Clinical Chemistry 55:6
1067–1075 (2009)

Review

Traceability in Laboratory Medicine

Hubert W. Vesper[1*] and Linda M. Thienpont[2]

BACKGROUND: In patient and population samples, generation of analytical results that are comparable and independent of the measurement system, time, and location is essential for the utility of laboratory information supplied in healthcare. Obtaining analytical measurement results with such characteristics is the aim of traceability in laboratory medicine. As awareness of the benefits of having traceable measurement results has increased, associated efforts have been directed toward making traceability a regulatory requirement and developing approaches to enable and facilitate the implementation of traceability. Although traceability has been a main focus of many laboratory standardization activities in the past, discussions are still ongoing with regard to traceability and its implementation.

CONTENT: This review provides information about the traceability concept and what needs can be fulfilled and benefits achieved by the availability of traceable measurement results. Special emphasis is given to the new metrological terminology introduced with this concept. The review addresses and describes approaches for technical implementation of traceable methods as well as the associated challenges. Traceability is also discussed in the context of other activities to improve the overall measurement process.

SUMMARY: Establishing metrological traceability of measurement results satisfies basic clinical and public health needs, thus improving patient care and disease control and prevention. Large advances have been made to facilitate the implementation of traceability. However, details in the implementation process, such as lack of available commutable reference materials and insufficient resources to develop new reference measurement systems continue to challenge the laboratory medicine community.

[1] Division of Laboratory Sciences, National Center for Environmental Health, Centers for Disease Control and Prevention, Atlanta, GA; [2] Laboratory for Analytical Chemistry, Faculty of Pharmaceutical Sciences, Gent University, Gent, Belgium and JCTLM-WG2 Reference Measurement Laboratories, Bureau International des Poids et Mesures, Sèvres Cedex, France.
* Address correspondence to this author at: 4770 Buford Hwy NE, MS-F25, Atlanta, GA, 30345. Fax 770-488-4192; e-mail HVesper@cdc.gov.
Received October 23, 2008; accepted February 13, 2009.
Previously published online at DOI: 10.1373/clinchem.2008.107052

Discussions on the subject of traceability over the past decade in the analytical laboratory community have had a profound effect in laboratory medicine, requiring manufacturers to reassess assay calibration procedures, challenging laboratory clinicians to develop tools and strategies to implement traceability, and calling for the general clinical laboratory community to adopt new terminology and to revisit the analytical measurement process for issues such as analyte definition and method calibration.

The term "traceability" originated in the metrological community, where it was first defined in 1993 in the International Vocabulary of General and Basic Terms in Metrology *(1)*. The same year, the Cooperation on International Traceability in Analytical Chemistry was formed to encourage the broad realization of traceability in analytical chemistry *(2)*. With the implementation of the European Union Directive on in vitro diagnostic devices *(3)*, establishing traceability of measurements performed with in vitro diagnostic devices became mandatory and had a worldwide effect on clinical laboratory measurements.

The aim of traceability in laboratory medicine is to link measurement results from a patient sample to a commonly accepted reference, making them comparable across measurement systems, location, and time *(4)*. The goals and principal approaches for establishing traceability are the same as those described over the last 50 years for standardizing clinical laboratory measurements *(5–9)*. Such activities include the CDC Lipid Standardization Program *(10)* and the National Glycohemoglobin Standardization Program *(11)*.

This review provides information on the traceability process, including previous and ongoing efforts to help implement traceability in laboratory medicine, establish the place of traceability in the overall measurement process, and meet the challenges and opportunities associated with its implementation.

Traceability Addresses Clinical and Public Health Needs

Measurement results that are linked to a common reference and are comparable across measurement systems, location, and time are essential for research translation, patient care, and disease prevention and control. Laboratory-based clinical and public health decisions are made by comparing measurement results

from individual patients or populations against a reference. This reference can be a clinical decision point, a reference population, or a set of values obtained from the same patient or population at an earlier time. Similarly, clinicians can translate research findings into patient care and disease prevention activities by comparing laboratory data across studies and formulating common clinical decisions based on these study data. Traceability facilitates these activities and helps to satisfy the basic needs and requirements of evidence-based laboratory medicine. Traceable measurement results allow for:

- Definition of generally accepted and usable reference intervals, rather than method-specific reference ranges;
- Application of consistent standards of medical care, best practice guidelines, evidence-based medicine, and laboratory medicine practice guidelines, and
- Pooling and comparison of data from various studies to facilitate medical research and research translation.

A recent US national report on laboratory medicine suggested that the lack of traceability limited these activities and benefits, a shortcoming that poses a current problem in the field *(12)*. Traceability of measurement results must be established for all clinical analytes. Clinical needs for comparable results among methods for certain analytes such as troponin I, B-type natriuretic peptide, and prostate-specific antigen have been described in a recent review *(13)*. The authors of this review stated that the lack of traceable results can lead to medical misinterpretation with the use of common decision-making criteria and may jeopardize patient safety. Using prostate-specific antigen as an example, they explained that nontraceable values may lead clinicians to perform unnecessary biopsies in some cases and to miss necessary biopsies in other cases. Other studies demonstrate the negative impact of measurement bias on early detection of chronic kidney disease, especially in pediatric patients *(14, 15)*. The authors of the latter studies stated the need for traceable measurements to reduce bias between methods and improve early detection of chronic kidney disease.

Clinical limitations of current testosterone measurements that prevent the assessment of androgen disorders in men and women have been pointed out for many years by researchers, physicians, and professional societies *(16–21)*, who have recommended standardizing testosterone measurements to overcome these problems.

Lack of traceability not only has direct consequences on patient care, disease prevention and control, and research translation, but it also affects healthcare and research costs. NIST estimated that measurement bias for calcium assays caused by nontraceable results leads to costs ranging from $60 million to $199 million per year *(22)*. According to the registry of federally and privately supported clinical trials conducted in the US and around the world, more than 600 open clinical trials deal with testosterone, vitamin D, serum creatinine, B-type natriuretic peptide, and troponin *(23)*. The data obtained in these studies will be of limited use for public health and patient care, because of lack of metrological traceability and standardization of these analytes. Considering today's cost of testing and of performing clinical trials, the clinical and public health benefits of having traceable measurement results are substantial. Unfortunately, researchers, physicians, and the general public frequently presume that we have these benefits. They often recognize the need for traceable measurements too late, after the costs or potential safety concerns of nontraceable results have become apparent. Both healthcare and research costs can be reduced by establishing traceability.

The Traceability Process

The concept of traceability of any analytical measurement results originated from the metrological community. Thus, metrological traceability requires understanding the basic metrological concepts and terminology as provided in the *International Vocabulary of Basic and General Terms in Metrology (VIM) (1)*: Metrological traceability is defined as the "property of a measurement result whereby the result can be related to a stated reference through a documented unbroken chain of calibrations, each contributing to the measurement uncertainty" *(1)*. Traceability refers to a measurement result, thus it applies to the calibration of a procedure and to its specificity. Numerically, traceability of a result can be assessed by its accuracy. The measurement result itself is defined as a "set of quantity values being attributed to a measurand together with any other available relevant information" *(1)*. This definition leads to the definition of a quantity, which is the "property of a phenomenon, body, or substance, to which a number can be assigned with respect to a reference" *(1)*. Quantities are designated in laboratory medicine by the format, "system-component (analyte); kind of quantity" (e.g., serum cortisol; amount-of-substance concentration equal to *x* nmol/L). Standardized templates for such designations of measurands can be found in an IFCC/International Union of Pure and Applied Chemistry database *(24)*. Thus, establishing traceability requires a reference measurement system consisting of:

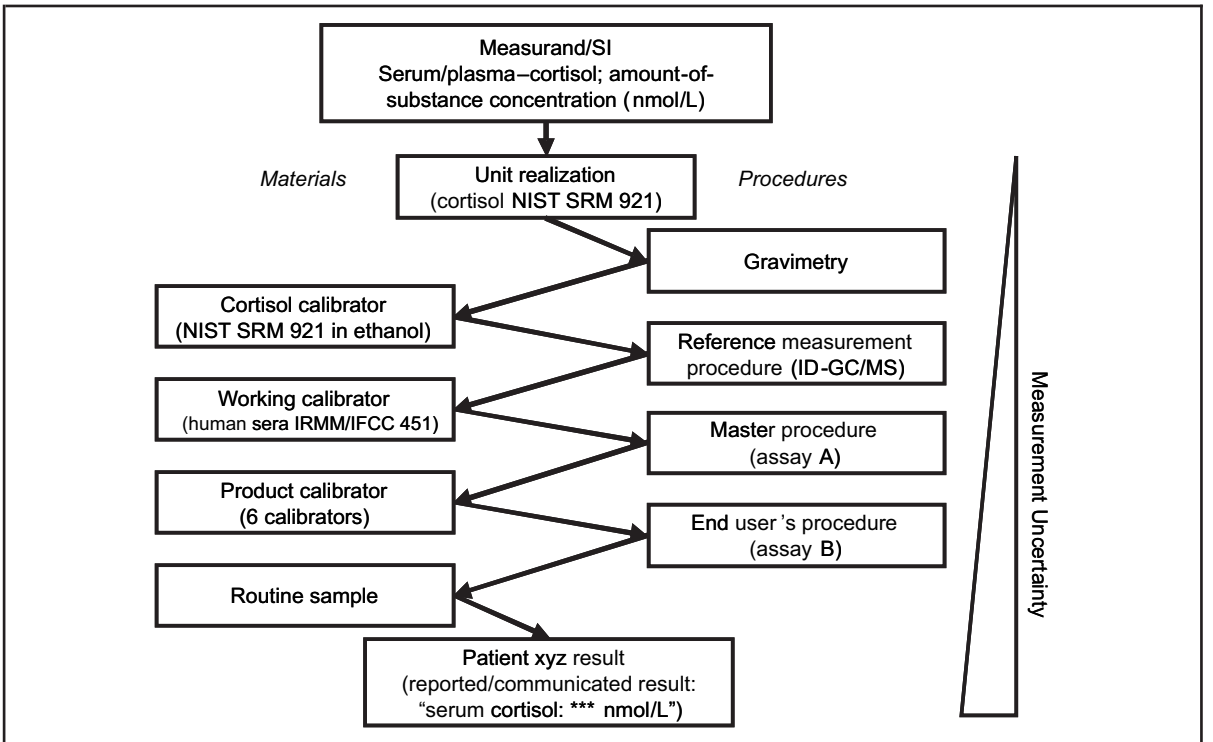

Fig. 1. Example of traceability chain developed for serum cortisol measurements.
Arrows pointing to the left indicate value assignment activity using the procedure; arrows pointing to the right indicate calibration activity using the material.

- Measurand definition,
- Stated reference procedure and/or reference material,
- Knowledge of the measurement uncertainties, and
- Unbroken chain of calibrations and value assignments.

Of note, the measurand is defined as the "quantity intended to be measured" *(1)*. Therefore, the measurand is not necessarily what is being measured. For example, when a Jaffe method is used the measurand is serum creatinine although alkaline-picrate reactive substances are actually measured. The full designation of the quantity actually being measured is intended for use by the laboratory and scientific communities and not for communicating routine clinical measurement results in patients. Typically, the system and component intended to be measured are communicated (e.g., serum creatinine).

A "stated reference" to which the measurement is made traceable must be identified or defined. This reference can be a pure compound, a measurement procedure, or matrix-based materials with defined characteristics, which will be described later.

The measurement result is accompanied by the measurement uncertainty, which is defined as a "nonnegative parameter characterizing the dispersion of the quantity values being attributed to a measurand" *(1)*. Thus the measurement result must be regarded as a best estimate of the value of the measurand, and all components of the measurement that contribute to the dispersion of results around the reported value should be considered when a value is assigned to a material. This process relies on appropriate characterization and validation to properly define the measurement uncertainty of the measurement procedure used for value assignment.

Finally, a calibration chain must be defined (see Fig. 1, discussed later), and adherence to that calibration chain allows for consistent traceability of results to the original reference and assures that the calibration is performed within the relevant concentration range.

Technical Realization

Traceability is realized by relating a measurement result to a stated reference through an unbroken chain of calibrations. Stated references may range from a cor-

Table 1. Main activities of organizations and groups in establishing metrological traceability.

Organization/group involved in establishing traceability	Activities/responsibilities of organization/group	Coordinating organizations
National metrology institutes	• Realization of SI units	Conférence Générale des Poids et Mesures
	• Calibration and measurement capabilities on the highest metrological level	
	• Reference materials with formal certificate	
Reference (calibration) laboratories	• Reference measurement services	Joint Committee for Traceability in Laboratory Medicine (Bureau International des Poids et Mesures, International Federation of Clinical Chemistry and Laboratory Medicine, International Laboratory Accreditation Cooperation)
	• Realization of non-SI units (i.e., IU as defined by WHO)	Reference laboratory networks
	• Value assignment to reference materials, calibration materials, and patient samples used for trueness control and calibration	
	• Accuracy assessments of measurement procedures and laboratories	
Routine testing laboratories	• Value assignment to patient samples	National entities
	• Calibration of routine measurement procedures	

porate standard to a certified primary reference material that embodies a unit of the Systéme International (SI).[3] These reference materials must have certain well-defined characteristics, such as homogeneity and stability, as described by the International Organization for Standardization (ISO) *(25)* and CLSI *(26)*. One important reference material characteristic that is frequently overlooked is its commutability, which is addressed later in this report.

Reference measurement procedures must be characterized, validated, and documented according to specifications outlined in documents such as ISO 15193 *(27)* or those created by the Joint Committee for Traceability in Laboratory Medicine *(28)*. The Joint Committee for Traceability in Laboratory Medicine acknowledges reference materials and measurement procedures that comply with the aforementioned standards and maintains a database of all such materials and methods that are used to establish traceability chains. Similarly, the Joint Committee for Traceability in Laboratory Medicine describes the criteria for laboratories performing reference measurements. In line with these activities, internationally recognized reference laboratory networks exist for certain analytes such as cholesterol and hemoglobin A_{1c}. These networks have established specific performance criteria and assure compliance through regular assessments.

The unbroken chain of calibrations applies a hierarchical order to materials and measurement procedures, as diagrammed in Fig. 1. It is an alternating process of assigning values to a material used to calibrate the next lower-order measurement procedure, which is then used to assign value to the next lower-order material. The hierarchy is defined by the degree of uncertainty associated with the measurement result. The measurement procedure with the lowest uncertainty, typically a reference procedure, is at the top, and that with the highest uncertainty, typically a routine clinical laboratory procedure, is at the bottom. The wedge on the right indicates that the uncertainty increases cumulatively at each step in the traceability chain. Certain portions of the chain can be established by various organizations or groups, as outlined in ISO 17511 *(4)* and further described in Table 1. For example, metrological institutions and reference laboratories can establish traceability of values assigned to a matrix-based reference material, whereas the assay manufacturer establishes traceability of the patient result to this matrix-based reference material through the product calibrator(s) for the routine measurement procedure.

The traceability chain for serum-cortisol to the SI concentration units of nanomoles per liter (Fig. 1)

[3] Nonstandard abbreviations: SI, Systéme International; ISO, International Organization for Standardization.

starts with definition of the measurand as "serum-cortisol; amount-of-substance concentration equal to *x* nmol/L." The unit is embodied in the Standard Reference Material 921 from NIST. This material consists of chemically purified cortisol and is used to prepare calibrators by dissolving defined amounts of this material in a defined volume of a solvent. These calibrators are suitable only for calibrating the isotope dilution–GC-MS reference measurement procedure. This procedure was tested and found suitable for measurement of cortisol accurately in calibrators as well as in sera. It is used to measure the cortisol concentration of a panel of native sera Institute for Reference Materials and Measurements /IFCC 451 (available as ERM-DA451® from the Institute for Reference Materials and Measurements). Those sera are used for calibrating the manufacturer's master procedure by a method comparison study. The manufacturer's master procedure is optimized for use of serum as sample matrix and thus would produce inaccurate results when calibrated with the calibrators used for the reference measurement procedure. The manufacturer's master procedure is used to assign calibrator values to the product calibrators used with the end user's routine measurement procedure. Results of measurement of routine clinical samples are thus traceable to the SI unit based on the pure cortisol standard. Importantly, the calibrators and measurement procedures used in the traceability chain must be suitable for the intended use.

The traceability chain shown in Fig. 1 illustrates a scenario in which all components are available to establish traceability to SI. Although establishing traceability to SI is the goal for all measurement results, for many measurements in the clinical laboratory the components to establish traceability to SI are not available. To account for such situations, alternate traceability chains were defined by ISO 17511 *(4)*. The 5 traceability chains defined in this standard are:

1. Measurements are traceable to an SI unit (i.e., mol). The calibration process follows, in principle, the procedure diagrammed in Fig. 1 and is further explained in detail below. In this scenario, the chemical and physical properties of the analyte are known, and primary reference measurement procedures and calibrators are available. Examples are analytes such as electrolytes, metabolites, glucose, cholesterol, steroid hormones and some thyroid hormones, and drugs.
2. Measurements are traceable to an international conventional reference measurement procedure (which is not primary) and to international conventional calibrator(s) without metrological traceability to SI. In this scenario, reference measurement procedures and calibrators are defined by convention or consensus. An example is total hemoglobin, which is measured by absorption spectrometry of its cyanide derivative and is calibrated by CRM 522 hemoglobin cyanide in bovine blood lysate from the Community Bureau of Reference.
3. Measurements are traceable to an international conventional reference measurement procedure (which is not primary), but there is no international conventional calibrator and no metrological traceability to SI. In this scenario, measurement procedures are defined by convention or consensus and no calibrators exist. Examples would be measurement procedures specified by the IFCC for the measurement of enzymes such as creatine kinase, lactate dehydrogenase, and aspartate aminotransferase [as described in a recent review *(29)*] or by the International Council for Standardization in Hematology for the measurement of the number concentration of erythrocytes and leukocytes in human blood.
4. Measurements have traceability to an international conventional calibrator (which is not primary), but there is no international conventional reference measurement procedure and no metrological traceability to SI. In this scenario, a reference material is defined by convention or consensus, and values are assigned to this material in arbitrary units such as International Units defined by WHO standards (e.g., human chorionic gonadotrophin).
5. Measurements have traceability to a manufacturer's selected measurement procedure but there is neither an international conventional reference measurement procedure nor an international conventional calibrator, and there is no metrological traceability to SI. In this scenario, neither reference materials nor calibrators are available. Examples would be analytes such as fibrin degradation products (D-dimer) and tumor markers such as cancer antigen 125 as well as antibodies to antigens such as *Chlamydia*. This scenario occurs when new biomarkers are identified and developed by individual research laboratories.

The assay manufacturer is responsible for establishing and documenting metrological traceability for commercially available methods. Individual clinical laboratories that use these commercial methods do not need to validate traceability as long as the manufacturer's instructions for use are followed. If a laboratory modifies a manufacturer's method or develops its own method, the laboratory is then responsible for validating traceability to the highest available reference *(4)*. This stipulation, however, does not imply that the manufacturer or laboratory must develop all components needed in the traceability chain. Individual reference laboratories and reference laboratory networks help link measurements to higher metrological standards. Several organizations have created documents

to help implement and document traceability, such as the general guide on traceability from Eurachem/Cooperation on International Traceability in Analytical Chemistry *(30)* and the specific guide on establishing traceability in laboratory medicine from the ISO [ISO 17511 *(4)* and 18153 *(31)*]. Other documents include supplemental information on implementing these ISO standards from the CLSI [CLSI report XR5 *(32)*] and guidance on calculating measurement uncertainty from the NIST *(33)*. Scientific publications that provide in-depth information about topics related to traceability are available in book format *(34)* and in volume 28 of the *Clinical Biochemist Reviews (35)*.

Traceability in the Context of the Overall Measurement Process and Measurement Performance

Traceability addresses assay calibration and associated assay specificity, which is only one aspect in the analytical phase of the overall measurement process. Other factors, such as assay selection, patient status, specimen collection and handling, and result reporting, are part of the preanalytical and postanalytical phases of the measurement process. The latter factors have profound influence on measurement results and data interpretation. A recent national status report found that the distribution of errors in laboratory medicine was 32%–75% in the preanalytical phase, 13%–23% in the analytical phase, and 9%–31% in the postanalytical phase *(12)*. These data suggest that one cannot neglect factors related to the pre- and postanalytical phases.

Major standardization projects such as the National Glycohemoglobin Standardization Program, the CDC Cholesterol Standardization Program, and the National Kidney Disease Education Program not only address metrological traceability but also assess the efficacy of its implementation and the standardization of pre- and postanalytical factors. Another essential aspect of these projects is the assurance that the results of these efforts are communicated and coordinated with the clinical, public health, research, and laboratory communities. Therefore, establishing metrological traceability is a key component in laboratory standardization efforts, which are comprehensive activities directed toward the overall improvement of the measurement process and of the use of measurement results.

Although the aforementioned procedures and activities lead to traceable results from individual measurement systems, clinically relevant differences in patient results can still occur owing to differences in assay specificity and operation or in value-transfer protocols. Thus, the efficacy of traceability implementation across measurement systems should be assessed on a continuous basis. Interlaboratory comparison studies are useful for such assessments, because they provide information about individual laboratories and their overall performance among participating laboratories. The latter is especially helpful when analytical performance goals have not been defined. The performance of such assessments at both the manufacturer and end-user levels helps assure traceability in reported patient results. Performance at the manufacturer level is commonly assessed by reference laboratories or laboratory networks through specific comparison studies *(36–38)*, and end-user performance can be assessed with accuracy-based external quality assurance programs *(39–41)*. Although these comparison studies are not required for establishing traceability, they are important for achieving and maintaining a desired overall measurement outcome in clinical settings.

Challenges and Possible Solutions

Establishing metrological traceability requires that each component used and step performed in this process be unequivocally defined and identified according to standards and guidance documents. This review cannot address all of the possible challenges encountered in this process. However, in the following we address issues that should be discussed and considered, technically and conceptually, in this undertaking.

The definition of the measurement system (e.g., blood, plasma, serum, urine) may seem obvious to many. In some cases, however, transferring the theoretical concept into the practice of measurement may be problematic *(42)*. Consider, for example, serum water (relevant for ionized electrolytes or free hormones). Many components in serum are bound to proteins or are complexed with other components. Therefore, their concentrations in serum water differ from their concentrations in serum. Currently, no measurement procedure can provide measurements of the "true" concentrations of components in serum water by directly measuring serum samples. Serum water must be separated from serum by ultrafiltration or dialysis, processes that may break the calibration chain required for metrological traceability. An example of the effects of such constraints is the required defining of the measurand for free thyroxine operationally as "equilibrium-dialysate from serum prepared under defined conditions-thyroxine (free); *x* pmol/L" *(43)*.

Defining the component to be measured may pose challenges *(42, 44)*. Consider follicle-stimulating hormone (FSH), for example *(45)*. Follicle-stimulating hormone is a glycoprotein hormone this is present in serum as a complex mixture. The current immunoassays used for follicle-stimulating hormone measurement are considered to be "blind" against the glycosylation forms and are presumed to measure these forms

in a more or less equimolar fashion (note, for equimolar assays, the measurement unit should be mol and not g). This issue is not yet resolved, however, and the profession still has not unequivocally defined the follicle-stimulating hormone component(s) *(45)*. Although different immunoassays use various epitopes for measurement, they all intend to quantify the same component (or component mixture). The measurand is intended to be the molecule of clinical and physiologic interest, and thus defining the measurand by the epitope used to measure it would be misleading.

Standards of certified purity can be prepared to embody an SI unit for measurands for which purified components are available (such as cortisol, glucose, and sodium). Such standards, however, are typically prepared in an aqueous matrix and are usually not commutable with native samples measured using routine measurement procedures; thus they cannot be used to directly calibrate routine procedures. Such standards can be used for reference measurement procedures because the specificity of the reference procedure makes it insensitive to the matrix of the measurand. A commutable standard material has the same performance with a measurement procedure in the traceability chain as is observed for native samples. Differences in analytical response between the standard material and the clinical samples will introduce a bias in the measurement. This bias is especially important for assays that measure the analyte of interest directly in the clinical sample matrix (e.g., serum and urine) without any prior isolation or purification steps (e.g., most routine measurement procedures). Consensus procedures to determine reference material commutability are currently being developed *(26)*. The relevance and assessment of commutability has been described in detail in the literature *(46–48)*. To overcome the lack of commutable reference materials, calibration can be done with panels of native samples measured with matrix-insensitive reference measurement procedures (see Fig. 1, cortisol example) *(42, 49)*.

Standards with certified purity cannot be prepared and SI units are not applicable for components for which the exact identity cannot be defined and/or the purity of standards is unknown. Therefore, the WHO developed the International Unit (IU) concept according to the following principle. An International Standard material is prepared with state-of-the-art purification and identification techniques, and its function is tested by the response of a biological system (a bioassay). The IU is then assigned by convention (e.g., 1 IU is assigned to 1 mg preparation). Once defined, the IU is passed to all further International Standard preparations by use of a traceability protocol. Many immunoassays are calibrated with such materials; however, the traceability to the WHO unit is often broken by the noncommutability of the WHO materials with native patient samples. This problem is reflected by differences in results among various methods that are nominally calibrated with the same WHO material *(50, 51)*. Measurand-specific reference measurement procedures and improved commutable reference materials must be developed to allow traceable calibration to be established for measurands in this category. Meanwhile, interim solutions, such as use of panels of patient samples with values assigned by consensus, are being discussed.

Traceability chains can change as new references become available. Reference measurement procedures based on more conventional technologies such as immunological methods or HPLC with ultraviolet/visible light detection may be replaced by procedures with higher specificity such as mass spectrometry. Improvement in reference measurement procedures can result in changes in values assigned to reference or calibration materials, which in turn may lead to changes in the results for clinical samples. Approaches should be developed in collaboration with the clinical communities to assure traceability to the highest available reference, while also ensuring that patient care remains consistent despite these changes.

Establishing traceability of measurement results should be linked to performance specifications that are based on clinical, research or public health needs. These specifications would provide information about the relevant calibration range, which is based on expected patient sample values, and about the allowable bias and imprecision to fit the intended clinical or research application (i.e., to distinguish a healthy from a diseased patient). Such performance specifications have been formulated for some analytes and clinical applications, for example, cholesterol and hemoglobin A_{1c}. However, they are missing for many other analytes *(52)*. The ongoing debate regarding goal-setting for analytical performance, however, is beyond the scope of this review *(53–56)*.

Traceability should be assured as early as possible in assay development and used before performing clinical trials and epidemiological studies. This will help to maximize the benefits and outcomes of these trials (i.e., by being able to pool data from different studies), as shown by the success of studies on cholesterol as a biomarker for cardiovascular diseases *(10)*. Costs and complications can also be minimized when traceability is established at an early stage as described above for prostate-specific antigen.

Addressing traceability of assays for new analytes and coping with standardization needs of existing measurements at the same time is challenging, given limited funding and the small and shrinking number of laboratories and individuals with appropriate expertise

in this area. Establishing and maintaining metrological traceability for all individual analytes is a resource-intensive task. Highly skilled, dedicated staff, expensive instrumentation, and a considerable budget are needed for developing, implementing, and maintaining a reference laboratory. Implementing traceability and standardizing measurements over time requires, in addition to the aforementioned resources, sustainable organizational structures. Major standardization programs such as those conducted by CDC for blood lipids and by the National Glycohemoglobin Standardization Program for hemoglobin A_{1c} provide these structures and resources. However, creating such programs for all clinical analytes does not appear practical. Nevertheless, any alternative approach still requires considerable capacities. Therefore, building capacity should become a priority for laboratory medicine so that current and future standardization needs can be better addressed.

The goal for traceability is to assure that results used for care of patients are accurate and comparable over time and location. Establishing metrological traceability satisfies the basic requirements of evidence-based laboratory medicine. Thus, it improves patient care, disease control and prevention, and saves money by allowing the pooling of clinical trial data rather than repeating studies.

Past efforts to establish metrological traceability not only have formalized the approaches and concepts used by standardization projects, but also have introduced a new terminology and nomenclature system that is now increasingly used in laboratory medicine. The challenges of implementing traceability have created a greater awareness of the components relevant to the measurement process itself and stimulated new activities to better define and execute laboratory measurements. Traceability continues to challenge laboratory medicine to assure that data are scientifically sound and are properly translated into information suitable for patient care and public health activities. Establishing traceability for measurement procedures requires considerable resources. The clinical laboratory community must address these needs through capacity building for future challenges in this area. Implementing traceability in laboratory medicine is a global effort led by the laboratory medicine community in collaboration with the clinical and public health communities.

Author Contributions: *All authors confirmed they have contributed to the intellectual content of this paper and have met the following 3 requirements: (a) significant contributions to the conception and design, acquisition of data, or analysis and interpretation of data; (b) drafting or revising the article for intellectual content; and (c) final approval of the published article.*

Authors' Disclosures of Potential Conflicts of Interest: *No authors declared any potential conflicts of interest.*

Role of Sponsor: The funding organizations played no role in the design of study, choice of enrolled patients, review and interpretation of data, or preparation or approval of manuscript.

Acknowledgments: L.M. Thienpont and H.W. Vesper are grateful to D. Stöckl for discussing the manuscript.

Disclaimer: The findings and conclusions in this manuscript are those of the author(s) and do not necessarily represent the views of the CDC/the Agency for Toxic Substances and Disease Registry.

References

1. International vocabulary of basic and general terms in metrology (VIM). ISO/IEC Guide 99:2007. Geneva: International Organization for Standardization; 2007.
2. Cooperation on International Traceability in Analytical Chemistry. Introducing CITAC. http://www.citac.cc (Accessed August 2008).
3. Directive 98/79/EC of the European Parliaments and of the Council of 27 October 1998 on in vitro diagnostic medical devices. L331. Off J Eur Communities 1998;41:1–37.
4. In vitro diagnostic medical devices—measurement of quantities in biological samples—metrological traceability of values assigned to calibrators and control materials. ISO 17511:2003. Geneva: International Organization for Standardization; 2003.
5. World Health Organization. Standardization of procedures for the study of glucose-6-phosphate dehydrogenase. World Health Organ Tech Rep Ser 1967;366:5–25.
6. Cali JP. Standardization of diagnostic materials; 2: problems of standardization in clinical chemistry. Bull World Health Organ 1973;48:721–6.
7. Cali JP. Worldwide reference methods [Letter]. Clin Chem 1975;21:1344.
8. Büttner J. Reference methods as a basis for accurate measuring systems. Eur J Clin Chem Clin Biochem 1991;29:223–35.
9. Tietz NW. A model for a comprehensive measurement system in clinical chemistry. Clin Chem 1979;25:833–9.
10. Rifai N, Cooper GR, Brown WV. Clinical Chemistry journal has contributed to progress in lipid and lipoprotein testing for fifty years. Clin Chem 2004;50:1861–70.
11. Little RR. Glycated hemoglobin standardization: National Glycohemoglobin Standardization Program (NGSP) perspective. Clin Chem Lab Med 2003;41:1191–8.
12. The Lewin Group. Laboratory medicine: a national status report. http://www.futurelabmedicine.org/reports/laboratory_medicine_-_a_national_status_report_from_the_lewin_group.pdf. (Accessed August 2008).
13. 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.
14. Myers GL, Miller WG, Coresh J, Fleming J, Greenberg N, Greene T, et al. Recommendations for improving serum creatinine measurement: a report from the Laboratory Working Group of the National Kidney Disease Education Program. Clin Chem 2006;52:5–18.
15. Delanghe JR, Cobbaert C, Galteau MM, Harmoinen A, Jansen R, Kruse R, et al. Trueness verification of actual creatinine assays in the European market demonstrates a disappointing variability that needs substantial improvement: an international study in the framework of the EC4 creatinine standardization working group. Clin Chem Lab Med 2008;46:1319–25.
16. Kane J, Middle J, Cawood M. Measurement of serum testosterone in women; what should we do?. Ann Clin Biochem 2007;44:5–15.
17. Stanczyk FZ, Lee JS, Santen RJ. Standardization of steroid hormone assays: why, how, and when?. Cancer Epidemiol Biomarkers Prev 2007;16:1713–9.
18. Wierman ME, Basson R, Davis SR. Androgen ther-

apy in women: an Endocrine Society Clinical Practice guideline. J Clin Endocrinol Metab 2006;91:3697–710.
19. Stanczyk FZ. Measurement of androgens in women. Semin Reprod Med 2006;24:78–85.
20. Matsumoto AM, Bremner WJ. Serum testosterone assays: accuracy matters. J Clin Endocrinol Metab 2004;89:520–4.
21. Herold DA, Fitzgerald RL. Immunoassays for testosterone in women: better than a guess?. Clin Chem 2003;49:1250–1.
22. National Institute for Standards and Technology. The impact of calibration error in medical decision making: planning report 04.1. http://www.nist.gov/director/prog-ofc/report04–1.pdf. (Accessed August 2008).
23. National Institutes of Health. ClinicalTrials.gov. www.clinicaltrials.gov. (Accessed August 2008).
24. IFCC-IUPAC. The IFCC-IUPAC coding system: properties and units in the Clinical Laboratory Science (Wash DC) Commission on nomenclature, properties and units (C-NPU). http://dior.imt.liu.se/cnpu/. (Accessed August 2008).
25. Reference materials—general and statistical principles for certification. ISO guide 35:2006. Geneva: International Organization for Standardization; 2006.
26. CLSI. Characterization and qualification of commutable reference materials for laboratory medicine; proposed guideline. CLSI document C53-P. Wayne (PA): CLSI; 2008.
27. In vitro diagnostic medical devices—measurement of quantities in samples of biological origin—presentation of reference measurement procedures. ISO 15193:2002. Geneva: International Organization for Standardization; 2002.
28. Bureau Interantional des Poids and Measures. http://www.bipm.org/ (Accessed October 2008).
29. Infusino I, Bonora R, Panteghini M. Traceability in clinical enzymology. Clin Bichem Rev 2007;28:155–61.
30. Eurachem/CITAC. Eurachem/CITAC guide: traceability in chemical measurements. http://www.citac.cc/. (Accessed August 2008).
31. In vitro diagnostic medical devices—measurement of quantities in biological samples—metrological traceability of values for catalytic concentration of enzymes assigned to calibrators and control materials. ISO 18153:2003. Geneva: International Organization for Standardization; 2003.
32. CLSI. Metrological traceability and its implementation; a report. CLSI document ×5-R. Wayne (PA): NCCLS; 2006.
33. National Institute for Standards and Technology. Uncertainty of measurement results. http://physics.nist.gov/cuu/Uncertainty/ (Accessed October 2008).
34. DeBievre P, Günzler H, eds. Traceability in chemical measurement. New York: Springer; 2005. 281 p.
35. Australasian Association of Clinical Biochemists. The clinical biochemist reviews. http://old.aacb.asn.au/pubs/cbr05ci.htm. (Accessed April 2009).
36. Thienpont LM, Van Uytfanghe K, Marriott J, Stokes P, Siekmann L, Kessler A, et al. Feasibility study of the use of frozen human sera in split-sample comparison of immunoassays with candidate reference measurement procedures for total thyroxine and total triiodothyronine measurements. Clin Chem 2005;51:2303–11.
37. Thienpont LM, Van Nuwenborg JE, Stöckl D. Intrinsic and routine quality of serum total potassium measurement as investigated by split-sample measurement with an ion chromatography candidate reference method. Clin Chem 1998;44:849–57.
38. Kimberly MM, Vesper HW, Caudill SP, Cooper GR, Rifai N, Dati F, Myers GL. Standardization of immunoassays for measurement of high-sensitivity C-reactive protein; phase I: evaluation of secondary reference materials. Clin Chem 2003;49:611–6.
39. Thienpont LM, Stöckl D, Kratochvíla J, Friedecký B, Budina M. Pilot External quality assessment survey for post-market vigilance of in vitro diagnostic medical devices and investigation of trueness of participants' results. Clin Chem Lab Med 2003;41:183–6.
40. Steele BW, Wang E, Klee GG, Thienpont LM, Soldin SJ, Sokoll LJ, et al. 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.
41. Miller WG, Myers GL, Ashwood ER, Killeen AA, Wang E, Ehlers GW, et al. State of the art in trueness and interlaboratory harmonization for 10 analytes in general clinical chemistry. Arch Pathol Lab Med 2008;132:838–46.
42. Thienpont LM, Van Uytfanghe K, De Leenheer AP. Reference measurement systems in clinical chemistry. Clin Chim Acta 2002;323:73–87.
43. Thienpont LM, Beastall G, Christofides ND, Faix JD, Ieiri T, Jarrige V, et al. Proposal of a candidate international conventional reference measurement procedure for free thyroxine in serum. Clin Chem Lab Med 2007;45:934–6.
44. Ekins R. Immunoassay standardization. Scand J Clin Lab Invest Suppl 1991;205:33–46.
45. Rose MP, Gaines Das RE, Balen AH. Definition and measurement of follicle stimulating hormone. Endocr Rev 2000;21:5–22.
46. Miller WG. Specimen materials, target values and commutability for external quality assessment (proficiency testing) schemes. Clin Chim Acta 2003;327:25–37.
47. Vesper HW, Miller WG, Myers GL. Reference materials and commutability. Clin Biochem Rev 2007;28:139–47.
48. Miller WG, Myers GL, Rej R. Why commutability matters. Clin Chem 2006;52:553–4.
49. Thienpont LM. A reference system for cortisol. Clin Biochem 1998;31:483–8.
50. Fierens C, Stockl D, Baetens D, De Leenheer AP, Thienpont LM. Standardization of C-peptide measurements in urine by method comparison with isotope-dilution mass spectrometry. Clin Chem 2003;49:992–4.
51. Marcovina S, Bowsher RR, Miller WG, Staten M, Myers G, Caudill SP, et al. Standardization of insulin immunoassays: report of the American Diabetes Association Workgroup. Clin Chem 2007;53:711–6.
52. Tonks DB. A study of the accuracy and precision of clinical chemistry determinations in 170 Canadian laboratories. Clin Chem 1963;9:217–33.
53. Klee GG. Clinical interpretation of reference intervals and reference limits: a plea for assay harmonization. Clin Chem Lab Med 2004;42:752–7.
54. Fraser CG, Kallner A, Kenny D, Petersen PH. Introduction: strategies to set global quality specifications in laboratory medicine. Scand J Clin Lab Invest 1999;59:477–8.
55. Stöckl D, Baadenhuijsen H, Fraser CG, Libeer JC, Petersen PH, Ricos C. Desirable routine analytical goals for quantities assayed in serum: discussion paper from the members of the External Quality Assessment (EQA) Working Group A on analytical goals in laboratory medicine. Eur J Clin Chem Clin Biochem 1995;33:157–69.
56. Fraser CG, Petersen PH. Desirable standards for laboratory tests if they are to fulfill medical needs. Clin Chem 1993;39:1447–53.