



Harmonization: the Sample, the Measurement, and the Report

W. Greg Miller, Ph.D.¹, Jillian R. Tate, M.Sc.², Julian H. Barth, M.D.³, and Graham R. D. Jones, MBBS.⁴

Department of Pathology¹, Virginia Commonwealth University, Richmond, VA, USA; Pathology Queensland², Department of Chemical Pathology, Royal Brisbane and Women's Hospital, Brisbane, Queensland, Australia; Blood Sciences³, Old Medical School, Leeds Teaching Hospitals Trust, Leeds, UK; SydPath⁴, Department of Chemical Pathology, St. Vincent's Hospital, Darlinghurst, NSW, and University of New South Wales, Kensington, Australia

Harmonization of clinical laboratory results means that results are comparable irrespective of the measurement procedure used and where or when a measurement was made. Harmonization of test results includes consideration of pre-analytical, analytical, and post-analytical aspects. Progress has been made in each of these aspects, but there is currently poor coordination of the effort among different professional organizations in different countries. Pre-analytical considerations include terminology for the order, instructions for preparation of the patient, collection of the samples, and handling and transportation of the samples to the laboratory. Key analytical considerations include calibration traceability to a reference system, commutability of reference materials used in a traceability scheme, and specificity of the measurement of the biomolecule of interest. International organizations addressing harmonization include the International Federation for Clinical Chemistry and Laboratory Medicine, the World Health Organization, and the recently formed International Consortium for Harmonization of Clinical Laboratory Results (ICHCLR). The ICHCLR will provide a prioritization process for measurands and a service to coordinate global harmonization activities to avoid duplication of effort. Post-analytical considerations include nomenclature, units, significant figures, and reference intervals or decision values for results. Harmonization in all of these areas is necessary for optimal laboratory service. This review summarizes the status of harmonization in each of these areas and describes activities underway to achieve the goal of fully harmonized clinical laboratory testing.

Key Words: Commutability, Harmonization, Standardization, Traceability

Received: February 3, 2014

Revision received: March 1, 2014

Accepted: March 7, 2014

Corresponding author: W. Greg Miller
Department of Pathology, Virginia Commonwealth University, P.O. Box 980286, Richmond, VA 23298-0286, USA
Tel: +1-804-828-0375
Fax: +1-804-828-0353
E-mail: gmliller@vcu.edu

© The Korean Society for Laboratory Medicine

This is an Open Access article distributed under the terms of the Creative Commons Attribution Non-Commercial License (<http://creativecommons.org/licenses/by-nc/3.0>) which permits unrestricted non-commercial use, distribution, and reproduction in any medium, provided the original work is properly cited.

INTRODUCTION

Guidelines for ordering and interpreting laboratory tests are typically created by national professional organizations, are increasingly based on international recommendations, and are derived, when possible, from clinical outcomes data. Fixed decision values for laboratory test results are frequently used in practice guidelines. Consequently, harmonized laboratory results are essential to enable appropriate use of clinical practice guidelines. When laboratory results are not harmonized among different measurement procedures, there is a risk that erroneous treatment decisions may be made that affect patient safety as well

as the cost to deliver healthcare.

The meaning of harmonization in this setting is that results are comparable irrespective of the measurement procedure used and where or when a measurement was made. Harmonization of test results includes consideration of pre-analytical, analytical, and post-analytical aspects. These considerations are being addressed by a number of organizations around the world. Unfortunately, the work of different organizations in different countries is not well coordinated for a variety of reasons including: different medical practice approaches, healthcare system funding priorities, longstanding practice customs, and inadequate communication among laboratory professionals practicing

ing in different countries. This review summarizes the current issues and approaches being implemented for harmonization of laboratory test results.

PRE-ANALYTICAL CONSIDERATIONS

1. Nomenclature

A standardized vocabulary for the name of a test procedure is becoming more important with the growing adoption of clinical provider order entry using computer systems. Common terminology avoids confusion regarding what test is to be ordered and provides an order catalog that will be the same in different health care organizations, thus enabling physicians to practice more efficiently when seeing patients in several institutions. It needs to be recognized that terminology for requesting is often different to that for resulting. For example, the single request "Liver Function Tests" may generate six or more individual results depending on the laboratory. In addition, a single request may require more than one sample type; for example, creatinine clearance requires both serum and 24 hr urine samples. Furthermore, adoption of international recommendations for testing requires unambiguous terminology.

Two international standards for terminology are Logical Observation Identifiers Names and Codes (LOINC) and SNOMED CT (currently only an acronym; formerly Systematic Nomenclature of Medicine and Clinical Terminology). These standards are also used for results reporting and will be described in the post-analytical considerations section.

2. Patient preparation and specimen collection

Interpretation of many laboratory results is influenced by the condition of the patient at the time of specimen collection and by the specimen collection process. Common examples of patient condition include: the duration of a fasting state for blood lipids, glucose and other biomarkers; the time of day for blood cortisol, urine albumin, and others; diet for several days preceding sample collection for blood glucose, creatinine, and others; posture for blood albumin, albumin bound molecules, and other proteins; time following last dose or change of dose for therapeutic drug monitoring; and time following challenge doses of oral glucose, intravenous dexamethasone, and other endocrine challenges.

Specimen collection can contribute to variability of a result and in some cases make a result sufficiently erroneous that it is not suitable for use for its intended clinical application. Glucose is a well known example because glucose declines in an unpre-

served whole blood specimen at a rate of approximately 5-7% per hour at room temperature [1]. However, the rate can be greater if the patient has a high white blood cell or platelet count. It is common practice to collect blood glucose in an oxalate fluoride container to inhibit glycolysis and slow the rate of glucose consumption by cells. However, it takes some time for diffusion into the mitochondria and inhibition of the Krebs cycle; consequently, there is inevitably some loss of glucose before the blood sample is centrifuged to obtain plasma for analysis unless the sample is chilled or other precautions are taken [2]. Another example is coagulation factor measurement where the correct volume of specimen must be collected with a specified amount of anticoagulant because the reagent for measurement is formulated to compensate for a specified concentration of anticoagulant in the specimen. A number of tests require special specimen collection and handling procedures to preserve the measurand to obtain a meaningful measurement result. Appropriate specimen transportation conditions to preserve the analyte are also required and are complicated because of different times, distances, and ambient temperatures that must be accommodated in different practice settings. Unfortunately, there is little agreement on what time intervals should be used.

Patient preparation and specimen collection requirements are listed in various books and compendiums [for example, 3]. However, the recommendations are frequently based on inadequate or incomplete literature reports making it difficult for a laboratory to collate and provide such information to providers of care and those who collect specimens for laboratory testing. In addition, there are substantial variations in practices and procedures to collect and transport specimens to a laboratory, which influence variability and uncertainty in the measured results.

3. Uncertainty of a laboratory result from pre-analytical components

Recommendations to estimate and then report or to make available information on uncertainty of a laboratory result have been discussed in recent years [4, 5]. Most recommendations limit uncertainty estimates to the result of an analytical measurement procedure. This limitation may be due in part to the highly variable nature of pre-analytical components, and thus the difficulty to adequately estimate their uncertainty as a numeric value. However, a clinical laboratory should be aware of and take measures to minimize the contribution of pre-analytical factors to uncertainty of laboratory results.

Appropriate control of pre-analytical factors is an important but difficult quality management area for laboratories and for

health care providers in general. Addressing harmonization of pre-analytical factors in laboratory testing is currently not coordinated on an international basis and is an opportunity for improvement in the laboratory medicine profession.

ANALYTICAL CONSIDERATIONS

1. Traceability

Harmonization of results among different clinical laboratory measurement procedures may be achieved by having the calibration of all procedures traceable to the same higher-level reference system. In addition to calibration traceability, it is necessary that the different clinical measurement procedures all measure the same measurand with suitable analytical specificity to prevent other substances that may be present in a clinical sample from influencing the result [6].

The international organization for standardization (ISO) standard 17511:2003, "In vitro diagnostic medical devices-measurement of quantities in biological samples-metrological traceability of values assigned to calibrators and control materials," provides the framework for calibration traceability in laboratory medicine [7]. In a complete reference system, Fig. 1, a pure substance primary reference material allows the SI unit to be realized in a well characterized solution using a primary reference measurement procedure such as gravimetry to prepare the solution. The solution of the primary reference material is used as a calibrator (frequently several concentrations are used) for a higher order secondary reference measurement procedure. Note that for some measurands, such as enzymes or coagulation factors, the primary reference measurement procedure defines the measurand and is the highest level in the traceability chain. The secondary reference measurement procedure is used to value assign a secondary reference material that may be a matrix based certified reference material or a panel of clinical samples. An important characteristic of a secondary reference measurement procedure is that it is not influenced by the difference in matrix between its calibrator and the secondary reference material that is value assigned by measurement with the secondary reference measurement procedure.

The secondary reference material may then be used as a common calibrator by manufacturers to calibrate an internal procedure that is used to value assign either a master lot of a working calibrator or the product calibrator that is provided to a clinical laboratory for use to calibrate the routine clinical laboratory procedure. A master lot is frequently used because it permits a manufacturer to maintain a consistent internal reference calibra-

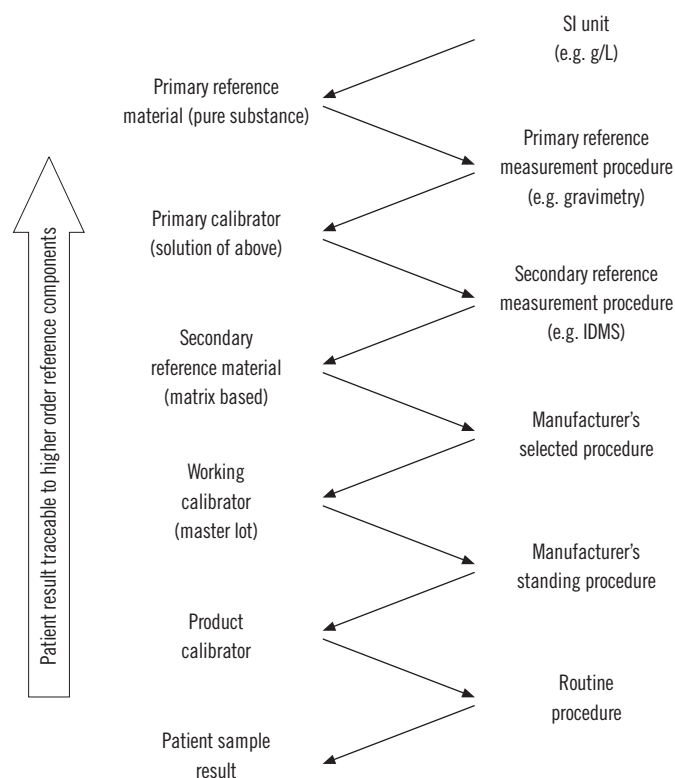


Fig. 1. Components of a complete reference system showing traceability of results from a routine measurement procedure to higher order reference system components based on international standardization for organization (ISO) standard 17511:2003 [7]. Abbreviations: SI, international system of units; IDMS, isotope dilution mass spectrometry.

tor to value assign sequential lots of product calibrators over extended time intervals. Alternatively, a secondary reference material with suitable commutability properties can be used as a true-ness control to confirm the accuracy of values assigned by a secondary reference measurement procedure, by a manufacturer's internal measurement procedure or by a routine procedure.

A calibration traceability hierarchy as shown in Fig. 1 provides a process to calibrate all clinical laboratory measurement procedures, irrespective of manufacturer, to a stable reproducible higher order reference system in a way that can be replicated over time and location to ensure consistent and harmonized results for measurements on clinical samples. The uncertainty in the calibration of the routine clinical laboratory procedure, and thus in the final result for a clinical sample, is the combined uncertainty of the values assigned to the sequence of materials at each step in the traceability chain. Consequently, reference laboratories and manufacturers use analytical approaches such as replication or gravimetric preparation of dilutions to reduce the imprecision of measurements at each step. The ISO standard

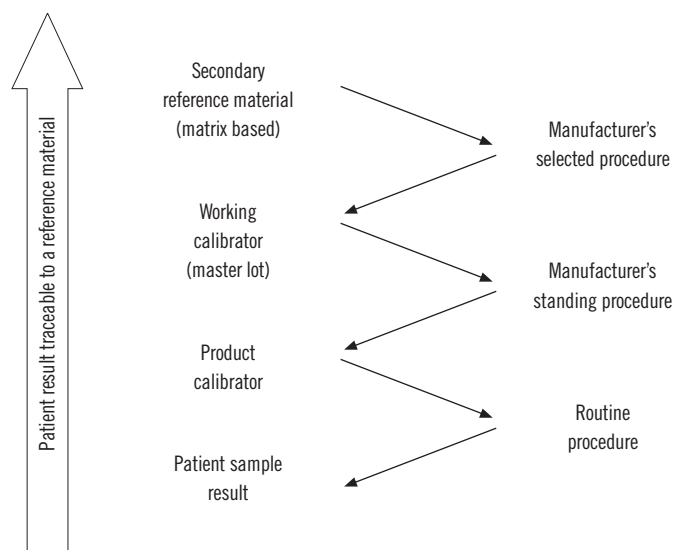


Fig. 2. Components of a reference system showing traceability of results from a routine measurement procedure that ends at a secondary reference material based on ISO Standard 17511:2003 [7].

does not specify the number of steps in a traceability chain for a particular measurement procedure; however, adopting the fewest number of steps necessary can contribute to smaller uncertainty in the final result.

In clinical laboratory medicine, there are only a relatively small number (perhaps 80-90) of measurands, for which higher order reference measurement procedures have been developed. The ISO standard includes a traceability category that has a secondary reference material as its highest order as shown in Fig. 2. The secondary reference material is used as a common calibrator by manufacturers of clinical laboratory measurement procedures. In this situation, the value assigned to the secondary reference material may be arbitrary, but as long as all clinical measurement procedures are calibrated to the same reference material, their results can be harmonized and thus be suitable for use with clinical practice guidelines. Various approaches have been used for value assignment such as: determining the dry mass of a protein [8], using the value of a selected measurement procedure or group of procedures that met defined specifications [9], or assigning an arbitrary value as units.

When there is no reference measurement procedure nor a reference material for a measurand, calibration of a routine clinical laboratory procedure is traceable only to a manufacturer's internal working calibrator. In this situation, which is the case for a large number of measurands, there is likely to be no harmonization of results among different measurement procedures. Interpretation of results is dependent on reference intervals or de-

cision values that are only applicable for the measurement procedure used to create them. Approaches to address this limitation include interpretation based on multiples of the upper reference limit for a given measurement procedure, as has been suggested for parathyroid hormone (PTH) [10], or on a procedure specific value associated with a specified percentile of a non-diseased population as has been recommended for troponin I [11]. Harmonization of measurement procedures in this category has been a challenge that has not been addressed until quite recently and no consensus approaches have been agreed. The International Consortium for Harmonization of Clinical Laboratory Results (ICHCLR) has posted a toolbox of proposed procedures to address harmonization when no reference material is available [12]. A proof of principle has been reported to assign values to clinical samples on the basis of all procedure trimmed mean [13].

2. Commutability

Commutability is an essential property of reference materials used as calibrators in the steps of a traceability chain. Commutability is particularly important for secondary reference materials used as common calibrators for different manufacturers' internal and routine measurement procedures. Commutability is a property of a reference material in which values measured for the reference material and for representative clinical samples have the same relationship between two, or more, measurement procedures for the same measurand [14, 15]. Fig. 3 illustrates the property of commutability. Panel A shows that commutable reference materials have the same relationship as clinical samples, and panel B shows that non-commutable materials have a different relationship. When a reference material is used as a common calibrator, it must be commutable with clinical samples for all of the measurement procedures for which it will be used. Fig. 4 shows that results for clinical samples will not be harmonized if a non-commutable material, such as illustrated in Fig. 3B, is used as a common calibrator for two measurement procedures [14, 16, 17].

Commutability has been recognized for more than 4 decades as an essential property for reference materials [18]. However, the clinical laboratory community has only recently embraced the importance of commutability in achieving harmonized results based on traceability to higher-order reference systems [16, 17]. Some of the historical reluctance to validate commutability of reference materials was based on the technical difficulty to perform the experiments because a panel of clinical samples is needed for the validation. However, it is now recognized that commuta-

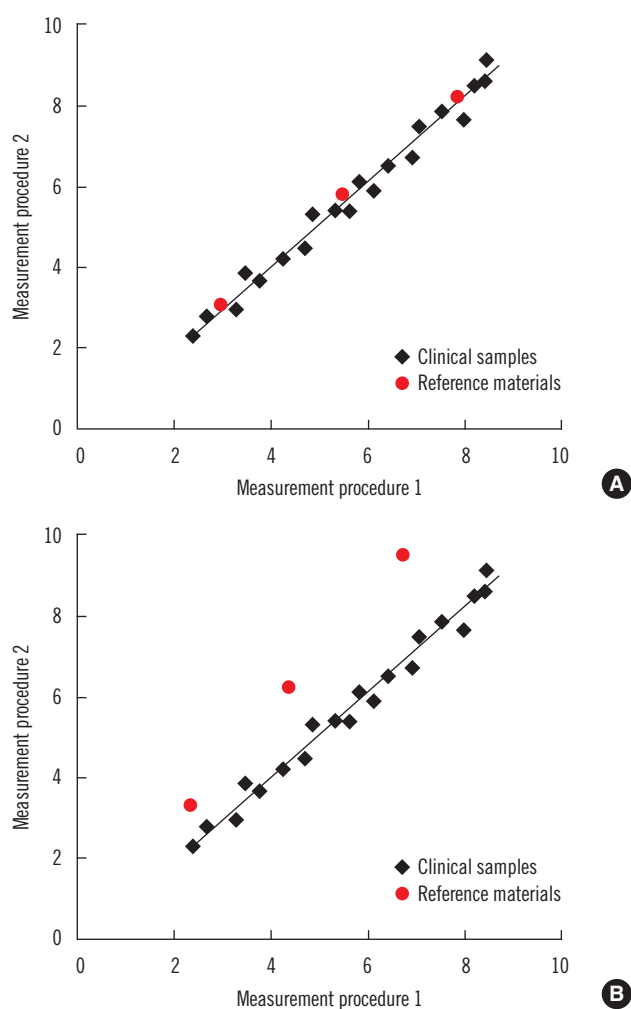


Fig. 3. Representation of the concept of commutability of reference materials with authentic clinical samples between two measurement procedures. Panel A shows that commutable reference materials have the same relationship as clinical samples, and panel B shows that non-commutable materials have a different relationship. Used with permission from reference 24.

bility validation is required for any new reference material and that many existing reference materials are not fit for purpose because they are non-commutable [16]. A consensus guideline for validating commutability is available from the Clinical and Laboratory Standards Institute [19]. External quality assessment programs that use commutable sample materials play an important role to identify measurands in need of harmonization and for surveillance of the success of a harmonization program.

3. Other analytical considerations

In addition to calibration traceability, a complete description of the quantity actually measured and adequate measurement procedure specificity for that quantity are important consider-

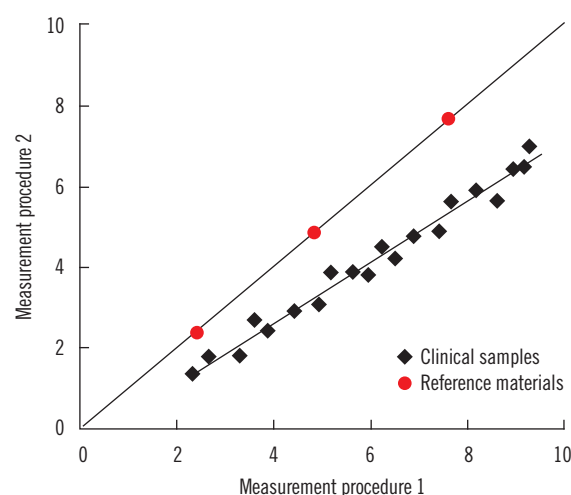


Fig. 4. Representation of differences in results for patient samples when measured by two different procedures that use a non-commutable reference material, such as shown in Fig. 3B, as a common calibrator.

ations for achieving harmonization of results among different measurement procedures. The measurand is defined by the current version of the International Vocabulary of Metrology as “the quantity intended to be measured” [15]. For many biomarkers, there is a difference between the molecule or biological substance intended to be measured and what is actually measured. Consider, for example, troponin I which exists in blood as a complex of several proteins and, furthermore, the composition of the complex changes at different times following a myocardial infarct [20]. Antibodies used in immunoassays for troponin I may bind to different epitopes of the protein’s three dimensional structure. In this example, the measurand, the quantity intended to be measured, is troponin I, but the quantity actually measured is defined by the epitopes that are recognized by the antibodies used in a given measurement procedure. Because different measurement procedures use different antibodies, the measurement procedures are not measuring the same quantity even though each claims to measure troponin I. This situation represents an inadequate definition of the measurand because the specific molecular entity that can be considered “troponin I” is not stated. An essential requirement to achieve harmonization is to adequately describe, at a molecular level, the quantity to be measured as the biomarker for a given clinical condition.

Another essential requirement for harmonization of results among different routine clinical laboratory measurement procedures is that all the procedures have adequate analytical specificity for the quantity to be measured. If a group of routine pro-

cedures do not measure the same quantity, it may not be possible for results from those procedures to be harmonized. Similarly, if a routine procedure is influenced by substances in a clinical sample other than the quantity being measured, it may not be possible for results from that procedure to be harmonized with results from other procedures that measure the same biomarker. Unfortunately, cost and throughput considerations may compromise the specificity of some routine measurement procedures. In such cases, it is necessary for the manufacturer to improve the performance of a measurement procedure before its results can be harmonized with those from other procedures.

4. International initiatives for harmonization of measurement results

The importance of harmonized laboratory results is recognized by many professional and public health organizations throughout the world. An international organization with a longstanding commitment to improve harmonization is the International Federation for Clinical Chemistry and Laboratory Medicine (IFCC). The IFCC Scientific Division has sponsored numerous committees and working groups to develop reference measurement procedures and reference materials for measurands of clinical importance [21]. In addition to reference systems, the IFCC sponsors initiatives to address reference intervals, nomenclature and other areas of laboratory medicine.

In 1998, the European Community (EC) Directive 98/79/EC on *in vitro* medical devices required measurement procedures sold in the European Union (EU) to have calibration traceable to higher order reference systems [22]. In response to the EU directive, the IFCC, the International Committee of Weights and Measures and the International Laboratory Accreditation Cooperation formed the Joint Committee for Traceability in Laboratory Medicine (JCTLM) in 2002. The JCTLM maintains lists of reference measurement procedures, reference measurement laboratories and reference materials that have been reviewed for conformance to the applicable ISO standards for each type of resource [23]. Listing by JCTLM has become a practical requirement for a reference system component to be used for calibration traceability by manufacturers of *in vitro* diagnostics. The JCTLM has updated the Quality Manual of Working Group 1 to require information on commutability for new reference material submissions that are intended for use as common calibrators for clinical laboratory measurement procedures.

The WHO provides reference materials referred to as “International Standards” and “Reference Reagents” that are used as common calibrators for measurement procedures in areas such

as blood safety, infectious diseases and endocrinology. Historically the WHO has not validated these materials for commutability and a number of reports have shown that traceability to these materials has not achieved harmonization [16] because they are non-commutable. In 2013, the WHO convened a conference to discuss commutability requirements and determined that it is necessary to consider commutability in future reference materials. Another limitation of WHO materials has occurred because there has not been a requirement to ensure consistency of value assignment when a depleted lot is replaced with a new lot. However, the situation can be complex because in some cases advances in science may make a replacement lot technically superior to an older one thus necessitating a change in value assignment.

The ICHCLR was formed in 2013 in response to recommendations from a conference organized by the American Association for Clinical Chemistry in 2010: “Improving Clinical Laboratory Testing through Harmonization: An International Forum.” The conference recommended that coordination and prioritization of harmonization activities needed improvement to best meet the needs of laboratory medicine worldwide [24]. The ICHCLR was formed to provide: a systematic approach for prioritization of measurands to be harmonized based on clinical importance and the technical feasibility to achieve harmonization; an information portal on global harmonization activities to avoid duplication of effort; and procedures to implement harmonization for measurands for which no reference measurement procedure was likely to be developed. The ICHCLR web site [25] provides information on harmonization activities being conducted by organizations throughout the world (in development), a toolbox of procedures to assess the status and feasibility to achieve harmonization, an on-line form to submit measurands to be reviewed and prioritized for a harmonization effort, and a list of prioritized measurands that will be developed as measurands are reviewed. The ICHCLR’s Harmonization Oversight Group will identify collaborators and sponsors for harmonization projects for the highest priority measurands. Interested stakeholders may join the Strategic Partners Group of the ICHCLR to become active participants in the work.

POST-ANALYTICAL CONSIDERATIONS

1. Nomenclature

A standardized vocabulary for the name of a test procedure result is becoming more important with the widespread use of electronic health record (EHR) systems. Results from several

Table 1. Examples of LOINC and SNOMED CT code values

Code value	Description
LOINC*	
1558-6	Fasting glucose [Mass/volume, mg/dL] in Serum or Plasma
14771	Fasting glucose [Moles/volume, mmol/L] in Serum or Plasma
1518-0	Glucose [Mass/volume, mg/dL] in Serum or Plasma --2 hr post 75 g glucose PO
14995-5	Glucose [Moles/volume, mmol/L] in Serum or Plasma --2 hr post 75 g glucose PO
2857-1	Prostate specific Ag [Mass/volume, ng/mL] in Serum or Plasma
35741-8	Prostate specific Ag [Mass/volume, µg/L] in Serum or Plasma by Detection limit < =0.01 ng/mL
19195-7	Prostate specific Ag [Units/volume, IU/L] in Serum or Plasma
33667-7	Prostate specific Ag protein bound [Mass/volume, ng/mL] in Serum or Plasma
10886-0	Prostate Specific Ag Free [Mass/volume, ng/mL] in Serum or Plasma
SNOMED CT†	
399068003	Malignant tumor of prostate (disorder)
399590005	Squamous cell carcinoma of prostate (disorder)
427492002	Hormone refractory prostate cancer
314969001	Local recurrence of malignant tumor of prostate (disorder)
369485004	Malignant tumor involving prostate by direct extension from bladder (disorder)

*LOINC information from <http://search.loinc.org>, accessed on 18 January 2014; †SNOMED CT information from <http://phinivads.cdc.gov/vads/http://phinivads.cdc.gov/vads/ViewCodeSystemConcept.action?oid=2.16.840.1.113883.6.96&code=399068003>, accessed on 18 January 2014.

different laboratories or obtained over extended time intervals, which could include a lifetime, are frequently aggregated into a common EHR to facilitate coherent care for a patient. Interoperability of electronic information requires harmonized terminology as well as harmonized protocols for data exchange.

Two international standards for terminology used for laboratory results are LOINC and SNOMED CT. A LOINC code identifies a laboratory test and may specify key clinical variables regarding patient preparation, collection conditions, specimen type, unit of measure, reference interval and some aspects of measurement methodology. Examples of LOINC codes for several laboratory tests are shown in Table 1. Examination of the prostate specific antigen examples indicates that the coding may not be adequate to discriminate when non-harmonized values are obtained by different measurement procedures. LOINC was initiated in 1994 by the Regenstrief Institute, a non-profit medical research organization. LOINC applies code values to medical laboratory and clinical terminology to facilitate electronic exchange and aggrega-

tion of results in an EHR. At least 10 countries define LOINC as their national standard and the system has been translated into 14 languages or dialects (e.g., American and British English are different dialects).

A SNOMED CT code specifies an anatomic site, diagnosis, organism and other parameters associated with a clinical observation. Examples of SNOMED codes are also shown in Table 1. SNOMED was initially developed in 1965 by the College of American Pathologists as the Systematized Nomenclature of Pathology and evolved into the Systematized Nomenclature of Medicine as it included other areas. SNOMED CT was created in 2002 by merging with the Clinical Terms coding system developed by the National Health Service of the United Kingdom (UK). SNOMED CT is now managed by the non-profit International Health Terminology Standards Development Organization, is used by 23 countries and is available in 7 languages or dialects.

LOINC and SNOMED work together to fulfill a request and a result for electronic data exchange using the Health Level 7 protocol. LOINC states the request and may fulfill the response for numeric or qualitative results. SNOMED fulfills the response for textual descriptive results. Taken together, these two coding systems provide a harmonization scheme to permit comparison of results in an EHR in a meaningful way to minimize confusion about comparability of results that may have originated from different laboratories, from different health centers or at different times in the course of a patient's disease progression. Unfortunately, the LOINC code system does not adequately recognize that result values may not be harmonized for all measurement procedures included in a single LOINC code.

An initiative in England (UK) has been the development of the National Laboratory Medicine catalogue. Although this catalogue is a national exercise, it is gaining momentum as a potential international standard. The aim is to produce a catalogue of all pathology tests in order to ensure that the nomenclature for all tests requesting and reporting is carried out in a standard, interoperable format [26].

2. Reporting units

Harmonized reporting units remain a challenge for the clinical laboratory community. In principle, the international system of units (SI) provides an agreed standard. However, in practice longstanding customs prevent countries from making a change. One argument is that the continuity of clinical care could be disrupted, and thus patient safety issues could occur if physicians had to adjust to a new system of units for laboratory tests with which they were already familiar and for which well established

interpretive guidelines were in place. At present, guidelines need to contain multiple sets of advice with different units [27]. Another argument is that physicians are accustomed to changes in laboratory test methodology with attendant changes in numeric values and reference intervals, and adapt quickly to the change. However, in a world where tests are transferred electronically, the argument for adopting a single standardized set of units is identical to that for test names. Clinicians typically consider test results by value alone and rarely consider that the units of measurement may be different. The experience in the UK of changing units of measurement for glycated hemoglobin suggests that, given appropriate education, clinicians and patients can manage a change in units without deterioration in clinical status [28]. Nonetheless, the goal of clinical laboratories should be to ensure harmonization of the numeric values and units from laboratory procedures to avoid changes in values used for interpretive criteria.

A reporting units issue is the variety of ways the same unit might be represented in an EHR or in a laboratory report. For example, one thousand cells per micro-liter might be expressed as: $10^3/\mu\text{L}$, $10^3/\mu\text{L}$, $1,000/\mu\text{L}$, $10^3/\text{L}$, $\text{Thou}/\mu\text{L}$, Th/mm^3 , or K/cumm . An approach to harmonize such unit representation issues is being developed by the Unified Code for Units of Measure (UCUM) program [29]. This code set is supported by LOINC and is intended to facilitate unambiguous electronic communication of quantities together with their units. The units in the previous example are represented as $10^3/\mu\text{L}$ in the UCUM. Absolute clarity of units is required when results may be subject to automated calculations based on electronic inputs and the use of UCUM provides a solid basis for this process.

A particularly troublesome aspect of lack of harmonization of units of measure is when the units cause a change in numeric values that may lead to misinterpretation by the physician and risk of harm to the patient. For example, the same concentration of digoxin might be expressed as 3 ng/mL or $0.3 \mu\text{g/dL}$; the latter value might be confused as an inadequate level causing an increase in dose and serious toxicity whereas the two concentrations are actually identical.

There have been national initiatives to recommend common units for laboratory test results. For example, the Pathology Harmony program in the UK has published a list of recommended units [30] and the Royal College of Pathologists of Australasia has done similarly [31]. The IFCC committee on Nomenclature, Properties and Units has a cooperative agreement with the International Union of Pure and Applied Chemistry that has worked since 1995 to develop terminology and a coding system

for laboratory results. This cooperative maintains a database of recommended nomenclature and units for laboratory tests [32]. More recently, this cooperative has a memorandum of understanding with the International Health Terminology Standards Development Organization and the Regenstrief Institute to develop a coherent coding scheme for nomenclature and units. However, these initiatives have not been widely adopted on an international scale partly because of embedded older systems and partly because these coding schemes include options for different units that are in current use thus failing to promote a single system. Despite substantial progress, standardization of both nomenclature and units remains a challenge for clinical laboratory medicine.

3. Reference intervals and decision values

While laboratories are well trained in method verification and validation to determine if assays are fit-for-purpose, they are less aware of the importance of selecting the most appropriate and evidence-based reference intervals for optimal interpretation of results, i.e., 'Right interpretation with the Right advice as to what to do next with the result' [33].

Studies have shown that the variation in reference intervals for clinical analytes may be much greater than the analytical inaccuracy of their measurements [34]. Laboratories that use the same platforms and same reagents but use different reference intervals/decision limits can give different result interpretations for the same values. This situation has the potential to worsen a patient's outcome by causing different clinical interpretation and unnecessary additional laboratory testing, with some risk to the patient of inappropriate investigation or treatment [35, 36].

Clinical care providers may not be aware of these methodological differences, especially if the result transfer from the laboratory to the general practice does not show the different interpretive criteria for the measurement procedures in use. There are implications for the amalgamation of results from different pathology providers and the use of electronic results messages being used to create a central EHR database of results for a patient.

One solution to the problem is to define "common" reference limits and decision points. Reference limits and decision points can be classified on the basis of their quality using the Stockholm hierarchy as a classification standard [37]. Decision thresholds based on clinical outcome studies constitute the highest level of quality, with the clinical expectation that all methods employed in the clinical setting are harmonized as is the case, for example, for glucose, cholesterol, and HbA1c. By contrast, reference limits based on a measurement procedure's kit insert data

constitute the lowest quality level, usually having the least harmonization with little possibility of shared reference limits.

In between these extremes there are other approaches to achieving harmonized reference limits as shown in Table 2. The Nordic Reference Interval Project (NORIP) established common reference intervals in apparently healthy adult populations from five Nordic countries for 25 of the most common clinical chemistry analytes [38]. The current project being run by the IFCC Working Group for Reference Intervals and Decision Limits is expected to report soon on the potential for global reference in-

Table 2. Scandinavian, United Kingdom, and Australian and New Zealand adult reference intervals

Analyte	Unit	NORIP*	Pathology Harmony†	AHRIA (proposed)‡
Calcium	mmol/L	2.17-2.47 (18-49 yr)	2.2-2.6	2.10-2.60
Phosphate	mmol/L	M: 0.75-1.65 (18-49 yr) F: 0.85-1.50 (18-49 yr)	0.8-1.5	0.75-1.50
Sodium	mmol/L	137-145	133-146	135-145
Potassium	mmol/L	3.6-4.6 (serum)	3.5-5.3	3.5-5.2 (serum)
Chloride	mmol/L	-	95-108	95-110
Bicarbonate	mmol/L	22-32	22-29	22-32
Creatinine	μmol/L	M: 60-100 F: 50-90	-	M: 60-110 (18-60 yr) F: 45-90 (18-60 yr)
Total Protein	g/L	62-78 (serum)	60-80	60-80
LDH [L to P]§	U/L	105-205 (18-69 yr) 115-255 (70+ yr)	-	120-250

*NORIP, Nordic Reference Interval Project [38]; †Pathology Harmony UK project [40]; ‡AHRIA, Australasian Harmonised Reference Intervals for Adults project [41, 42]; §LDH [L to P], lactate dehydrogenase [lactate to pyruvate].

tervals [39]. Importantly, both studies have used results that were traceable to higher-order reference systems for much of their work.

In the UK, harmonized reference limits were established by a formal process involving a survey of reference intervals followed by a consensus agreement [40], and this agreement was then endorsed by all the national laboratory medicine associations. In Australia and New Zealand, an initiative is currently underway to achieve harmonized reference limits through an evidence-based approach and understanding the various physiological factors that affect reference limits and the pathological factors that affect decision limits [41, 42]. A checklist assessment process is

Table 3. Checklist criteria used in Australian and New Zealand Harmonized Reference Intervals project*

1. Define analyte (measurand)
2. Define measurement procedure used, accuracy base, analytical specificity
3. Consider important pre-analytical differences, actions in response to interference
4. Define distribution of reference interval values (e.g., central 95%, 99%, etc.)
5. Describe evidence for merging of reference intervals
 - data sources (literature, lab surveys, manufacturer)
 - data mining
 - bias goal as quality criterion for acceptance
6. Consider partitioning based on age, sex, etc.
7. Define degree of rounding
8. Clinical considerations of the reference interval
9. Consider use of common reference interval
10. Document and implement

*Checklist was derived from reference 43.

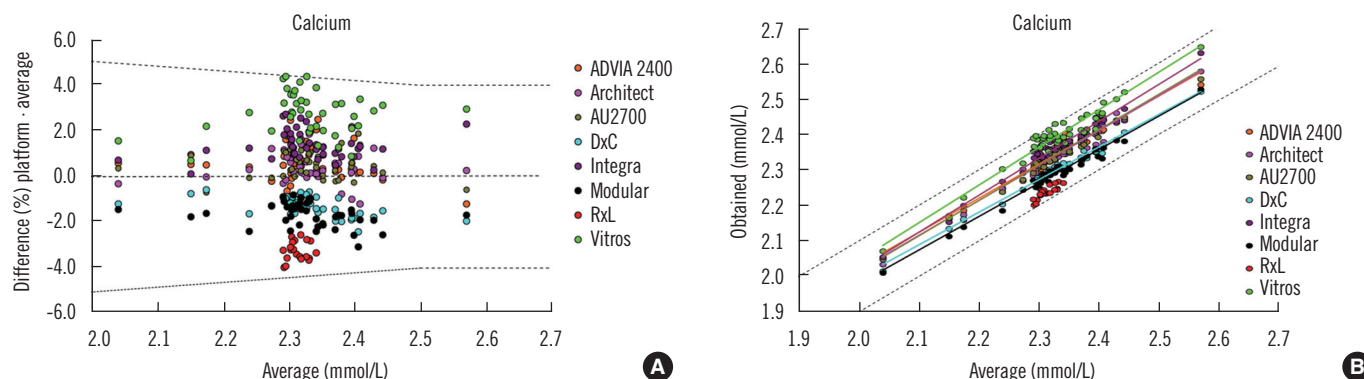


Fig. 5. Assessment of suitability of a common reference interval for different routine measurement procedures for calcium using data from 33 reference interval subjects measured by 24 laboratories using 8 platforms (at least 3 laboratories participated per platform) and acceptance criteria from the Royal College of Pathologists of Australasia Quality Assurance Program [45]. (A) almost all results for calcium fell within the allowable limits of agreement (± 0.1 mmol/L up to 2.5 mmol/L and $\pm 4\%$ when > 2.5 mmol/L variation from the all methods mean) (B) the regression lines were all within the allowable limits of performance for the eight routine measurement procedures that were evaluated. (A) is used with permission from reference 42. (B) is used with permission from a study performed by the Harmonisation Group of the Australasian Association of Clinical Biochemists (www.aacb.asn.au/professionaldevelopment/harmonisation).

being used to assess the evidence for the use of common reference intervals (Table 3), and data are recorded in a structured spreadsheet template format [43].

One important aspect of harmonized reference intervals is development of the criteria for a laboratory to use a common reference interval, i.e. the allowable bias and imprecision of a measurement procedure compared to an accepted reference. Information about method comparability may be determined through an assessment of the between-method bias using commutable biological samples. In the Australasian study, the approach taken when analysing the data was to compare the average result for each measurement procedure with the mean of all results from the Royal College of Pathologists of Australasia Quality Assurance Program. The allowable limits of performance were used to determine whether bias would prevent the use of a common reference interval [44, 45]. Fig. 5 shows an example of assessment of suitability of a common reference interval for different routine measurement procedures for calcium. Panel A shows that almost all results fell within the allowable limits of agreement for the analyte, and panel B shows that the regression lines were all within the allowable limits of performance for the eight routine measurement procedures that were evaluated [42]. The final responsibility for adopting a proposed common reference interval lies with the laboratory director, and local validation of the common reference intervals is recommended to ensure their appropriate use for the population served by a clinical laboratory [46].

The analytical quality of the measurement procedures will ultimately determine which analytes can share common or harmonized reference intervals. Reference intervals can be transferred between laboratories and between measurement procedures provided that the measurement procedures used produce results traceable to the same reference system, there is verification of similar pre-analytical conditions, and the demographics of populations being tested are similar [47].

CONCLUSIONS AND RECOMMENDATIONS

The increased use of standardized clinical practice guidelines developed from evidence-based outcomes research requires clinical laboratory test results to be harmonized. Such guidelines are becoming international thus requiring global cooperation among laboratory medicine stakeholders to achieve harmonized results. The rapidly expanding use of EHR as well as an increasingly mobile population is also driving the importance of harmonized laboratory results. Substantial efforts are underway to ad-

vance the science and the technical processes to achieve harmonization. These efforts need to be supported and organized on a global scale to deliver cost effective and clinically optimized laboratory medicine services.

REFERENCES

1. Sacks DB. In: Burtis CA, Ashwood ER, Bruns DE, ed. Tietz textbook of clinical chemistry and molecular diagnostics. 5th eds. Philadelphia: Elsevier, 2012:719.
2. Sacks DB, Arnold M, Bakris GL, Bruns DE, Horvath AR, Kirkman MS, et al. Guidelines and recommendations for laboratory analysis in the diagnosis and management of diabetes mellitus. Clin Chem 2011;57:e1-e47.
3. Wu AHB. Tietz clinical guide to laboratory tests. 4th ed. St. Louis: WB Saunders, 2006.
4. Clinical and Laboratory Standards Institute. Expression of measurement uncertainty in laboratory medicine; Approved Guideline. CLSI document EP29-A. Wayne, PA: Clinical and Laboratory Standards Institute, 2012.
5. White GH. Basics of estimating measurement uncertainty. Clin Biochem Rev 2008;29(S1):S53-60.
6. Vesper HW and Thienpont LM. Traceability in laboratory medicine. Clin Chem 2009;55:1067-75.
7. ISO 17511:2003. *In vitro* diagnostic medical devices -measurement of quantities in biological samples-metrological traceability of values assigned to calibrators and control materials. 2003 International Organization for Standardization Geneva, Switzerland, 2003.
8. Blirup-Jensen S. Protein standardization II: dry mass determination procedure for the determination of the dry mass of a pure protein preparation. Clin Chem Lab Med 2001;39:1090-7.
9. Zegers I, Keller T, Schreiber W, Sheldon J, Albertini R, Blirup-Jensen S, et al. Characterization of the new serum protein reference material ERM-DA470k/IFCC: value assignment by immunoassay. Clin Chem 2010; 56:1880-8.
10. Sturgeon CM, Sprague SM, Metcalfe W. Variation in parathyroid hormone immunoassay results—a critical governance issue in the management of chronic kidney disease. Nephrol Dial Transplant 2011;26: 3440-5.
11. Thygesen K, Alpert JS, Jaffe AS, Simoons ML, Chaitman BR, White HD; Joint ESC/ACCF/AHA/WHF Task Force for the Universal Definition of Myocardial Infarction. Third universal definition of myocardial infarction. Circulation 2012;126:2020-35.
12. Weycamp C, Eckfeldt J, Vesper H, Thienpont L, Burns C, Caliendo A, et al. Eds. Toolbox of technical procedures to be considered when developing a process to achieve harmonization for a measurand. http://www.harmonization.net/Resource/Documents/Tool_Box_2013.pdf; (Updated on Jun 2013).
13. Van Houcke SK, Van Aelst S, Van Uytendange K, Thienpont LM. Harmonization of immunoassays to the all-procedure trimmed mean - proof of concept by use of data from the insulin standardization project. Clin Chem Lab Med 2013;51:e103-5.
14. Miller WG, Myers GL, Rej R. Why commutability matters. Clin Chem 2006;52:553-4.
15. The Joint Committee for Guides in metrology. International vocabulary of metrology- Basic and general concepts and associated terms. VIM. 3rd edn. Se'vres, France: JCGM 200: 2008. See http://www.bipm.org/utis/common/documents/jcgm/JCGM_200_2012.pdf (accessed on Jan 2014).

16. Miller WG and Myers GL. Commutability still matters. *Clin Chem* 2013; 59:1291-3.
17. Vesper HW, Miller WG, Myers GL. Reference materials and commutability. *Clin Biochem Rev* 2007;28:139-47.
18. Fasce CF Jr, Rej R, Copeland WH, Vanderlinde RE. A discussion of enzyme reference materials: applications and specifications. *Clin Chem* 1973;19:5-9.
19. Clinical and Laboratory Standards Institute. Characterization and qualification of commutable reference materials for laboratory medicine; Approved guideline. CLSI Document EP30-A. Wayne, PA: CLSI, 2010.
20. Wu AH, Feng YJ, Moore R, Apple FS, McPherson PH, Buechler KF, et al. Characterization of cardiac troponin subunit release into serum after acute myocardial infarction and comparison of assays for troponin T and I. *Clin Chem* 1998;44:1198-208.
21. Gillery P and Young IS. Progress toward standardization: an IFCC Scientific Division Perspective. *Clin Chem Lab Med* 2013;51:915-8.
22. 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(Dec 7);L331:1-37.
23. Joint Committee for Traceability in Laboratory Medicine (JCTLM). <http://www.bipm.org/jctlm/> (Accessed on Jan 2014).
24. Miller WG, Myers GL, Gantzer ML, Kahn SE, Schönbrunner ER, Thienpont LM, et al. Roadmap for harmonization of clinical laboratory measurement procedures. *Clin Chem* 2011;57:1108-17.
25. American Association of Clinical Chemistry (AACC), International consortium for harmonization of clinical laboratory results (ICHCLR). <http://www.harmonization.net/Pages/default.aspx> (Updated on Aug 2013).
26. Jones R, Batstone G, Gutteridge C, Croal B, Barnes I. NLMC-more than just for pathology. *Bull Roy Coll Pathol* 2012;160:233-6.
27. Funder JW, Carey RM, Fardella C, Gomez-Sanchez CE, Mantero F, Stowasser M, et al. Case detection, diagnosis, and treatment of patients with primary aldosteronism: an endocrine society clinical practice guideline. *J Clin Endocrinol Metab* 2008;93:3266-81.
28. Kilpatrick ES, Rigby AS, Atkin SL, Barth JH. Glycemic control in the 12 months following a change to SI hemoglobin A1c reporting units. *Clin Chem* 2013;59:1457-60.
29. The UCUM Organization, The Unified Code for Units of Measure. <http://unitsofmeasure.org> (Updated on Oct 2013).
30. Pathology Harmony, UK. <http://www.pathologyharmony.co.uk> (Accessed on Jan 2014).
31. Royal College of Pathologists of Australasia, Pathology terminology and information standardization downloads. <http://www.rcpa.edu.au/Library/Practising-Pathology/PTIS/APUTS-Downloads> (Updated on Nov 2013).
32. International Federation of Clinical Chemistry and Laboratory Medicine, Nomenclature, Properties and Units (C-NPU) in collaboration with International Union of Pure and Applied Chemistry (IUPAC). <http://www.ifcc.org/ifcc-scientific-division/sd-committees/c-npu/> (Accessed on Jan 2014).
33. Plebani M. Harmonization in laboratory medicine: the complete picture. *Clin Chem Lab Med* 2013;51:741-51.
34. Jones GR, Barker A, Tate J, Lim CF, Robertson K. The case for common reference intervals. *Clin Biochem Rev* 2004;25:99-104.
35. Jassam N, Yundt-Pacheco J, Jansen R, Thomas A, Barth JH. Can current analytical quality performance of UK clinical laboratories support evidence-based guidelines for diabetes and ischaemic heart disease? --A pilot study and a proposal. *Clin Chem Lab Med* 2013;51:1579-84.
36. Jones GR. Laboratory reporting of urine protein and albumin. *Clin Biochem Rev* 2011;32:103-7.
37. Sikaris K. Application of the Stockholm hierarchy to defining the quality of reference intervals and clinical decision limits. *Clin Biochem Rev* 2012;33:141-8.
38. Rustad P, Felding P, Franzson L, Kairisto V, Lahti A, Mårtensson A, et al. The Nordic Reference Interval Project 2000: recommended reference intervals for 25 common biochemical properties. *Scand J Clin Lab Invest* 2004;64:271-84.
39. Ozarda Y, Ichihara K, Barth JH, Klee G; Committee on Reference Intervals and Decision Limits (C-RIDL), International Federation for Clinical Chemistry and Laboratory Medicine. Protocol and standard operating procedures for common use in a worldwide multicenter study on reference values. *Clin Chem Lab Med* 2013;51:1027-40.
40. Berg J and Lane V. Pathology Harmony; a pragmatic and scientific approach to unfounded variation in the clinical laboratory. *Ann Clin Biochem* 2011;48:195-7.
41. Australasian Association of Clinical Biochemists, Australasian Association of Clinical Biochemists Harmonisation Activity. www.aacb.asn.au/professionaldevelopment/harmonisation (Accessed on Jan 2014).
42. Koerbin G, Sikaris KA, Jones GR, Ryan J, Reed M, Tate J; On behalf of the AACB Committee for Common Reference Intervals. Evidence-based approach to harmonized reference intervals. *Clin Chim Acta* 2013; doi: 10.1016/j.cca.2013.10.021. [Epub ahead of print].
43. Jones G and Barker A. Reference intervals. *Clin Biochem Rev* 2008; 29(S1):S93-7.
44. Jones GR, Sikaris K, Gill J. 'Allowable Limits of Performance' for External Quality Assurance Programs - an Approach to Application of the Stockholm Criteria by the RCPA Quality Assurance Programs. *Clin Biochem Rev* 2012;33:133-9.
45. Allowable limits of performance for the Royal College of Pathologists of Australasia Quality Assurance Program are available at: <http://www.rcpaqap.com.au/wp-content/uploads/2013/06/chempath/Allowable%20Limits%20of%20Performance.pdf>.
46. Jones GR. Validating common reference intervals in routine laboratories. *Clin Chim Acta* 2013; doi: 10.1016/j.cca.2013.10.005. [Epub ahead of print].
47. Panteghini M and Ceriotti F. Obtaining reference intervals traceable to reference measurement systems: is it possible, who is responsible, what is the strategy? *Clin Chem Lab Med* 2011;50:813-7.