Standardization of measurements of 25-Hydroxyvitamin D3 and D2

Linda M. Thienpont[1], **Hedwig C. M. Stepman**[1], and **Hubert W. Vesper**[2]

[1]Laboratory for Analytical Chemistry, Faculty of Pharmaceutical Sciences, Ghent University, Harelbekestraat 72, B-9000 Ghent, Belgium

[2]Centers for Disease Control and Prevention (CDC), Division of Laboratory Sciences, 4770 Buford Highway, MS F25, Atlanta, GA 30341, United States

Abstract

The vitamin D status is increasingly assessed/monitored in different populations, research cohorts and individual patients. This is done by measuring the liver metabolites 25-hydroxyvitamin D3 and D2 as biomarkers. Recommendations for using specific serum concentrations of these biomarkers to assess a person's vitamin D status were done. This requires current vitamin D assays to be sufficiently accurate over time, location and laboratory procedures. In view of the fact that several studies demonstrated that current 25(OH)D measurement methods do not meet this prerequisite, standardization is needed. This paper rehearses the basic concept of standardization, in particular applied to measurements of 25-hydroxyvitamin D. Progress has been made by establishing a reference measurement system consisting of reference methods and reference materials. Coordinated efforts to improve the accuracy and standardize measurements are being performed by organizations such as the U.S. NIH, the CDC and Prevention, the NIST together with their national and international partners. Beyond describing the available reference measurement system and its use as calibration hierarchy to establish traceability of measurements with routine laboratory methods to the SI-unit, this report will also focus on other aspects considered essential for a successful and sustainable standardization, such as analytical issues related to the definition of the measurand and analytical performance goals.

Correspondence: L. M. Thienpont Laboratory for Analytical Chemistry, Faculty of Pharmaceutical Sciences, Ghent University, Harelbekestraat 72, B-9000 Ghent, Belgium. linda.thienpont@ugent.be.

Disclaimer

The findings and conclusions in this report are those of the author(s) and do not necessarily represent the views of the Centers for Disease Control and Prevention/the Agency for Toxic Substances and Disease Registry.

Declaration of interest: The authors report no conflicts of interest. The authors alone are responsible for the content and writing of the paper.

Introduction

As a result of the emerging consensus on the implication of severe vitamin D deficiency on people's health, there is an increasing need to assess and monitor the vitamin D status in different populations, research cohorts and individual patients [1–4]. The circulating metabolites 25-hydroxyvitamin D3 and D2 (25(OH)D3 and 25(OH)D2) are currently used as biomarkers. Numerous different immunological, mass spectrometry-based and spectrophotometric methods are currently being used to measure these compounds in serum [5].

The development of evidence-based guidelines for diagnosis and management of diseases related to impaired vitamin D status requires aggregation and interpretation of data generated in different research and epidemiological studies using different analytical methods. In addition, the implementation of such guidelines in everyday patient care requires measurements performed on individual patients being comparable to those used to develop the clinical guidelines (see Figure 1). It needs to be pointed out that considerable time can pass between the generation of research data, the development of guidelines and the use of these guidelines in patient care. Therefore, it is highly important that measurements are comparable over time, location and laboratory procedures.

Several studies found considerable variability in results of 25(OH)D measurements between analytical methods be it those based on (radio)immunochemistry, high performance liquid chromatography with ultraviolet detection (HPLC-UV) or isotope dilution-liquid chromatography/tandem mass spectrometry (ID-LC/tandem MS) [5,7–10]. The sometimes huge between-method discrepancy is known for many years from data obtained in dedicated international proficiency surveys, such as the Vitamin D External Quality Assessment Scheme (DEQAS) [11, 12]. In 2010, a National Institute of Health (NIH) roundtable with different federal organizations and researchers also pointed out considerable fluctuations of assays over time [13]. It was found that the variability between laboratories as well as within assays can lead to misdiagnosis of patients [14] and misinterpretation of population data for public health policy making [13–16]. This problem was highlighted in a recent report from the Institute of Medicine on dietary reference intakes for calcium and vitamin D, in which specific serum vitamin D concentrations were suggested to evaluate a person's vitamin D status and at the same time it was stated that "a single individual might be deemed deficient or sufficient, depending on the laboratory where the blood is tested" [17]. To overcome these problems, the need for standardization of 25(OH)D measurements was stated by many organizations and scientists.

In 2011, the Office of Dietary Supplements (ODS) of the NIH held a meeting on "international standardization of 25-hydroxyvitamin D concentration measurements in national health surveys" to discuss possible approaches to standardize these measurements in national health surveys. One outcome of this meeting was the Vitamin D Standardization Program (VDSP), which is conducted as a collaboration between NIH/ODS, the Centers for Disease Control and Prevention (CDC) and the National Institute of Standards and Technology (NIST) [18].

Author Manuscript

The present paper will "rehearse" the basic concept of standardization, in particular applied to 25(OH)D concentration measurements. Beyond the technical aspects of standardization, it will also deal with critical points that are essential for full success and sustainability.

The process for standardizing measurements of 25(OH)D concentration

The core component of standardization is the "establishment of metrological traceability". According to the International Vocabulary of Basic and General Terms in Metrology, traceability is defined as *"property of a measurement result whereby the result can be related to a reference through a documented unbroken chain of calibrations, each contributing to the measurement uncertainty"* [19].

In practice, the standardization process can be structured into 3 basic steps (see Figure 2). In the first step a so called reference basis or reference measurement system is established. Also the measurand and units for expression of measurement results are defined and the different hierarchical levels of materials and measurement procedures used for calibration are created. In the second step, assays are calibrated using the established calibration hierarchy. In this top down process a material is used for calibration of the measurement procedure at the same level; the latter is then used to assign a value to the material at the level below. This calibration process is described in greater detail in a comprehensive review [20]. As a result measurements of hierarchically lower order become traceable to the unit at the top and its realization in the primary calibrator within stated uncertainty constraints. In Table I, an overview of the elements of a Système International d'Unités (SI)-reference measurement system is presented. Note that traceability to the SI-unit is intended to be equivalent to accuracy (trueness and precision). However, establishing traceability does not ensure that the measurement uncertainty is adequate for a given clinical purpose or that there is absence of errors. This aspect will be dealt with in greater detail below.

While the top portion of the traceability chain is commonly established by dedicated groups and organizations, the bottom portion commonly is performed by the assay manufacturer and the laboratory measuring patient samples. Because the ultimate goal in clinical laboratory standardization is the trueness, precision and applicability of patient results, it is important to validate traceability and applicability of the measurement result at the end-user level. This is accomplished in the third step in the standardization process. Here the reference measuring system consisting of reference methods and materials is used to assess the trueness and applicability of the measurements performed on patient samples across measurement procedures.

It needs to be pointed out that reagents, calibrators and analytical methods change over time and with it, measurement accuracy and other important parameters can change. Therefore, it is important to perform the standardization process continuously to assure accurate and reliable measurements over time.

For clinical measurements of 25(OH)D in serum or plasma a reference measuring system has been put in place recently. It is represented in Figure 3. It defines the measurand as "amount of substance concentration of total 25(OH)D ("total" because of 2 components:

25(OH)D2 and 25(OH)D3), in serum/plasma". Because of this specific definition of the measurand, SI-units apply for expression of measurement results, i.e., "nmol/L". For realization of the SI-units, certified 25(OH)D2 and 25(OH)D3 primary calibration solutions are available, i.e., the Standard Reference Materials (SRMs) 2972 (ethanolic 25(OH)D2 and 25(OH)D3 solutions) from NIST [21]. These are used for direct calibration of ID-LC/ tandem MS reference measurement procedures (RMPs), as available today from NIST and the University of Ghent [22,23]. The SRMs and RMPs have been approved to be conform with International Organization for Standardization (ISO) and are listed in the database of the Joint Committee for Traceability in Laboratory Medicine [24].

NIST developed matrix-based reference materials (SRM 972, 4 concentrations, level 1 through 4), intended for traceability establishment or trueness assessment of lower order measurement procedures [25]. Some levels of SRM 927 were found to be non-commutable and thus not useful with immunoassays [26]. However, a new SRM 972a is under development and is currently being assessed for commutability. In addition, panels of single-donation sera, as described in Figure 3 can be used at the level of the working calibrator for combination with the master procedure. The use of a panel of native sera has several advantages: (i) it circumvents non-commutability problems typically encountered with materials that are processed or artificial to some extent (e.g., due to supplementation, because of non-human origin of the matrix material, etc.); (ii) it also allows assessing the intrinsic quality of a method (see below) in the validation process and finally using single-donor panels that are as similar as possible to regular patient samples providing valuable information about the usefulness of results of measurements with a particular method.

When implementing the described process, the following items require special consideration.

Definition of the measurand and related analytical items

As described above, the component in the measurand is "total 25(OH)D", comprising 25(OH)D2 and 25(OH)D3. Consequently, 25(OH)D measurement comprises two analytes, which has the following implications: the measurement procedures must either distinguish between the 2 components and quantitate them separately or measure their concentration together in a quantitative ("equimolar") manner. Procedures applying chromatographic separation and/or MS detection (LC/tandem MS) have the potential to meet both requirements. For procedures based on the immunochemistry principles, equimolar measurement requires that the quantitation is done in a manner that accurately reflects the concentration of both 25(OH)D2 and 25(OH)D3. Some immunoassays may use antibodies with different affinities to 25(OH)D2 and 25(OH)D3 and therefore may not measure 25(OH)D in an equimolar manner.

Accurate measurement of 25(OH)D2 and 25(OH)D3 concentrations implies that all measurement procedures have sufficient specificity in terms of chromatographic separation and detection, or antibody specificity. This means that measurements are not affected by potentially interfering compounds, e.g., the 3-epi form of 25(OH)D. Measurement procedures need to be able to distinguish between the 3-epi form and the measurand,

because the current definition of the measurand does not include the 3-epi forms of 25(OH)D.

The C-3 epimerization of the A-ring of 25(OH) D3 is indeed a common pathway for all major metabolites of vitamin D3 [27]. In 2006, Singh et al. demonstrated for the first time that 3-epi-25(OH)D3 was present in significant concentrations in serum from infants [28]. Recently, it has also been shown that in some adults 3-epi-25(OH) D3 is present in considerable amounts [29–32]. This observation initiated further investigations about the possible biological role of the 3-epi metabolites of vitamin D. Pending the outcome of these investigations, measurement of 3-epi metabolite concentrations might become relevant in patient care and public health. Recommendations regarding the measurement of concentration of the 3-epi form were made by other organizations [33].

Currently, some immunoassays claim that they do not capture the 3-epi form with their antibody [28], likewise, some routine LC/tandem MS methods perform the chromatographic separation that allows the separation of the 3-epi form from the analytes [31,34]. However, the 3-epi form is yet not commonly discriminated in all routine measurement procedures. Assessment of cross reactivity and interfering compounds can efficiently be performed through comparison studies with a RMP, which, per definition, is capable to measure the 25(OH)D2 and 25(OH)D3 separately with sufficient specificity.

The second implication is that measurement results expressed in "nmol/L" is the most advantageous unit for expression of measurement results. Because of differences in molecular weight, correct conversions of molar concentrations to weight will require separate conversion calculations for 25(OH)D2 and 25(OH)D3. While this can be achieved with methods that quantitate 25(OH)D2 and 25(OH)D3 separately (i.e., LC/tandem MS), such conversions cannot be made correctly with methods that cannot distinguish between both analytes (i.e., immunoassays). A simple calculation clarifies that nmol/L and ng/mL cannot be interchanged: 10 ng/ml 25(OH)D2 = 24.2 nmol/L, 10 ng/mL 25(OH) D3 = 25.0 nmol/L, thus 20 ng/mL total 25(OH)D can either be 48.4 nmol/L, or 50 nmol/L or 49.2 nmol/L dependent on whether it represents the concentration of 25(OH)D2, 25(OH)D3 or a mixture of both. For the purpose of standardizing measurements of 25(OH)D, it is planned to use "nmol/L" as units.

Finally, 25(OH)D2 can occur in the general population at concentrations that can be below or near the limit of quantitation for some measurement procedures. Therefore, differences in the ability to quantitate small concentrations of 25(OH)D2 can lead to measurement bias. Recommendations about specific limits of quantitation that are required for appropriate patient care and public health activities have not been formulated.

Use of the 25(OH)D reference measuring system for establishing traceability or validation of hierarchical lower methods

Through the use of a reference measuring system, hierarchically lower methods become traceable to the SI-unit and its realization in the primary calibrator. Assuring accurate measurements of patient samples requires careful calibration and choice of calibration materials.

A primary calibrator with defined amounts of 25(OH)D2 and 25(OH)D3 dissolved in ethanol is available from NIST (SRM 2972) [21]. While this material can be used by chromatographic methods as calibrators, it is known that the antigen-antibody reaction prevents direct calibration of immunological methods with this calibration solution. Therefore, serum-based materials are needed to calibrate immunological methods. The use of a panel of sera assigned with values by a RMP as working calibrator in combination with their master procedure is already used by the in-vitro diagnostic industry. In this way, immunoassays become indirectly traceable to the SRM 2972 and compliant with Conformité Européenne-marketing requirements [35].

Laboratory developed methods, such as HPLC-UV and in particular ID-LC/tandem MS can directly be calibrated with the SRM 2972, be it through direct use or after addition to a matrix-based solution (stripped serum, albumin- or buffer solution). By doing so, these methods establish metrological traceability as described in ISO 17511 [36]. However, as data from, for example, DEQAS and College of American Pathologists (CAP) show, measurement results between these measurement procedures are still highly variable [11,12,37]. This observed variability can have multiple sources e.g., some methods only use a very basic sample preparation, do very fast chromatographic separations with the risk for matrix effects and interference; others are not properly evaluated for ion suppression, interference from analogs or metabolites, or calibration stability. To control and minimize these sources of variability and assure accurate patient results, it is highly advantageous to use serum-based materials with values assigned by RMPs to calibrate HPLC-UV and ID-LC/tandem MS measurement procedures or verify calibration.

Calibration and trueness control using serum-based materials can be achieved using pooled serum-based materials and panels of single donor sera with values assigned by a RMP. When using pooled and otherwise modified materials, the commutability of the material has to be assessed first for the intended measurement procedure. Further, since pooled materials frequently are available at only few concentrations, dilutions of these materials are needed to appropriately establish calibration curves. Thus, commutability of dilutions of these serum materials needs to be assessed as well. Commutable reference materials from pooled sera are effective tools to calibrate measurement procedures or to assess trueness of measurements. However, frequently materials may not be commutable or commutability is unknown.

Alternatively, a set of single donor sera can directly be used for calibration and calibration verification without the need to assess commutability. Further, individual sera help identify and rectify effects related to specimen matrix that may not be detectable in pooled materials and thus assures that individual patient samples are accurate. E.g., a recent study found that concentrations of vitamin D binding protein could affect measurement results [9]. Such influence factors could be more effectively accounted for by using individual patient samples. Further, single donor serum samples can help identify interference and other sample specificity problems. Thus panels of single donor samples provide important information that is not easily obtainable with pooled materials. For these reasons, CDC will provide single donor sera with values assigned by a RMP to calibrate measurement procedures, and to verify calibration and assess calibration over time as part of its performance certification program. Pooled materials for assessing trueness and identifying

other measurement problems are available from proficiency testing providers such as the CAP's accuracy-based surveys [37].

Analytical performance goals

As already mentioned before, traceability of measurement does not warrant that the uncertainty is adequate for a given purpose or that there is absence of errors. To assure that measurement results for 25(OH)D are useful in patient care and public health, specifications for trueness and precision need to be defined for measurements performed with patient samples [38]. Approaches for deriving such specifications have been described [39]. It needs to be noted that general recommendations for bioanalytical methods are available [40,41]. However, these recommendations were not intended for measurement procedures used in patient care and therefore are not applicable, especially regarding specifications for trueness and precision.

Since traceability is established/verified with a reference measuring system involving materials and measurement procedures combined in a hierarchical structure, the specifications should be interrelated between the different levels. For serum/plasma 25(OH)D measurements, Stöckl et al. proposed a concept to derive specifications for trueness and precision of the reference measurement system [42]. Shortly described, the concept first derives specifications for hierarchically lower measurements used for patient samples and tailors the goals for hierarchically higher measurements and materials based on these initial goals. The following relationship was proposed: limit for the imprecision of a RMP, half the limit for a routine method; limit for the bias of a RMP, one third of the maximum bias for a routine measurement, and limit for the expanded uncertainty of the primary calibrator, one third of the allowable bias for a RMP. On this basis, numerical goals were derived according to 4 scientific models [39]. To finally end up with achievable goals, those numerical values were retained that hold the balance between desirable quality, state-of-the-art performance and certification capabilities. The recommended specifications are summarized in Table II.

Another analytical performance goal that seems relevant for 25(OH)D measurements is measurement of the limit of quantitation as discussed earlier. However, concepts and approaches to derive the appropriate limit of quantitation for 25(OH)D2 and 25(OH) D3 measurements have not been established yet.

Implementation of standardization

To effectively implement standardization of 25(OH)D an initial thorough assessment of the analytical performance of each measurement procedure is highly recommended. Such an assessment should be designed that it provides information on all aspects of the measurement procedure that could potentially affect measurement performance and reliability of patient results. For 25(OH)D measurements, the assessment of e.g., calibration consistency and reagent lot-to-lot variability in addition to trueness, precision and limit of quantitation appear relevant parameters. Another parameter that could be of relevance is the variability related to specimen matrix effects commonly referred to as 'sample-related effects'. These can best be judged after minimizing calibration bias through recalibration of

the data, minimizing of imprecision through use of means of replicate measurements on the same sample and presentation of data in a difference plot. The scatter of data remaining after minimizing bias and imprecision can be attributed to interfering compounds and general specimen matrix characteristics (i.e., lipid content, viscosity, optical density). Information from this initial assessment will help in effectively identify, address and overcome variability between measurement procedures and thus facilitate the implementation of standardization. Procedures for performing these initial assessments that could be adopted for 25(OH)D measurements have been developed and successfully applied for thyroid function tests [43–45]. In 2011, the NIH/ODS performed an interlaboratory comparison study of laboratories performing 25(OH)D measurements for national surveys together with CDC and its partners from the clinical laboratory community and assay manufactures [18]. The study design followed the protocols used for assessing assay performances of thyroid hormone tests [43–45] and for assessing commutability of high-sensitivity C-reactive protein [46,47]. The findings from this study will provide information that will greatly facilitate the implementation of standardization.

After the initial assessment of measurement performance, the measurement procedure can be recalibrated as needed and any other problems affecting measurement performance can be addressed. CDC is working with clinical laboratories and assay manufacturers addressing potential measurement problems. Measurement procedures that achieved the desired measurement performance criteria can enroll in a formal program administered by CDC, in which the participants are challenged 4 times per year using 10 single-donor patient samples per challenge that are measured in two independent measurements (Figure 4). Measurement results obtained from these 4 challenges will then be compared against pre-defined performance criteria such as those stated above. This program is part of CDC's Hormones Standardization Program [48]. Because reagents, tests and technologies change over time, it is important to assure standardization of measurements over long time periods. The standardization services provided by CDC are available to all laboratories, assay manufactures and organizations involved in assessing the performance of 25(OH)D measurements on an ongoing basis.

Conclusion

Today tools are available for establishing SI-traceability of 25(OH)D measurements by hierarchically lower methods. However, for a successful implementation of the described process, items such as analytical issues related to the definition of the measurand, analytical performance goals to warrant the extent of traceability commensurate with a method's intended use, sustainability of traceability over time, require additional attention. Other items still need to be clarified, such as specific limits of quantitation of a method that are required for appropriate patient care and public health activities.

Questions and Answers

J van den Ouweland, *Netherlands*

How do the NIST and Ghent methods compare when you analyse real patients' samples? I ask because I think the NIST method uses APCI with different transitions, so a different

methodology to your procedure. I believe there are some dangers with using APCI technology because it uses specific transitions with water loss. You also have the risk of deuterium scrambling.

Secondly, if yours is a reference procedure, why are you using deuterium-labelled internal standards and not ^{13}C?

L Thienpont

In the vitamin D standardisation programme, NIST and our laboratory analysed 50 samples in parallel and independent from each other. I have seen a preliminary comparison of the two methods and the results look very good. Where you can show agreement, there is proof of reliability of the procedure.

In answer to your second question, regarding deuterium versus ^{13}C, any difficulties depend on where the ^{2}H atoms are in the molecule, and whether they are in positions where they can be changed. I think in the case of the available ^{2}H-labelled 25(OH)D, there does not appear to be a problem.

Acknowledgments

Abbreviations

25(OH)D2	25-hydroxyvitamin D2
25(OH)D3	25-hydroxyvitamin D3
CDC	Centers for Disease Control and Prevention
CAP	College of American Pathologists
HPLC-UV	high performance liquid chromatography-ultraviolet detection
ISO	International Organization of Standardization
ID-LC/tandem MS	isotope dilution-liquid chromatography/tandem mass spectrometry
NIH	National Institute of Health
NIST	National Institute of Standards and Technology
ODS	Office of Dietary Supplements
RMP	reference measurement procedure

SRM	Standard Reference Material
SI	Système International d'Unités
DEQAS	Vitamin D External Quality Assessment Scheme
VDSP	Vitamin D Standardization Program

References

1. U.S. Department of Health and Human Services. [Accessed February 2012] Agency for Healthcare Research and Quality. Vitamin D and calcium: a systematic review of health outcomes. http://www.ahrq.gov/clinic/tp/vitadcaltp.htm

2. Holick MF. Vitamin D deficiency. N Engl J Med. 2007; 357:266–81. [PubMed: 17634462]

3. Cavalier E, Delanaye P, Chapelle JP, Souberbielle JC. Vitamin D: current status and perspectives. Clin Chem Lab Med. 2009; 47:120–7. [PubMed: 19099529]

4. Audran M, Briot K. Critical reappraisal of vitamin D deficiency. Joint Bone Spine. 2010; 77:115–9. [PubMed: 20097593]

5. Wallace AM, Gibson S, de la Hunty A, Lamberg-Allardt C, Ashwell M. Measurement of 25-hydroxyvitamin D in the clinical laboratory: current procedures, performance characteristics and limitations. Steroids. 2010; 75:477–88. [PubMed: 20188118]

6. Price, CP., Christenson, RH., editors. Evidence-Based Laboratory Medicine: Principles, Practice and Outcomes. 2. Washington DC: AACC Press; 2007. p. 29-31.

7. Moon HW, Cho JH, Hur M, Song J, Oh GY, Park CM, Yun YM, Kim JQ. Comparison of four current 25-hydroxyvitamin D assays. Clin Biochem. 2012; 45:326–330. [PubMed: 22244986]

8. Farrell CJL, Martin S, McWhinney B, Straub I, Williams P, Herrmann M. State-of-the-art vitamin D assays: a comparison of automated immunoassays with liquid chromatography-tandem mass spectrometry methods. Clin Chem. 2012; 58:531–42. [PubMed: 22230812]

9. Heijboer AC, Blankenstein MA, Kema IP, Buijs MM. Accuracy of 6 routine 25-hydroxyvitamin D assays; influence of vitamin D binding protein concentration. Clin Chem. 2012; 58:543–548. [PubMed: 22247500]

10. Binkley N, Krueger DC, Morgan S, Wiebe D. Current status of clinical 25-hydroxyvitamin D measurement: an assessment of between-laboratory agreement. Clin Chim Acta. 2010; 411:1976–82. [PubMed: 20713030]

11. Carter GD, Carter R, Jones J, Berry J. How accurate are assays for 25-hydroxyvitamin D? Data from the international for Vitamin D External Quality Assessment Scheme. Clin Chem. 2004; 50:2195–7. [PubMed: 15375018]

12. Carter GD, Berry JL, Gunter E, Jones G, Jones JC, Makin HLJ, et al. Proficiency testing of 25-hydroxyvitamin D (25-OHD) assays. J Steroid Biochem Mol Biol. 2010; 121:176–9. [PubMed: 20302938]

13. Yetley EA, Pfeiffer CM, Schleicher RL, Phinney KW, Lacher DA, Christakos S, et al. Vitamin D roundtable on the NHANES monitoring of serum 25(OH)D: assay challenges and options for resolving them. NHANES monitoring of serum 25-hydroxyvitamin D: a roundtable summary. J Nutr. 2010; 140:2030S–45S. [PubMed: 20881084]

14. Binkley N, Krueger D, Cowgill CS, Plum L, Lake E, Hansen KE, et al. Assay variation confounds the diagnosis of hypovitaminosis D: a call for standardization. J Clin Endocrinol Metab. 2004; 89:3152–7. [PubMed: 15240586]

15. Barake M, Daher RT, Salti I, Cortas NK, Al-Shaar L, Habib RH, et al. 25-Hydroxyvitamin D assay variations and impact on clinical decision making. J Clin Endocrinol Metab. Jan 11.2012 [Epub ahead of print] ad. doi: 10.1210/jc.2011-2584

16. Nowak M, Harrison SL, Buettner PG, Kimlin M, Porter D, Kennedy L, et al. Vitamin D status of adults from tropical Australia determined using two different laboratory assays: implications for public health messages. Photochem Photobiol. 2011; 87:935–43. [PubMed: 21534976]

17. IOM (Institute of Medicine). Dietary Reference Intakes for Calcium and Vitamin D. Washington, DC: The National Academies Press; 2011. http://www.iom.edu/Reports/2010/Dietary-Reference-Intakes-for-Calcium-and-Vitamin-D.aspx [Accessed February 2012]

18. Notice of Vitamin D Standardization Program. [Accessed February 2012] A notice by the National Institutes of Health on 03/02/2011. http://www.federalregister.gov/articles/2011/03/02/2011-4603/notice-of-vitamin-d-standardization-program

19. JCGM 200:2008. International vocabulary of metrology – Basic and general concepts and associated terms (VIM). Geneva: International Organization for Standardization; 2007. http://www.bipm.org/en/publications/guides/vim.html [Accessed February 2012]

20. Vesper HW, Thienpont LM. Traceability in laboratory medicine. Clin Chem. 2009; 55:1067–75. [PubMed: 19359540]

21. Certificate of analysis, Standard Reference Material 2972: 25-hydroxyvitamin D2 and D3 calibration solutions. Gaithersburg, MD: Standard Reference Materials Program, NIST; 2009.

22. Tai SS, Bedner M, Phinney KW. Development of a candidate reference measurement procedure for the determination of 25-hydroxyvitamin D3 and 25-hydroxyvitamin D2 in human serum using isotope-dilution liquid chromatography-tandem mass spectrometry. Anal Chem. 2010; 82:1942–8. [PubMed: 20136128]

23. Stepman HCM, Vanderroost A, Van Uytfanghe K, Thienpont LM. Candidate reference measurement procedures for serum 25-hydroxyvitamin D3 and 25-hydroxyvitamin D2 using isotope dilution-liquid chromatography-tandem mass spectrometry. Clin Chem. 2011; 57:441–8. [PubMed: 21248072]

24. [Accessed February 2012] Joint Committee for Traceability in Laboratory Medicine database: Laboratory medicine and in-vitro diagnostics. http://www.bipm.org/jctlm/

25. Certificate of analysis, standard reference material 972: Vitamin D in human serum. Gaithersburg, MD: Standard Reference Materials Program, NIST; 2009.

26. Horst RL. Exogenous versus endogenous recovery of 25-hydroxyvitamins D2 and D3 in human samples using high-performance liquid chromatography and the DiaSorin LIAISON Total-D assay. J Steroid Biochem Mol Biol. 2010; 121:180–2. [PubMed: 20214981]

27. Kamao M, Tatematsu S, Hatakeyama S, Sakaki T, Sawada N, Inouye K, et al. C-3 epimerization of vitamin D3 metabolites and further metabolism of C-3 epimers: 25-hydroxyvitamin D3 is metabolized to 3-epi-25-hydroxyvitamin D3 and subsequently metabolized through C-1alpha or C-24 hydroxylation. J Biol Chem. 2004; 279:15897–907. [PubMed: 14757768]

28. Singh RJ, Taylor RL, Reddy GS, Grebe SK. C-3 epimers can account for a significant proportion of total circulating 25-hydroxyvitamin D in infants, complicating accurate measurement and interpretation of vitamin D status. J Clin Endocrinol Metab. 2006; 91:3055–61. [PubMed: 16720650]

29. Stepman HCM, Vanderroost A, Stöckl D, Thienpont LM. Full-scan mass spectral evidence for 3-epi-25-hydroxyvitamin D3 in serum of infants and adults. Clin Chem Lab Med. 2011; 49:253–6. [PubMed: 21143012]

30. Lensmeyer G, Poquette M, Wiebe D, Binkley N. The C-3 epimer of 25-hydroxyvitamin D(3) is present in adult serum. J Clin Endocrinol Metab. 2012; 97:163–8. [PubMed: 22013102]

31. van den Ouweland JM, Beijers AM, van Daal H. Fast separation of 25-hydroxyvitamin D3 from 3-epi-25-hydroxyvitamin D3 in human serum by liquid chromatography-tandem mass spectrometry: variable prevalence of 3-epi-25-hydroxyvitamin D3 in infants, children, and adults. Clin Chem. 2011; 57:1618–9. [PubMed: 21914787]

32. Strathmann FG, Sadilkova K, Laha TJ, Lesourd SE, Bornhorst JA, Hoofnagle AN, et al. 3-epi-25 hydroxyvitamin D concentrations are not correlated with age in a cohort of infants and adults. Clin Chim Acta. 2012; 413:203–6. [PubMed: 21983164]

33. de la Hunty A, Wallace AM, Gibson S, Viljakainen H, Lamberg-Allardt C, Ashwell M. UK Food Standards Agency Workshop Consensus Report: the choice of method for measuring 25-hydroxyvitamin D to estimate vitamin D status for the UK National Diet and Nutrition Survey. Br J Nutr. 2010; 104:612–9. [PubMed: 20712915]

34. Schleicher RL, Encisco SE, Chaudhary-Webb M, Paliakov E, McCoy LF, Pfeiffer CM. Isotope dilution ultra performance liquid chromatography-tandem mass spectrometry method for

simultaneous measurement of 25-hydroxyvitamin D2, 25-hydroxyvitamin D3 and 3-epi-25-hydroxyvitamin D3 in human serum. Clin Chim Acta. 2011; 412:1594–9. [PubMed: 21601563]

35. 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.

36. In vitro diagnostic medical devices - Measurement of quantities in biological samples - Metrological traceability of values assigned to calibrators and control materials. ISO 17511. 2003

37. Paxton, A. [Accessed February 2012] Accuracy-based surveys carve higher QA profile. CAP Today. Oct. 2010 http://www.cap.org/apps/cap.portal?_nfpb=true&cntvwrPtlt_actionOverride=/portlets/contentViewer/show&_windowLabel=cntvwrPtlt&cntvwrPtlt%7bactionForm.contentReference%7d=cap_today/1010/1010e_accuracy_based_surveys.html&_state=maximized&_pageLabel=cntvwr

38. Stöckl D, Baadenhuijsen H, Fraser CG, Libeer JC, Petersen PH, Ricós 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. Erratum in: Eur J Clin Chem Clin Biochem 1995;33. [PubMed: 7605829]

39. Petersen PH, Fraser CG, Kallner A, Kenny D. Strategies to set global quality specifications in laboratory medicine. Scand J Clin Lab Invest. 1999; 59:475–585. [PubMed: 10667681]

40. US FDA Technical Review Guide: Validation of Chromatographic Methods. Center for Drug Evaluation and Research (CDER); Rockville, MD: 1993.

41. Shah VP, Midha KK, Dighe S, McGilveray IJ, Skelly JP, Yacobi A, et al. Analytical methods validation: bioavailability, bioequivalence and pharmacokinetic studies. Conference report. J Pharmac Sci. 1992; 81:309–12.

42. Stöckl D, Sluss PM, Thienpont LM. Specifications for trueness and precision of a reference measurement system for serum/plasma 25-hydroxyvitamin D analysis. Clin Chim Acta. 2009; 408:8–13. [PubMed: 19563791]

43. Thienpont LM, Van Uytfanghe K, Beastall G, Faix JD, Ieiri T, Miller WG, et al. for the IFCC Working Group on Standardization of Thyroid Function Tests. Report of the IFCC Working Group for Standardization of Thyroid Function Tests, part 1: thyroid-stimulating hormone. Clin Chem. 2010; 56:902–11. [PubMed: 20395624]

44. Thienpont LM, Van Uytfanghe K, Beastall G, Faix JD, Ieiri T, Miller WG, et al. for the IFCC Working Group on Standardization of Thyroid Function Tests. Report of the IFCC Working Group for Standardization of Thyroid Function Tests, part 2: free thyroxine and free triiodothyronine. Clin Chem. 2010; 56:912–20. [PubMed: 20395623]

45. Thienpont LM, Van Uytfanghe K, Beastall G, Faix JD, Ieiri T, Miller WG, et al. for the IFCC Working Group on Standardization of Thyroid Function Tests. Report of the IFCC Working Group for Standardization of Thyroid Function Tests, part 3: total thyroxine and total triiodothyronine. Clin Chem. 2010; 56:921–9. [PubMed: 20395622]

46. Kimberly MM, Vesper HW, Caudill SP, Cooper GR, Rifai N, Dati F, et al. Standardization of immunoassays for measurement of high-sensitivity C-reactive protein. Phase I: evaluation of secondary reference materials. Clin Chem. 2003; 49:611–6. [PubMed: 12651814]

47. Kimberly MM, Caudill SP, Vesper HW, Monsell EA, Miller WG, Rej R, et al. Standardization of high-sensitivity immunoassays for measurement of C-reactive protein; II: Two approaches for assessing commutability of a reference material. Clin Chem. 2009; 55:342–50. [PubMed: 19074518]

48. Vesper HW, Botelho JC. Standardization of testosterone measurements in humans. J Steroid Biochem Mol Biol. 2010; 121:513–9. [PubMed: 20302935]
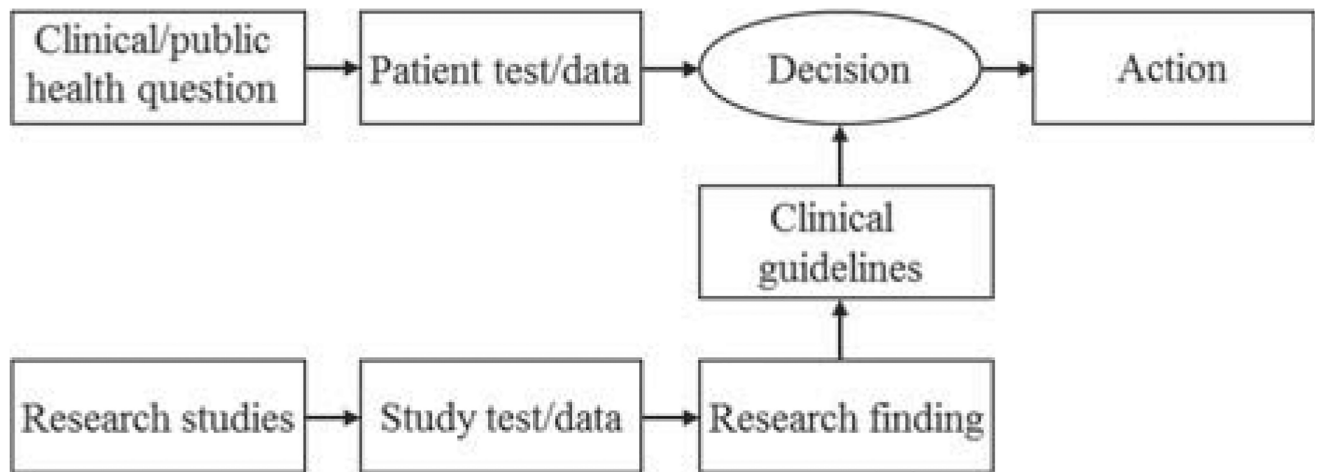
Figure 1.

The development and implementation of evidence-based clinical guidelines depends on the comparability of tests performed in research and patient care. Adopted (and modified) from CP Price and RH Christenson (eds): [6].
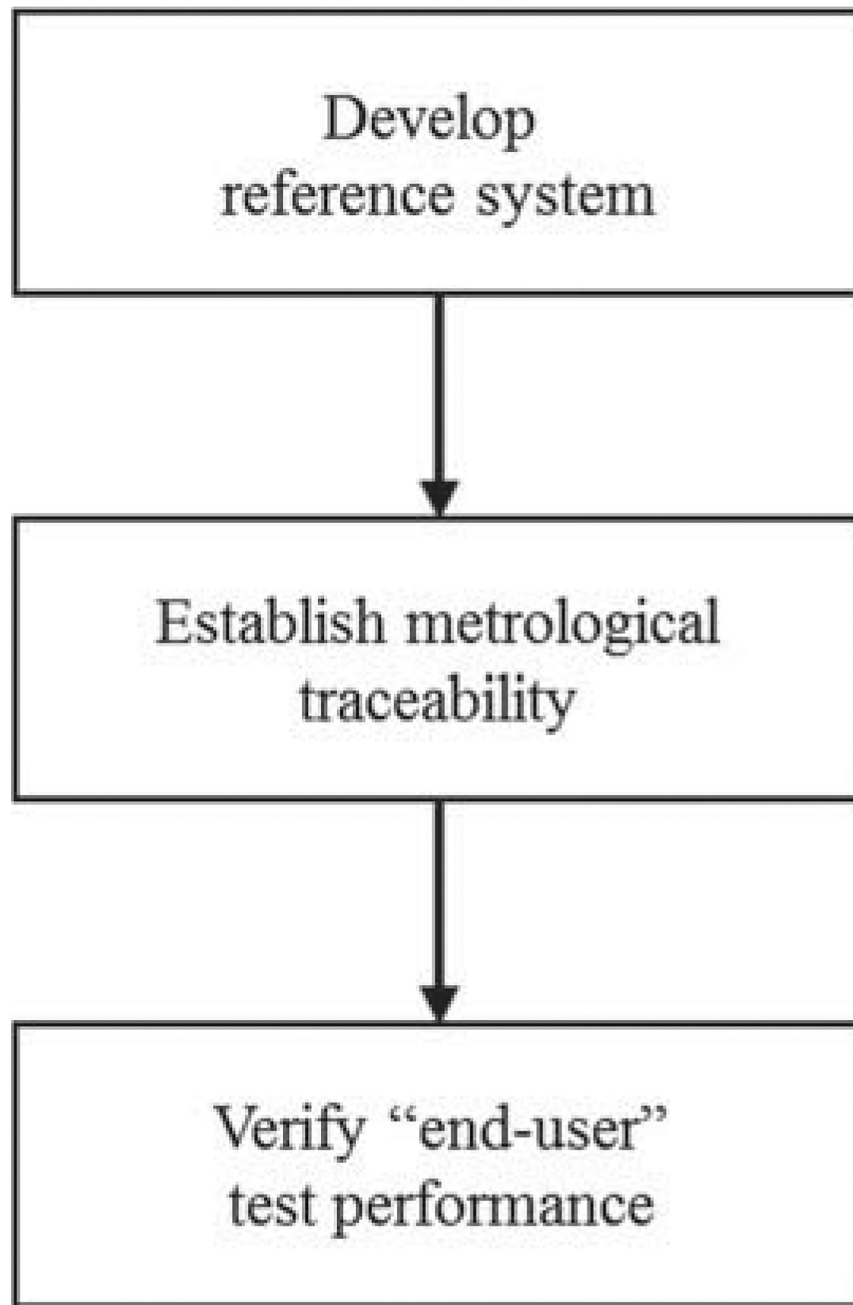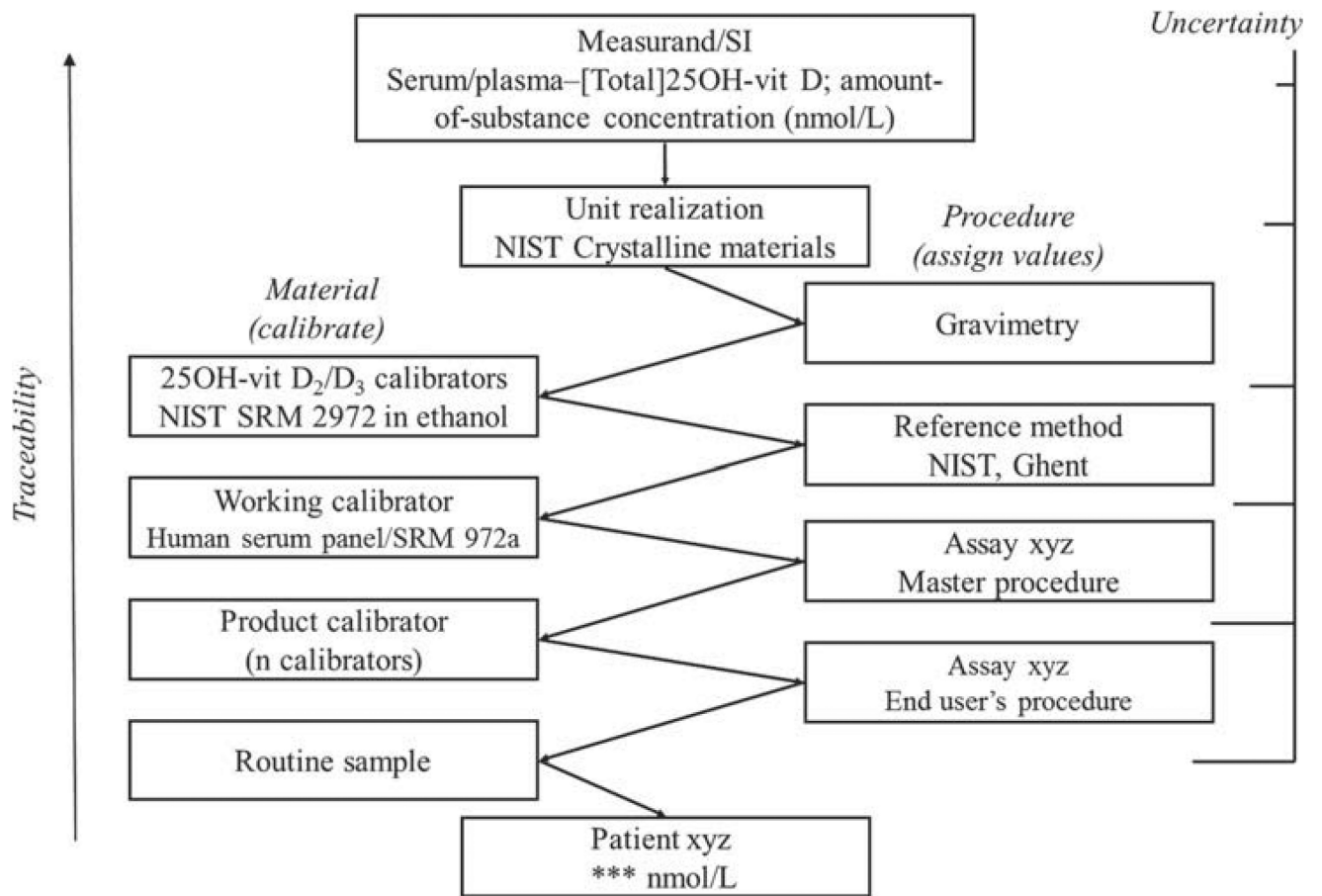
Figure 2.
Stevps to laboratory standardization.

Figure 3.
Outline of the reference measurement system for serum/plasma 25(OH)D.
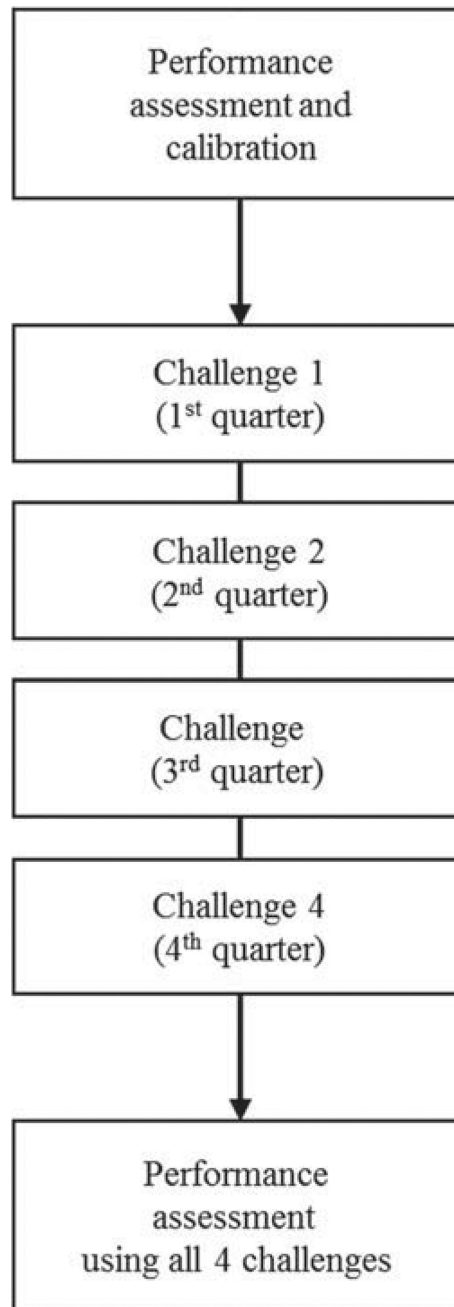
Figure 4.
Procedure for standardization and performance certification.

Table I

Overview of the elements of an SI-reference measurement system, with indication of the related tasks and responsibilities.

Element	Organization	Task
SI-Unit	Conférence Générale des Poids et Mesures (CGPM)	Establishment of a coherent system of units (mol)
Component (analyte)	IFCC	Definition of the relevant component
Reference material	National Metrology Institutes; IRMM, NIST	Realization of SI units. Production and certification of reference material
Reference measurement procedure	Reference laboratory or other competent analytical laboratory	Development and validation of reference measurement procedures
Reference laboratory	No representative organization ("Networks")	Application of the reference measurement procedures

Table II

Recommendation for maximum imprecision (CV), bias and expanded uncertainty (U) of measurements and primary calibrators in a 25(OH)D reference measurement system [41].

Routine measurements	CV: 10 %	Bias: 5 %
Reference measurements	CV: 5 %	Bias: 1.7 %
Primary calibrators	U: 0.6 %	