by establishing metrological traceability, which ensures that measurement procedures measure the same quantity and that the calibration of measurement procedures is traceable to a common reference system consisting of reference methods and materials. Whereas standardization ensures traceability to the International System of Units, harmonization ensures traceability to a reference system agreed on by convention. This article provides an overview of standardization and harmonization with an emphasis on commutability as an important variable that affects testing accuracy. Commutability of reference materials is required to ensure that traceability is established appropriately and that laboratory results are comparable. The use of noncommutable reference materials leads to inaccurate results. Whereas procedures and protocols for standardizing measurements are established and have been successfully applied in efforts such as the Hormones Standardization Program of the CDC, harmonization activities require new, more complex procedures and approaches. The American Association for Clinical Chemistry, together with its domestic and international partners, formed the International Consortium for Harmonization of Clinical Laboratory Results to coordinate harmonization efforts. Reference systems, as well as procedures and protocols to establish traceability of clinical laboratory tests, have been established and continue to be developed by national and international groups and organizations. Serum tests of thyroid function, including those for the thyroid hormones thyroxine and triiodothyronine, are among the clinical procedures for which standardization efforts are well under way. Approaches to the harmonization of measurement procedures for serum concentrations of thyroid-stimulating hormone are likewise under development. *Am J Clin Nutr* 2016;104(Suppl):907S–12S.

Keywords: clinical laboratory tests, harmonization, iodine status, standardization, thyroid function tests

INTRODUCTION

Clinical laboratory measurements are used to identify individuals with diseases or those at increased risk of disease, guide treatment decisions, monitor the success of treatment, and assess the risk of disease recurrence. Laboratory data obtained from epidemiologic studies in patient and nonpatient populations are essential for developing public health strategies and clinical practice guidelines for disease. To achieve these objectives, it is crucial that different measurement procedures produce results that are comparable as well as reliable (i.e., equivalent within clinically meaningful limits for the same patient samples), regardless of which laboratory produced the results or when they were produced. Note that in the present article, the term "patient samples" denotes clinical samples from patients in treatment or research settings as well as nonpatients in research or public health settings.

The importance of comparability of measurement results becomes obvious when results across studies are compared or results obtained throughout the length of a long-term investigational study are examined. Noncomparable results can make research findings from different studies appear inconsistent, and incorrect conclusions may be drawn. Furthermore, it is common for patients to be seen in a variety of health care settings in which different laboratories may use different measurement procedures. In addition, the measurement procedures used by a given laboratory may change while patients receiving treatment are being monitored. When comparing a patient's test results with clinical decision points described in evidence-based clinical practice guidelines, results that are not comparable across measurement procedures can lead to inconsistent assessment of a patient and, in some situations, incorrect treatment. In addition, electronic health records intended to facilitate the interpretation of patient data across health care systems are likely to become less useful in the aforementioned scenario.

In the absence of the standardization or harmonization of measurement procedures, patient sample results differ substantially across commercial assays for a number of serum tests of thyroid function that may be useful for assessing the benefits and

[1] Presented at the workshop "Assessment of Iodine Intake: Analytical Methods and Quality Control" held by the NIH Office of Dietary Supplements in Rockville, MD, 22–23 July 2014.

[2] The authors reported no funding received for this study.

[3] The findings and conclusions in this article are those of the authors and do not necessarily represent the views of the CDC.

*To whom correspondence should be addressed. E-mail: hvesper@cdc.gov.

First published online August 17, 2016; doi: 10.3945/ajcn.115.110387.

risks of iodine supplementation (1). Serum tests that have been studied methodically across commercial assays include the concentration of thyroid-stimulating hormone (TSH)[7] and concentrations of the thyroid hormone thyroxine (T4) as total T4 and free (i.e., unbound) T4 (FT4). This article provides an overview of key principles and activities used in laboratory medicine, clinical research, and public health to standardize or harmonize measurement methods and procedures for the purpose of ensuring comparable and reliable results.

ESTABLISHING TRACEABILITY OF RESULTS

To achieve comparable results, all measurement procedures must measure the same quantity. In addition, the calibration of all measurement procedures should be traceable to a common reference system consisting of reference methods and materials. The process for achieving these requirements is commonly referred to as "establishing metrological traceability." Metrological traceability is defined by the International Organization for Standardization (2) as "the property of the result of a measurement or the value of a standard whereby it can be related to stated references, usually national or international standards, through an unbroken chain of comparisons all having stated uncertainties." A diagram of the steps necessary to establish traceability is provided in **Figure 1**; the application of this approach to the evaluation of insulin immunoassay performance has been described elsewhere (3). Further details about the process of establishing metrological traceability are provided in recent reviews (4–6).

The terms "harmonization" and "standardization" are used to describe the 2 principal approaches for establishing metrological traceability. The term "standardization" is used when both of the following conditions prevail: *1*) the measurand (the analyte to be measured) is clearly defined and *2*) agreement of test results is achieved by establishing traceability to a higher-order reference measurement procedure or pure-substance reference material that can be defined by using the International System of Units [Systèm International (SI)]. Higher-order reference measurement procedures are well-characterized analytical methods that are intended for assigning target values to reference materials and have a level of accuracy, precision, and specificity that is higher than that typically observed with routine clinical measurement procedures. As described by Long et al. (7) in this supplement issue, the National Institute of Standards and Technology has been working with the NIH Office of Dietary Supplements for several years to develop higher-order reference measurement procedures and Standard Reference Materials to support the validation of new analytical methods relevant to the assessment of iodine status and thyroid function.

Unfortunately, the number of measurands that can be standardized is fairly limited in relation to the hundreds of tests performed in laboratory medicine; the reasons for this are the lack of clearly defined measurands, reference methods, and/or reference materials. Although standardization, as described above,

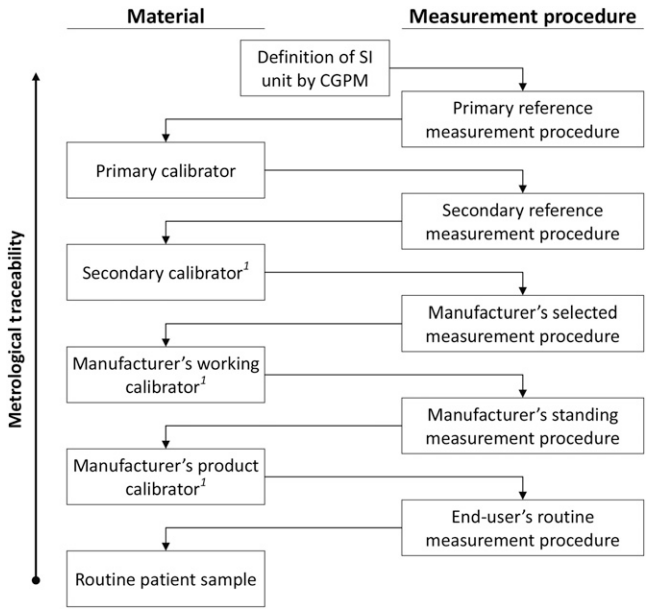

FIGURE 1 Diagram of the metrological traceability chain linking results obtained for patient samples (bottom) to SI units (top) through a series of calibrations of measurement procedures and value assignments of calibrator materials. [1]The calibrator can be a material with a matrix resembling those of the samples of human origin to be measured by the end-user's routine measurement procedure. CGPM, Conférence Générale des Poids et Mesures; SI, Systèm International (International System of Units).

cannot be achieved for many analytes, agreement among measurement procedures can still be obtained through the process called "harmonization," in which there is a reference system consisting of methods and materials that are not traceable to the SI but are agreed upon to act as references. For harmonization, a single method (called a "designated comparison method") can be selected or a set of different methods can be used to assign an "all-methods mean" to a reference material or materials. The terms "all-method trimmed mean" and "all-procedure trimmed mean" are also used to describe this summary statistic. Reference materials can be prepared from purified biomarkers, a set of single-donor blood materials, or pooled patient samples. In special cases, a manufacturer's calibrator can be designated as a reference material. Both standardization and harmonization activities aim to achieve comparable and reliable measurement results. Whereas standardization ensures traceability to the SI, harmonization ensures traceability to a reference system agreed upon by convention.

The principal steps for achieving comparable results through standardization and harmonization can be summarized as follows:

1) Establishing a reference system consisting of reference methods and materials

2) Calibrating measurement procedures using the reference system established in the first step

3) Verifying comparability of measurements performed for patient care and research, usually by measuring a set of authentic patient samples to assess the uniformity of results across different methods

Establishing metrological traceability, which is achieved by performing step 2, does not always lead to comparable measurement

[7] Abbreviations used: AACC, American Association for Clinical Chemistry; CLSI, Clinical and Laboratory Standards Institute; FT4, free thyroxine; IFCC, International Federation of Clinical Chemistry and Laboratory Medicine; SI, Systèm International (International System of Units); TSH, thyroid-stimulating hormone; T4, thyroxine.

results. Such lack of comparability can be explained by changes in instrument performance and reagent quality over time, inconsistent calibration of measurement procedures, the use of reference materials that are unsuitable for the intended use, and other factors. Therefore, programs to standardize or harmonize clinical laboratory measurements should appropriately address all 3 steps of the process to ensure comparable measurement results for patient care and public health research. With respect to the assessment of thyroid function in relation to iodine status, step 3 could be achieved, for example, through participation in standardization programs or accuracy-based proficiency-testing programs in which single-donor patient serum samples covering a range of concentrations of total T4 and/or FT4 are analyzed.

The European Union 1998 directive on in vitro diagnostic medical devices (8) requires that traceability of the calibration of measurement procedures be ensured. Current regulatory requirements, in addition to addressing growing clinical and public health needs for accurate and reliable measurement results, have produced a profound increase in standardization and harmonization activities.

PROGRESS IN STANDARDIZATION AND HARMONIZATION

A wide range of activities directed toward the standardization and harmonization of clinical laboratory testing are under way around the world; these are being conducted by international entities such as the International Federation of Clinical Chemistry and Laboratory Medicine (IFCC), the WHO, and the European Union Joint Research Centre Institute for Reference Materials and Methods. National entities such as the CDC, the National Institute for Standards and Technology, and offices and institutes of the NIH are also furthering standardization and harmonization efforts, as are various academic institutions.

The 2 longest-standing and most comprehensive standardization programs that address all 3 steps of the standardization process are the CDC's Lipid Standardization Program for cholesterol and blood lipids (9) and the National Glycohemoglobin Standardization Program for hemoglobin A1c (10).

National and international organizations and research groups are developing and establishing the components needed for standardization or harmonization of TSH, total T4, FT4, and other clinical analytes relevant to the assessment of iodine status and thyroid function. For example, the University of Ghent in collaboration with the IFCC has established reference methods and materials for thyroid hormone testing. As discussed by Faix and Miller (11) in this supplement issue, the IFCC Committee for Standardization of Thyroid Function Tests has established a conventional reference measurement procedure for FT4 and an approach to harmonization (rather than standardization) for TSH. Reference methods and materials that meet criteria outlined in standards from the International Organization for Standardization are compiled and maintained in a database by the Joint Committee for Traceability in Laboratory Medicine (12).

The CDC Hormone Standardization Program is one example of a standardization effort developed in response to an acknowledged need. The program, which started in 2006, was initiated because inaccurate and unreliable measurements for testosterone and estradiol prevented the implementation of research findings in patient care and hindered correct treatment. These problems were discussed at a CDC workshop in 2006 and summarized in a special issue of the journal *Steroids* (13). Further details about problems and challenges in testosterone and estradiol testing are described in position statements by the Endocrine Society and the American Urology Association (14–16). The CDC developed reference measurement procedures (17) and panels of single-donor sera to assist laboratories in the operation of laboratory-developed tests and to assist assay manufacturers with both calibration and calibration verification. To ensure that the measurement accuracy of the laboratory-developed tests and assays operated by manufacturers is maintained over time, the CDC is assessing the participants quarterly with 10 single-donor sera; measurement accuracy is evaluated by combining the data obtained from 4 consecutive quarters using established protocols such as the Clinical and Laboratory Standards Institute (CLSI) protocol EP 9, "Method Comparison and Bias Estimation Using Patient Samples" (18). Participants meeting predefined analytical performance criteria are listed on the CDC website (9).

In addition, the CDC is collaborating with proficiency-testing providers such as the College of American Pathologists to assess the performance of measurements conducted in medical laboratories through accuracy-based surveys. These activities in the context of metrological traceability are shown diagrammatically in **Figure 2** and further described in a recent review (19). Recently, the CDC successfully applied the same approach and procedures in its Standardization Certification Program for serum 25-hydroxyvitamin D (20).

Procedures to achieve traceability through harmonization, which are more complex than those required for standardization, are being developed by several groups around the world. Both standardization and harmonization efforts require long-term commitments with respect to technical and financial resources. Until recently, no central organizing body existed for coordinating the various standardization and harmonization activities conducted worldwide to minimize duplication of effort or to optimize the use of resources. In 2010, an international leadership forum of the American Association for Clinical Chemistry (AACC) recommended that an infrastructure be created to coordinate worldwide efforts to harmonize measurands (21). In 2013, the AACC, working with its domestic and international partners, formed the International Consortium for Harmonization of Clinical Laboratory Results to implement these recommendations, including prioritizing analytes that require harmonization and coordinating harmonization work among

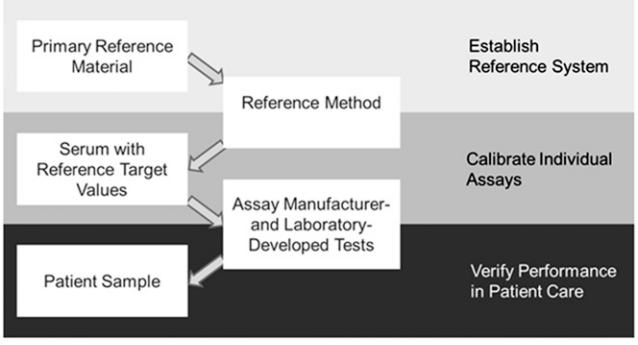

FIGURE 2 Abbreviated diagram of the 3 basic steps performed in standardization programs, showing the traceability chain.

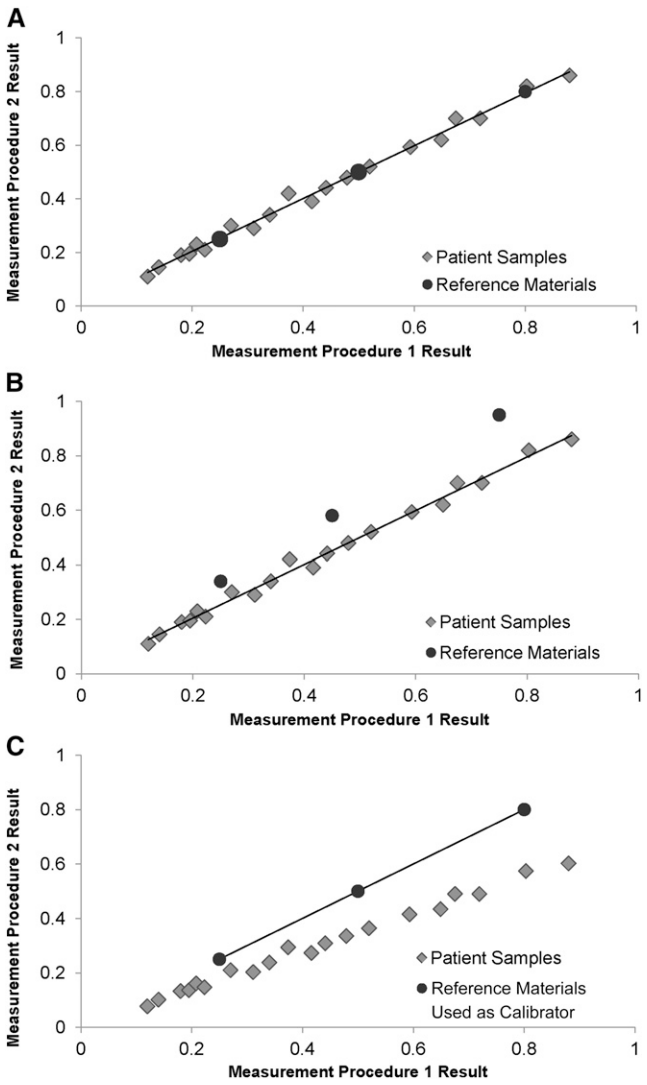

FIGURE 3 Scatter plots comparing the hypothetical results of 2 measurement procedures. The plots illustrate how test outcomes for 20 patient samples are affected by the commutability of RMs. In panel A, the 3 RMs are commutable and thus, when analyzed by measurement procedures 1 and 2, yield results consistent with those of patient samples. In panel B, the 3 RMs are noncommutable and thus, when analyzed by measurement procedures 1 and 2, yield results inconsistent with those of patient samples. In panel C, the 2 measurement procedures are calibrated by using 3 noncommutable RMs; because the RMs are noncommutable, using them to calibrate the 2 measurement procedures leads to discrepant, inaccurate results for patient samples. In panels A and B, the line shown is the regression line for the comparison of patient sample results across the 2 measurement procedures. In panel C, the line shown is the regression line for the comparison of 3 noncommutable RMs results across the 2 measurement procedures. RM, reference material.

different organizations. The governing body is a council made up of organizations from around the world that contribute financially to support the administration of the program. A list of the analytes currently being addressed, which includes TSH, total T4, and FT4, is available from the AACC website (22).

COMMUTABILITY

Reference materials used as calibrators or to verify the accuracy of a measurement procedure are key components for establishing metrological traceability. Like reference measurement procedures, reference materials must meet certain criteria for key characteristics, including homogeneity and stability. In recent years, commutability has come to be recognized as another critically important characteristic of reference materials (23). The importance of commutability became apparent in situations in which measurement procedures were calibrated to be traceable to the same reference material but patient samples measured using these methods showed substantial differences in results (3, 24–32). In these situations, the reference material is considered "noncommutable" and thus unsuitable for establishing metrological traceability.

Commutability is an essential property of reference materials used in a traceability chain (33). Commutability is defined by the International Vocabulary of Metrology 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 2 given measurement procedures, and the relation obtained among the measurement results for other specified materials" (34). This definition can be restated in the context of clinical laboratory measurement procedures as the closeness of agreement between 2 relations: the procedure 1 vs. procedure 2 results relation observed for a reference material and the procedure 1 vs. procedure 2 results relation observed for patient samples. **Figure 3**A shows the concept of commutability, in which the relation of the results between 2 measurement procedures for patient samples and the analogous relation for reference materials are quite similar. Figure 3B shows the concept of noncommutability, in which the relation of the results between 2 measurement procedures for patient samples is quite different from that for reference materials. Figure 3C shows that the use of the noncommutable reference materials shown in Figure 3B as calibrators will cause inaccurate measurement of patient samples. As per the above definition of commutability, a reference material is described as commutable (or noncommutable) with patient samples among a group of measurement procedures. It is important to point out that commutability is a property of a reference material; it is incorrect to use the term "commutability" to

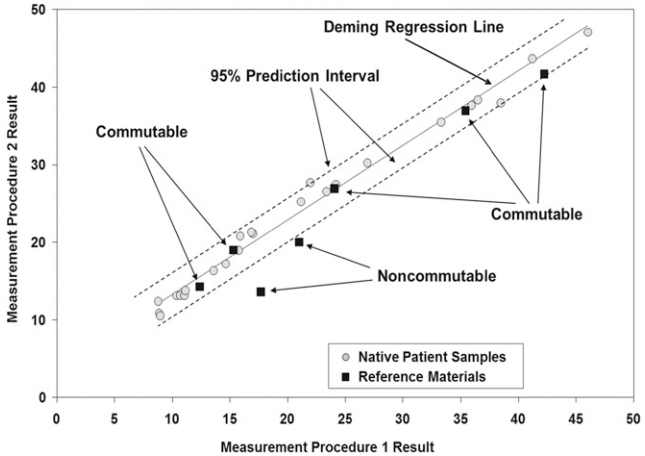

FIGURE 4 Example of the use of regression analysis with a 95% prediction interval to evaluate the commutability of reference materials, as described in the Clinical and Laboratory Standards Institute's document EP-30A (34).

refer to the agreement among results from a group of measurement procedures.

Higher-order reference measurement procedures are designed to have a high degree of analytical specificity for the substance being measured and a level of accuracy that is independent of the specimen matrix. Consequently, commutability is rarely an issue for reference materials used to calibrate higher-order reference measurement procedures. Nor is commutability an issue for higher-order reference measurement procedures when assigning target values to a secondary reference material, which usually has a matrix similar to that of pertinent clinical samples. However, commutability of secondary reference materials becomes a critical consideration when those secondary reference materials are intended for use with less-specific measurement procedures, such as routine clinical laboratory procedures and the steps through which manufacturers assign values to the calibrators used in routine clinical laboratory procedures. In these situations, commutability of secondary reference materials is an essential requirement for ensuring accurate and reliable measurements in patient samples. A study investigating the accuracy of 16 TSH measurement procedures, all claiming to be traceable to the same WHO reference material, showed that 3 of these tests were inaccurate because the reference material used for calibration traceability was not commutable for these measurement procedures (35). Unfortunately, there are a substantial number of secondary reference materials in use today for which commutability has not been examined; findings similar to those described for TSH have been reported for other analytes (24–32).

The CLSI has published a consensus guideline to facilitate validation of the commutability of reference materials (36). One of the commonly used approaches described in the CLSI guideline is shown in **Figure 4**. In this example, a linear relation between 2 measurement procedures is obtained for the patient samples. The relation can be described with a regression equation and the 95% prediction interval around the regression line. Reference materials with results that fall within the 95% prediction interval are considered to be in close agreement with the relation for patient samples and thus are commutable, whereas reference materials with results falling outside this interval are considered noncommutable.

Limitations of current approaches for commutability evaluation include the following: *1*) the dependence of the evaluation criteria on patient samples being representative of the disease condition or conditions for which a laboratory test is typically used and *2*) the analytical precision and measurand specificity of the measurement procedures. Suitable analytical specificity implies that all measurement procedures measure the same biomarker without influence from interfering substances that may be present in the patient samples. For example, the use of inappropriate patient samples and/or excessively imprecise or nonspecific methods can lead to very wide prediction intervals around the regression line. In this situation, values for reference materials may fall within the prediction interval, and although they fulfill the criteria for commutability, the apparent closeness of agreement with widely scattered patient samples may be inadequate to allow the use of such reference materials as calibrators in a traceability scheme. At the other extreme, appropriate patient samples measured by highly precise and specific methods could lead to very narrow prediction intervals. In this situation, reference materials could be slightly outside the prediction interval, and thus considered noncommutable, when they are actually suitable for use as calibrators in a traceability scheme, because the bias introduced would not be clinically relevant. These examples show that conclusions from an assessment of commutability need to be drawn in the context of the overall method performance and the clinical requirements for measurement accuracy. The IFCC's Working Group on Commutability is advancing the approaches used for commutability assessment by addressing the limitations in current approaches.

SUMMARY

Standardization and harmonization are collaborative efforts conducted by national and international organizations and institutions. These efforts require long-term commitments of technical and financial resources.

A clearly defined measurand and appropriate implementation of metrological traceability with suitable reference methods and commutable reference materials are fundamental to achieving comparability of measurement results independent of time, place, and measurement procedure. Procedures and protocols that exist or are in development in the area of thyroid function testing include several for defining measurands clearly and for implementing appropriate reference methods and commutable reference materials in standardization and harmonization programs.

We acknowledge contributions to the article made by Gay Goodman, Iodine Initiative Consultant to the NIH Office of Dietary Supplements, in the course of providing expert technical review.

The authors' responsibilities were as follows—HWV: had primary responsibility for the final content; and all authors: participated in researching and writing the manuscript and read and approved the final manuscript. WGM received a research grant from Abbott Diagnostics. The authors reported no conflicts of interest related to the study.

REFERENCES

1. Steele BW, Wang E, Klee GG, Thienpont LM, Soldin SJ, Sokoll LJ, Winter WE, Fuhrman SA, Elin RJ. Analytic bias of thyroid function tests: analysis of a College of American Pathologists fresh frozen serum pool by 3900 clinical laboratories. Arch Pathol Lab Med 2005;129:310–7.
2. International Organization for Standardization. In vitro diagnostic medical devices: measurement of quantities in biological samples—metrological traceability of values assigned to calibrators and control materials. ISO 17511. Geneva (Switzerland): International Organization for Standardization; 2003.
3. Miller WG, Thienpont LM, Van Uytfanghe K, Clark PM, Lindstedt P, Nilsson G, Steffes MW; Insulin Standardization Work Group. Toward standardization of insulin immunoassays. Clin Chem 2009;55:1011–8.
4. White GH. Metrological traceability in clinical biochemistry. Ann Clin Biochem 2011;48:393–409.
5. Vesper HW, Thienpont LM. Traceability in laboratory medicine. Clin Chem 2009;55:1067–75.
6. Miller WG, Tate JR, Barth JH, Jones GR. Harmonization: the sample, the measurement and the report. Ann Lab Med 2014;34:187–97.
7. Long SE, Catron BL, Boggs ASP, Tai S, Wise SA. Development of Standard Reference Materials to support assessment of iodine status for nutritional and public health purposes. Am J Clin Nutr 2016;104(Suppl): 902S–6S.
8. European Union. Directive 98/79/EC of the European Parliament and of the Council of 27 October 1998 on in vitro diagnostic medical devices. Official Journal of the European Communities 1998;L 331:1–37. [cited 2015 Apr 1]. Available from: http://eur-lex.europa.eu/legal-content/EN/ALL/?uri=CELEX%3A31998L0079.
9. CDC. Laboratory quality assurance and standardization programs. 29 July 2014 [cited 2014 Dec 1]. Available from: http://www.cdc.gov/labstandards/.
10. National Glycohemoglobin Standardization Program. c2010 [cited 2014 Dec 1]. Available from: http://www.ngsp.org/.

11. Faix JD, Miller WG. Progress in standardizing and harmonizing thyroid function tests. Am J Clin Nutr 2016;104(Suppl):913S–7S.
12. Joint Committee for Traceability in Laboratory Medicine. JCTLM database: laboratory medicine and in vitro diagnostics [cited 2014 Dec 1]. Available from: http://www.bipm.org/jctlm/.
13. Rosner W, Vesper H. Preface. CDC workshop report on improving steroid hormone measurements in patient care and research translation. Steroids 2008;73:1285.
14. Rosner W, Hankinson SE, Sluss PM, Vesper HW, Wierman ME. Challenges to the measurement of estradiol: an endocrine society position statement. J Clin Endocrinol Metab 2013;98:1376–87.
15. Rosner W, Auchus RJ, Azziz R, Sluss PM, Raff H. Position statement: utility, limitations, and pitfalls in measuring testosterone: an Endocrine Society position statement. J Clin Endocrinol Metab 2007;92:405–13.
16. Paduch DA, Brannigan RE, Fuchs EF, Kim ED, Marmar JL, Sandlow JI. The laboratory diagnosis of testosterone deficiency. Urology 2014;83:980–8.
17. Wang Y, Gay GD, Botelho JC, Caudill SP, Vesper HW. Total testosterone quantitative measurement in serum by LC-MS/MS. Clin Chim Acta 2014;436:263–7.
18. Clinical and Laboratory Standards Institute. Method comparison and bias estimation using patient samples: approved guideline. 2nd ed. Wayne (PA): Clinical and Laboratory Standards Institute; 2002. CLSI Document No.: EP9-A2.
19. Vesper HW, Botelho JC, Wang Y. Challenges and improvements in testosterone and estradiol testing. Asian J Androl 2014;16:178–84.
20. CDC. Laboratory quality assurance and standardization programs: Hormone Standardization Program. 25 March 2015 [cited 2015 Apr 1]. Available from: http://www.cdc.gov/labstandards/hs.html/.
21. Miller WG, Myers GL, Gantzer ML, Kahn SE, Schönbrunner ER, Thienpont LM, Bunk DM, Christenson RH, Eckfeldt JH, Lo SF, et al. Roadmap for harmonization of clinical laboratory measurement procedures. Clin Chem 2011;57:1108–17.
22. American Association for Clinical Chemistry. Clinical laboratory test harmonization. 2014 [cited 2014 Dec 1]. Available from: http://www.harmonization.net/.
23. Miller WG, Myers GL. Commutability still matters. Clin Chem 2013;59:1291–3.
24. Rose MP. Follicle stimulating hormone international standards and reference preparations for the calibration of immunoassays and bioassays. Clin Chim Acta 1998;273:103–17.
25. Stephan C, Klaas M, Müller C, Schnorr D, Loening SA, Jung K. Interchangeability of measurements of total and free prostate-specific antigen in serum with 5 frequently used assay combinations: an update. Clin Chem 2006;52:59–64.
26. Sokoll LJ, Rosenwald S, Lyons J, Elliott DJ, Chan DW. Is the WHO 90:10 prostate-specific antigen (PSA) first international reference standard really 90% α1-antichymotrypsin-bound PSA and 10% free PSA? Clin Chem 2011;57:1776–7.
27. Little RR, Rohlfing CL, Tennill AL, Madsen RW, Polonsky KS, Myers GL, Greenbaum CJ, Palmer JP, Rogatsky E, Stein DT. Standardization of C-peptide measurements. Clin Chem 2008;54:1023–6.
28. Sturgeon CM, Berger P, Bidart J-M, Birken S, Burns C, Norman RJ, Stenman U-H. Differences in recognition of the 1st WHO international reference reagents for hCG-related isoforms by diagnostic immunoassays for human chorionic gonadotropin. Clin Chem 2009;55:1484–91.
29. Caliendo AM, Shahbazian MD, Schaper C, Ingersoll J, Abdul-Ali D, Boonyaratanakornkit J, Pang XL, Fox J, Preiksaitis J, Schönbrunner ER. A commutable cytomegalovirus calibrator is required to improve the agreement of viral load values between laboratories. Clin Chem 2009;55:1701–10.
30. Tate JR, Bunk DM, Christenson RH, Katrukha A, Noble JE, Porter RA, Schimmel H, Wang L, Panteghini M; IFCC Working Group on Standardization of Troponin I. Standardisation of cardiac troponin I measurement: past and present. Pathology 2010;42:402–8.
31. Bachmann LM, Nilsson G, Bruns DE, McQueen MJ, Lieske JC, Zakowski JJ, Miller WG. State of the art for measurement of urine albumin: comparison of routine measurement procedures to isotope dilution tandem mass spectrometry. Clin Chem 2014;60:471–80.
32. Wieringa GE, Sturgeon CM, Trainer PJ. The harmonisation of growth hormone measurements: taking the next steps. Clin Chim Acta 2014;432:68–71.
33. Vesper HW, Miller WG, Myers GL. Reference materials and commutability. Clin Biochem Rev 2007;28:139–47.
34. Joint Committee for Guides in Metrology. International vocabulary of metrology (VIM) - Basic and general concepts and associated terms. 3rd ed. Joint Committee for Guides in Metrology 2012;200 [cited 2015 Apr 1]. Available from: http://www.bipm.org/en/publications/guides/.
35. Thienpont LM, Van Uytfanghe K, Beastall G, Faix JD, Ieiri T, Miller WG, Nelson JC, Ronin C, Ross HA, Thijssen JH, et al.; IFCC Working Group on Standardization of Thyroid Function Tests. Report of the IFCC Working Group for Standardization of Thyroid Function Tests; part 1: thyroid-stimulating hormone. Clin Chem 2010;56:902–11.
36. Clinical and Laboratory Standards Institute. Characterization and qualification of commutable reference materials for laboratory medicine: approved guideline. Wayne (PA): Clinical and Laboratory Standards Institute; 2010. CLSI Document No.: EP30-A (formerly C53-A).