Analytical and Bioanalytical Chemistry
https://doi.org/10.1007/s00216-024-05143-w

CRITICAL REVIEW

Evolution and impact of Standard Reference Materials (SRMs) for determining vitamin D metabolites

Stephen A. Wise[1] · **Adam J. Kuszak**[1] · **Johanna E. Camara**[2]

Received: 6 December 2023 / Revised: 8 January 2024 / Accepted: 10 January 2024

Abstract

The National Institute of Standards and Technology (NIST), in collaboration with the National Institutes of Health, Office of Dietary Supplements (NIH ODS), introduced the first Standard Reference Material® (SRM) for determining vitamin D metabolites in 2009 motivated by significant concerns about the comparability and accuracy of different assays to assess vitamin D status. After 14 years, a suite of five serum matrix SRMs and three calibration solution SRMs are available. Values were also assigned for vitamin D metabolites in five additional SRMs intended primarily to support measurements of other clinical diagnostic markers. Both the SRMs and the certification approach have evolved from significant exogenous serum content to primarily endogenous content and from value assignment by combining the results of multiple analytical methods to the use of measurements exclusively from reference measurement procedures (RMPs). The impact of the availability of these SRMs can be assessed by both the distribution information (sales) and by reports in the scientific literature describing their use for method validation, quality control, and research. In this review, we describe the development of these SRMs, the evolution in design and value assignment, the expansion of information reported, and SRM use in validating analytical methods and providing quality assurance within the vitamin D measurement community.

Keywords 25-Hydroxyvitamin D_2 · 25-Hydroxyvitamin D_3 · Total 25-hydroxyvitamin D · 3-*epi*-25-Hydroxyvitamin D_3 · Vitamin D binding protein (VDBP) · Reference measurement procedure (RMP)

Introduction

For over 40 years, the National Institute of Standards and Technology (NIST), formerly the National Bureau of Standards (NBS), has developed Standard Reference Materials® (SRMs) for clinically relevant marker compounds in human serum. SRMs are certified reference materials (CRMs) produced by NIST. In 1980, NBS issued the first human serum matrix SRM for clinical diagnostic markers, SRM 909 Human Serum, with certified values for cholesterol, creatinine, glucose, urea, uric acid, and inorganic electrolytes (Ca, Li, Mg, K, Na, and Cl) [1]. During the next four decades, SRMs were issued for additional organic clinical marker compounds including cortisol, estradiol, homocysteine, thyroid hormones, and testosterone. SRMs specifically for cholesterol, glucose, and creatinine gained the most attention with multiple renewals and improvements for each material over time including a transition from a storage convenient freeze-dried serum matrix SRM to fresh frozen serum matrices like actual clinical samples. In 2000, NIST published an economic impact study [2] for the NIST cholesterol SRMs during their first two decades, which calculated an estimated benefit-to-cost ratio of 4.5 and concluded that the SRMs "played an important economic role in support of a national effort to monitor, measure, and control cholesterol levels, thereby contributing to reduced levels of cardiovascular disease" [2].

In 2009, NIST issued the first SRM for the measurement of clinically relevant diagnostic markers for vitamin D status, SRM 972 Vitamin D in Human Serum. In many respects, the development of SRM 972 was part of a response to a "perfect storm" scenario for the need for a

Published in the topical collection *Advances in (Bio-)Analytical Chemistry: Reviews and Trends Collection 2024*.

✉ Stephen A. Wise
stephen.wise@nih.gov

[1] Office of Dietary Supplements (ODS), National Institutes of Health (NIH), Bethesda, MD 20817, USA

[2] Chemical Sciences Division, National Institute of Standards and Technology (NIST), Gaithersburg, MD 20899, USA

Published online: 18 January 2024

clinical diagnostic marker SRM [1], which began several years earlier when the National Institutes of Health, Office of Dietary Supplements (NIH ODS) provided significant funding to NIST to support activities to improve the comparability and quality of measurements to assess vitamin D status. NIH ODS established and coordinated the Vitamin D Standardization Program (VDSP) [3, 4] to promote and support improved comparability of analytical measurements for total serum 25-hydroxyvitamin D [25(OH)D], which is defined as the sum of 25-hydroxvitamin D_2 [25(OH)D_2] and 25-hydroxvitamin D_3 [25(OH)D_3] and is the current clinical marker to assess vitamin D status. The VDSP catalyzed the implementation of a measurement system to determine total 25(OH)D that included development of (1) reference measurement procedures (RMPs), (2) SRMs, and (3) an accuracy-based quality assurance program, and eventually support to transition the long-running international Vitamin D External Quality Assessment Scheme (DEQAS) to an accuracy-based program. This critical review will focus on the development, evolution, and impact of the SRMs developed as part of this measurement system to improve the accuracy and comparability of measurements used to assess vitamin D status.

Assessment of vitamin D status

The major metabolites of vitamin D_2 (ergocalciferol) and vitamin D_3 (cholecalciferol) are 25(OH)D_2 and 25(OH)D_3, respectively, with 25(OH)D_3 as the predominant metabolite unless supplementation with ergocalciferol has occurred. Epimers of 25(OH)D_2 and 25(OH)D_3 are present in human serum but with only the 3-*epi*-25(OH)D_3 significant at 3 to 6% of the 25(OH)D_3 concentration [5, 6]. The primary clinical marker for assessing vitamin D status is total 25(OH)D, which is defined as the sum of 25(OH)D_2 and 25(OH)D_3 excluding the 3-*epi*-25(OH)D_3. Circulating 25(OH)D is primarily bound to proteins, with 85 to 90% strongly bound to vitamin D binding protein (VDBP), which is a specific transport protein for vitamin D; 10 to 15% is loosely bound to serum albumin; and only 0.02 to 0.04% is available in a free, unbound form [7–9]. Total 25(OH)D (both free and protein bound) is the measurand which is traceable to the International System of Units (SI) through measurements of each of the two metabolites.

Analytical methods for total 25(OH)D

There are currently two primary analytical methodologies for the determination of 25(OH)D: (1) ligand binding assays, mainly immunoassays, that provide a response for total 25(OH)D and (2) liquid chromatography with tandem mass spectrometry detection (LC–MS/MS) that provides individual measurements of 25(OH)D_2 and 25(OH)D_3 that are summed for a total 25(OH)D result. There have been several reviews describing the historical development, evolution, and current status of analytical methods for the determination of 25(OH)D, with particular emphasis on LC–MS/MS [10–15]. These reviews highlight methods for total 25(OH)D with early assays based on competitive protein binding, radioimmunoassay (RIA), and eventually a transition to fully automated immunoassays and the chromatography-based methods and their eventual shift exclusively to LC–MS/MS. The characteristics of 12 commercial competitive binding assays and immunoassays with comments on their performance are provided in a table by Alteri et al. [12]. There are currently over 30 commercial assays for 25(OH)D with DiaSorin, Roche, Siemens, IDS-iSYS, and Abbott identified as the most frequently represented ligand binding assays reported in recent DEQAS exercises [16] and with 19 non-LC–MS/MS assays certified by the CDC Vitamin D Standardization-Certification Program [17]. Volmer and coworkers [10, 18–20] reviewed the evolution of LC–MS/MS as the "gold standard" for determining 25(OH)D_2 and 25(OH)D_3 and for profiling other vitamin D metabolites. Volmer et al. [19] provide a table summarizing 32 papers published between 2001 and 2013 reporting LC–MS/MS methods for 25(OH)D_2 and 25(OH)D_3, often including other vitamin D metabolites. Alexandridou et al. [10] followed up with a continuation of the summary of LC–MS/MS methods with a table containing 16 LC–MS/MS methods through 2020. These LC–MS/MS methods generally are based on the use of isotopically labeled internal standards for quantification which is denoted as isotope dilution (ID). Müller and Volmer [20] included a table with the most important recent studies quantifying multiple vitamin D species (3 to 8 species) including characteristics such as internal standards used, derivatization, LC column, *m*/*z* transitions, linear range, and precision.

A critical focus of LC–MS/MS method development in the first decade of the emergence of LC–MS/MS as the gold standard was the need to chromatographically separate 3-*epi*-25(OH)D_3 from 25(OH)D_3. Initially, significant amounts of 3-*epi*-25(OH)D_3 were observed only in infants [21]; however, 3-*epi*-25(OH)D_3 is now known to be present in adult human serum typically at 3 to 6% of the content of 25(OH)D_3 [5, 6, 22, 23]. Because both analytes have the same molecular mass and recorded mass transitions, LC–MS/MS measurements would be positively biased by this amount if the epimer is not chromatographically separated from the 25(OH)D_3 [22]. Early LC–MS/MS methods relied on the use of the popular C_{18} stationary phase, which could not separate the epimer. For the majority (21 of 32) of the LC–MS/MS methods from 2001 to 2013 reviewed by Volmer et al. [19], the epimer was not separated. In 2006, Lensmeyer et al. [24] reported an LC method using a cyanopropyl (CN) stationary phase; however, only later did

Lensmeyer et al. [6] recognize that the CN column was also capable of separating the epimer. In 2009, a new stationary phase, pentafluorophenyl (PFP), appeared for the separation of vitamin D metabolites including the 3-epimer [25–28]. Currently, the PFP phase is the most widely used stationary phase providing a separation of the 25(OH)D_3 and 3-*epi*-25(OH)D_3. In a 2022 interlaboratory comparison of 25(OH)D assays including 15 custom LC–MS/MS assays, 6 methods did not separate the epimer; however, currently there is no valid technical reason that an LC–MS/MS should not provide this separation [22].

Controversies with assays for 25(OH)D

The U.S. National Health and Nutrition Examination Study (NHANES III) first included serum 25(OH)D measurements as part of the samples collected in 1988–1994 and they have continued since 2000. Measurements for 25(OH)D were performed in NHANES III (1988–1994) and NHANES 2000–2004 using the DiaSorin RIA kit. Looker et al. [29] published a comparison of the 25(OH)D results from the two studies that "suggests that a decline in measured vitamin D status may have occurred in the population over the past 10 – 15 y….." [29]. However, after adjustments for assay differences, the authors concluded that "….most of the observed difference in serum 25(OH)D between NHANES III and NHANES 2000–2004 appears to have been an artifact of assay changes rather than an actual decline in serum 25(OH)D concentrations" [29]. Based on differences in assay results in NHANES, the NIH ODS funded NIST in 2007 to develop RMPs and SRMs to improve the comparability of measurements for total serum 25(OH)D.

Before the RMPs and an SRM were available, additional concerns about the reliability of 25(OH)D measurements were raised after a report that a large commercial testing laboratory using an LC–MS/MS assay had admitted that some erroneous results had been reported because of problems with calibration and a lack of following proper procedures [30]. Just as the first SRM was released, NIH ODS convened a roundtable "NHANES Monitoring of Serum 25(OH)D: Assay Challenges and Options for Resolving Them" [31] to discuss the following: (1) options for addressing 25(OH)D assay fluctuations in NHANES, (2) approaches for transitioning from the RIA used between 1988 and 2006 to LC–MS/MS for future NHANES studies, and (3) approaches for incorporating the new SRM 972 into NHANES. In an editorial [32], Graham D. Carter, DEQAS organizer, opined on the unfortunate event with the commercial testing laboratory, the potential negative impact for LC–MS/MS methods, the need for quality assurance, and the potential of the soon-to-be available RMPs and SRM to provide a way to standardize 25(OH)D measurements.

Development of reference measurement procedures (RMPs) for vitamin D metabolites

NIST developed the first RMPs for the determination of 25(OH)D_2 and 25(OH)D_3 [33]. Within the clinical chemistry community, an RMP is "accepted as providing measurement results fit for their intended use in assessing measurement trueness of measured quantity values obtained from other measurement procedures for quantities of the same kind, in calibration, or in characterizing reference materials" [34, 35]. In practice, an RMP is a higher order method based on specified criteria [36, 37] and recognized by the Joint Committee on Traceability in Laboratory Medicine (JCTLM). Typically, an RMP for a small, well-defined molecule clinical marker (i.e., Type A analyte as described by Panteghini [34]) would typically be based on using isotopically labeled analogues of the marker as internal standards in ID gas chromatography (GC)- or LC–MS/MS methods. When used with a traceable primary calibrator, ID MS/MS is considered a primary ratio measurement method (i.e., measuring the value of a ratio of the unknown to a standard of the same quantity [38]) providing measurement results in complex matrices that are traceable to the SI.

In 2010, Tai et al. [33] published details of the NIST RMPs for 25(OH)D_2 and 25(OH)D_3. The University of Ghent followed in 2011 [39] with the second RMP, and the CDC published the third RMP in 2015 [40] (see Table 1 for details and comparison of the three RMPs). For the development of the RMPs, NIST recognized the need to chromatographically separate the 3-*epi*-25(OH)D_3 from the 25(OH)D_3 and was guided by the paper of Lensmeyer et al. [24] using a CN stationary phase rather than the commonly used C_{18} phase (Fig. 1A). The NIST, Ghent, and CDC RMPs are all based on ID LC–MS/MS using LC columns and conditions that separate the 25(OH)D_3 and 3-*epi*-25(OH)D_3 and using isotopically labeled analogs of 25(OH)D_2 and 25(OH)D_3 for quantification (Table 1). The Ghent and CDC RMPs use shorter LC columns with smaller diameter particles to provide a separation in 12 min compared with the NIST RMP separation in 30 min (Fig. 1B). Another critical element in the development of the RMPs was the evaluation of potential interfering compounds, particularly structural analogues of 25(OH)D_2 and 25(OH)D_3 including the 3-epimers. While NIST only reported testing for interference from 3-*epi*-25(OH)D_2 and 3-*epi*-25(OH)D_3 [33], Ghent and CDC evaluated 10 and 8 related compounds, respectively, including dihydroxyvitamin D isomers [39, 40].

Because 25(OH)D in serum is bound to proteins, it must be liberated for accurate measurement. Sample preparation for the NIST RMPs consists of addition of ethanolic

Table 1 Reference measurement procedures (RMPs) for the determination of $25(OH)D_2$ and $25(OH)D_3$

RMP (date)	Sample preparation	LC column/separation	Internal standards	Calibrants	MS/MS transitions	LoD	Precision
NIST (2010) [33]	Serum sample amount: 2 g; ethanolic IS solution added to serum; equilibrated at RT for 1 h; pH adjusted with carbonate buffer; extracted (2x) with 8 mL of hexane/ethyl acetate (50:50 volume fraction); shake 10 min (2nd extraction 3 min); combined extracts dried and reconstituted in 120 µL of MeOH	Zorbax SB-CN (250 × 4.6 mm, 5 µm); isocratic 34/66 H_2O/MeOH at 30 °C; separation of $25(OH)D_2$, $25(OH)D_3$, and epimers in 40 min Later Modifications: Ascentis Express F5 (pentafluorophenyl, PFP) column (150 × 4.6 mm, 2.7 µm) isocratic at 27:73 H_2O/MeOH at 0.75 mL/min; separation in 30 min (+ 10 min wash and 12 min re-equilibration for run time 52 min); Zorbax SB-CN column (150 × 4.6 mm, 3.5 µm) at 0.75 mL/min; separation in 30 min (+ 10 min wash and 12 min re-equilibration for run time 52 min)	$25(OH)D_3$-d_3 and $25(OH)D_2$-d_3; match IS to analyte in 1:1 ratio based on preliminary analysis of samples Modification: $25(OH)D_3$-d_6, 3-*epi*-$25(OH)D_3$-d_3, $25(OH)D_2$-d_3	$25(OH)D_2$ and $25(OH)D_3$ used as primary standards assessed for purity at NIST using multiple techniques; 6 calibrants gravimetrically prepared from 3 working solutions; IS solution added for 1:1 isotope ratio (±35%) (labeled to unlabeled) Modifications: SRM 2972a used to prepare calibrants fort $25(OH)D_2$	MS/MS: positive atmospheric pressure chemical ionization (APCI); *m/z* 401 → 383 for $25(OH)D_3$ and *m/z* 404 → 386 for $25(OH)D_3$-d_3; *m/z* 413 → 395 for $25(OH)D_2$ and *m/z* 416 → 398 for $25(OH)D_2$-d_3 Modification: Three runs (1) CN column—*m/z* 401 → 383 for $25(OH)D_3$ and *m/z* 407 → 389 for $25(OH)D_3$-d_6; (2) PFP column—*m/z* 413 → 395 for $25(OH)D_2$ and *m/z* 416 → 398 for $25(OH)D_2$-d_3, and (3) CN column—*m/z* 401 → 383 for 3-*epi*-$25(OH)D_3$ and *m/z* 407 → 389 for 3-*epi*-$25(OH)D_3$-d_3	LoD for 25(OH)D was 0.015 ng/g (40 pg on column) LoQ = 0.5 ng/g or 1.2 nmol/L for each analyte	CVs = 0.2 to 0.6% for 25(OH)D > 1 ng/g and 2% for < 1 ng/g
Univ. of Ghent (2011) [39]	Serum sample amount: 250 µL (500 µL max); serum diluted to 1 mL with 0.9% NaCl solution; IS solution added; mixture alkalinized, extracted with *n*-hexane, and fractionated with Sephadex LH-20 chromatography	2-dimensional LC system: (1) BEH300 C4 (50 × 2.1 mm) (2) BEHC18 (50 × 2.1 mm, 1.7 µm) for $25(OH)D_2$ and Zorbax SB-CN (250 × 2.1 mm, 5 µm, 80Å) for $25(OH)D_3$; separation of $25(OH)D_2$, $25(OH)D_3$, and epimers within 12 min	$25(OH)D_3$-d_6 and $25(OH)D_2$-d_6; match IS to analyte in 1:1 ratio	3 working solutions with exact values assigned by comparison to SRM 2972; 1-point calibration at 1:1 Isotope ratio (±25%)	MS/MS: positive electrospray mode: *m/z* 401.3 → 159.3 [$25(OH)D_3$ and 3-*epi*-$25(OH)D_3$; *m/z* 407.3 → 159.3 [$25(OH)D_3$-d_6]; *m/z* 413.4 → 159.4 [$25(OH)D_2$]; *m/z* 419.4 → 159.4 [$25(OH)D_2$-d_6]	LoQ and LoD: $25(OH)D_2$ = 1.2 nmol/L and 0.05 pmol/L; $25(OH)D_3$ = 1.1 nmol/L and 0.09 pmol/L	CV = 2.1% for $25(OH)D_3$; 2% for $25(OH)D_2$

Table 1 (continued)

RMP (date)	Sample preparation	LC column/separation	Internal standards	Calibrants	MS/MS transitions	LoD	Precision
CDC (2015) [40]	Serum sample amount: 0.5 g; sample added to H_2O to avoid protein precipitation with addition of ethanolic IS solution; aqueous Na_2CO_3 added to release metabolites from binding protein; liquid-liquid extraction with *n*-hexane (2x); extract dried and reconstituted in 73% MeOH/water	Ascentis F5 (PFP) (150 × 2.1 mm, 2.7 µm) @ 27 °C; isocratic elution; separation of 25(OH)D_2, 25(OH)D_3 and epimers in 12 min	25(OH)D_3-d_6 and 25(OH)D_2-d_3	25(OH)D_3 solid used to prepare 3 working solutions that are compared to SRM 2972 to assign exact concentration; SRM 2972 used for 25(OH)D_2; 8 calibrants from four independent working solutions prepared with mass isotope ratios of 0.25, 0.50, 1.25, and 2.5	MS/MS: positive atmospheric pressure chemical ionization (APCI); *m/z* 383.3 → 365.1 (quantitation) and *m/z* 383.3 → 105.0 (confirmation) for 25(OH)D_3; *m/z* 389.3 → 371.1 for 25(OH)D_3-d_6; *m/z* 395.3 → 377.3 (quantitation) and *m/z* 395.3 → 209.1 (confirmation) for 25(OH)D_2; *m/z* 398.3 → 380.3 25(OH)D_2-d_3	25(OH)D_2 = 0.13 nmol/L; 25(OH)D_3 = 4.61 nmol/L	
NIST (modified for SRM 1949 [53])[a]	Serum sample amount: 750 µL; Ethanolic IS solution added to serum sample; equilibrated at RT for 1 h; pH adjusted with carbonate buffer; extracted (2x) with 5 mL of hexane/ethyl acetate (50:50 fraction); shake 10 min (2nd extraction 3 min); combined extracts dried and reconstituted in 150 µL of MeOH	Ascentis Express F5 (pentafluorophenyl, PFP) column (150 × 4.6 mm, 2.7 µm) isocratic at 20 °C at 29:71 H_2O/MeOH at 1.1 mL/min with run of 20 min; Zorbax SB-CN column (150 × 4.6 mm, 3.5 µm) isocratic at 30 °C 32/68 H_2O/MeOH at 0.75 mL/min with run time 55 min	25(OH)D_3-$^{13}C_5$, 25(OH)D_2-$^{13}C_3$, 3-*epi*-25(OH)D_3-d_6 added assuming normal 25(OH)D levels to 1:1 mass ratio	SRM 2972 used to prepare 9 calibrant solutions with isotope ratios from 0.2 to 1.8	MS/MS: positive atmospheric pressure chemical ionization (APCI); *m/z* 401.4 → 383.2 for 25(OH)D_3 and *m/z* 406.4 → 388.2 for 25(OH)D_3-$^{13}C_5$; PFP column—*m/z* 401.4 → 383.2 for 3-*epi*-25(OH)D_3 and *m/z* 404.4 → 386.2 for 3-*epi*-25(OH)D_3-d_3; *m/z* 413.4 → 395.2 for 25(OH)D_2 and *m/z* 416.4 → 398.2 for 25(OH)D_2-$^{13}C_3$	LoQ = 0.5 ng/g	CV = < 1.5% for 25(OH)D_3; < 4% for 3-*epi*-25(OH)D_3; 3 to 13% for 25(OH)D_2 at levels near LoQ

[a]This method is not the NIST RMP. This method was developed as a high-throughput method using the basic approach of the RMP with modifications to the separation and internal standards

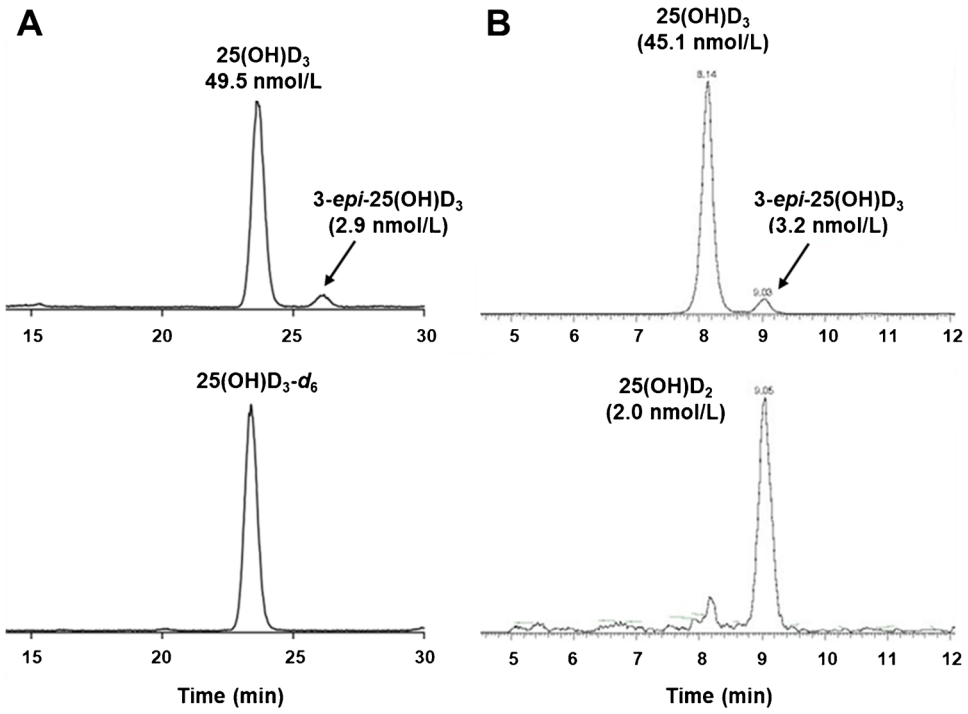

Fig. 1 Selected reaction monitoring chromatograms illustrating the separations achieved with **A** the NIST RMP for SRM 972a L3 and **B** the CDC RMP for SRM 972a L2. The separation shown in **A** was obtained using the CN column, whereas the separation in **B** was obtained using the PFP column. **A** Adapted from Phinney et al. [54] and **B** adapted from Mineva et al. [40]

internal standard solution, equilibration, and pH adjustment followed by extraction (2x) with 50:50 (volume fraction) hexane:ethylacetate to remove the 25(OH)D_2 and 25(OH)D_3. The Ghent and CDC RMPs follow similar approaches to isolate the analytes with the Ghent RMP including an additional clean up step using Sephadex chromatography. To assess the accuracy of the NIST RMPs, fortification experiments were used to determine recoveries of the 25(OH)D_2 and 25(OH)D_3. However, with the availability of SRM 972, Ghent and CDC used SRM 972 to validate their RMPs. The NIST and Ghent RMPs use a CN stationary phase, while the CDC RMP developed 5 years later employs a PFP stationary phase column (Fig. 1B). During the 14 years since the NIST RMPs were developed, they have evolved with minor modifications including the use of a second stationary phase, the PFP column. The NIST RMPs are now three separate chromatographic runs on two different columns: (1) the CN column for determination of 25(OH)D_3 using 25(OH)D_3-d_6, (2) the PFP column for determination of 25(OH)D_2 using 25(OH)D_2-d_3 as the internal standard, and (3) the CN column for determination of 3-*epi*-25(OH)D_3 using 3-*epi*-25(OH)D_3-d_6 as the internal standard. The three ID LC–MS/MS methods recognized as RMPs for 25(OH)D_2 and 25(OH)D_3 include the measurement of 3-*epi*-25(OH)D_3. None of the RMPs, however, included 3-*epi*-25(OH)D_3 as a RMP measurand for recognition by the JCTLM because the focus of these RMPs was determining total 25(OH)D based on the sum of individual measurements of 25(OH)D_2 and 25(OH)D_3. Despite the lack of RMP status for 3-*epi*-25(OH)D_3, determining 3-*epi*-25(OH)D_3 content using these methods provides high-quality, RMP-like results. The RMPs at NIST, Ghent, and CDC are the foundation for the reference measurement system for total 25(OH)D. NIST uses the RMPs to (1) assign certified values to serum-based SRMs, (2) assign target values to important sample sets (e.g., commutability studies), and (3) provide an accuracy basis for quality assurance programs.

Development of SRM 972 and SRM 972a

Design of SRM 972

With input from vitamin D measurement stakeholders (e.g., research and clinical laboratories, assay manufacturers, and CDC), SRM 972 was designed as four pools of human serum with different concentrations of total 25(OH)D and varying concentrations of the individual vitamin D metabolites, i.e., normal concentration of 25(OH)D_3, low concentration of 25(OH)D_3, high concentration of 25(OH)D_2, and high concentration of 3-*epi*-25(OH)D_3. The resulting SRM 972 levels were as follows: (L1) endogenous normal level of 25(OH)D_3; (L2) low level total 25(OH)D serum pool produced by diluting by a factor of 2 with horse serum, which contains no 25(OH)D; (L3) normal level serum pool with added exogenous 25(OH)D_2; and (L4) normal serum pool with added exogenous 3-*epi*-25(OH)D_3. SRM 972 was released in 2009 and sold over 700 units/year resulting in depletion of the inventory after only 2 years. Even with the high demand,

there were concerns about the design of SRM 972 with some customers claiming that the use of horse serum to produce the low level of 25(OH)D and the exogenous 25(OH)D_2 to produce the elevated level of 25(OH)D_2 affected the performance of various immunoassays. Although commutability (see discussion below) of SRM 972 was never evaluated, the use of horse serum in the low level 25(OH)D material probably negatively impacted the use of this level with some immunoassays.

To address these concerns, a significantly improved SRM 972a Vitamin D Metabolites in Frozen Human Serum was issued in 2013 with three endogenous concentration levels of vitamin D metabolites, including a normal 25(OH)D pool (within "sufficient" health status range of 20 to 30 ng/mL as defined by the Institute of Medicine [41, 42]), a low level 25(OH)D pool (within "insufficient" range of 12 to 20 ng/mL), and a high 25(OH)D_2 pool, which had been achieved with donors supplemented with ergocalciferol to provide similar concentrations of 25(OH)D_2 and 25(OH)D_3, and a high concentration of total 25(OH)D (Table 2 and Figure S1, Electronic Supplemental Material, ESM). Unfortunately, due to difficulty in obtaining donors with endogenous levels of 3-*epi*-25(OH)$D_3 > 2$ ng/mL, SRM 972a L4 still contained exogenous 3-*epi*-25(OH)D_3.

Assignment of certified values in SRM 972 and SRM 972a

The assignment of certified values in NIST SRMs has evolved over the past four decades [1, 43–46]. However, the general approach is the use of a reference method or multiple, independent analytical methods [38, 47, 48]. For clinical marker compounds in serum-based SRMs, NIST has historically used a reference method or higher order method to assign certified values (e.g., [49, 50]). Because the NIST LC–MS/MS methods had not been designated as RMPs by the JCTLM when SRM 972 was completed, NIST used results from multiple analytical methods to assign certified values. In addition to using results from the candidate RMPs, measurements were also obtained using two ID LC–MS methods using LC stationary phases with different separation selectivity (i.e., C_{18} and CN phases) and an ID LC–MS/MS method performed at CDC using a C_{18} column [51]. The NIST LC–MS/MS candidate RMPs and the LC–MS (CN) method both successfully separated the 25(OH)D_3 and the 3-*epi*-25(OH)D_3, whereas the NIST LC–MS (C_{18}) method and the CDC LC–MS/MS method did not resolve the 25(OH)D_3 and 3-*epi*-25(OH)D_3, and therefore included a small contribution from 3-*epi*-25(OH)D_3. Since the NIST LC–MS/MS method provided the amount of 3-*epi*-25(OH)D_3 present, these results were used to adjust the 25(OH)D_3 measurements from the NIST LC–MS (C_{18}) method and the CDC LC–MS/MS method. While the use of adjusted results is less than ideal, ignoring this known difference between methods would unnecessarily bias the resulting certified value if based on combining all the various method results (Fig. 2A and B). Ultimately, certified values were assigned for 25(OH)D_3 in all four levels of SRM 972, for 25(OH)D_2 in three levels, and for 3-*epi*-25(OH)D_3 in only one level (i.e., the fortified level) with reference values assigned for 25(OH)D_2 and 3-*epi*-25(OH)D_3 in the remaining levels (Table 2). The uncertainties associated with the certified values for 25(OH)D_3 ranged from 2.5 to 6.1% and for 25(OH)D_2 ranged from 4.8 to 8.9% indicating good agreement of the measurements from the four different analytical methods.

With the production of a replacement for SRM 972, the goal was not only to improve the material (i.e., less exogenous content) but also to improve on the assignment of certified values for the vitamin D metabolites (i.e., reduce the uncertainties). For the certification of SRM 972a, results from ID LC–MS and ID LC–MS/MS at NIST and ID LC–MS/MS at CDC were used. A significant modification in all the methods was the incorporation of the PFP column to resolve the small but significant amount of 3-*epi*-25(OH)D_3 from 25(OH)D_3. Previously, NIST had used a CN stationary phase in both LC–MS and LC–MS/MS methods that provided a separation of 25(OH)D_3 and the epimer; however, the chromatographic run time was 1 h with column wash and re-equilibration time. After extensive evaluation of the separation of 25(OH)D_2, 25(OH)D_3, and 3-*epi*-25(OH)D_3 on C_{18}, CN, and PFP stationary phases by Bedner and Phinney [25], the PFP column and the CN column were included in the certification measurements for the NIST LC–MS and LC–MS/MS methods. The CDC LC–MS/MS method also used the PFP column. The use of two LC stationary phases (i.e., the PFP and CN) with differing selectivity (relative retention) minimizes the likelihood of measurement bias arising from potential undetected interferences [45]. In the evaluation of C_{18}, CN, and PFP stationary phases [25], chromatograms from the analysis of SRM 972 and SRM 909c are provided to illustrate the separation of these three metabolites.

Improvements were incorporated into the NIST methods for the quantification of 3-*epi*-25(OH)D_3 with the use of an isotopically labeled analog as an internal standard. For SRM 972, only L4, which had been fortified with 3-*epi*-25(OH)D_3, was assigned a certified value using 25(OH)D_3-d_6 as the internal standard. For SRM 972a, 3-*epi*-25(OH)D_3-d_3 was used as an internal standard to quantify the epimer in both the LC–MS (CN and PFP) and LC–MS/MS methods, thereby improving the precision of these measurements even at low levels, particularly with the LC–MS/MS (Table 2). By the time that SRM 972a was analyzed to assign values, the NIST ID LC–MS/MS methods had been recognized by the JCTLM as RMPs, and the modified CDC ID LC–MS/MS method now resolved 25(OH)D_3 and 3-*epi*-25(OH)D_3 using

Table 2 Values assigned for vitamin D metabolites and VDBP in NIST SRMs

SRMs	Concentration (ng/mL)[a]									
	25(OH)D_2	%[b]	25(OH)D_3	%[b]	3-*epi*-25(OH)D_3	%[b]	Total 25(OH)D	%[b]	24R,25(OH)$_2$$D_3$	%[b]
SRM 972[c]										
Level 1	0.60 ± 0.20		**23.9 ± 0.8**	**3.4**	1.39 ± 0.04		24.5 ± 0.8[d]			
Level 2	**1.71 ± 0.08**	**4.8**	**12.3 ± 0.6**	**5.0**	0.76 ± 0.02		14.0 ± 0.6[d]			
Level 3	**26.4 ± 1.9**	**7.4**	**18.5 ± 1.1**	**6.1**	1.06 ± 0.03		44.9 ± 2.2[d]			
Level 4	**2.40 ± 0.21**	**8.9**	**33.0 ± 0.8**	**2.5**	**37.7 ± 1.2**	**3.0**	35.4 ± 0.8[d]			
SRM 972a										
Level 1	0.54 ± 0.06		**28.8 ± 1.1**	**3.9**	**1.81 ± 0.10**	**5.6**	29.3 ± 1.1[d]		**2.66 ± 0.10**	**3.8**
Level 2	**0.81 ± 0.06**	**11.2**	**18.1 ± 0.4**	**2.3**	**1.28 ± 0.09**	**7.2**	**18.9 ± 0.4**	**2.2**	**1.41 ± 0.05**	**3.6**
Level 3	**13.2 ± 0.3**	**2.3**	**19.8 ± 0.4**	**2.1**	1.17 ± 0.14		**33.0 ± 0.5**	**1.5**	**1.62 ± 0.06**	**3.8**
Level 4	0.55 ± 0.10		**29.4 ± 0.9**	**3.1**	**26.0 ± 2.2**	**8.5**	30.0 ± 0.9[d]		**2.64 ± 0.09**	**3.5**
SRM 2973	0.65 ± 0.02		**39.4 ± 0.8**	**2.1**	2.10 ± 0.08		40.1 ± 0.8[d]		**3.13 ± 0.11**	**3.6**
SRM 2969	**2.01 ± 0.05**	**2.5**	**11.9 ± 0.3**	**2.5**			**13.9 ± 0.3**	**2.2**	0.57 ± 0.01[e]	
SRM 2970	**23.5 ± 0.3**	**1.3**	**9.63 ± 0.31**	**3.2**			**33.1 ± 0.4**	**1.2**	0.73 ± 0.01[e]	
SRM 1950	0.52 ± 0.17		**24.78 ± 0.77**	**3.1**			25.30 ± 0.79[d]			
SRM 909c			19.65 (0.42)[f]							
SRM 968e[c]										
Level 1			7.09 ± 0.14							
Level 2			12.9 ± 0.3							
Level 3			19.9 ± 0.4							
SRM 968f										
Level 1	0.849 ± 0.051[g]		12.32 ± 0.20[g]		0.720 ± 0.033[g]		13.17 ± 0.21[g]			
Level 2	0.167 ± 0.014[g]		15.64 ± 0.20[g]		1.07 ± 0.14[g]		15.81 ± 0.20[g]			
SRM 1949										
Non-Pregnant	0.67 ± 0.03[h]		24.98 ± 0.28		1.32 ± 0.06		25.65 ± 0.28[d]			
1st Trimester	1.20 ± 0.05		26.01 ± 0.22		1.43 ± 0.02		27.21 ± 0.23[d]			
2nd Trimester	0.514 ± 0.037		30.00 ± 0.50		1.87 ± 0.07		30.51 ± 0.50[d]			
3rd Trimester	0.897 ± 0.057		29.43 ± 0.41		1.87 ± 0.04		30.33 ± 0.41[d]			
SRM 1949	**VDBP[i] (µg/mL)**		**VDBP[i] (µmol/L)**							
Non-Pregnant	211.5 ± 2.8		4.01 ± 0.05							
1st Trimester	286.7 ± 3.8		5.43 ± 0.06							
2nd Trimester	349.7 ± 4.3		6.64 ± 0.07							
3rd Trimester	383.4 ± 5.1		7.28 ± 0.08							
SRM 1950	175 ± 18		3.33 ± 0.33[j]							

[a]Bold font values are denoted as certified values; normal font values are designated as reference or noncertified values. For certified and reference/noncertified values, the uncertainty provided is generally an expanded uncertainty calculated as $U = ku_c$, where u_c is the combined uncertainty and k is a coverage factor corresponding to approximately 95% confidence ($k = 2$). For specific details for each SRM, see the current COA at www.nist.gov/SRMs and search by number to find the COA

[b]Relative uncertainty as % for certified values only

[c]SRM is no longer available

[d]Value for total 25(OH)D is not on the COA; value determined from sum of 25(OH)D_2 and 25(OH)D_3 and a combined expanded uncertainty

[e]Value for 24,25(OH)$_2$$D_3$ not reported on the COA; value from certification report by Hahm et al. [57] and converted from ng/g to ng/mL

[f]Value not on COA; value reported in Bedner and Phinney [25]; value in parentheses is SD ($n = 10$)

[g]Value on COA is in ng/g; value converted to ng/mL by multiplying by the density of the serum 1.0180 g/mL for L1 and 1.0202 g/mL for L2

[h]Value is not on COA

[i]*VDBP*, vitamin D binding protein

[j]Value not reported on COA in µmol/L; value calculated from average mass and density as reported in Kilpatrick et al. [67]

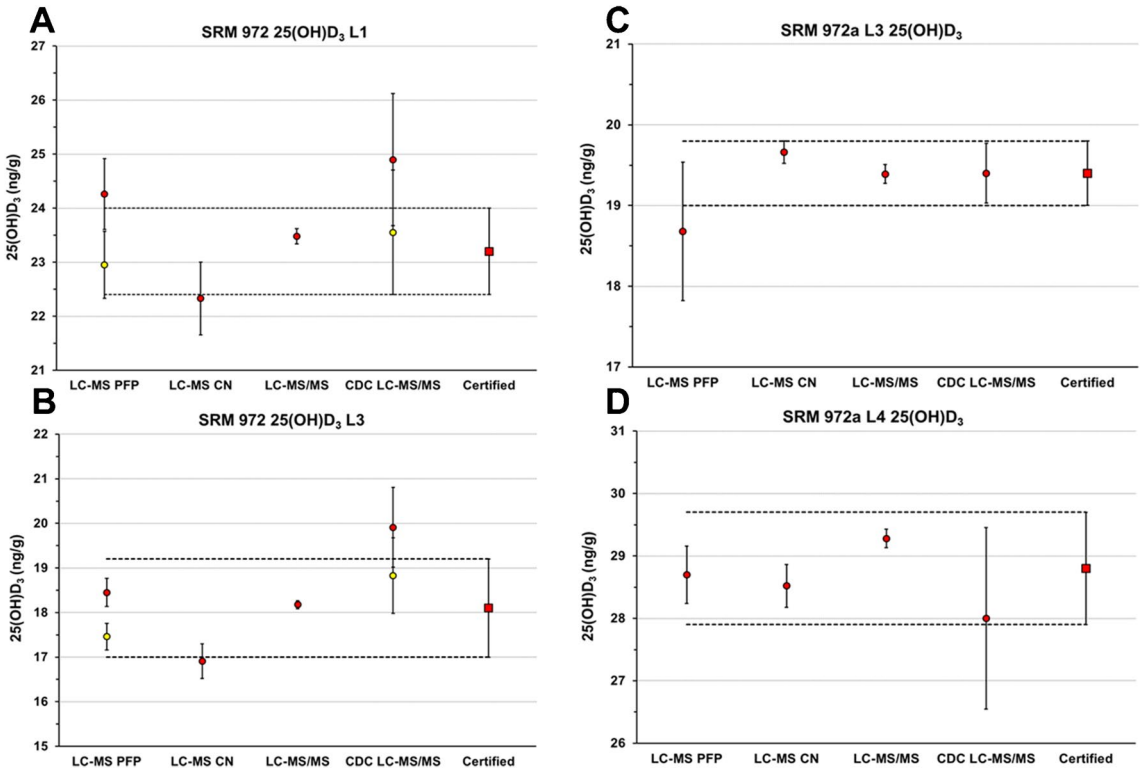

Fig. 2 Results from the different analytical methods used to assign the certified values for 25(OH)D_3 in **A** SRM 972 L1, **B** SRM 972 L3, **C** SRM 972a L3, and **D** SRM 972a (L4). For both SRM 972 and SRM 972a, the red circles are the mean results of the analytical method with error bars representing the standard deviation of the measurements. The red square is the certified value with error bars representing the uncertainty as described on the Certificate of Analysis. For SRM 972 (**A** and **B**), the yellow circle represents the result for 25(OH)D_3 after adjustment by subtracting the amount of 3-*epi*-25(OH)D_3. Plots are based on results from Phinney et al. [51, 54]

the PFP column [27] (Fig. 2C and D and Figures S2 and S3, ESM). Because of these method improvements and the better agreement of the results, certified values were assigned for 25(OH)D_3 in all levels of SRM 972a with uncertainties ranging from 2.1 to 3.9% compared with 2.5 to 6.1% in SRM 972. For 25(OH)D_2, the uncertainty for the elevated level was reduced from 7.4 to 2.3% and a certified value was assigned to a lower level than previously (i.e., 0.81 ng/mL ± 0.06 ng/mL). For 3-*epi*-25(OH)D_3, certified concentrations were assigned in three levels compared to only one level previously (Table 2).

Development of multiple SRMs for determining vitamin D metabolites

Based on the success of SRM 972, efforts during the next decade focused on preparing additional serum matrix SRMs with various levels of vitamin D metabolites and assigning values for vitamin D metabolites in other serum/plasma-matrix SRMs with unique features. After the unanticipated early depletion of the inventory of SRM 972 at the end of 2011, NIST assigned values for 25(OH)D_2 and 25(OH)D_3 in SRM 968e Fat-Soluble Vitamins and Carotenoids in Frozen Human Serum and SRM 1950 Metabolites in Human Plasma and directed customers to use these materials until SRM 972a became available in early 2013. SRM 968e consisted of three levels, all of which were assigned certified values for 25(OH)D_3 with concentrations from about 7 to 20 ng/mL. SRM 968e was replaced in late 2017 by SRM 968f Fat-Soluble Vitamins in Frozen Human Serum which consists of two levels with reference values for 25(OH)D_2, 25(OH)D_3, and 3-*epi*-25(OH)D_3. In both levels of SRM 968f, the total 25(OH)D content is relatively low (nominally 13 ng/mL and 15 ng/mL). SRM 1950 is a unique plasma-based material, which became available in 2011 [52], with values assigned for 90 metabolites including 25(OH)D_2 and 25(OH)D_3 with a total 25(OH)D concentration similar to SRM 972 L1 (Table 2).

A novel SRM was produced in 2017 using human serum pools from female donors of reproductive age who were not pregnant or were pregnant and identified as in the first, second, or third trimester of pregnancy. While SRM 1949 Frozen Prenatal Human Serum [53] is intended primarily for determining thyroid hormones, i.e., total thyroxine (T_4), and total triiodothyronine (T_3), reference values were assigned

for 25(OH)D_2, 25(OH)D_3, and 3-*epi*-25(OH)D_3 using an ID LC–MS/MS method intended as a modified, higher throughput version of the RMPs as described by Boggs et al. [53] (Table 1). Note that this ID LC–MS/MS method incorporated ^{13}C-labeled internal standards for the quantification of 25(OH)D_2 and 25(OH)D_3.

Even though SRM 972a addressed some of the limitations of SRM 972 [54], the concentration of total 25(OH)D in the four levels had a range of nominally 19 to 33 ng/mL (Table 2). Total serum 25(OH)D concentrations in the US population typically range from 40 to 75 nmol/L (16 to 30 ng/mL) [55]; however, about 36% of the population have total 25(OH)D concentrations greater than 75 nmol/L (30 ng/mL) [55]. To complement SRM 972a, NIST developed three additional materials: (1) SRM 2973 Vitamin D Metabolites in Frozen Human Serum (High Level) [56], (2) SRM 2969 Vitamin D Metabolites in Frozen Human Serum (Total 25-Hydroxyvitamin D Low Level) [57], and (3) SRM 2970 Vitamin D Metabolites in Frozen Human Serum (25-Hydroxyvitamin D_2 High Level) [57].

The first of these materials, SRM 2973, was released in 2017 [56] with a concentration of total 25(OH)D that was 21% higher than any level in 972a and 34% higher in 25(OH)D_3 content. In 2021, SRM 2969 was issued providing a lower level of total 25(OH)D (13.9 ng/mL versus 18.9 ng/mL in SRM 972a). The bar graph in Fig. 3 illustrates the range of concentrations of total 25(OH)D for the 20 SRM levels with values assigned for 25(OH)D_2 and 25(OH)D_3. With the addition of SRM 2969 and SRM 2973, SRMs are now available that span a range of total 25(OH)D from 13.9 to 40.1 ng/mL, thereby nearly doubling the concentration range. SRM 2970, also introduced in 2021, has a high concentration of 25(OH)D_2 and a low concentration of 25(OH)D_3, which was produced using serum from donors who were supplementing their diet with vitamin D_2. SRM 2970 has the lowest content of 25(OH)D_3 now expanding the range for 25(OH)D_3 from 9.6 to 39.4 ng/mL (SRM 2973) (Table 2 and Fig. 3).

Certification approach for SRM 2969, SRM 2970, and SRM 2973

For SRM 2969, SRM 2970, and SRM 2973, the certification approach had evolved to assigning certified values based only on results from the NIST ID LC–MS/MS RMPs, i.e., a single method approach used for many of the previous SRMs for clinical markers (e.g., cholesterol and glucose). The certified concentrations for 25(OH)D_3 in these three recent materials have uncertainties ranging from 3.2% for the lowest concentration (SRM 2970) to 2.1% for the highest concentration (SRM 2973) available in the current suite of SRMs (Table 2).

Calibration Solution CRMs

In parallel with the development of serum matrix SRMs for vitamin D metabolites, NIST, as well as commercial vendors, have developed solution CRMs for 25(OH)D_2, 25(OH)D_3, 3-*epi*-25(OH)D_3, and 24,25(OH)$_2$$D_3$ for use as calibrants. Issued in 2009, SRM 2972 25-Hydroxyvitamin D_2 and 25-Hydroxyvitamin D_3 Calibration Solutions consisted of individual ethanolic solutions of 25(OH)D_2 and 25(OH)D_3 at nominal mass fractions of 300 ng/g and 400 ng/g, respectively. SRM 2972 was updated in 2014 to contain four solutions, i.e., the original two solutions of 25(OH)D_2 and 25(OH)D_3 with revised mass fractions, a

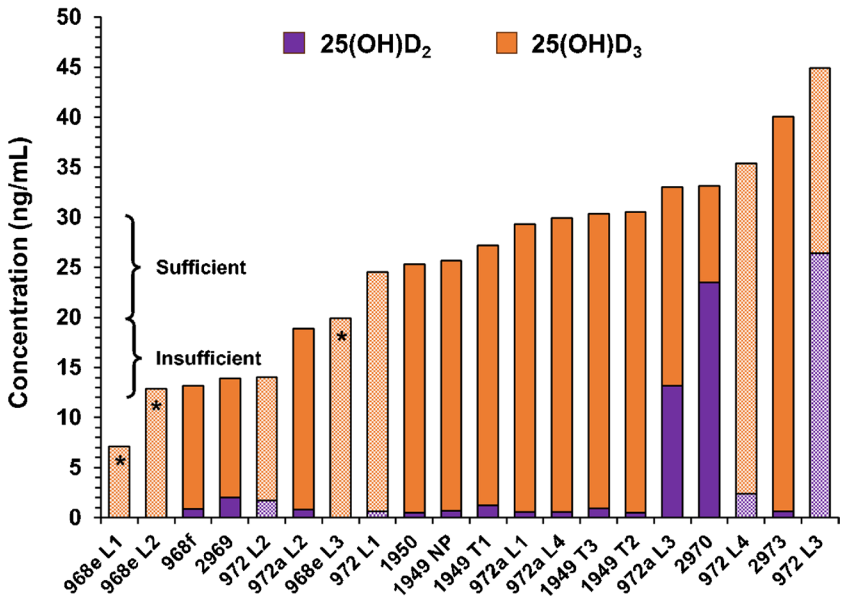

Fig. 3 Comparison of the concentrations of 25(OH)D_3 (orange) and 25(OH)D_2 (purple) ordered based on increasing total 25(OH)D concentration (ng/mL). Bars that are textured color represent SRMs that are no longer available. For bars with an asterisk, no concentration for 25(OH)D_2 was determined. Brackets indicate the range for vitamin D "sufficiency" and "insufficiency" as defined by the Institute of Medicine [41, 42]

second solution of 25(OH)D_3 at nominal mass fraction of 800 ng/g, and a solution of 3-*epi*-25(OH)D_3 at nominal mass fraction 300 ng/g. Based on customer feedback, the updated material, denoted as SRM 2972a 25-Hydroxyvitamin D Calibration Solutions, was intended to provide a more convenient, ready-to-use form for calibrant preparation and use. In 2022, the 3-*epi*-25(OH)D_3 solution was removed from SRM 2972a to create a new SRM 2968 3-*epi*-25-Hydroxyvitamin D_3 Calibration Solution, and the remaining solutions from SRM 2972a became SRM 2972b 25-Hydroxyvitamin D_2 and 25-Hydroxyvitamin D_3 Calibration Solutions. To improve comparability of 24,25-dihydroxyvitamin D_3 [24,25(OH)$_2D_3$] measurements, SRM 2971 24R,25-Dihydroxyvitamin D_3 Calibration Solution was developed in 2018 with a nominal mass fraction of 1000 ng/g (Table 3). In 2017, Cerilliant (now part of MilliporeSigma) initiated the development of solution CRMs for vitamin D metabolites and isotopically labeled analogues to support the measurement of vitamin D metabolites. The CRMs from Cerilliant (Table 3) are at considerably higher concentrations (about 100x) compared to the NIST SRMs.

Addition of new information to existing SRMs

Values for 24,25(OH)$_2D_3$

Because of growing interest in 24,25(OH)$_2D_3$ [58], NIST developed a RMP for this metabolite in 2015 as reported by Tai et al. [59]. Using the results from the candidate RMP only, NIST assigned values for 24,25(OH)$_2D_3$ in SRM 972a and SRM 2973 as shown in Table 2. The values were initially designated as reference values and eventually changed to certified values after the method was recognized by the JCTLM as a RMP in 2017. The concentrations of 24,25(OH)$_2D_3$ range from 1.41 to 3.13 ng/mL with associated uncertainties consistently at 3.5 to 3.8% (Table 2). However, since the four pools of serum used in SRM 972a have similar levels of 25(OH)D_3 in L1 and L4 (same serum pool) and L2 and L3 (different pools with similar concentrations of 25(OH)D_3), there are only two nominal concentrations for 24,25(OH)$_2D_3$ in these four pools (i.e., 1.4 ng/mL and 2.6 ng/mL). SRM 2969 and SRM 2970 have relatively low concentrations of 25(OH)D_3, and therefore, the

Table 3 Calibration solution CRMs for vitamin D metabolites

SRM/CRM	Unit	Mass fraction (ng/g)[a]							
		25(OH)D_2	%[b]	25(OH)D_3	%[b]	3-epi-25(OH)D_3	%[b]	24R,25(OH)$_2D_3$	%[b]
NIST SRMs[c]									
SRM 2972b	15 × 1.2 mL	293.6 ± 9.1	3.1						
Level 1				410.0 ± 14.9	3.6				
Level 2				812.0 ± 29.2	3.6				
SRM 2968	5 × 1.2 mL					293.4 ± 13.5	4.6		
SRM 2971	5 × 1.0 mL							1054.4 ± 19.0	1.8
MilliporeSigma CRMs[c]		**Concentration (µg/mL)[d]**							
25-(OH)D_2 (H-073)	1 × 1 mL	50.0 ± 0.3	0.6						
25-(OH)D_3 (H-083)	1 × 1 mL			100.0 ± 0.6	0.6				
3-*epi*-25(OH)D_3 (E-086)	1 × 1 mL					50.0 ± 0.3	0.6		
								1α, 25(OH)$_2$D	
1α, 25(OH)$_2D_2$ (H-089)	1 × 1 mL							5.00 ± 0.03	0.6
1α, 25(OH)$_2D_3$ (H-090)	1 × 1 mL							5.000 ± 0.028	0.6
1α, 25(OH)$_2D_3$-$^{13}C_3$ (H-107)	1 × 1 mL							5.000 ± 0.028	0.6
D_6-25(OH)D_3 (H-074)	1 × 1 mL			50.0 ± 0.3	0.6				

[a]The uncertainty provided with each value is an expanded uncertainty about the weighted mean to cover the measurand with approximately 95% confidence. The expanded uncertainty is calculated as $U = ku_c$, where u_c incorporates the observed difference between the results from the methods and their respective uncertainties, as well as uncertainties related to purity estimation and possible degradation of the solution over time and k is the coverage factor ($k = 2$) corresponding to approximately 95% confidence. For specific details for each SRM, see the current COA at www.nist.gov/SRMs and search by number to find the COA

[b]Relative uncertainty as %

[c]All solutions are in ethanol

[d]Uncertainty is expressed as an expanded uncertainty in accordance with ISO 17034 at the approximate 95% confidence interval using a coverage factor of $k = 2$. For more details for each CRM, see the current COA at www.cerilliant.com/products/catalog.aspx and search by number to find the COA

concentrations of 24,25$(OH)_2D_3$ in SRM 2969 and SRM 2970 are low (nominal 0.57 ng/mL and 0.73 ng/mL, respectively) at levels about half of the lowest levels in SRM 972a (Table 2). These values were not reported on the COAs for SRM 2969 and SRM 2970; however, they are provided in the certification report [57].

Values for vitamin D binding protein and free 25(OH)D

Release of the 25(OH)D from VDBP and albumin is a critical step during sample preparation and extraction prior to LC–MS/MS, and variations in VDBP content could affect extraction recovery. Studies have shown that VDBP concentration increases during pregnancy [60–63]; therefore, SRM 1949 provides an excellent material for use to investigate the relationship of VDBP and 25(OH)D during pregnancy and to assess the accuracy of VDBP measurements. VDBP is typically measured using an ELISA; however, recent efforts to quantify VDBP using LC–MS/MS have been reported [64, 65]. Kilpatrick and Phinney [65] developed an approach based on measurement of isoform-specific peptides using LC–MS/MS and provided a preliminary value for VDBP in SRM 1950. In 2020, Kilpatrick et al. [66] expanded the development of an LC–MS/MS method for quantification of VDBP using isotopically labeled peptides as internal standards and implemented the method for value assignment of VDBP content in SRM 1949 (Table 2).

The results for VDBP in SRM 1949 demonstrated the increase in VDBP concentration during pregnancy (see Figure S4, ESM). Using the same LC–MS/MS method, Kilpatrick et al. [67] assessed the VDBP content in SRM 1950 as part of an interlaboratory comparison between NIST [65] and the University of Washington [64]. Each laboratory used different sample preparation and quantification approaches but followed a similar measurement protocol resulting in VDBP concentration of 3.17 nmol/g (%CV = 3.44) (NIST) and 3.50 nmol/g (%CV = 2.68) (University of Washington), which were combined to provide a target value for SRM 1950 of 3.33 nmol/g ± 0.33 nmol/g (175 µg/mL ± 18 µg/mL) [67]. The availability of reference values for VDBP content in SRM 1949 and SRM 1950 allows researchers to validate the assays used in their laboratories for determination of VDBP.

Recent studies suggest that the measurement of free 25(OH)D may be a better indicator of vitamin D status than total 25(OH)D content [9, 68, 69]. At present, the principle method for determination of free 25(OH)D is an ELISA based on monoclonal anti-25(OH)D antibodies that uses a specific incubation buffer that allows capture of only the free fraction of 25(OH)D [70]. As an initial step toward harmonization of free 25(OH)D measurements, Sempos et al. [71] reported target values of free 25(OH)D in three existing SRMs, i.e., SRM 972a, SRM 2973, and SRM 1949, determined using this ELISA. For this study, two laboratories used the only commercial free 25(OH)D ELISA [70] to measure the free 25(OH)D in the three SRMs as summarized in Table 4. The linear regression for free 25(OH)D and total 25(OH)D in the nine levels of the three SRMs is shown in Fig. 4A indicating that the relationship in the three pregnancy trimester pools is significantly different from the other SRMs that represent non-pregnant donors. The relationship between free 25$(OH)D_3$ and VDBP (Fig. 4B) illustrates the decrease in free 25(OH)D with the increase in VDBP as pregnancy advances [60–63]. Although the values for free 25(OH)D are not reported on the SRM COAs, the results reported in Sempos et al. [71] can be used as target values by researchers for use as controls in free 25(OH)D measurements.

Commutability studies for vitamin D SRMs

CRMs for clinical diagnostic markers in human serum, which are typically prepared as pooled and/or processed serum samples, should undergo a commutability assessment to support their equivalent performance to that of patient samples [72, 73]. The International Vocabulary of Metrology (VIM) defines commutability as "a property

Table 4 Summary of percent free 25$(OH)D_3$ in NIST SRMs

	Free 25(OH)D_3 (pg/mL)		Total 25(OH)D (ng/mL)		Percent Free 25 (OH)D[a]
	Mean	SD	Mean	Uncertainty[b]	
SRM 972a Level 1	7.75	0.59	29.3	1.1	0.0264
SRM 972a Level 2	4.88	0.58	18.9	0.4	0.0258
SRM 972a Level 3	7.83	0.54	33.0	0.5	0.0237
SRM 972a Level 4	7.90	0.53	30.0	0.9	0.0263
SRM 2973	10.00	0.58	40.1	0.8	0.0249
SRM 1949 NP[c]	6.15	0.38	25.3[d]	0.28[e]	0.0243
SRM 1949 1T[c]	4.70	0.42	27.2[d]	0.23[e]	0.0173
SRM 1949 2T[c]	4.40	0.48	30.5[d]	0.50[e]	0.0144
SRM 1949 3T[c]	3.76	0.36	30.3[d]	0.41[e]	0.0124

[a]Free 25(OH)D/total 25(OH)D × 100

[b]Uncertainty from the COA, if provided

[c]*NP*, non-pregnant; *1 T*, 1st trimester of pregnancy; *2 T*, 2nd trimester of pregnancy; *3 T*, 3rd trimester of pregnancy

[d]Values calculated (sum of 25$(OH)D_2$ and 25$(OH)D_3$) for this study and not assigned by NIST on the SRM Certificate of Analysis (COA)

[e]Uncertainty is the combined expanded uncertainty calculated from the sum of the values reported for 25$(OH)D_2$ and 25$(OH)D_3$

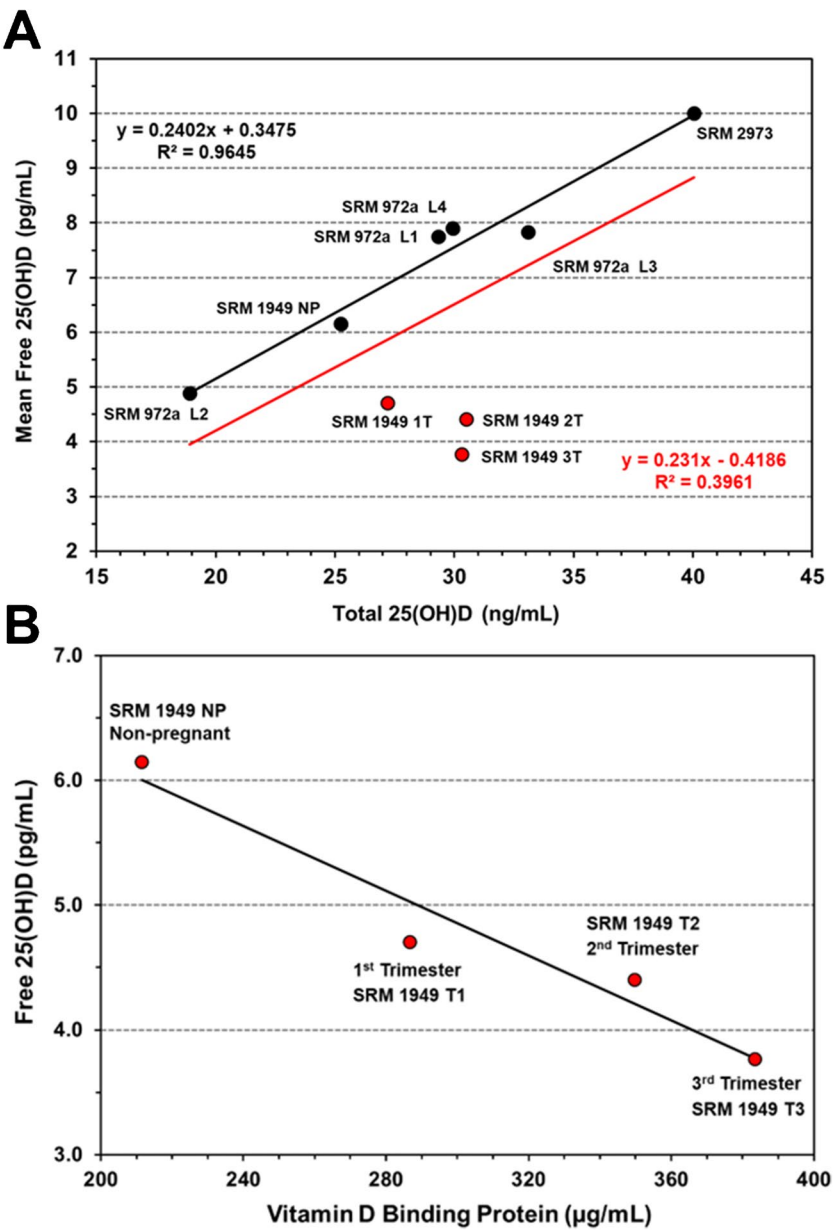

Fig. 4 Least squares linear regression for free (25(OH)D and total 25(OH)D for nine SRM levels **A** and for 25(OH)D versus vitamin D binding protein for various levels of SRM 1949 **B**. Adapted from Sempos et al. [71]

of a reference material, demonstrated by the closeness of agreement between the relationship among the measurement results for a stated quantity and the relationship obtained among measurement results for other specified materials" (e.g., patient samples in clinical laboratory medicine) [74]. As a practical definition, commutability is the property of the RM whereby the measurement response for the RM is the same as for an individual clinical sample with the same concentration of the analyte when measured using two or more measurement systems. Ideally, one of the measurement methods is an RMP.

The SRMs intended specifically for vitamin D metabolites have been included in commutability studies through collaborations among NIH ODS, NIST, CDC, University of Ghent, assay manufacturers, and research laboratories. To date, three commutability studies have assessed SRM 972a, SRM 2973, SRM 2969, SRM 2970, and SRM 1949. The first commutability study for SRM 972a was conducted in 2011 [75], a second study for SRM 972a and SRM 2973 was conducted in 2016 [76], and the most recent study for SRM 2969, SRM 2970, and SRM 1949 was conducted in 2022 [77]. For the first study [75], 18 laboratories provided results (14 immunoassays and 4 LC–MS/MS); unfortunately, only 9 laboratories gave permission for publication of their results thereby limiting the value of the study. In this first study, all four levels of SRM 972a were assessed as commutable for the six immunoassays reporting results and for the three LC–MS/MS methods with two methods providing

"uncertain" commutability for two different levels of SRM 972a. For the second study [76], results were received from 28 laboratories using 20 immunoassays (12 unique assays) and 14 LC–MS/MS methods for the assessment of SRM 972a again and SRM 2973. For this study, SRM 972a L3, with a high endogenous concentration of 25(OH)D_2, was assessed as noncommutable for several immunoassays. SRM 972a L4 was assessed as noncommutable using six assays due to inadequate separation of the 3-epimer. However, even though commutability is defined as a property of the RM, the problem in this instance was not the SRM but the lack of specificity of the LC–MS/MS method to provide the quantity values for the intended measurand, i.e., the method was not fit for the intended purpose.

The 2022 commutability study was designed to assess the commutability of the three most recent SRMs representing unique properties that had not been assessed previously (i.e., low total 25(OH)D, 25(OH)D_2, and serum from pregnant women) and to evaluate additional assays not evaluated previously. For this third commutability study, 17 unique immunoassays and 9 LC–MS/MS assays were used to assess SRM 2969, SRM 2970, and SRM 1949, and the results were evaluated using both the traditional Clinical Laboratories Standards Institute (CLSI) approach using 95% prediction intervals that was used in the two previous studies [75, 76] and the recent International Federation of Clinical and Laboratory Chemistry (IFCC) approach based on differences in bias [78]. The results of the 2022 commutability study have been evaluated and are expected to be published soon [77]. Results from both the 2016 and 2022 commutability studies for 11 SRM levels are combined using the Abbott ARCHITECT assay in Fig. 5. For SRM 972a L3 and SRM 2970, which have high endogenous concentrations of 25(OH)D_2 (13.2 ng/mL and 23.5 ng/mL, respectively), both SRMs were noncommutable (i.e., outside the pre-set criterion of 8.8%) using the Abbott ARCHITECT assay (Fig. 5).

Efforts to assess and improve accuracy and comparability of 25(OH)D measurements

Quality assurance programs for 25(OH)D

The improvement in accuracy and comparability of 25(OH)D measurements during the past 15 years can be attributed to several quality assurance efforts including an expanded use of SRMs and quality assurance programs. DEQAS was

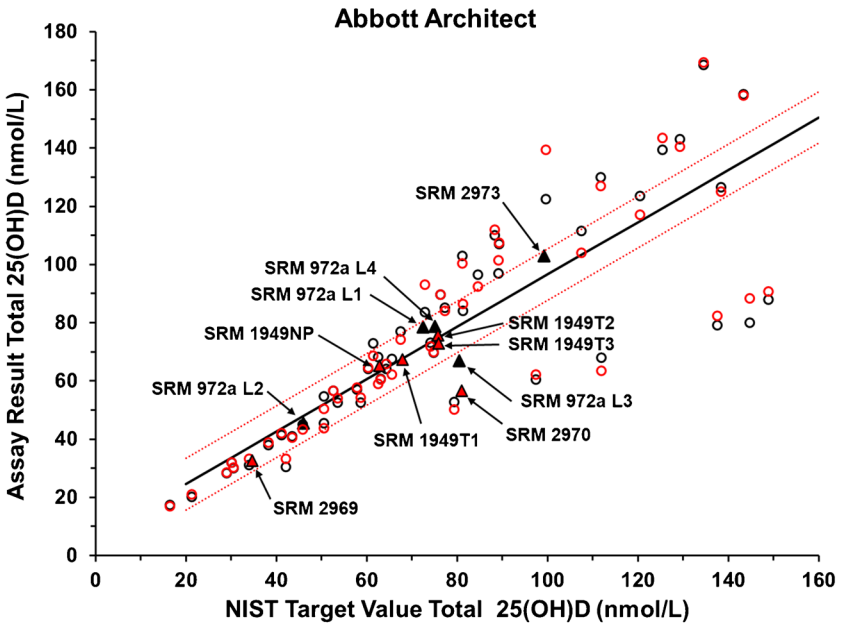

Fig. 5 Commutability assessment using the CLSI approach with a pre-set criterion of 8.8% of 11 NIST SRM levels for total 25(OH)D using the Abbott ARCHITECT assay in two different commutability studies. Black circles are the Abbott ARCHITECT assay results and the NIST-assigned target values for 25(OH)D in 50 single-donor samples in 2016 commutability study. NIST values assigned using the RMPs for 25(OH)D_2 and 25(OH)D_3. Red circles are the Abbott ARCHITECT assay results from the 2022 commutability study for the same 50 single-donor samples. Black triangles are the SRMs assessed in the 2016 study, and the red triangles are the SRMs assessed in the 2022 study. SRM 972a L3 and SRM 2970 have high concentrations of 25(OH)D_2. The solid black line is the Deming linear regression of the 2016 study results (black circles). The red dashed lines are the 8.8% criterion from the Deming line used for the assessment of commutability using the CLSI approach. Results for 2016 commutability study [76] and 2022 commutability study [77] are published elsewhere. Figure adapted from [87]

established in 1989 "to ensure the analytical reliability of 25(OH)D assays" [79] and continues today with approximately 1000 participants worldwide receiving quarterly sample sets that are analyzed to determine 25(OH)D using about 30 different methods (2017) [16]. The accuracy of DEQAS participants' results was originally assessed using the consensus All-laboratory Trimmed Mean (ALTM). The first accuracy-based quality assurance program for 25(OH)D, known as the Vitamin D Metabolites Quality Assurance Program (VitDQAP), was initiated in 2009 by NIST, in collaboration with NIH ODS, using the RMPs and SRMs to establish trueness of measurements [4, 80, 81]. The VitDQAP conducted 12 exercises from 2009 through 2016 with a total of 99 participating organizations. During these exercises, SRM 972a and SRM 2973 were distributed to participants as unknown samples as well as other SRMs that eventually had values assigned for 25(OH)D_2 and 25(OH)D_3, e.g., SRM 968e, SRM 968f, and SRM 1950. Target value assignment for 25(OH)D_2 and 25(OH)D_3 by NIST using their RMPs allowed DEQAS to become an "accuracy-based" scheme in 2013 [82]. In late 2018, DEQAS transitioned from NIST to CDC providing target values using their RMP. With the conversion of the long-running DEQAS program to an accuracy basis and the establishment of the College of American Pathologists (CAP) accuracy-based vitamin D program (ABVD) [83], NIH ODS and NIST support for accuracy-based performance testing through VitDQAP was sunset. The results of each of the 12 VitDQAP exercises were published in NIST reports (see, e.g., [84]). Two papers highlight the results of the exercises [4, 81] and a final report summarizes the findings and impact of the program [80].

CDC vitamin D standardization – certification program (CDC-VDSCP)

The CDC Vitamin D Standardization – Certification Program [17] is intended to provide reference measurements for total 25(OH)D, assess the accuracy and precision of vitamin D tests, and monitor their performance over time. Laboratories participating in the CDC VDSCP demonstrate that their 25(OH)D assay performance on quarterly sets of samples meets performance criterion of 5% mean bias compared to the CDC RMP and imprecision of < 10%. Participants meeting these criteria are listed on the CDC VDSCP website as "Certified Total 25(OH)D Assays." As of September 2023, 42 assays are listed as certified, equally distributed between LC–MS/MS and immunoassays [17]. The CDC VDSCP is part of the CDC Vitamin D Reference Laboratory [85], which states that the measurements are based on the "highly accurate and precise reference method" (i.e., the CDC RMP), which is calibrated using SRM 2972a providing traceability to the SI.

International intercomparison for determination of 25(OH)D_2 and 25(OH)D_3

In 2015, candidate SRM 2973 and a sample from the VitDQAP with a high endogenous content of 25(OH)D_2 were used as unknown test samples in an international comparison conducted among national metrology institutes (NMI) as part of the Consultative Committee on the Amount of Substance (CCQM) [86, 87]. CCQM intercomparison exercises are intended to demonstrate the equivalence of measurements among NMIs and CRMs [86] using a Key Comparison Reference Value (KCRV) with an associated uncertainty based on means of qualified results from the participating laboratories. Seven NMI laboratories analyzed these two serum materials using ID LC–MS/MS methods to determine 25(OH)D_2 and 25(OH)D_3 as shown in Fig. 6. For this study, the KCRV for 25(OH)D_3 for SRM 2973 (Fig. 6A) and the VitDQAP sample (Fig. 6B) were 37.85 ng/g ± 0.65 ng/g and 25.31 ng/g ± 0.65 ng/g, respectively, compared with the NIST certified value for SRM 2973 (38.6 ng/g ± 0.8 ng/g) and the NIST target value for VitDQAP sample (25.7 ng/g ± 0.3 ng/g), both determined with the NIST RMPs, demonstrating agreement with the KCRV. For the measurements of 25(OH)D_2 in VitDQAP sample, the KCRV was 6.22 ng/g ± 0.3 ng/g compared with NIST measurements of 6.33 ng/g ± 0.12 ng/g.

Intended uses of CRMs

Matrix CRMs are intended for the following uses: (1) analytical method development and new method validation, (2) to serve as a control material during routine measurements, (3) to assign values for in-house quality control materials, and (4) to provide metrological traceability of measurement results. Practical guides for the use of CRMs have been published [88–90]. Use of serum-based CRMs, ideally with endogenous levels of the measurand of interest, is critical for new method development and validation to assess accuracy (trueness) of the complete analytical process (i.e., extraction, sample extract clean up, chromatographic separation and detection, and quantification) for determining clinical diagnostic markers.

During routine measurements of patient serum samples to determine vitamin D status, CRMs, or in-house control materials value assigned using a CRM, should be included in every batch of samples to assess the accuracy of the results and to validate that the measurements are in control. Control charts are often used to monitor the quality of measurements in a laboratory over time allowing for comparison of measurements in different studies and at different times. An example of a control chart is illustrated in Fig. 7 where SRM 972a L3 was analyzed quarterly at NIST for assigning target values for DEQAS and

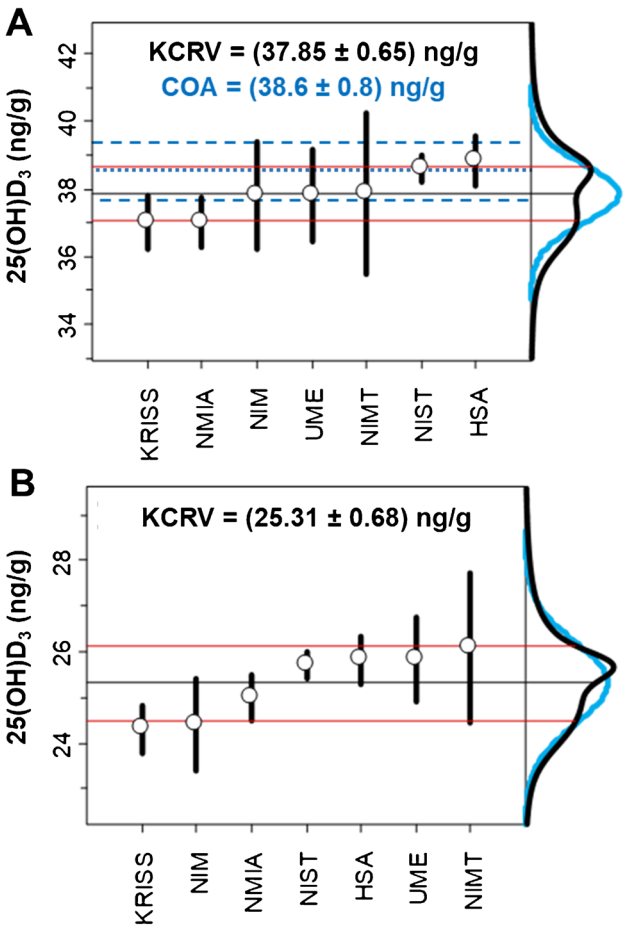

Fig. 6 Results of an international comparison of measurements for $25(OH)D_3$ in **A** candidate SRM 2973 and **B** a VitDQAP sample. The KCRV (solid black line) and associated uncertainty (solid red line) are compared with the certified value for SRM 2973 (blue dotted line) and uncertainty from the COA (dashed blue line). The *x*-axis lists the NIM participating in the study: Korea Research Institute of Standards and Science (KRISS) National Measurement Institute Australia (NMIA), National Institute of Metrology China (NIM), National Metrology Institute of Turkey (UME), National Institute of Metrology Thailand (NIMT), NIST, and Health Sciences Authority Singapore (HSA)

for other measurement campaigns. The control chart (Fig. 7) is based on measurements over 6 years and performed by five different analysts using the NIST RMPs. Based on the low control result for July 2018, the samples were re-analyzed with SRM 972a L1 as an additional control, and the results were found to be within the certified value.

Impact of vitamin D metabolite SRMs

The impact of the availability of both serum matrix SRMs and calibration solution SRMs that have been developed over the past 15 years can be assessed by both sales' information and by reports in the scientific literature describing their use for method validation, quality control, and research. Sales of the SRMs are a good indicator of their use and acceptance within the vitamin D measurement community (see Figure S4A). Sales of SRM 972 exceeded expectations with the total inventory of over 1600 units depleted in 2 years. A detailed discussion of the sales of SRM 972 and SRM 972a from 2009 through 2016 was published previously [4] including the distribution of sales among different customer sectors, i.e., commercial testing laboratories, instrument manufacturers, hospitals and medical centers, universities, clinical diagnostic manufacturers, resellers, and government. Sales of SRM 2972 (and its updates SRM 2972a and 2972b) have been consistent at a rate of 256 units/year (Figure S4A).

Literature reports of use of vitamin D SRMs

The serum matrix and calibration solution SRMs for vitamin D metabolites have been widely used to validate new analytical methods and to assure the quality of 25(OH)D measurements using both LC–MS/MS methods and immunoassays. Since 2010, over 90 papers have been published where the authors report the use of these SRMs to validate their methods and to provide quality assurance for their measurements as summarized in Table S1 (ESM).

Review papers recommending SRM use Several workshop recommendations looked forward to the highly anticipated release of SRM 972 and its impact on standardization of measurements for 25(OH)D [31, 91]. In the reviews by Volmer and coworkers [10, 19], the first paper [19] describes the need for and the beginnings of the efforts to standardize 25(OH)D measurements, and in the second, Alexandridou et al. [10] outline the progress that has been made including the availability of SRMs and concludes with the statement "…..successful harmonization and standardization of mass analytical procedures has strongly improved reliability and comparability of vitamin D assays on a global scale." [10]. Other recent review papers on the status of measurements of vitamin D metabolites [11, 13, 14] have included extensive sections on quality assurance and have attributed improvements in comparability and accuracy of results to quality assurance activities including the availability of SRMs. In the review by Volmer et al. [19] citing 32 papers using LC–MS/MS for 25(OH)D, the authors indicated that 10 of these papers report using SRM 972/972a for quality control and validation of their methods.

Method development and validation In the early 2010s, chromatography method development for $25(OH)D_2$ and $25(OH)D_3$ often focused on the separation of 3-*epi*-$25(OH)D_3$ and $25(OH)D_3$. Several of these method validation studies found that their LC-UV and LC–MS/MS methods were deficient due

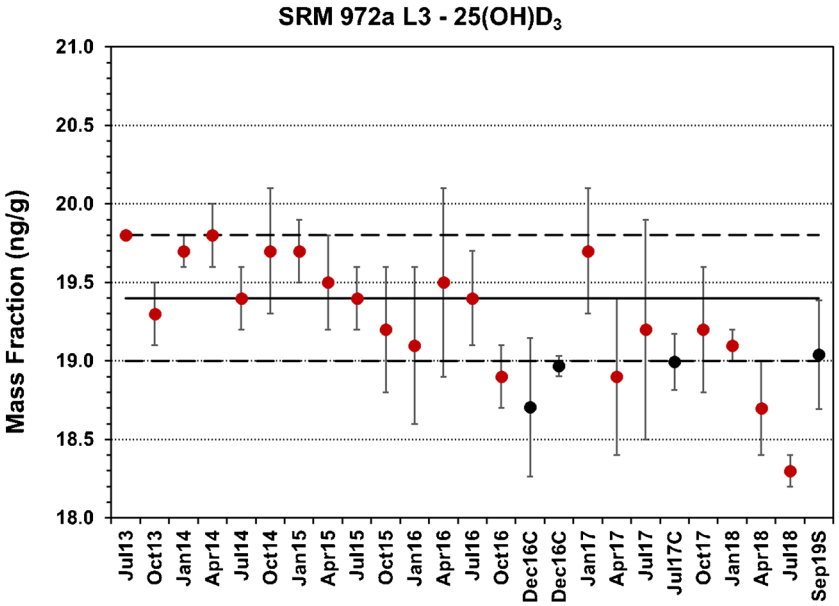

Fig. 7 Control chart for the NIST determination of 25(OH)D_3 in SRM 972a L3 for quarterly DEQAS exercises from July 2013 to July 2018 (red dots) and a commutability study and SRM certification measurements (black dots). Error bars are ± SD for duplicate analyses of SRM in DEQAS measurements and the commutability study and for 4 replicates in SRM certification measurements. The solid line is the certified value, and the dashed line is the uncertainty of the certified value. The *x*-axis are the dates (month and year) for the quarterly DEQAS exercises and dates ending with C and S refer to the commutability study and the SRM measurements, respectively. Adapted from Burdette et al. [82] and expanded with additional results

to the lack of separation of the 3-*epi*-25(OH)D_3 and 25(OH)D_3, which was often demonstrated by using SRM 972 L4, containing a high exogeneous concentration of 3-*epi*-25(OH)D_3. Hymøller and Jensen [92], Bogusz et al. [93], Adamec et al. [94], and Mochizuki et al. [95] developed LC-UV or LC–MS/MS methods for 25(OH)D_2 and 25(OH)D_3 and used SRM 972 for method validation achieving good agreement with the certified values except for L4 due to their inability to separate the 3-epimer. Strathmann et al. [96] investigated the significance of the contribution of 3-*epi*-25(OH)D_3 in measurements of total 25(OH)D and attributed the lack of agreement of immunoassay and LC–MS/MS results to the presence of 3-*epi*-25(OH)D_3; they used SRM 972 "To rule out the possibility that the calibration could have resulted in the differences we observed…which demonstrated 100–108% (mean 103.3%) recovery of non-epimeric 25(OH)D_3" [96].

A critical part of the development and validation of the University of Ghent and CDC RMPs for 25(OH)D_2 and 25(OH)D_3 was the use of SRM 2972 as a calibrator and SRM 972 and SRM 972a as precision and/or accuracy assessment tools. For the Ghent RMP, Stepman et al. [39] used SRM 2972 to value assign the calibrants for 25(OH)D_2 and 25(OH)D_3 through a direct comparison denoted as the "calibration transfer protocol." SRM 972 was used as part of the precision, trueness/accuracy, and limits of quantification assessments for the RMP. Using 10 replicates of SRM 972, the trueness, expressed as % recovery compared to certified values, was determined to be 101.0 to 101.7% (95% CI) for 25(OH)D_2 for three SRM levels and 99.4 to 101.9% for 25(OH)D_3 for all four SRM levels. Later, as part of a commutability study for SRM 972a [97], the NIST and Ghent RMPs were compared directly for a set of 50 single-donor samples and the combined results were used to assign target values. The results demonstrated a slight difference between the RMPs for total 25(OH)D with the NIST RMPs providing slightly lower results, i.e., − 0.52 ng/mL ($u = 0.04$ ng/mL), corresponding to a relative average difference between methods of < 0.8% based on the median concentration of total 25(OH)D in the samples [97].

Before developing the CDC RMP, Schleicher et al. [27] developed an ID LC–MS/MS method for 25(OH)D_2, 25(OH)D_3, and 3-*epi*-25(OH)D_3 with calibration traceable to SRM 2972 and validation using SRM 972. Using data from 44 replicates of all four levels of SRM 972, the bias (95% CI) for 25(OH)D_2, 25(OH)D_3, and 3-*epi*-25(OH)D_3 compared to the certified values was − 2 to 5% (three levels), − 2 to 0% (four levels), and 4% (one level), respectively. The authors pointed out that when compared to assigned reference values the bias was significantly higher, i.e., − 18% for 25(OH)D_2 and 18 to 29% for 3-*epi*-25(OH)D_3, indicating the importance of assigning certified values, if possible. The CDC RMP is based on the method developed by Mineva et al. [40], and a key element of the method validation was using

SRM 972/972a for accuracy assessment. Based on 10 independent measurements of the SRM 972 (L1 and L3) and SRM 972a (L2, L3, and L4), trueness was assessed by the bias compared to certified values for 25(OH)D_2 at 98.5 to 101.4% (three levels), for 25(OH)D_3 at 99.0 to 101.7% (four levels), and for 3-*epi*-25(OH)D_3 at − 1.0% (one level). As another approach to assessing method trueness, the authors compared their candidate RMP to the Ghent RMP using 40 patient samples and observed a mean bias of − 0.9% for 25(OH)D_3, 2.3% for 25(OH)D_2, and 0.2% for total 25(OH)D and a linear regression line with slope of 1.0188.

Dowling et al. [98] developed an LC–MS/MS method for 25(OH)D_2 and 25(OH)D_3 with a focus on including 3-*epi*-25(OH)D_3 and 24,25$(OH)_2D_3$ in the method, which often interfere in many LC–MS/MS assays and immunoassays. Validation of the method for all four metabolites was accomplished using SRM 972a with observed bias of − 1.7%, − 1.5%, − 4.9%, and 12.0% for 25(OH)D_2, 25(OH)D_3, 3-*epi*-25(OH)D_3, and 24,25$(OH)_2D_3$, respectively. The paper by Dowling et al. [98] in 2017 was the first to report using SRM 972a to validate measurements of 24,25$(OH)_2D_3$, which had only recently had values assigned for this metabolite.

Mondello and coworkers [99, 100] reported method validation using SRM 972a for two newly developed methods based on GC–MS/MS and UPLC-MS/MS. Micalizzi et al. [100] developed a novel GC–MS/MS method using derivatization and mild ionization conditions for determining vitamin D metabolites in less than 12 min. Bias compared to the certified values for 25(OH)D_3 was within 5% for three of the levels and the fourth was at 12%. For the higher level of 25(OH)D_2 in L3, the GC–MS/MS method was within 10% of the certified values; however, for the low levels in the other three SRM levels, the GC–MS/MS method provided higher results (116 to 392%). For 24,25$(OH)_2D_3$, the agreement with the certified values was excellent for the two levels above the LoQ (L1 and L4). For 3-*epi*-25(OH)D_3, only L4 had a value above the LoQ with bias 4% lower than the certified value. The authors provided examples of GC–MS/MS chromatograms for the four levels of SRM 972a to illustrate the separation and analysis. Donnarumma et al. [99] described a UPLC-MS/MS method for total lipidome analysis and for targeted quantification of five vitamin D species. For 25(OH)D_3, the results were biased from − 0.2 to 7.9% compared with SRM 972a. For 25(OH)D_2, only L3 was above the LoQ and results agreed within − 2%. For the 24,25$(OH)_2D_3$, only L1 and L4 were above the LoQ, and the measurements were 11% and 15% higher than the certified values.

Because of the wide diversity of immunoassays and their response for 25(OH)D, the serum-based SRMs have also been valuable in assessing the comparability of measurements among these assays. Using SRM 972, Moon et al. [101] compared bias results for three immunoassays (Roche Elecsys, Siemens Centaur, and DiaSorin Liaison) and an LC–MS/MS assay. Only the LC–MS/MS assay had acceptable bias (mean bias of 1.9%) whereas the DiaSorin Liaison and Siemens ADVIA Centaur showed mean negative bias of − 14.7% and − 11.6%, respectively. The Roche Elecsys had a mean bias of 9.1%; however, individual levels of SRM 972 exhibited bias from 23 to 30% for L1 and L2 to 58% for L4, and − 25% for L3 with elevated level of exogeneous 25(OH)D_2. Janssen et al. [102] compared an LC method, LC–MS/MS method, a protein binding method (Roche), and five immunoassays (DiaSorin Liaison and RIA, IDS iSYS, Siemens ADVIA Centaur, and Abbott Architect) analyzing 60 patient samples for total 25(OH)D including SRM 972 to assess accuracy. The LC–MS/MS and the LC method were successful in achieving within 10% of the certified values except for L4 since neither of these chromatographic methods separated the epimer. For the RIA, protein binding assay, and immunoassays, only the RIA and protein binding assay achieved within 10% for the unmodified L1. However, for L2, L3, and L4, which were modified with horse serum or exogeneous vitamin metabolites, none of the non-chromatographic assays provided results within 10% of the certified values. Interestingly, the authors commented that since it had been suggested that the modified levels of SRM 972 were inappropriate when using binding assays, they analyzed a DEQAS sample with high endogenous 25(OH)D_2 and found that it gave results like SRM 972 L3.

Use of SRMs and standardization protocols within national survey programs

Changes in measurements of 25(OH)D over time within NHANES catalyzed the development of SRMs and other efforts to standardize these measurements. NHANES 2007–2010 was the first NHANES to report 25(OH)D concentrations determined using a "standardized" LC–MS/MS assay [55] based on the method of Schleicher et al. [27] and incorporating SRM 972 for QC. Over 15,000 serum samples were analyzed with in-house QC pools ($n = 250$ to 260) and SRMs ($n = 37$) included. As part of the Electronic Supplemental Material, Schleicher et al. [27] demonstrate accuracy as % bias from certified values for 25(OH)D_3 of − 2.8 to 4.1% (four levels); for 25(OH)D_2 of 0.5% and 1.9% (L4 and L3), and 15% (L2 lowest certified concentration); and for 3-*epi*-25(OH)D_3 of 1.4% (L1 only certified value). The current CDC Laboratory Procedures Manual for NHANES measurements of 25(OH)D_2, 25(OH)D_3, and 3-*epi*-25(OH)D_3 using LC–MS/MS [103] states that SRM 972a is tested four times per year and SRM 2972a is used to verify stock solutions as needed. The CDC Reference Method Laboratory

for Vitamin D indicates on their website that their method is traceable to SRM 2972a [85].

Traceability to CRMs/RMPs is a critical aspect of a VDSP protocol to standardize serum 25(OH)D measurements retrospectively described by Durazo-Arvizu et al. [104]. The protocol requires that a specified number of stored samples from the previous study be re-analyzed using a laboratory's current assay (preferably an LC–MS/MS assay), in conjunction with 40 to 50 single-donor samples with values assigned by or traceable to RMPs and CRMs [104]. Using this approach, the NHANES measurements from 1988 to 2010 were retrospectively standardized as reported in Schleicher et al. [105]. Accuracy of the LC–MS/MS measurements was demonstrated by bias relative to SRMs during the course of re-analysis of specimens, and for the NHANES 2007–2010 analyses, bias was "minimal < 1% for 25(OH)D_3 and 25(OH)D_2 > 2 nmol/L" [105].

Cashman et al. [106] reported the first of several retrospective standardizations of 25(OH)D results using the Irish National Adult Nutrition Survey. In this case study, the samples were originally analyzed using the IDS ELISA and were re-analyzed with an LC–MS/MS assay traceable to NIST and Ghent RMPs through the use of a 50 single-donor sample set (as in the Schleicher et al. [105] study) from the first VDSP commutability study [23, 75]. The standardized results showed that the year-round prevalence rate for serum 25(OH)D concentrations < 30 nmol/L indicating deficiency increased from 6.5 to 11.4% [106]. Subsequent retrospective standardizations of 25(OH)D results from past surveys by Cashman et al. [107, 108] have relied on their institution's "certified" and "standardized" LC–MS/MS method traceable to the NIST RMPs and the CDC VDSCP.

The VDSP standardization approach was also implemented for national surveys in Canada [109] and Germany [110]. For the Canadian Health Measures Survey (CHMS), Sirafin et al. [109] standardized results from 2007 to 2011 surveys generated using the DiaSorin Liaison assay available in 2011 to the DiaSorin Liaison assay available in 2015 through analysis of the 50 single-donor samples from the VDSP commutability study [75, 97] to gain traceability to the NIST and Ghent RMPs. Standardization of the 25(OH)D measurements allowed two sets of CHMS results to be combined providing a more thorough investigation of the factors that affect 25(OH)D status in Canadians by increasing the total number of survey participants. Overall, the standardized values for total 25(OH)D were lower thereby increasing the percentage of participants with 25(OH)D concentrations < 40 nmol/L. The standardized results also reversed an apparent time-dependent decrease in 25(OH)D status emphasizing the benefit of evaluating data from long-term surveys, which the authors identified as a "critical step in the development of government policies, requiring an understanding of time-dependent changes in status and differences in racial-ethnic groups that make up populations" [109].

Three previous German national health surveys were standardized using the VDSP protocol as reported by Rabenberg et al. [110]. In the original studies, the samples were analyzed using the DiaSorin assay and a subset of samples was re-analyzed using the certified University College Cork LC–MS/MS assay. The authors state that this work "highlights how standardization of 25(OH)D data has a substantial impact on estimates of vitamin D status in Germany including higher mean levels, higher prevalence of vitamin D sufficiency, and lower prevalence of vitamin D deficiency overall as well as in age- and sex-specific analyses" [110]. According to Durazo-Arvizu et al. [104], a total of 23 studies have undergone retrospective standardization using the VDSP protocol. A review by Frazer et al. [13] comments that "Standardization of 25(OH)D assays has begun to show an impact on the estimation of the prevalence of vitamin D deficiency on a global scale…….The high prevalence of vitamin D deficiency across the vast geographical footprint represents a major public health concern on an alarming scale."

Traceability to CRMs

The examples above describe how researchers have used the SRMs in validating and assuring the quality of their 25(OH)D measurements. However, several papers claim only that their measurements are traceable to the CRM/SRM and/or to NIST (see Table S1) without providing any specific details and/or evidence of the traceability linkage. Several commercial LC–MS/MS assay manufacturers claim traceability of their methods to NIST SRMs including Chromsystems (Chromsystems Instruments & Chemicals GmbH, Gräfelfing, DE) and ClinMass (RECIPE Chemicals & Instruments GmbH, Munich, Germany). Commercial producers of calibrators or serum-based quality control materials often state that they are traceable to SRM 2972a or SRM 972a including Golden West Diagnostics LLC (Temecula, CA, USA) (e.g., denoted in ref. [111] as "NIST-compliant"), Waters Mass Trak [112], and Quantimetrix Complete D 25-OH Vitamin D Control (Redondo Beach, CA, USA) [113]. Metrological traceability to the SI can be achieved through appropriate use of a CRM; however, the concept of traceability to NIST may be inappropriate in these statements [114].

Next generation of vitamin D metabolite SRMs

SRM 972a has been available for over a decade and the inventory of the four-level material will be exhausted in several years. SRM 2973, SRM 2969, and SRM 2970 have been developed to complement SRM 972a and should be available for the next 5 to 10 years. SRM 972a has been

extremely popular, significant amounts of information have been published related to it, and users have provided feedback on its use, thereby raising the question—what should the next generation of SRMs for vitamin D metabolites look like? Should it continue as multiple levels spanning a similar range of concentrations of 25(OH)D or should there be multiple single-level SRMs? Is the SRM 972a L4 with the high concentration of exogenous 3-*epi*-25(OH)D_3 still useful? An SRM with a high concentration of 3-*epi*-25(OH)D_3 was useful in identifying LC–MS/MS methods that were biased because they did not separate the epimer; however, use of such LC–MS/MS methods should be mostly in the past. These are questions that need to be addressed in the design of the next generation of SRMs for vitamin D metabolites.

Conclusions

The five current serum matrix SRMs consisting of 11 concentration levels of 25(OH)D, including different levels of 25(OH)D_2, 25(OH)D_3, 3-*epi*-25(OH)D_3, and 24,25$(OH)_2D_3$, and information on VDBP and free 25(OH)D content, are valuable tools for determining accuracy and comparability of measurements to assess vitamin D status. The NIST and commercial vendor CRM calibration solutions also provide a readily useable, critical tool for establishing traceability of calibrants to the SI. These SRMs are critical for new method validation and routine quality control as demonstrated by sales, reported use in the literature, and claims of traceability for commercial assays and calibrators. These SRMs and the RMPs provide a foundation for programs to assure the quality of vitamin D metabolite measurements, including DEQAS, CAP ABVD, and CDC-VDSCP. Standardization of results from previous national surveys has allowed broader and more precise assessments of vitamin D deficiency and sufficiency among various populations.

Supplementary Information The online version contains supplementary material available at https://doi.org/10.1007/s00216-024-05143-w.

Acknowledgements The authors acknowledge Paul M. Coates (NIH ODS, retired), Joseph M Betz (NIH ODS, retired), and Mary Francis Picciano (NIH ODS, deceased) for their foresight and support in the development of vitamin D metabolite SRMs. We acknowledge Christopher T. Sempos (NIH-ODS, retired) for his leadership of the Vitamin D Standardization Program. We also acknowledge the NIST chemists involved in vitamin D metabolite measurements: Mary Bedner, Carolyn Q. Burdette, Grace Hahm, and particularly Susan C. Tai for her exceptional work in the development of RMPs and the certification of SRMs. We acknowledge Karen W. Phinney (NIST) for leadership in the development of SRM 972 and SRM 972a and David Duewer (NIST) for statistical consultation. We acknowledge Rosemary L. Schleicher (CDC, retired) and Christine M. Pfieffer (CDC) and CDC chemists for their measurements contributing to the certification of SRM 972 and SRM 972a and for coordination and participation in commutability studies.

Declarations

Disclaimer The opinions expressed in this article are the authors' own and do not reflect the view of the National Institutes of Health, the Department of Health and Human Services, the National Institute of Standards and Technology, the Department of Commerce, or the United States government. Any citations of commercial organizations and trade names in this report do not constitute an official NIH or NIST endorsement of approval of the products or services of these organizations.

Conflict of interest S. A. Wise is an Editor of the journal *Analytical and Bioanalytical Chemistry* and was not involved in peer reviewing this manuscript. There are no financial or nonfinancial conflicts of interest for any of the coauthors.

References

1. Wise SA. From urban dust and marine sediment to ginkgo biloba and human serum - A top ten list of Standard Reference Materials (SRMs). Anal Bioanal Chem. 2022;414:31–52.
2. Leech DP. The Economic Impacts of NIST's Cholesterol Standards Program. National Institute of Standards and Technology (NIST) Gaitherburg, MD, 2000.
3. Sempos CT, Vesper HW, Phinney KW, Thienpont LM, Coates PM. VDSP, Vitamin D status as an international issue: National surveys and the problem of standardization. Scand J Clin Lab Invest. 2012;72:32–40.
4. Wise SA, Tai SSC, Burdette CQ, Camara JE, Bedner M, Lippa KA, Nelson MA, Nalin F, Phinney KW, Sander LC, Betz JM, Sempos CT, Coates PM. Role of the National Institute of Standards and Technology (NIST) in support of the vitamin D initiative of the National Institutes of Health, Office of Dietary Supplements. J AOAC Int. 2017;100:1260–76.
5. 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.
6. Lensmeyer G, Poquette M, Wiebe D, Binkley N. The C-3 epimer of 25-hydroxyvitamin D3 Is present in adult serum. J Clin Endocrinol Metab. 2012;97:163–8.
7. Bikle DD, Malmstroem S, Schwartz J. Current controversies are free vitamin metabolite levels a more accurate assessment of vitamin D status than total levels? Endocrinol Metabol Clin North Amer. 2017;46:901–18.
8. Heureux N. Vitamin D testing-where are we and what is on the horizon? In: Makowski GS, editor. Advances in Clinical Chemistry, vol. 78. San Diego: Elsevier Academic Press Inc; 2017. p. 59–101.
9. Tsuprykov O, Buse C, Skoblo R, Hocher B. Comparison of free and total 25-hydroxyvitamin D in normal human pregnancy. J Steroid Biochem Mol Biol. 2019;190:29–36.
10. Alexandridou A, Schorr P, Stokes CS, Volmer DA. Analysis of vitamin D metabolic markers by mass spectrometry: Recent progress regarding the "gold standard" method and integration into clinical practice. Mass Spectrom Rev. 2023;42:1647–87.
11. Alonso N, Zelzer S, Eibinger G, Herrmann M. Vitamin D metabolites: Analytical challenges and clinical relevance. Calcif Tissue Int. 2023;112:158–77.
12. Altieri B, Cavalier E, Bhattoa HP, Perez-Lopez FR, Lopez-Baena MT, Perez-Roncero GR, Chedraui P, Annweiler C, Della Casa S, Zelzer S, Herrmann M, Faggiano A, Colao A, Holick MF.

Vitamin D testing: advantages and limits of the current assays. Eur J Clin Nutr. 2020;74:231–47.
13. Fraser WD, Tang JCY, Dutton JJ, Schoenmakers I. Vitamin D measurement, the debates continue, new analytes have emerged, developments have variable outcomes. Calcif Tissue Int. 2020;106:3–13.
14. Herrmann M. Assessing vitamin D metabolism - four decades of experience. Clin Chem Lab Med. 2023;61:880–94.
15. Volmer DA, Muller MJ, Lammert F, Stokes CS. Measurement of vitamin D Metabolic markers in human samples. Anticancer Res. 2015;35:3607–8.
16. Carter GD, Berry J, Durazo-Arvizu R, Gunter E, Jones G, Jones J, Makin HLJ, Pattni P, Sempos CT, Twomey P, Williams EL, Wise SA. Hydroxyvitamin D assays: An historical perspective from DEQAS. J Steroid Biochem Mol Biol. 2018;177:30–5.
17. Centers for Disease Control and Prevention (CDC). Vitamin D Standardization — Certification Program, Centers for Disease Control and Prevention, Atlanta, GA, USA. https://www.cdc.gov/labstandards/vdscp.html. Accessed 8 Jan 2024.
18. Stokes CS, Lammert F, Volmer DA. Analytical methods for quantification of vitamin D and implications for research and clinical practice. Anticancer Res. 2018;38:1137–44.
19. Volmer DA, Mendes L, Stokes CS. Analysis of vitamin D metabolic markers by mass spectrometry: Current techniques, limitations of the "gold standard" method and anticipated future directions. Mass Spectrom Rev. 2015;34:2–23.
20. Muller MJ, Volmer DA. Mass spectrometric profiling of vitamin D metabolites beyond 25-hydroxyvitamin D. Clin Chem. 2015;61:1033–48.
21. Singh RJ, Taylor RL, Reddy GS, Grebe SKG. 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.
22. Wise SA, Camara JE, Sempos CT, Burdette CQ, Hahm G, Nalin F, Kuszak AJ, Merkel J, Durazo-Arvizu R, Hoofnagle AN, Williams EL, Ivison F, Fischer R, Van den Ouweland JMW, Ho CS, Law EWK, Simard J-N, Gonthier R, Holmquist B, Meadows S, Cox L, Robyak K, Creer MH, Fitzgerald R, Clarke MW, Breen N, Lukas P, Cavalier E. Interlaboratory comparison of 25-hydroxyvitamin D assays: Vitamin D Standardization Program (VDSP) intercomparison study 2 – Part 1 Liquid chromatography – tandem mass spectrometry (LC-MS/MS) assays – Impact of 3-epi-25-hydroxyvitamin D3 on assay performance. Anal Bioanal Chem. 2022;414:333–49.
23. Wise SA, Phinney KW, Tai SSC, Camara JE, Myers GL, Durazo-Arvizu R, Tian L, Hoofnagle AN, Bachmann LM, Young IS, Pettit J, Caldwell G, Liu A, Brooks SPJ, Sarafin K, Thamm M, Mensink GBM, Busch M, Rabenberg M, Cashman KD, Kiely M, Kinsella M, Galvin K, Zhang JY, Oh K, Lee SW, Jung CL, Cox L, Goldberg G, Guberg K, Prentice A, Carter GD, Jones J, Brannon PM, Lucas RM, Crump PM, Cavalier E, Merkel J, Betz JM, Sempos CT. Baseline assessment of 25-hydroxyvitamin D assay performance: A Vitamin D Standardization Program (VDSP) interlaboratory comparison study. J AOAC Int. 2017;100:1244–52.
24. Lensmeyer GL, Wiebe DA, Binkley N, Drezner MK. HPLC method for 25-hydroxyvitamin D measurement: Comparison with contemporary assays. Clin Chem. 2006;52:1120–6.
25. Bedner M, Phinney KW. Development and comparison of three liquid chromatography-atmospheric pressure chemical ionization/mass spectrometry methods for determining vitamin D metabolites in human serum. J Chromatogr A. 2012;1240:132–9.
26. Newman MS, Brandon TR, Groves MN, Gregory WL, Kapur S, Zava D. A liquid chromatography/tandem mass spectrometry method for determination of 25-hydroxy vitamin D2 and 25-hydroxy vitamin D3 in dried blood spots: A potential adjunct to diabetes and cardiometabolic risk screening. J Diabetes Sci Technol. 2009;3:156–62.
27. 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.
28. van den Ouweland JMW, 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.
29. Looker AC, Pfeiffer CM, Lacher DA, Schleicher RL, Picciano MF, Yetley EA. Serum 25-hydroxyvitamin D status of the US population: 1988–1994 compared with 2000–2004. Am J Clin Nutr. 2008;88:1519–27.
30. Pollack A. Quest acknowledges errors in vitamin D tests. New York Times. 2009.
31. Yetley EA, Pfeiffer CM, Schleicher RL, Phinney KW, Lacher DA, Christakos S, Eckfeldt JH, Fleet JC, Howard G, Hoofnagle AN, Hui SL, Lensmeyer GL, Massaro J, Peacock M, Rosner B, Wiebe D, Bailey RL, Coates PM, Looker AC, Sempos C, Johnson CL, Picciano MF. NHANES monitoring of serum 25-hydroxyvitamin D: A roundtable summary. J Nutr. 2010;140:2030S-2045S.
32. Carter GD. 25-Hydroxyvitamin D assays: The quest for accuracy. Clin Chem. 2009;55:1300–2.
33. Tai SSC, 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.
34. International Bureau of Weights and Measures (BIPM). Joint Committee for Guides in Metrology (JCGM), International Vocabulary of Metrology - Basic and General Concepts and Associated Terms (VIM) 3rd Ed, 2012. Available bipm.org/en/committees/jc/jcgm/publications. Accessed 6 Jan 2024.
35. Panteghini M. Traceability, reference systems and result comparability, Biochimica. Clinica. 2007;31:247–53.
36. International Organization for Standardization (ISO). In Vitro Diagnostic Devices - Measurement of quantities in samples of biological origin - Requirements for content and presentation of reference measurement procedures, 2009.
37. Schimmel H, Zegers I. Performance criteria for reference measurement procedures and reference materials. Clin Chem Lab Med. 2015;53:899–904.
38. May WE, Parris RM, Beck II CM, Fassett JD, Greenberg RR, Guenther FR, Kramer GW, Wise SA, Gills TE, Gettings R, MacDonald B. Definitions of terms and modes used at NIST for value assignment of reference materials for chemical measurment, NIST Special Publication 260–136, National Institute of Standards and Technology, Gaithersburg, MD, USA, 2000.
39. Stepman HCM, Vanderroost A, Van Uytfanghe K, Thienpont LM. Candidate reference measurement procedures for serum 25-hydroxyvitamin D3 and 25-hydroxyvitamin D2 by using isotope dilution liquid chromatography-tandem mass spectrometry. Clin Chem. 2011;57:441–8.
40. Mineva EM, Schleicher RL, Chaudhary-Webb M, Maw KL, Botelho JC, Vesper HW, Pfeiffer CM. A candidate reference measurement procedure for quantifying serum concentrations of 25-hydroxyvitamin D3 and 25-hydroxyvitamin D2 using isotope-dilution liquid chromatography-tandem mass spectrometry. Anal Bioanal Chem. 2015;407:5615–24.

41. Ross AC, Taylor CL, Yaktine AL, Del Valle HB. Dietary Reference Intakes for Calcium and Vitamin D. Washington DC: National Academies Press; 2011.
42. Sempos CT, Binkley N. 25-Hydroxvitamin D assay standardization and vitamin D guidelines paralysis. Public Health Nutr. 2020;23:1153–64.
43. Epstein MS. The independent method concept for certifying chemical composition reference materials, Spectrochimica Acta Part B-Atomic. Spectroscopy. 1991;46:1583–91.
44. Wise SA, Phillips MM. Evolution of reference materials for the determination of organic nutrients in food and dietary supplements - a critical review. Anal Bioanal Chem. 2019;411:97–127.
45. Wise SA, Phinney KW, Sander LC, Schantz MM. Role of chromatography in the development of Standard Reference Materials for organic analysis. J Chromatogr A. 2012;1261:3–22.
46. Wise SA, Poster DL, Kucklick JR, Keller JM, VanderPol SS, Sander LC, Schantz MM. Standard reference materials (SRMs) for determination of organic contaminants in environmental samples. Anal Bioanal Chem. 2006;386:1153–90.
47. Beauchamp CRC, Camara JE, Carney J, Choquette SJ, Cole KD, DeRose PC, Duewer DL, Epstein MS, Kline MC, Lippa KA, Lucon E, Malloy J, Nelson MA, Phnney KW, Polakoski M, Possolo A, Sander LC, Schiel JE, Sharpless KE, Toman B, Winchester MR, Windover D. Metrological tools for the reference materials and reference instruments of the NIST Materials Measurement Laboratory. Washington, DC: NIST Special Publication 260–136, U.S. Government Printing Office; 2021.
48. Beauchamp CRC, Carney JE, Choquette SJ, Cole KD, DeRose PC, Epstein MS, Kline MC, Lippa KA, Lucon E, Phnney KW, Polakoski M, Sharpless KE, Sieber JR, Toman B, Winchester MR, Windover D. Metrological tools for the reference materials and reference instruments of the NIST Materials Measurement Laboratory. Washington, DC: NIST Special Publication 260–136, U.S. Government Printing Office; 2020.
49. Ellerbe P, Meiselman S, Sniegoski LT, Welch MJ, White E. Determination of serum-cholesterol by a modification of the isotope-dilution mass-spectrometric definitive method. Anal Chem. 1989;61:1718–23.
50. Welch MJ, Cohen A, Hertz HS, Ng KJ, Schaffer R, Vanderlijn P, White E. Determination of serum creatinine by isotope dilution mass-spectrometry as a candidate definitive method. Anal Chem. 1986;58:1681–5.
51. Phinney KW, Bedner M, Tai SSC, Vamathevan VV, Sander LC, Sharpless KE, Wise SA, Yen JH, Schleicher RL, Chaudhary-Webb M, Pfeiffer CM, Betz JM, Coates PM, Picciano MF. Development and certification of a Standard Reference Material for vitamin D metabolites in human serum. Anal Chem. 2012;84:956–62.
52. Phinney KW, Ballihaut G, Bedner M, Benford BS, Camara JE, Christopher SJ, Davis WC, Dodder NG, Eppe G, Lang BE, Long SE, Lowenthal MS, McGaw EA, Murphy KE, Nelson BC, Prendergast JL, Reiner JL, Rimmer CA, Sander LC, Schantz MM, Sharpless KE, Sniegoski LT, Tai SSC, Thomas JB, Vetter TW, Welch MJ, Wise SA, Wood LJ, Guthrie WF, Hagwood CR, Leigh SD, Yen JH, Zhang NF, Chaudhary-Webb M, Chen HP, Fazili Z, LaVoie DJ, McCoy LF, Momin SS, Paladugula N, Pendergrast EC, Pfeiffer CM, Powers CD, Rabinowitz D, Rybak ME, Schleicher RL, Toombs BMH, Xu M, Zhang M, Castle AL. Development of a Standard Reference Material for metabolomics research. Anal Chem. 2013;85:11732–8.
53. Boggs ASP, Kilpatrick LE, Burdette CQ, Tevis DS, Fultz ZA, Nelson MA, Jarrett JM, Kemp JV, Singh RJ, Grebe SKG, Wise SA, Kassim BL, Long SE. Development of a pregnancy-specific reference material for thyroid biomarkers, vitamin D, and nutritional trace elements in serum. Clin Chem Lab Med. 2021;59:671–9.
54. Phinney KW, Tai SSC, Bedner M, Camara JE, Chia RRC, Sander LC, Sharpless KE, Wise SA, Yen JH, Schleicher RL, Chaudhary-Webb M, Maw KL, Rahmani Y, Betz JM, Merkel J, Sempos CT, Coates PM, Durazo-Arvizu RA, Sarafin K, Brooks SPJ. Development of an improved Standard Reference Material for vitamin D metabolites in human serum. Anal Chem. 2017;89:4907–13.
55. Schleicher RL, Sternberg MR, Looker AC, Yetley EA, Lacher DA, Sempos CT, Taylor CL, Durazo-Arvizu RA, Maw KL, Chaudhary-Webb M, Johnson CL, Pfeiffer CM. National estimates of serum total 25-hydroxyvitamin D and metabolite concentrations measured by liquid chromatography-tandem mass spectrometry in the US population during 2007–2010. J Nutr. 2016;146:1051–61.
56. Tai SSC, Nelson MA, Bedner M, Lang BE, Phinney KW, Sander LC, Yen JH, Betz JM, Sempos CT, Wise SA. Development of Standard Reference Material (SRM) 2973 vitamin D metabolites in frozen human serum (high level). J AOAC Int. 2017;100:1294–303.
57. Hahm G, Nelson M, Camara J, Toman B. Certification of Standard Reference Materials 2969 and 2970: Vitamin D Metabolites in Frozen Human Serum (Total 25-Hydroxyvitamin D Low Level) and (25-Hydroxyvitamin D_2 High Level). National Institute of Standards and Technology (NIST), Gaithersburg, MD, 2021.
58. Cashman KD, Hayes A, Galvin K, Merkel J, Jones G, Kaufmann M, Hoofnagle AN, Carter GD, Durazo-Arvizu RA, Sempos CT. Significance of serum 24,25-dihydroxyvitamin D in the assessment of vitamin D status: A double-edged sword? Clin Chem. 2015;61:636–45.
59. Tai SSC, Nelson MA. Candidate reference measurement procedure for the determination of (24R),25-dihydroxyvitamin D3 in human serum using isotope dilution liquid chromatography-tandem mass spectrometry. Anal Chem. 2015;87:7964–70.
60. Best CM, Pressman EK, Queenan RA, Cooper E, O'Brien KO. Longitudinal changes in serum vitamin D binding protein and free 25-hydroxyvitamin D in a multiracial cohort of pregnant adolescents. J Steroid Biochem Mol Biol. 2019;186:79–88.
61. Jones KS, Assar S, Prentice A, Schoenmakers I. Vitamin D expenditure is not altered in pregnancy and lactation despite changes in vitamin D metabolite concentrations. Sci Rep. 2016;6:26795.
62. Tsuprykov O, Buse C, Skoblo R, Hocher B. Free 25 (OH) vitamin D, but not total 25 (OH) vitamin D, is strongly correlated with gestational age and calcium in normal human pregnancy. J Bone Miner Res. 2017;32:S323.
63. Zhang JY, Lucey AJ, Horgan R, Kenny LC, Kiely M. Impact of pregnancy on vitamin D status: a longitudinal study. Br J Nutr. 2014;112:1081–7.
64. Henderson CM, Lutsey PL, Misialek JR, Laha TJ, Selvin E, Eckfeldt JH, Hoofnagle AN. Measurement by a novel LC-MS/MS methodology reveals similar serum concentrations of vitamin D-binding protein in blacks and whites. Clin Chem. 2016;62:179–87.
65. Kilpatrick LE, Phinney KW. Quantification of total vitamin-D-binding protein and the glycosylated isoforms by liquid chromatography isotope dilution mass spectrometry. J Proteome Res. 2017;16:4185–95.
66. Kilpatrick LE, Boggs ASP, Davis WC, Long SE, Yen JH, Phinney KW. Assessing a method and reference material for quantification of vitamin D binding protein during pregnancy. Clinical Mass Spectrometry. 2020;16:11–7.
67. Kilpatrick LE, Bouillon R, Davis WC, Henderson CM, Hoofnagle AN, Pauwels S, Vanderschueren D, Waelkens E, Wildiers H, Yen JH, Phinney KW. The influence of proteoforms: assessing the accuracy of total vitamin D-binding protein

quantification by proteolysis and LC-MS/MS. Clin Chem Lab Med. 2023;61:78–85.
68. Bikle D, Bouillon R, Thadhani R, Schoenmakers I. Vitamin D metabolites in captivity? Should we measure free or total 25(OH)D to assess vitamin D status? J Steroid Biochem Mol Biol. 2017;173:105–16.
69. Tsuprykov O, Elitok S, Buse C, Chu C, Kramer BK, Hocher B. Opposite correlation of 25-hydroxyvitamin D and 1,25-dihydroxyvitamin D metabolites with gestational age, bone and lipid-biomarkers in pregnant women. Sci Rep. 2021;11:10.
70. Heureux N, Lindhout E, Swinkels L. A direct assay for measuring free 25-hydroxyvitamin D. J AOAC Int. 2017;100:1318–22.
71. Sempos CT, Lindhout E, Heureux N, Hars M, Parkington DA, Dennison E, Durazo-Arvizu R, Jones KS, Wise SA. Towards harmonization of directly measured free 25-hydroxyvitamin D using an enzyme-linked immunosorbent assay. Anal Bioanal Chem. 2022;414:7793–803.
72. Miller WG, Myers GL. Commutability still matters. Clin Chem. 2013;59:1291–3.
73. Vesper HW, Miller WG, Myers GL. Reference materials and commutability. Clin Biochem Rev. 2007;28:139–47.
74. International Standards Organization (ISO). In vitro diagnostic medical devices — Measurement of quantities in samples of biological origin — Requirements for certified reference materials and the content of supporting documentation, Geneva, Switzerland, 2009.
75. Phinney KW, Sempos CT, Tai SSC, Camara JE, Wise SA, Eckfeldt JH, Hoofnagle AN, Carter GD, Jones J, Myers GL, Durazo-Arvizu R, Miller WG, Bachmann LM, Young IS, Pettit J, Caldwell G, Liu A, Brooks SPJ, Sarafin K, Thamm M, Mensink GSM, Busch M, Rabenberg M, Cashman KD, Kiely M, Galvin K, Zhang JY, Kinsella M, Oh K, Lee SW, Jung CL, Cox L, Goldberg G, Guberg K, Meadows S, Prentice A. Baseline assessment of 25-hydroxyvitamin D reference material and proficiency testing/external quality assurance material commutability: A Vitamin D Standardization Program Study. J AOAC Int. 2017;100:1288–93.
76. Camara JE, Wise SA, Hoofnagle AN, Williams EL, Carter GD, Jones J, Burdette CQ, Hahm G, Nalin F, Kuszak AJ, Merkel J, Durazo-Arvizu RA, Lukas P, Cavalier E, Popp C, Beckert C, Schultess J, Van Slooten G, Tourneur C, Pease C, Kaul R, Villarreal A, Ivison F, Fischer R, van den Ouweland JMW, Ho CS, Law EWK, Simard JN, Gonthier R, Holmquist B, Batista MC, Pham H, Bennett A, Meadows S, Cox L, Jansen E, Khan DA, Robyak K, Creer MH, Kilbane M, Twomey PJ, Freeman J, Parker N, Yuan JY, Fitzgerald R, Mushtaq S, Clarke MW, Breen N, Simpson C, Sempos CT. Assessment of serum total 25-hydroxyvitamin D assay commutability of Standard Reference Materials and College of American Pathologists Accuracy-Based Vitamin D (ABVD) Scheme and Vitamin D External Quality Assessment Scheme (DEQAS) materials: Vitamin D Standardization Program (VDSP) Commutability Study 2. Anal Bioanal Chem. 2021;413:5067–84.
77. Wise SA, Cavalier E, Camara JE, Lukas P, Peeters S, Le Goff C, Minerva E, Pfeiffer CM, Vesper H, Briggs LE, Williams EL, Popp C, Beckert C, W. K., Tourneur C, Pease C, Osterritter D, Fischer R, Saida B, Dou C, Satoshi K, Weiler H, Bielecki A, Pham H, Bennett A, You S, Ghoshal AK, Vogel C, Freeman J, Parker N, Yang P, Cheek J, Li J, Tsukamoto H, Galvan K, Cashman KD, Lawrence L, Hoofnagle AN, Boggs ASP, Burdette CQ, Hahm G, Nalin F, Kuszak AJ. Commutability assessment of new Standard Reference Materials for serum total 25-hydroxyvitamin D using ligand binding assays and liquid chromatography – tandem mass spectrometry (LC-MS/MS). Anal Bioanal Chem. to be submitted 2024.
78. Nilsson G, Budd JR, Greenberg N, Delatour V, Rej R, Panteghini M, Ceriotti F, Schimmel H, Weykamp C, Keller T, Camara JE, Burns C, Vesper HW, MacKenzie F, Miller WG, Commutability IWG. IFCC working group recommendations for assessing commutability part 2: Using the difference in bias between a reference material and clinical samples. Clin Chem. 2018;64:455–64.
79. DEQAS, Vitamin D External Quality Assessment Scheme (DEQAS);www.deqas.org. Accessed 8 Jan 2024.
80. Bedner M, Lippa KA, Tai SS-C, Burdette CQ. NIST/NIH Vitamin D Metabolites Quality Assurance Program (VitDQAP): Final Report, NISTIR 8293, National Institute of Standards and Technology (NIST), Gaithersburg, MD, 2020.
81. Bedner M, Lippa KA, Tai SSC. An assessment of 25-hydroxyvitamin D measurements in comparability studies conducted by the Vitamin D Metabolites Quality Assurance Program. Clin Chim Acta. 2013;426:6–11.
82. Burdette CQ, Camara JE, Nalin F, Pritchett J, Sander LC, Carter GD, Jones J, Betz JM, Sempos CT, Wise SA. Establishing an accuracy basis for the Vitamin D External Quality Assessment Scheme (DEQAS). J AOAC Int. 2017;100:1277–87.
83. Erdman P, Palmer-Toy DE, Horowitz G, Hoofnagle A. Accuracy-based vitamin D survey six years of quality improvement guided by proficiency testing. Arch Pathol Lab Med. 2019;143:1531–8.
84. Bedner M, Lippa KA, STai SS-C. NIST/NIH Vitamin D Metabolites Quality Assurance Program Report of Participant Results: Winter 2012 Comparability Study (Exercise 5), National Institute of Standards and Technology (NIST), Gaithersburg, MD, 2013.
85. Centers for Disease Control and Prevention (CDC). Vitamin D Reference Method Laboratory, Atlanta, GA, https://www.cdc.gov/labstandards/csp/vdscp_laboratory.html. Accessed 8 Jan 2024.
86. Teo TL, Lippa KA, Mackay L, Yong S, Liu QD, Camara JE, Delatour V, Lee TK, Lalere B, O'Connor G, Henrion A, Kato M, Numata M, Kwon HJ, Jeong JS, Xu B, Song DW, Nammoonnoy J, Wollinger W. Enhancing the accuracy of measurement of small molecule organic biomarkers. Anal Bioanal Chem. 2019;411:7341–55.
87. Wise SA, Tai SSC, Duewer DL, Bedner M, Camara JE, Lippa KA, Qinde L, Kang D, Kim B, Quan C. CCQM-K132: low-polarity analytes in a biological matrix: vitamin D metabolites in human serum. Metrologia. 2017;54:08027.
88. Linsinger T. Comparison of a measurement result with the certified value, Application Note 1, Jan 2010, available at https://crm.jrc.ec.europa.eu/graphics/cms_docs/erm1_english.pdf. Accessed 15 Jan 2024.
89. Sharpless KE, Lippa KA, Duewer DL, Rukhin AL. The ABC's of Using Standard Reference Materials in the Analysis of Foods and Dietary Supplements: A Practical Guide. NIST Special Publication 260–136, National Institute of Standards and Technology (NIST), Gaithersburg, MD, USA, 2014.
90. S. Wood, Using reference materials to establish metrological traceability, Application Note 3, Feb. 2008 avialable at https://crm.jrc.ec.europa.eu/graphics/cms_docs/erm3_english.pdf. Accessed 15 Jan 2024.
91. 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.
92. Hymoller L, Jensen SK. Vitamin D analysis in plasma by high performance liquid chromatography (HPLC) with C-30 reversed phase column and UV detection - Easy and acetonitrile-free. J Chromatogr A. 2011;1218:1835–41.
93 Bogusz MJ, Al Enazi E, Tahtamoni M, Jawaad JA, Al Tufail M. Determination of serum vitamins 25-OH-D2 and 25-OH-D3 with liquid chromatography-tandem mass spectrometry using atmospheric pressure chemical ionization or electrospray source and

core-shell or sub-2 um particle columns: a comparative study. Clin Biochem. 2011;44:1329–37.
94. Adamec J, Jannasch A, Huang JJ, Hohman E, Fleet JC, Peacock M, Ferruzzi MG, Martin B, Weaver CM. Development and optimization of an LC-MS/MS-based method for simultaneous quantification of vitamin D2, vitamin D3, 25-hydroxyvitamin D2 and 25-hydroxyvitamin D3. J Sep Sci. 2011;34:11–20.
95. Mochizuki A, Kodera Y, Saito T, Satoh M, Sogawa K, Nishimura M, Seimiya M, Kubota M, Nomura F. Preanalytical evaluation of serum 25-hydroxyvitamin D3 and 25-hydroxyvitamin D2 measurements using LC-MS/MS. Clin Chim Acta. 2013;420:114–20.
96. Strathmann FG, Sadilkova K, Laha TJ, LeSourd SE, Bornhorst JA, Hoofnagle AN, Jack R. 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.
97. Phinney KW, Camara JE, Tai SSC, Sander LC, Wise SA, De Grande LAC, Thienpont LM, Possolo AM, Toman B, Sempos CT, Betz JM, Coates PM. Value assignment of vitamin D metabolites in Vitamin D Standardization Program serum samples. J AOAC Int. 2017;100:1253–9.
98. Dowling KG, Hull G, Sundvall J, Lamberg-Allardt C, Cashman KD. Improved accuracy of an tandem liquid chromatography-mass spectrometry method measuring 24R,25-dihydroxyvitamin D3 and 25-hydroxyvitamin D metabolites in serum using unspiked controls and its application to determining cross-reactivity of a chemiluminescent microparticle immunoassay. J Chromatogr A. 2017;1497:102–9.
99. Donnarumma D, Arena A, Trovato E, Rigano F, Zoccali M, Mondello L. A miniaturized comprehensive approach for total lipidome analysis and vitamin D metabolite quantification in human serum. Anal Bioanal Chem. 2023;415:4579–90.
100. Micalizzi G, Vento F, Buzzanca C, Salerno TMG, Mondello L. Fast gas chromatography-tandem mass spectrometry under milder electron ionization conditions for the assay of vitamin D metabolites in human serum, J Chromatogr B-Anal Technol Biomed Life Sci. 2023;1227:123813.
101. 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;45:326–30.
102. Janssen MJW, Wielders JPM, Bekker CC, Boesten LSM, Buijs MM, Heijboer AC, van der Horst FAL, Loupatty FJ, van den Ouweland JMW. Multicenter comparison study of current methods to measure 25-hydroxyvitamin D in serum. Steroids. 2012;77:1366–72.
103. Pirkle JL. Laboratory Procedure Manual 25-Hydroxyvitamin D3, C3-epimer-25-Hydroxyvitamin D3 and 25-Hydroxyvitamin D2 in Serum by High Performance Liquid Chromatography - Tandem Mass Spectrometry (Method 4027.07), Centers for Disease Control and Prevention (CDC), Atlanta, GA, USA, 2020.
104. Durazo-Arvizu RA, Tian L, Brooks SPJ, Sarafin K, Cashman KD, Kiely M, Merkel J, Myers GL, Coates PM, Sempos CT. The Vitamin D Standardization Program (VDSP) manual for retrospective laboratory standardization of serum 25-hydroxyvitamin D data. J AOAC Int. 2017;100:1234–43.
105. Schleicher RL, Sternberg MR, Lacher DA, Sempos CT, Looker AC, Durazo-Arvizu RA, Yetley EA, Chaudhary-Webb M, Maw KL, Pfeiffer CM, Johnson CL. The vitamin D status of the US population from 1988 to 2010 using standardized serum concentrations of 25-hydroxyvitamin D shows recent modest increases. Am J Clin Nutr. 2016;104:454–61.
106. Cashman KD, Kiely M, Kinsella M, Durazo-Arvizu RA, Tian L, Zhang Y, Lucey A, Flynn A, Gibney MJ, Vesper HW, Phinney KW, Coates PM, Picciano MF, Sempos CT. Evaluation of Vitamin D Standardization Program protocols for standardizing serum 25-hydroxyvitamin D data: a case study of the program's potential for national nutrition and health surveys. Am J Clin Nutr. 2013;97:1235–42.
107. Cashman KD, Dowling KG, Skrabakova Z, Gonzalez-Gross M, Valtuena J, De Henauw S, Moreno L, Damsgaard CT, Michaelsen KF, Molgaard C, Jorde R, Grimnes G, Moschonis G, Mavrogianni C, Manios Y, Thamm M, Mensink GBM, Rabenberg M, Busch MA, Cox L, Meadows S, Goldberg G, Prentice A, Dekker JM, Nijpels G, Pilz S, Swart KM, van Schoor NM, Lips P, Eiriksdottir G, Gudnason V, Cotch MF, Koskinen S, Lamberg-Allardt C, Durazo-Arvizu RA, Sempos CT, Kiely M. Vitamin D deficiency in Europe: pandemic? Am J Clin Nutr. 2016;103:1033–44.
108. Cashman KD, Dowling KG, Skrabakova Z, Kiely M, Lamberg-Allardt C, Durazo-Arvizu RA, Sempos CT, Koskinen S, Lundqvist A, Sundvall J, Linneberg A, Thuesen B, Husemoen LLN, Meyer HE, Holvik K, Gronborg IM, Tetens I, Andersen R. Standardizing serum 25-hydroxyvitamin D data from four Nordic population samples using the Vitamin D Standardization Program protocols: Shedding new light on vitamin D status in Nordic individuals. Scand J Clin Lab Invest. 2015;75:549–61.
109. Sarafin K, Durazo-Arvizu R, Tian L, Phinney KW, Tai S, Camara JE, Merkel J, Green E, Sempos CT, Brooks SPJ. Standardizing 25-hydroxyvitamin D values from the Canadian Health Measures Survey. Am J Clin Nutr. 2015;102:1044–50.
110. Rabenberg M, Scheidt-Nave C, Busch MA, Thamm M, Rieckmann N, Durazo-Arvizu RA, Dowling KG, Skrabakova Z, Cashman KD, Sempos CT, Mensink GBM. Implications of standardization of serum 25-hydroxyvitamin D data for the evaluation of vitamin D status in Germany, including a temporal analysis. BMC Public Health. 2018;18:14.
111. Miyamoto H, Kawakami D, Hanafusa N, Nakanishi T, Miyasaka M, Furutani Y, Ikeda Y, Ito K, Kato T, Yokoyama K, Arakawa S, Saito M, Furusho T, Matsuura T, Ochi S. Determination of a serum 25-hydroxyvitamin D reference ranges in Japanese adults using fully automated liquid chromatography - tandem mass spectrometry. J Nutr. 2023;153:1253–64.
112. Breen N, Wardle R, Davey L, Lambert P, Roche D, Crushell M, Calton LJ. Performance characteristics of the Mass Trak vitamin D solution, Application Note, Waters Corporation, 2019, revised April 2022, available at https://www.waters.com/nextgen/us/en/library/application-notes/2017/performance-characteristics-of-the-masstrak-vitamin-d-solution.html. Accessed 15 Jan 2024.
113. Quantimetrix, Complete D 25-OH Vitamin D Control, Quantimetrix, Redondo Beach, CA, USA, avialable at https://quantimetrix.com/product/complete-d-25-oh-vitamin-d-control/. Accessed 15 Jan 2024.
114. National Institue of Standards and Technology (NIST), NIST Policy on Metrological Traceability, 2021, available at NIST.gov/calibrations/traceability. Accessed 6 Jan 2024.