**NISTIR 8293**

# NIST/NIH Vitamin D Metabolites Quality Assurance Program (VitDQAP): Final Report

Mary Bedner
Katrice A. Lippa
Susan S.-C. Tai (retired)
Carolyn Q. Burdette

This publication is available free of charge from:
https://doi.org/10.6028/NIST.IR.8293

National Institute of Standards and Technology
U.S. Department of Commerce

**NISTIR 8293**

# NIST/NIH Vitamin D Metabolites Quality Assurance Program (VitDQAP): Final Report

Mary Bedner
Katrice A. Lippa
Susan S.-C. Tai (retired)
Carolyn Q. Burdette
*Chemical Sciences Division*
*Material Measurement Laboratory*

This publication is available free of charge from:
https://doi.org/10.6028/NIST.IR.8293

September 2020


U.S. Department of Commerce
*Wilbur L. Ross, Jr., Secretary*

National Institute of Standards and Technology
*Walter Copan, NIST Director and Undersecretary of Commerce for Standards and Technology*

**National Institute of Standards and Technology Interagency or Internal Report 8293**
**Natl. Inst. Stand. Technol. Interag. Intern. Rep. 8293, 39 pages (September 2020)**

**This publication is available free of charge from:**
**https://doi.org/10.6028/NIST.IR.8293**

## Abstract

The National Institute of Standards and Technology (NIST) established a Vitamin D Metabolites Quality Assurance Program (VitDQAP) in collaboration with the National Institutes of Health Office of Dietary Supplements (NIH-ODS) in 2009. The VitDQAP at NIST administered twelve interlaboratory comparison exercises through 2016 for the measurement of vitamin D metabolites in human serum and plasma. The comparability of participant measurements for target analytes improved over time through the development and promulgation of robust measurement technologies, identification and production of suitable reference materials, isolation and identification of measurement system biases, and support and encouragement of within-laboratory measurement quality control efforts. This VitDQAP Final Report summarizes the program and provides highlights of the twelve exercises conducted during the program's eight-year lifetime.



## Keywords

25-hydroxyvitamin $D_2$, 25-hydroxyvitamin $D_3$, 25-hydroxyvitamin $D_{Total}$, 3-epi-25-hydroxyvitamin $D_3$, human serum and plasma, interlaboratory comparison study

# Contents



# Tables

# Figures



# Acronyms and Symbols

| | |
|---|---|
| 25(OH)$D_2$ | 25-hydroxyvitamin $D_2$ |
| 25(OH)$D_3$ | 25-hydroxyvitamin $D_3$ |
| 25(OH)$D_{Total}$ | total 25-hydroxyvitamin D, the sum of 25(OH)$D_2$ and 25(OH)$D_3$ |
| 3-epi-25(OH)$D_3$ | 3-epimer form of 25-hydroxyvitamin $D_3$ |
| CAP | College of American Pathologists |
| CDC | Centers for Disease Control and Prevention |
| CLEIA | chemiluminescence enzyme immunoassay |
| CLIA | chemiluminescence immunoassay |
| COA | Certificate of Analysis |
| CV% | coefficient of variation, expressed as a percentage |
| EC | electrochemical |
| EIA | enzyme immunoassay |
| DEQAS | Vitamin D External Quality Assessment Scheme |
| FAQAP | Fatty Acid Quality Assurance Program |
| HAMQAP | Health Assessment Measurements Quality Assurance Program |
| IA | immunoassay |
| ID | isotope dilution |
| JCTLM | Joint Committee for Traceability in Laboratory Medicine |
| L1 | Level 1 of a multi-level SRM |
| L2 | Level 2 of a multi-level SRM |
| L3 | Level 3 of a multi-level SRM |
| L4 | Level 4 of a multi-level SRM |
| LC | liquid chromatography |
| LOQ | limit of quantitation for 25(OH)$D_2$, ≈0.5 ng/mL |
| $MAD_e$ | median absolute deviation, adjusted to estimate a standard deviation |
| MMQAP | Micronutrients Measurement Quality Assurance Program |
| MS | mass spectrometry |
| $MS^n$ | tandem mass spectrometry |

| | |
|---|---|
| MRM | multiple reaction monitoring |
| N | number of quantitative results |
| NIH-ODS | National Institutes of Health – Office of Dietary Supplements |
| NIST | National Institute of Standards and Technology |
| NISTIR | NIST Internal Report |
| QAP | Quality Assurance Program |
| RIA | radioimmunoassay |
| RMP | reference measurement procedure |
| R&D | research and development |
| SD | standard deviation |
| SRM | Standard Reference Material |
| $U_{95}$ | 95 % confidence expanded uncertainty |
| UV | ultraviolet absorbance |
| VDSP | Vitamin D Standardization Program |
| VitDQAP | Vitamin D Metabolites Quality Assurance Program |

## 1. INTRODUCTION

Over the past 15 years, methods used in the assessment of vitamin D status have been subject to a high level of scrutiny because of clinical studies that questioned testing accuracy. In 2010, the National Institutes of Health – Office of Dietary Supplements (NIH-ODS) established the Vitamin D Standardization Program (VDSP) in collaboration with the National Institute of Standards and Technology (NIST) and the Centers for Disease Control and Prevention (CDC) with the goal of standardizing laboratory measurements of vitamin D status in clinical health assessments worldwide. Before and after the formal establishment of the VDSP, NIST developed both higher-order analytical methods for the measurement of vitamin D metabolites in clinical samples and Standard Reference Materials (SRMs) for use as primary calibration materials and quality control samples.

In 2009, NIH-ODS and NIST partnered to develop the first accuracy-based Quality Assurance Program (QAP) for the determination of total serum 25-hydroxyvitamin D ($25(OH)D_{Total}$), the Vitamin D Metabolites Quality Assurance Program (VitDQAP). NIST assigned quantitative $25(OH)D_{Total}$ values to the samples used in VitDQAP exercises. The VitDQAP was conducted in parallel with the long-established consensus-based Vitamin D External Quality Assessment Scheme (DEQAS) program. In 2011, a separate College of American Pathologists (CAP) accuracy-based vitamin D measurement program was enacted that utilized CDC measurements for the value assignment of the study samples. In 2012, NIST began providing value assignment for the samples in the DEQAS program, which then advanced to become an accuracy-based program akin to the VitDQAP. With the evolution of DEQAS and CAP programs to adopt an accuracy-based approach to clinical determination of vitamin D status, NIST and NIH-ODS sunset the VitDQAP in 2016. A complete history of NIST's role in the support of the vitamin D initiative of the NIH-ODS is provided in Wise et al. [1].

In addition to the CAP and DEQAS programs, the Vitamin D measurement community is now served through the Health Assessment Measurements Quality Assurance Program (HAMQAP), although vitamin D studies are not the sole focus of the program and are not conducted with the same regularity as in the former VitDQAP. HAMQAP, in part a collaboration with the NIH-ODS, represents NIST's ongoing and future QAP support of the communities previously served by the Dietary Supplements QAP (DSQAP), Micronutrients Measurement Quality Assurance Program (MMQAP), and Fatty Acid Quality Assurance Program (FAQAP), as well as the VitDQAP. For the vitamin D metabolite measurement community, the HAMQAP has focused studies on both emerging measurement needs such as 24R,25-dihydroxyvitamin $D_3$ as well as the 'traditional' 25-hydroxyvitamin D metabolites.

## 2. SUMMARY OF EXERCISES

From 2009 to 2016, the VitDQAP conducted 12 exercises. Each exercise was designed for participants to measure 25-hydroxyvitamin $D_2$ ($25(OH)D_2$), 25-hydroxyvitamin $D_3$ ($25(OH)D_3$) and $25(OH)D_{Total}$ (the sum of $25(OH)D_2$ and $25(OH)D_3$) in a range of challenge serum and plasma-based samples. Reliable measurements of $25(OH)D_2$, $25(OH)D_3$, and $25(OH)D_{Total}$ are necessary for the medical community to make accurate clinical and health-care decisions.

VitDQAP participants used the method of their choice, and all reported using either liquid chromatography (LC)-based methods or immunoassay (IA) platform-based assays. LC-based methods enable quantitation of $25(OH)D_2$ and $25(OH)D_3$ separately through adequate chromatographic separation and/or specific detection of the metabolites, whereas IA methods measure $25(OH)D_{Total}$ without differentiating between the $25(OH)D_2$ and $25(OH)D_3$ forms. Most of the LC methods employed mass spectrometric (MS) detection, but in some cases ultraviolet absorbance (UV) or electrochemical (EC) detection was used. IA methods included radioimmunoassay (RIA), enzyme immunoassay (EIA), chemiluminescence immunoassay (CLIA), and chemiluminescence enzyme immunoassay (CLEIA). Both IA and LC participants provided results for $25(OH)D_{Total}$ in each study material. Even though most LC participants also provided values for $25(OH)D_2$ and $25(OH)D_3$, only the results for $25(OH)D_{Total}$ can be used to compare the laboratory performance of IA and LC techniques.

The 3-epimer form of 25-hydroxyvitamin $D_3$ (3-epi-$25(OH)D_3$) was an optional reporting metabolite because it is not included in $25(OH)D_{Total}$. As the program evolved, an increasing number of participants employing LC separations with tandem MS detection ($LC\text{-}MS^n$) methods reported results for 3-epi-$25(OH)D_3$ in select study materials. When a suitable number of data were reported, consensus values were determined and were disseminated along with NIST values.

For each study, a report of results was provided to the participants and served as the basis for a corresponding NIST Internal Report (NISTIR). A summary of the NISTIRs associated with the twelve VitDQAP Exercises is provided in Table 1.

In addition to the individual NISIRs for each exercise, results from the VitDQAP were published in two peer-reviewed manuscripts. A summary of key results from VitDQAP Exercises 1 to 6 is published in Bedner et al. [14]. A summary of key results from VitDQAP Exercises 7 to 12 is published in Wise et al. [1]. A complete summary of the VitDQAP exercises including the number of samples, control samples used, number of participants and the number of results is also provided in Wise et al. [1].

Lastly, this VitDQAP Final Report (NISTIR 8293) provides a comprehensive program overview including summaries of the exercises, participants, participant methods, study and control materials, data analysis and reporting, and key results.

**Table 1**: Summary of VitDQAP Exercises and NISTIRs

| VitDQAP | Internal Report Title | Publication Link [Reference] |
|---|---|---|
| Exercise 1 | Winter 2010 Comparability Study | https://doi.org/10.6028/NIST.IR.7890 [2] |
| Exercise 2 | Summer 2010 Comparability Study | https://doi.org/10.6028/NIST.IR.7891 [3] |
| Exercise 3 | Winter 2011 Comparability Study | https://doi.org/10.6028/NIST.IR.7892 [4] |
| Exercise 4 | Summer 2011 Comparability Study | https://doi.org/10.6028/NIST.IR.7893 [5] |
| Exercise 5 | Winter 2012 Comparability Study | https://doi.org/10.6028/NIST.IR.7894 [6] |
| Exercise 6 | Summer 2012 Comparability Study | https://doi.org/10.6028/NIST.IR.7895 [7] |
| Exercise 7 | Summer 2013 Comparability Study | https://doi.org/10.6028/NIST.IR.8000 [8] |
| Exercise 8 | Winter 2014 Comparability Study | https://doi.org/10.6028/NIST.IR.8133 [9] |
| Exercise 9 | Summer 2014 Comparability Study | https://doi.org/10.6028/NIST.IR.8141 [10] |
| Exercise 10 | Winter 2015 Comparability Study | https://doi.org/10.6028/NIST.IR.8142 [11] |
| Exercise 11 | Summer 2015 Comparability Study | https://doi.org/10.6028/NIST.IR.8143 [12] |
| Exercise 12 | Summer 2016 Comparability Study | https://doi.org/10.6028/NIST.IR.8169 [13] |



## 3. SUMMARY OF PARTICIPANTS

Over the course of the VitDQAP, there were 99 total participating organizations that enrolled in the program and returned results for at least one exercise.

Participants in the VitDQAP encompassed different types of organizations and were characterized in five major categories, including:

- **Hospital** and clinic laboratories providing patient testing services
- **Academic** institutions, typically conducting research
- **University hospital** and clinic laboratories conducting both patient testing and research
- **Government** (Gov.) organizations with a range of functions such as testing, research and development (R&D), and standards development
- **Industry** and non-profits companies providing testing, R&D, and assay and/or control material development and manufacturing (mfg.)

Figure 1 displays the total distribution of the participants by type of organization, and the breakout of the industry participants by function (testing, R&D, and assay and control manufacturing).

Figure 2 shows the distribution of participants with respect to lab type over the course of the program. For Figure 1 and Figure 2, the characterizations of organizational type were defined by NIST for descriptive purposes, and the participants were not asked to self-identify their lab type during the program. A complete list of the program participants and their assigned type is provided in Appendix A

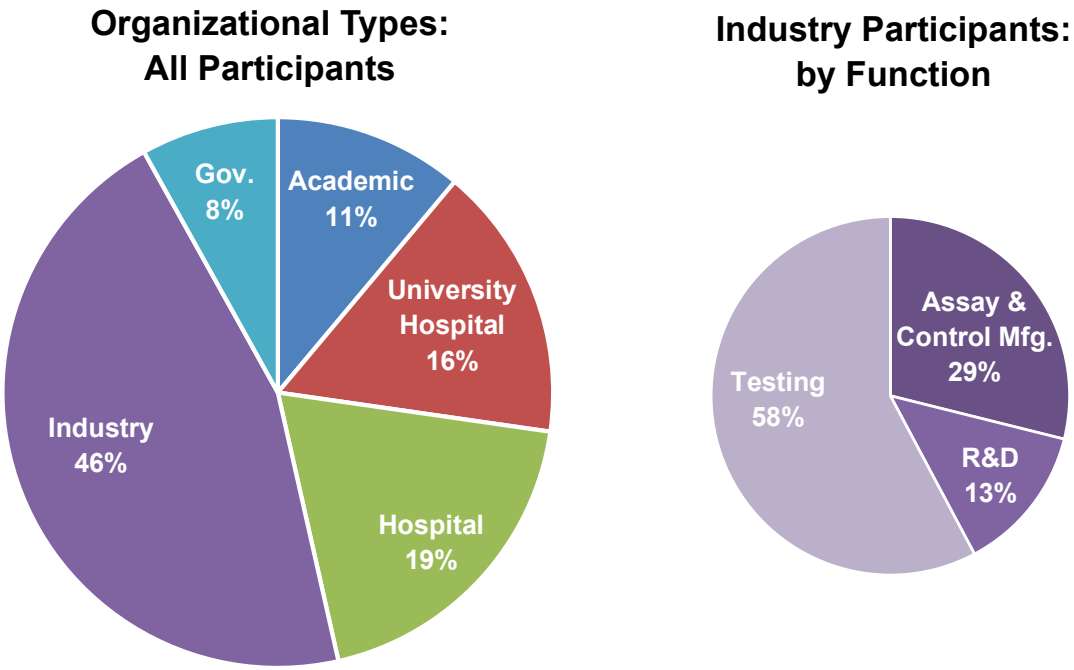



**Figure 1**: Overview of VitDQAP Participants by Organizational Type

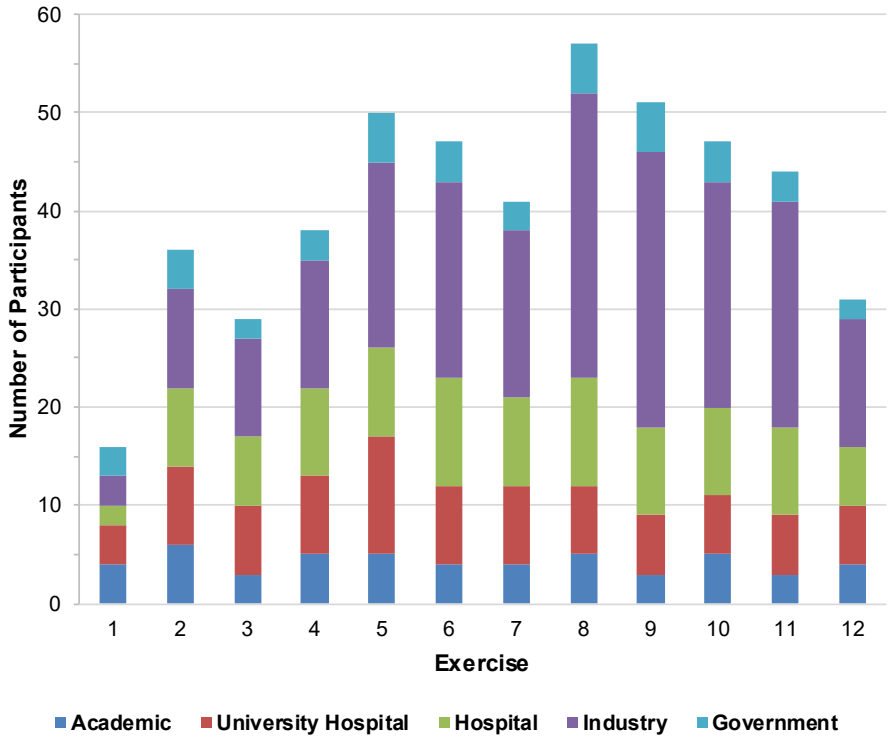



**Figure 2**: Distribution of Participants by Organizational Type Across All Exercises

Figure 2 also depicts the variation in the total number of participants in each exercise across the entire program. While there were 99 total participating organizations in the program, the number that enrolled and returned results in any given exercise ranged from 16 (Exercise 1; 2010) to a peak of 57 (Exercise 8; 2014). In early 2016, participants were informed that the program was ending and were encouraged to seek involvement in either the CAP or DEQAS accuracy-based programs, which likely influenced the decreased participation for the final exercise.

The VitDQAP was an international program that included participants from 20 different countries representing all major continents (except Antarctica). However, almost two-thirds (60 of 99) of the participants were from USA-based organizations. The distribution of VitDQAP participants by country is presented in Figure 3.

The USA-based participants represented almost half of the states (24 of 50) and the District of Columbia (DC), as depicted in Figure 4. The wide geographical range in participation reflects the domestic and global interest in improving the comparability of vitamin D metabolite measurements, as well as the widespread dissemination and impact of VitDQAP resources.

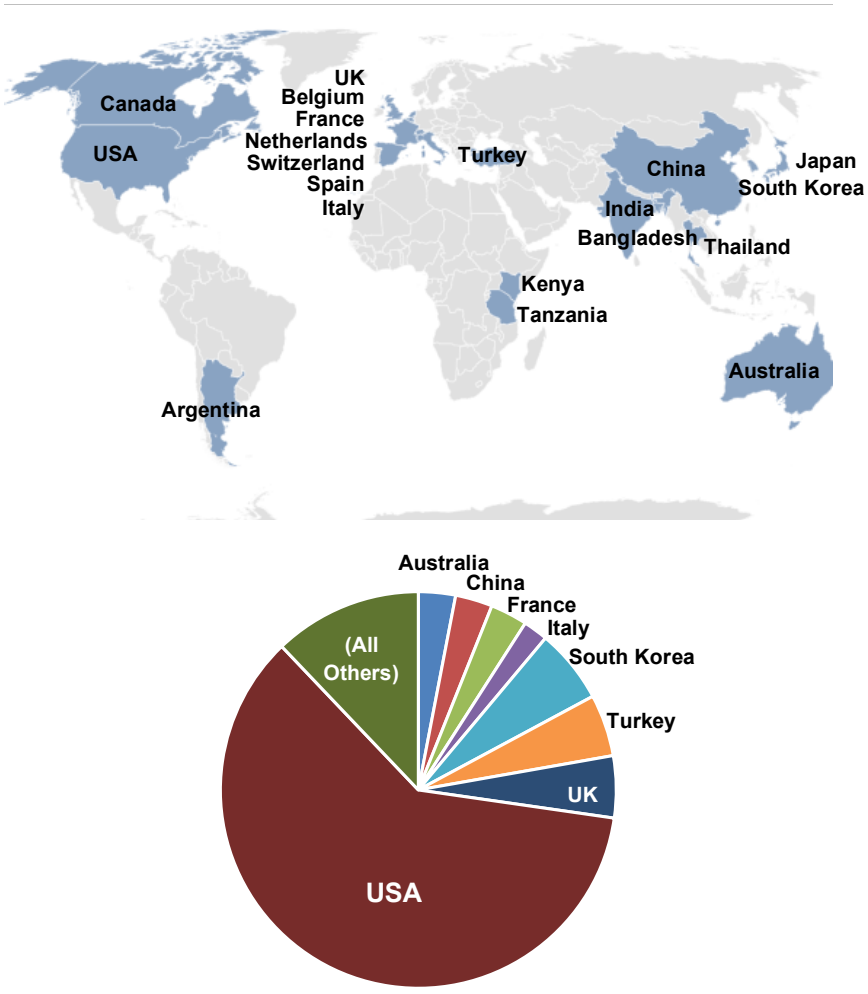



**Figure 3**: Distribution of VitDQAP Participants by Country

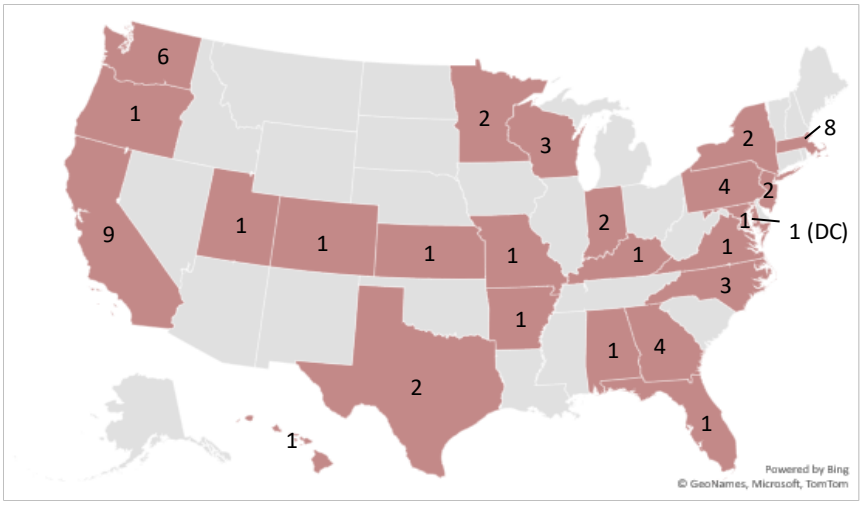



**Figure 4**: Number of Participants in the USA by State (Total = 60)

## 4. DESCRIPTION OF MATERIALS

Before the study materials were sent to the participants, NIST evaluated human serum and plasma materials in-house to determine if they were suitable for distribution as VitDQAP study samples. Often NIST SRMs were selected for study samples because they meet the rigorous requirements for samples for interlaboratory comparison studies: homogeneity and analyte stability, which was reinforced through temperature control of the samples both during storage and shipping under dry ice conditions. However, there were limited human serum and plasma SRMs certified for 25(OH)$D_2$, 25(OH)$D_3$, or 25(OH)$D_{Total}$ available to cover all the exercises for the entire VitDQAP. Therefore, additional study materials and/or value assignment measurements were frequently required to establish NIST values to support the accuracy-based VitDQAP.

For all the VitDQAP comparability studies, the study samples were minimally processed, fresh frozen pooled human plasma or serum materials. The total number of study samples ranged from two to four for each comparability study, and in some cases included blinded replicates of the individual study sample. In the first six exercises, study samples were sent in duplicate or triplicate to assess the laboratory repeatability for the same material. However, it became evident that repeatability on a single material was not as informative as the variability of participant results for materials of different levels. Beginning in Exercise 7, the design was changed to sending single samples of two different blinded study materials and one sample of SRM 968d Level 1 (L1), the measurement control. This study design was used for the remainder of the program. As part of the sampling redesign, additional study materials were needed, particularly those with high or high normal levels of 25(OH)$D_3$ and those with measurable levels of endogenous 25(OH)$D_2$ and 3-epi-25(OH)$D_3$. While the program typically relied on SRMs to meet the need for study samples, unique materials were acquired to meet the evolving needs of the program.

A control material, either SRM 2972 or SRM 968d L1, was also provided in each study. All control and study sample materials were provided to the participants free of charge. A graphical representation of the sampling scheme for the controls and samples used across all exercises of the VitDQAP is provided in Figure 5.

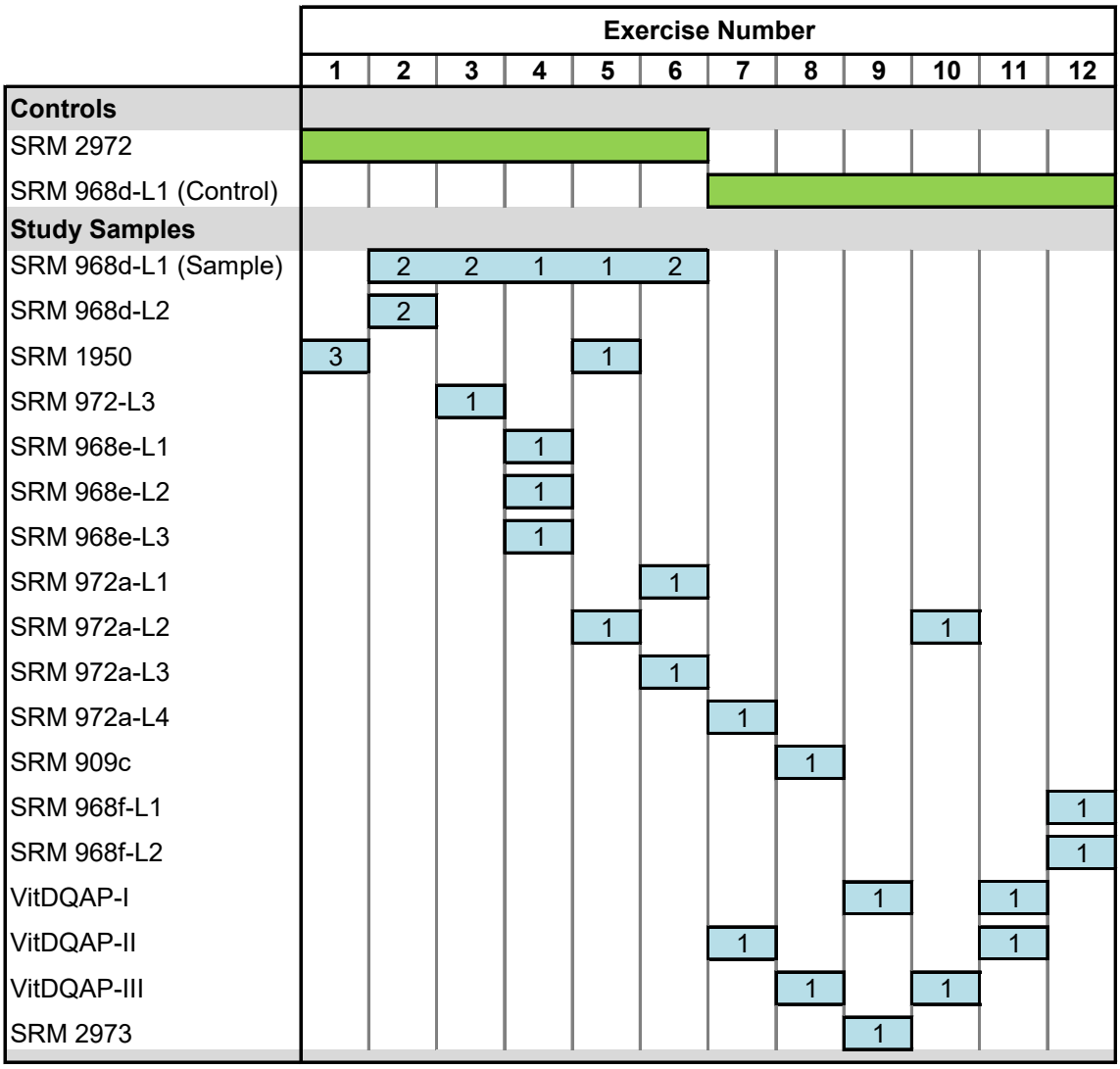



**Figure 5**: Sampling Scheme for VitDQAP Control and Study Materials

The colored boxes indicate which exercises the materials were used as study samples or controls, and the numbers within the boxes indicate the number of individual vials that were sent to participants as blinded samples.

All control and study sample materials used in the VitDQAP are summarized in Table 2 with detailed descriptions provided below. The mass concentration values in units of ng/mL and their 95 % confidence expanded uncertainties ($U_{95}$) for 25(OH)$D_2$ and 25(OH)$D_3$ are consistent with what has been reported on the Certificate of Analysis (COA) for the SRMs, where applicable. The $U_{95}$ for these values includes components for measurement variability and the uncertainty associated with the density of the materials. The 25(OH)$D_{Total}$ values were determined as the sum of the individual values for 25(OH)$D_3$ and, when the value was above the limit of quantitation (LOQ) of ≈0.5 ng/mL, 25(OH)$D_2$. The $U_{95}$ for the 25(OH)$D_{Total}$ values incorporate the

uncertainties for the two analytes. Information regarding the replications and to which exercise each material was applied is also included, reiterating the information summarized in Figure 5.

Each material listed in Table 2 is accompanied by a brief description. For the purposes of this report, 'low' level materials are roughly defined as having $25(OH)D_{Total}$ concentrations ≤ 20 ng/mL. 'Normal' levels have concentrations roughly between 20 ng/mL and 50 ng/mL. None of the study materials had levels near this upper bound. A subcategory of 'high normal' was used to describe those materials with values ≥ 30 ng/mL.

**Table 2**: Description of VitDQAP Control and Study Samples

| | | Mass Concentration, ng/mL | | |
|---|---|---|---|---|
| *Controls:* | | $25(OH)D_2$ | $25(OH)D_3$ | $25(OH)D_{Total}$ |
| SRM 2972 | calibration solutions (ethanol) | 231.3 ± 7.2*† | 323.0 ± 11.7*† | Not applicable |
| SRM 968d-L1 | human serum; endogenous, low level | <LOQ | 12.4 ± 0.3 | 12.4 ± 0.3 |
| *Study Samples:* | | | | |
| SRM 968d-L1 | human serum; endogenous, low level | <LOQ | 12.4 ± 0.3 | 12.4 ± 0.3 |
| SRM 968d-L2 | human serum; endogenous for $25(OH)D_{Total}$ (exogenous for carotenoids), low level | <LOQ | 10.4 ± 0.2 | 10.4 ± 0.2 |
| SRM 1950 | human plasma; endogenous, normal level | 0.52 ± 0.17 | 24.78 ± 0.77* | 25.3 ± 0.8 |
| SRM 972-L3 | human serum; exogenous $25(OH)D_2$, endogenous $25(OH)D_3$, high normal level | 26.4 ± 2.0* | 18.5 ± 1.1* | 44.9 ± 2.3 |
| SRM 968e-L1 | human serum; endogenous, low level | <LOQ | 7.09 ± 0.14* | 7.09 ± 0.14 |
| SRM 968e-L2 | human serum; endogenous, low level | <LOQ | 12.9 ± 0.3* | 12.9 ± 0.3 |
| SRM 968e-L3 | human serum; endogenous, normal level | <LOQ | 19.9 ± 0.4* | 19.9 ± 0.4 |
| SRM 972a-L1 | human serum; endogenous, normal level | 0.54 ± 0.06 | 28.8 ± 1.1* | 29.3 ± 1.1 |
| SRM 972a-L2 | human serum; endogenous, low level | 0.81 ± 0.06* | 18.1 ± 0.4* | 18.9 ± 0.4* |
| SRM 972a-L3 | human serum; endogenous, high normal level | 13.3 ± 0.3* | 19.8 ± 0.4* | 33.2 ± 0.6* |
| SRM 972a-L4 | human serum; endogenous $25(OH)D_2$ and $25(OH)D_3$; high normal level exogenous 3-epi-$25(OH)D_3$, | 0.55 ± 0.10 | 29.4 ± 0.9* | 30.0 ± 1.0 |
| SRM 909c | human serum; endogenous, normal level | <LOQ | 20.7 ± 0.7 | 20.7 ± 0.7 |
| SRM 968f-L1 | human serum; endogenous, low level | 0.85 ± 0.05 | 12.3 ± 0.2 | 13.2 ± 0.5 |
| SRM 968f-L2 | human serum; endogenous, low level | 0.17 ± 0.01 | 15.6 ± 0.2 | 15.8 ± 0.5 |
| VitDQAP-I | human serum; endogenous, high normal level | 0.68 ± 0.06 | 31.3 ± 0.8 | 32.0 ± 0.8 |
| VitDQAP-II | human serum; endogenous, high normal level | 0.44 ± 0.04 | 37.1 ± 0.9 | 37.5 ± 0.9 |
| VitDQAP-III | human serum; endogenous, high normal level | 6.5 ± 0.2 | 26.2 ± 0.6 | 32.7 ± 0.7 |
| SRM 2973 | human serum; endogenous, high normal level | 0.65 ± 0.02 | 39.4 ± 0.8* | 40.1 ± 0.8 |

* Certified values

† Converted from ng/g on COA to ng/mL for use in the VitDQAP

*4.1 Control SRM 2972 Calibration Solutions*

SRM 2972 25-Hydroxyvitamin $D_2$ and $D_3$ Calibration Solutions consisted of two separate solutions of the vitamin D metabolites 25(OH)$D_2$ and 25(OH)$D_3$ in ethanol. The intended use for SRM 2972 was for calibration of instruments and techniques employed for the determination of these metabolites.

This VitDQAP control material was originally characterized at NIST using both gravimetry and LC-MS at the time of these studies but was later characterized using isotope dilution LC-MS (ID-LC-MS) [15] and LC with absorbance detection when the solutions were reissued as SRM 2972a. Each solution of SRM 2972 was certified for its respective metabolite, 25(OH)$D_2$ or 25(OH)$D_3$, in mass fraction units of ng/g. A certified value is a value for which NIST has the highest confidence in its accuracy. For use in the VitDQAP, the certified values were converted to mass concentration units of ng/mL by multiplying by 0.78775 g/mL, the density of ethanol at 22 °C. Participants were provided these values both on the shipping package and within the data reporting sheet so that they could qualify their methods prior to analyzing the study samples.

SRM 2972 was used as a VitDQAP control sample for Exercises 1 to 6.

*4.2 Control SRM 968d Level 1 Human Serum*

SRM 968d Fat-Soluble Vitamins, Carotenoids, and Cholesterol in Human Serum was intended for use in validating methods for determining fat-soluble vitamins, carotenoids, and cholesterol in human serum and plasma and for quality assurance when assigning values to in-house control materials. This SRM was originally prepared with multiple levels, some of which contained exogenously spiked levels of carotenoids. The spiked-serum materials were deemed unsatisfactory with respect to carotenoid homogeneity after evaluations conducted within the MMQAP [16]. When SRM 968d was superseded by SRM 968e, SRM 968d-L1, a completely endogenous material, was repurposed as a VitDQAP control material.

As the vitamin D metabolites were not value assigned in SRM 968d-L1, the NIST values for 25(OH)$D_3$ in SRM 968d-L1 were obtained using an ID-LC-MS/MS reference measurement procedure (RMP) [17] recognized by the Joint Committee for Traceability in Laboratory Medicine (JCTLM) that was developed at NIST for certification of 25(OH)$D_2$ and 25(OH)$D_3$ in other vitamin D-specific SRMs. The NIST value for 25(OH)$D_2$ was not determined for SRM 968d-L1 as the estimated concentration value was well below the LOQ. For the Control SRM 968d-L1, the participants were provided the NIST target values within the data reporting sheet so that they could qualify their methods prior to analyzing the study samples.

SRM 968d-L1 was used as a VitDQAP control material for Exercises 7 to 12.

*4.3 Study Sample SRM 1950 Human Plasma*

SRM 1950 Metabolites in Frozen Human Plasma is currently available and is intended primarily for validation of methods for determining metabolites such as fatty acids, electrolytes, vitamins, hormones, and amino acids in human plasma and similar materials. This SRM can also be used for comparison of measurement technologies used in metabolomic studies and for quality assurance when assigning values to in-house reference materials. This SRM is intended to represent “normal” human plasma.

The $25(OH)D_3$ was determined as a certified value representing results from both the NIST ID-LC-MS method [15] and the NIST ID-LC-MS/MS RMP [17]. As the level of $25(OH)D_2$ was approaching the LOQ, the value assigned is the best estimate of the true value based on available data but was not considered sufficiently reliable to justify certification. More information can be found in the Certificate of Analysis for SRM 1950 [18].

SRM 1950 was used as a blinded study sample in triplicate in Exercise 1 and as a single sample in Exercise 5.

*4.4 Study Samples SRM 968d Level 1 and Candidate SRM 968d Level 2 Human Serum*

SRM 968d Fat-Soluble Vitamins, Carotenoids, and Cholesterol in Human Serum was intended for use in validating methods for determining fat-soluble vitamins, carotenoids, and cholesterol in human serum and plasma and for quality assurance when assigning values to in-house control materials. This SRM was originally prepared with multiple levels, some of which contained exogenously spiked levels of carotenoids. The spiked-serum materials were deemed unsatisfactory with respect to carotenoid homogeneity after evaluations conducted within the MMQAP [16]. Both SRM 968d-L1, a completely endogenous material, and candidate Level 2 (968d-L2), one of the materials spiked with carotenoids that did not become part of SRM 968d, were repurposed as study samples for the VitDQAP. While 968d-L2 contained exogenous levels of carotenoids, the $25(OH)D_2$ and $25(OH)D_3$ levels in that material were endogenous, which is why it was deemed to be an appropriate study material. SRM 968d-L1 and 968d-L2 were both evaluated as study samples prior to establishing SRM 968d-L1 as a VitDQAP control for Exercises 7 to 12.

The $25(OH)D_2$ and $25(OH)D_3$ levels in 968d-L2 were determined as previously described for SRM 968d-L1 as the control material.

The NIST values for the vitamin D metabolites in SRM 968d-L1 when used as blinded study sample (Table 4) were identical to those when it was used as a control. SRM 968d-L1 was used as a blinded study sample in duplicate for Exercises 2, 3 and 6 and in singlicate in Exercises 4 and 5. 968d-L2 was used as a blinded study sample in duplicate for Exercise 2.

*4.5 Study Sample SRM 972 Level 3 Human Serum*

SRM 972 Level 3 (SRM 972-L3) Vitamin D in Human Serum was intended for use as an accuracy control in the critical evaluation of methods for determining the amount of substance concentration of vitamin D metabolites in human serum. This SRM could also be used for assigning values of vitamin D metabolite mass concentrations to in-house control materials. SRM 972-L3 was a "normal" human serum pool that was spiked with $25(OH)D_2$, but the $25(OH)D_3$ was endogenous.

The NIST values for $25(OH)D_2$ and $25(OH)D_3$ in SRM 972-L3 were determined as certified values representing NIST results from both ID-LC-MS (15) and the ID-LC-MS/MS RMP (17) as well as CDC results from an independent ID-LC-MS/MS method. SRM 972 was depleted around 2011 and was succeeded by SRM 972a, which is currently available.

SRM 972-L3 was used as a blinded study sample in singlicate in Exercise 3.

*4.6 Study Samples SRM 968e Levels 1, 2, and 3 Human Serum*

SRM 968e Fat-Soluble Vitamins, Carotenoids, and Cholesterol in Human Serum was intended for use in validating methods for determining fat-soluble vitamins, carotenoids, and cholesterol in human serum and plasma and could also be used for quality assurance when assigning values to in-house control materials for these constituents. From the lessons learned in the preparation of the preceding SRM 968d, the three concentration levels of SRM 968e were obtained from endogenous sources of serum, and then blended to provide three materials with different concentration levels.

The $25(OH)D_3$ was determined as a certified value in all three levels (Table 4) of SRM 968e using the NIST ID-LC-MS/MS RMP [17]. The level of $25(OH)D_2$ was below the LOQ in all three levels of SRM 968e, and therefore was not determined.

SRM 968e Levels 1, 2, and 3 (SRM 968e-L1, SRM 968e-L2, and SRM 968e-L3) were each used as a blinded study sample in singlicate in Exercise 4.

*4.7 Study Samples SRM 972a Levels 1, 2, 3, and 4 Human Serum*

SRM 972a Vitamin D Metabolites in Frozen Human Serum is currently available and is intended for use as an accuracy control in the critical evaluation of methods for determining the amount-of-substance concentration of vitamin D metabolites in human serum and can also be used as a quality assurance tool for assigning values to in-house control materials for vitamin D metabolites. SRM 972a replaced SRM 972 when it became depleted. Levels 1, 2, and 3 of SRM 972a (SRM 972a-L1, SRM 972a-L2, and SRM 972a-L3) were prepared from pools of human serum with endogenous concentrations of $25(OH)D_3$. SRM 972a Level 4 (SRM 972a-L4) was prepared from a pool of human serum that was fortified with 3-epi-$25(OH)D_3$.

The $25(OH)D_3$ in all four levels and $25(OH)D_2$ for two of the four levels (SRM 972a-L2 and SRM 972a-L3) were determined as certified values representing NIST results from both ID-LC-MS [15] and the ID-LC-MS/MS RMP [17] as well as CDC results from an independent ID-LC-MS/MS method. The level of $25(OH)D_2$ was at the LOQ in SRM 972a-L1 and SRM 972a-L4, and these values were not considered certifiable. For SRM 972a-L2 and SRM 972a-L3, the $25(OH)D_{Total}$ was also a certified value, representing the sum of the individual certified values for $25(OH)D_3$ and $25(OH)D_2$. More information can be found in the Certificate of Analysis for SRM 972a [19] and in a peer-reviewed manuscript [20].

SRM 972a-L1 was used as a blinded study sample in singlicate in Exercise 6. SRM 972a-L2 was used as a blinded study sample in singlicate in Exercises 5 and 10. SRM 972a-L3 was used as a blinded study sample in singlicate in Exercise 6. SRM 972a-L4 was used as a blinded study sample in singlicate in Exercise 7.

*4.8 Study Sample SRM 909c Human Serum*

SRM 909c Frozen Human Serum is currently available and is intended for use in validating analytical methods for the determination of specified constituents in human serum. This SRM can also be used for quality assurance when assigning values to in-house control materials. SRM 909c is considered a "normal" human serum material.

Even though not included on the Certificate of Analysis for SRM 909c, the NIST value for 25(OH)$D_3$ was obtained using results from both the NIST ID-LC/MS method (15) and the ID-LC-MS/MS RMP [17]. This value is considered of the same quality as other 25(OH)$D_3$ certified values. The NIST value for 25(OH)$D_2$ was not determined for SRM 909c as the estimated concentration value was below the LOQ.

SRM 909c was used a blinded study sample in singlicate in Exercise 8.

*4.9 Study Samples SRM 968f Levels 1 and 2 Human Serum*

SRM 968f Fat-Soluble Vitamins in Frozen Human Serum is currently available and is intended for use in validating methods for determining fat-soluble vitamins in human serum and plasma and qualifying control materials produced in-house and analyzed using those methods. The two concentration levels of SRM 968f were obtained from endogenous sources of sera, and then blended to result in two materials with differing concentration levels.

The NIST values for 25(OH)$D_3$ and 25(OH)$D_2$ were determined using the NIST ID-LC-MS/MS RMP [17] and are reported as non-certified values in the COA for SRM 968f [21].

SRM 968f Levels 1 and 2 (SRM 968f-L1 and SRM 968f-L2) were each used as a blinded study sample in singlicate in Exercise 12.

*4.10 Study Samples VitDQAP-I, -II, and -III Human Serum*

A series of three pooled human serum materials consisting of high normal levels of endogenous 25(OH)$D_{Total}$ were specially prepared for the VitDQAP exercises to address the gap from most materials having low or low-normal levels. The VitDQAP materials were prepared using the same rigorous protocols established for NIST SRMs.

The values for 25(OH)$D_3$ in the three VitDQAP study materials were obtained using the NIST ID-LC-MS/MS RMP [17]. The NIST values for 25(OH)$D_2$ were also obtained using the RMP.

VitDQAP-I was used as a blinded study sample in singlicate in Exercises 9 and 11. VitDQAP-II was used as a blinded study sample in singlicate in Exercises 7 and 11. VitDQAP-III was used as a blinded study sample in singlicate in Exercises 8 and 10.

*4.11 Study Sample SRM 2973 Human Serum*

SRM 2973 Vitamin D Metabolites in Frozen Human Serum (High Level) is currently available and is intended for use as an accuracy control in the critical evaluation of methods for determining the amount-of-substance concentration of vitamin D metabolites in human serum and can also be used as a quality assurance tool for assigning values to in-house control materials for vitamin D metabolites. SRM 2973 is comprised of a single material and is intended to address the need for a high-normal level of 25(OH)$D_{Total}$.

The 25(OH)$D_3$ in SRM 2973 was determined as a certified value, and 25(OH)$D_2$ was determined as a non-certified value. Both values were obtained using the NIST ID-LC-MS/MS RMP [17]. These individual values along with the non-certified value for 25(OH)$D_{Total}$ are reported in the COA for SRM 2973 [22] and in a peer-reviewed manuscript [23].

SRM 2973 was used as a blinded study sample in singlicate in Exercise 9 where it was coded 'VitDQAP-IV.' The material was identified as SRM 2973 in the final report, NISTIR 8141 [10].

## 5. INSTRUCTIONS TO PARTICIPANTS

Participants were required to enroll for each study of the VitDQAP. Prior to each exercise, an invitation letter and participation form were sent by email to current program participants, who were also encouraged to forward the notice to other organizations that might be interested in participating. Once enrolled, participants were notified via email prior to distributing the samples and during the transit of the shipment. When samples were received by the participants, delivery of the shipment was confirmed by returned packing slips, e-mails, and/or fax-based receipts.

For each exercise, participants were instructed to provide a single value for each control and study sample material. The data file template given to participants for reporting data included a single sheet for the participants to provide results and list information on analytical methods.

Participants were asked to provide individual concentration values for $25(OH)D_2$ and $25(OH)D_3$ along with $25(OH)D_{Total}$ for each study sample. For Exercises 1 to 6 in which the control was the SRM 2972 calibration solutions, participants were asked to report $25(OH)D_2$ and $25(OH)D_3$ separately. For Exercises 7 to 12 in which the control was SRM 968d-L1, participants were asked to report $25(OH)D_2$ and $25(OH)D_3$ along with $25(OH)D_{Total}$. Participants were not required to report values for all measurands listed on the reporting sheet but were required to report $25(OH)D_{Total}$.

Participants were also invited to report any measurable concentration values for 3-epi-$25(OH)D_3$ for the serum-based control SRM 968d-L1 or the serum- and plasma-based study samples.

Throughout each exercise, participants were encouraged to communicate with the study coordinators regarding their measurement performance or with any technical inquiries to successfully complete their analyses.

## 6. PARTICIPANT METHODS

For each of the exercises, participants were asked to use the analytical protocols currently employed in their laboratory to analyze the control and sample materials. The methods used by participants were categorized into two main groups: 1) IA, including EIA, CLIA, RIA, and CLEIA; and 2) LC, including MS, MS/MS, UV, and EC detection. Due to the comparability of the methods, the LC-MS and LC-MS/MS results in the program were collectively referred to as $LC\text{-}MS^n$ to simplify the data reporting and analysis. In some cases, participating organizations provided results for multiple methods employed in their laboratory. Of the 99 organizations that participated in at least one exercise of the VitDQAP, 77 used one method, 19 used 2 methods, and 3 used 3 methods for a total of 124 methods. The distribution of all methods used to report results by the participants is presented in Figure 6. Almost half of the VitDQAP participants (N=60) used $LC\text{-}MS^n$ methods, with LC-MS/MS predominating (N=58).

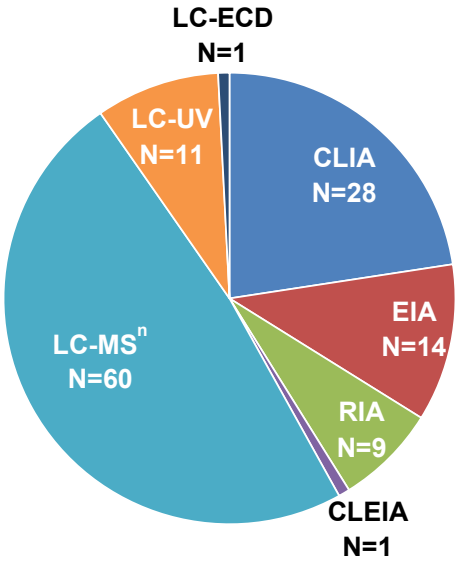



**Figure 6**: Summary of All Methods Used by Participants in the VitDQAP (Total = 124)

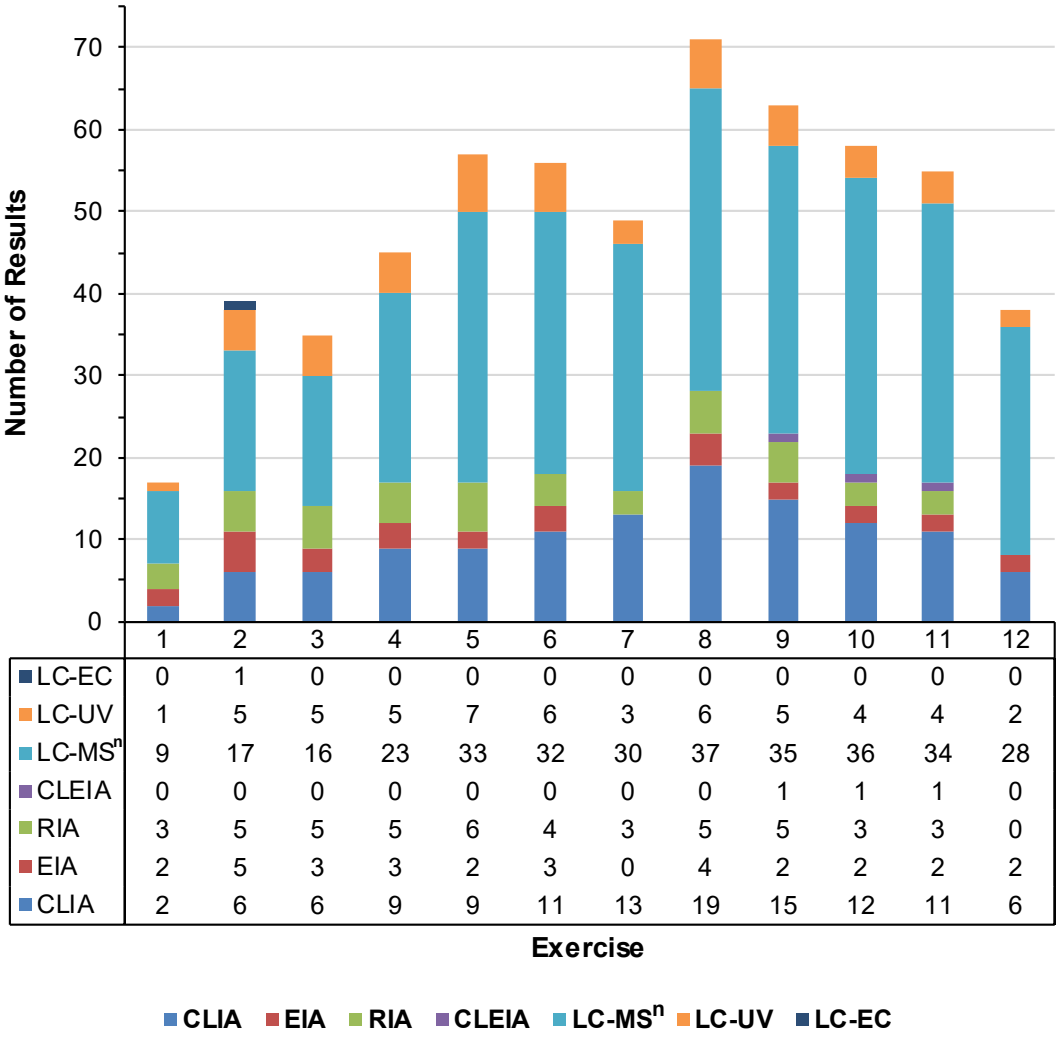
While there were 99 participants in the VitDQAP, the number that enrolled in each exercise and hence the number of reported results was typically lower, ranging from 17 (Exercise 1; 2010) to 71 (Exercise 8; 2014). A summary of the number of results reported by participants for each method across all exercises of the VitDQAP is shown in Figure 7. A full timeline with a breakout for each participating organization, including the methods used, is provided in Appendix B.

Individual details for each of the reported participant methods are summarized in the Appendices of the NISTIRs associated with each Exercise [2-13]. In summary, the following types of information were generally reported:

For IA methods, the sample preparation and limited detection details (when available) were provided by participants. Information regarding kit vendors for the IA methods was collected but not reported by the VitDQAP, as it is NIST policy to neither endorse nor potentially discriminate against any vendor and/or company.

For LC-$MS^n$ methods, the internal standards used, sample preparation, chromatographic conditions, and mass spectroscopic detection ion details were provided by participants. Most of the $MS^n$ participants reported using MS/MS methods with at least one stable isotope labeled internal standard and multiple reaction monitoring (MRM) to differentiate the mass transitions for 25(OH)$D_2$ and 25(OH)$D_3$. For the LC-UV methods, the internal standards used, sample preparation, chromatographic conditions, and wavelength detection details were provided by participants. In a few rare cases, participants opted not to provide complete descriptions of their LC methods.

For all study materials, the immunoassay methods reported values for 25(OH)$D_{Total}$ only, whereas LC method participants provided values for 25(OH)$D_2$, 25(OH)$D_3$, and 25(OH)$D_{Total}$, and in a few cases, 3-epi-25(OH)$D_3$ and 24R,25-dihydroxyvitamin $D_3$.




| | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| LC-EC | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| LC-UV | 1 | 5 | 5 | 5 | 7 | 6 | 3 | 6 | 5 | 4 | 4 | 2 |
| LC-MS$^n$ | 9 | 17 | 16 | 23 | 33 | 32 | 30 | 37 | 35 | 36 | 34 | 28 |
| CLEIA | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 1 | 1 | 0 |
| RIA | 3 | 5 | 5 | 5 | 6 | 4 | 3 | 5 | 5 | 3 | 3 | 0 |
| EIA | 2 | 5 | 3 | 3 | 2 | 3 | 0 | 4 | 2 | 2 | 2 | 2 |
| CLIA | 2 | 6 | 6 | 9 | 9 | 11 | 13 | 19 | 15 | 12 | 11 | 6 |

**Figure 7**: Number of Results Reported by Participants for Each Method Across All Exercises

## 7. OVERVIEW OF DATA ANALYSIS AND REPORTING

Participant results were anonymized using participant codes in data tables and graphs in all exercise reports. The same code was used for laboratories that participated in multiple exercises across the duration of the VitDQAP. The VitDQAP also placed a considerable emphasis on using graphs to help participants visualize their results, relative to both the performance of their peers in the measurement community as well as to NIST measurement values.

### *7.1 Tables of Individual Results, Consensus Data, and NIST Values*

The individual results for each control and study sample provided by participants were compiled and evaluated by NIST. The consensus results included the median values for $25(OH)D_{Total}$, $25(OH)D_2$, $25(OH)D_3$, or 3-epi-$25(OH)D_3$; the median absolute deviation estimate ($MAD_e$, a robust estimate of the standard deviation (SD)); and the coefficient of variation expressed as a percentage (CV%).

For $25(OH)D_{Total}$ in serum or plasma materials, the consensus statistics were determined for all reported methods, the IA methods only, and the LC methods only. A separate consensus value was also provided for the LC-$MS^n$ results as this was the predominant method used by participants.

For $25(OH)D_2$, $25(OH)D_3$, or 3-epi-$25(OH)D_3$ in serum or plasma materials, the consensus summary statistics were determined for LC methods and LC-$MS^n$ methods, as only LC techniques can differentiate the metabolites. The SRM 2972 control materials used for exercise 1 through 6 were comprised of separate ethanolic solutions of the two metabolites, $25(OH)D_2$ and $25(OH)D_3$. For these solutions, consensus statistics were determined for all methods, LC methods only, IA methods, and LC-$MS^n$ methods. However, for data sets where the number of results for consensus analysis was less than ≈ 10 (e.g., IA, LC-$MS^n$), summary statistics were not determined.

For each exercise, a summary table of the individual participant results was provided together with the consensus results and the NIST value for each material. Table 3 provides an example of a summary table of data showing individual participant results, consensus values, and NIST values for $25(OH)D_{Total}$ in serum study materials and the control. The results are sorted by the anonymous code assigned to each participant. Participants that reported results from different methods have letters appended to their code numbers (a, b, or c) to differentiate the results.

### *7.2 Consensus Data Plots*

For each study material, consensus data plots reported the participant results together with the consensus ranges for the two major techniques, IA or LC, and the NIST value range. The participant values were displayed in ascending order within the methods. Figure 8 is an example and description of a consensus data plot where the NIST value range falls within the consensus ranges for both major techniques. This indicates that there is no major method bias for this sample.

By using data in both tabular form (Table 3) and graphical form (Figure 8), participants were able to readily assess their performance for each study material when compared to 1) all other labs measuring $25(OH)D_{Total}$; 2) other labs using the same major technique, IA or LC; 3) other labs using the same method (e.g., LC-$MS^n$); and 4) the true value as determined by NIST. While individual labs were not assigned a 'pass' or fail' based on their performance relative to the NIST value, providing true values was critical in helping participants assess their own method biases.

## **Table 3**: Example of Participant Data and Consensus Values (Exercise 8)

Results are for 25(OH)$D_{Total}$ for the various methods in three study samples (VitDQAP-III, SRM 909c and SRM 968d-L1) from the Winter 2014 Comparability Study (Exercise 8). All results are in units of ng/mL.

| | | VitDQAP-III | SRM 909c | SRM 968d L1 |
|---|---|---|---|---|
| Lab | Method | Vial A | Vial B | Control |
| 017 | CLIA | 31.0 | 16.2 | 13.9 |
| 026 | LC-MS/MS | 33.7 | 21.7 | 13.4 |
| 030a | RIA | 40.5 | 24.8 | 15.3 |
| 030b | LC-MS/MS | 24.5 | 23.9 | 13.2 |
| 056a | LC-MS/MS | 34.5 | 20.3 | 12.8 |
| 056b | LC-MS/MS | 34.0 | 20.3 | 12.7 |
| 060 | LC-MS/MS | 33.4 | 22.0 | 13.0 |
| 110 | LC-UV | 21.8 | 79.6 | 12.8 |
| 116 | LC-MS/MS | 34.2 | 20.7 | 12.8 |
| 119 | LC-MS/MS | 39.5 | 27.3 | 15.6 |
| 139 | LC-UV | 24.7 | 51.9 | 13.1 |
| 150 | LC-MS/MS | 31.2 | 17.7 | 12.1 |
| 161 | CLIA | 30.6 | 14.4 | 11.8 |
| 180 | RIA | 29.1 | 19.5 | 14.9 |
| 185a | LC-MS/MS | 38.4 | 36.4 | 17.3 |
| 185b | CLIA | 16.4 | 33.3 | 14.5 |
| 187 | LC-MS/MS | 30.3 | 19.0 | 13.1 |
| 188 | CLIA | 36.1 | 18.4 | 14.1 |
| 189 | LC-UV | 27.6 | 38.2 | 16.5 |
| 193 | EIA | 33.9 | 22.3 | 18.6 |
| 194 | LC-MS/MS | 33.0 | 15.8 | 13.0 |
| 196 | CLIA | 30.7 | 15.1 | 14.1 |
| 197 | LC-MS/MS | 34.2 | 22.0 | 13.1 |
| 198a | LC-MS/MS | 30.8 | 18.5 | 12.6 |
| 198c | CLIA | 29.1 | 19.7 | 8.3 |
| 199 | LC-MS/MS | 38.2 | 29.6 | 13.7 |
| 200 | RIA | 27.7 | 18.0 | 13.8 |
| 204a | CLIA | 27.5 | 14.0 | 12.2 |
| 204b | LC-MS/MS | 33.8 | 20.2 | 12.8 |
| 209 | LC-MS/MS | 33.9 | 22.9 | 12.5 |
| 210a | RIA | 31.4 | 17.9 | 12.4 |
| 210b | CLIA | 38.7 | 25.0 | 9.1 |
| 211 | LC-MS/MS | 36.0 | 18.3 | 13.6 |
| 212 | LC-MS/MS | 35.7 | 22.3 | 12.3 |
| 213a | CLIA | 30.4 | 19.4 | 6.5 |
| 213b | EIA | 26.3 | 17.3 | 12.9 |
| 214a | RIA | 27.5 | 31.3 | 11.5 |
| 214b | CLIA | 31.2 | 14.4 | 13.0 |
| 214c | LC-MS/MS | 31.7 | 20.5 | 12.2 |
| 215 | LC-MS/MS | 37.2 | 21.2 | 13.6 |
| 216 | LC-MS/MS | 32.9 | 21.9 | 12.8 |
| 218a | CLIA | 33.7 | 14.5 | 14.5 |
| 218b | LC-MS/MS | 31.8 | 23.8 | 12.2 |
| 220 | LC-MS/MS | 36.0 | 23.0 | 15.0 |
| 221a | LC-MS/MS | 36.7 | 31.7 | 15.0 |
| 221b | LC-UV | 37.5 | 36.3 | 10.7 |
| 222 | CLIA | 35.2 | 23.5 | 11.1 |
| 225 | LC-MS/MS | 31.0 | 14.8 | 10.4 |
| 228a | LC-MS/MS | 35.4 | 21.5 | 12.7 |
| 231 | LC-UV | 30.2 | 30.8 | 8.5 |
| 241 | LC-MS/MS | 28.7 | 17.2 | 12.3 |
| 242 | LC-MS/MS | 31.1 | 18.7 | 12.2 |
| 243a | LC-UV | 27.8 | 19.6 | 11.5 |
| 243b | LC-MS/MS | 28.2 | 21.1 | 12.4 |
| 244 | LC-MS/MS | 33.0 | 25.0 | 13.0 |
| 247a | CLIA | 31.9 | 16.7 | 11.7 |
| 247b | EIA | 31.0 | 22.3 | 15.7 |
| 249 | LC-MS/MS | 33.3 | 21.7 | 12.6 |
| 251 | LC-MS/MS | 34.4 | 21.6 | n/r |
| 253 | LC-MS/MS | 35.5 | 21.0 | 13.1 |
| 255 | LC-MS/MS | 32.9 | 20.2 | 12.8 |
| 256 | CLIA | 30.4 | 17.3 | 18.5 |
| 257 | CLIA | 31.9 | 22.4 | 6.2 |
| 258 | CLIA | 32.8 | 17.5 | 11.4 |
| 259 | LC-MS/MS | 26.1 | 25.5 | 14.8 |
| 260 | EIA | 32.0 | 22.0 | 18.2 |
| 261 | CLIA | 32.3 | 16.2 | 10.3 |
| 262 | CLIA | 34.1 | 52.1 | 16.5 |
| 263 | CLIA | 28.7 | 13.4 | 12.5 |
| 264 | LC-MS/MS | 41.4 | 29.0 | 11.4 |
| 265 | LC-MS/MS | 39.0 | 24.0 | 14.0 |

n/r = not reported or not determined

| | | VitDQAP-III | SRM 909c | SRM 968d L1 |
|---|---|---|---|---|
| | | Vial A | Vial B | Control |
| All methods | N | 71 | 71 | 70 |
| | Median | 32.3 | 21.2 | 12.8 |
| | MADe | 3.1 | 4.3 | 1.4 |
| | CV% | 9.5 | 20 | 11 |
| IA methods | N | 28 | 28 | 28 |
| | Median | 31.1 | 18.2 | 13.0 |
| | MADe | 3.0 | 5.6 | 2.3 |
| | CV% | 9.5 | 31 | 18 |
| LC methods | N | 43 | 43 | 42 |
| | Median | 33.4 | 21.7 | 12.8 |
| | MADe | 3.6 | 3.3 | 0.7 |
| | CV% | 11 | 15 | 5.8 |
| LC-$MS^n$ | N | 37 | 37 | 36 |
| | Median | 33.8 | 21.6 | 12.8 |
| | MADe | 3.1 | 2.1 | 0.7 |
| | CV% | 9.0 | 9.6 | 5.3 |
| NIST Value | | 32.7 | 20.7 | 12.5 |
| *U* | | 0.7 | 0.7 | 0.4 |

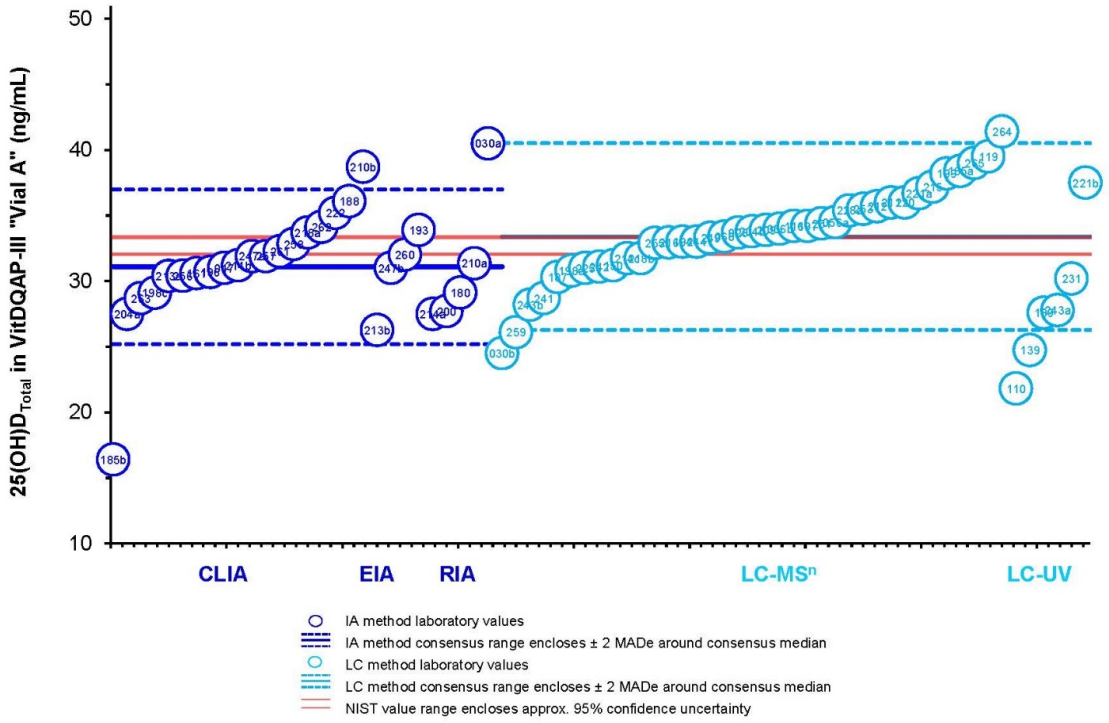



**Figure 8**: Example and Description of a Consensus Data Plot (Sample VitDQAP-III)

Results are for 25(OH)$D_{Total}$, Vial A, in the Winter 2014 Comparability Study (Exercise 8). The solid lines (——) and (——) represent the consensus medians, and the dashed lines (- - - - -) and (- - - - -) represent the approximate 95 % confidence intervals (2 × $MAD_e$) of the major techniques, IA or LC. Participants with results that fall between the two dashed lines are within the consensus variability area for their technique, IA or LC. The red lines (——) represent the NIST value and its associated 95 % expanded uncertainty (i.e., value $\pm U_{95}$) for the study sample (in this case, VitDQAP-III). NIST has confidence that the "true" value for the material lies within this interval. The NIST lines fall within the consensus ranges for both IA and LC for this sample; however, when they do not, there may be method bias.

### *7.3 Youden Plots*

Laboratory values and consensus results between multiple materials were also compared using Youden plots to help identify trends of performance among the participants and to identify any specific technical issues within a method. Youden plots are commonly utilized graphical techniques for analyzing interlaboratory data to compare the performance of a single participant's performance on two or more study materials. Figure 9 presents an example and description of a Youden plot.

For this example, the Youden line runs through the center of both the IA and LC consensus boxes, illustrating that most of the IA and LC results agree with each other and with the NIST results for these materials.

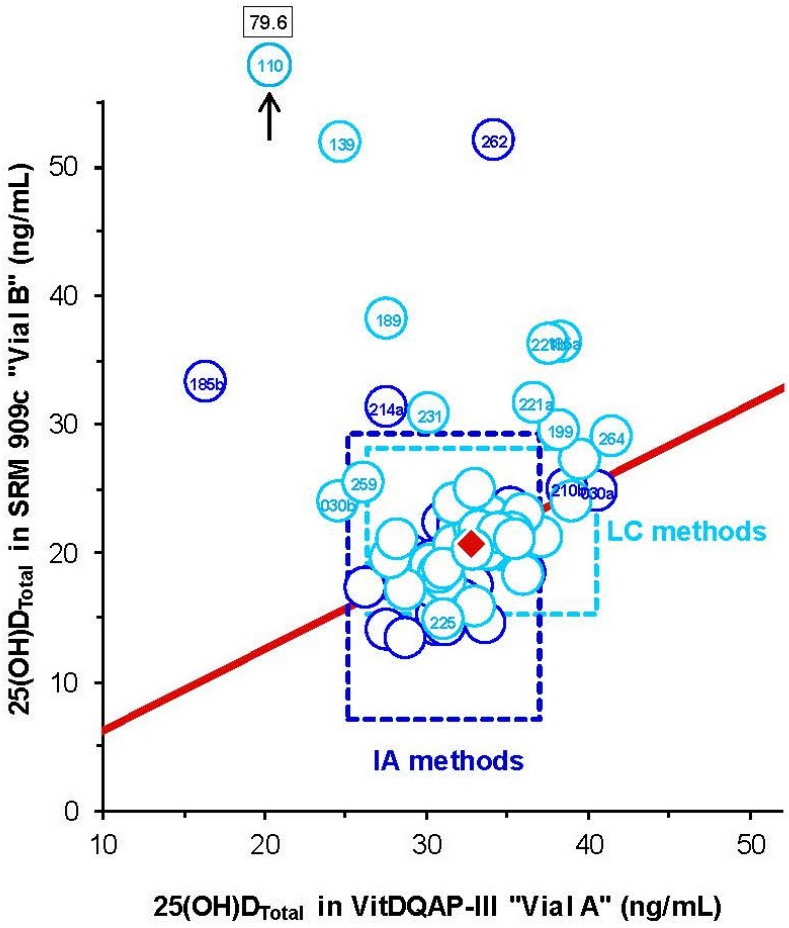



**Figure 9**: Example and Description of a Youden Plot for 25(OH)$D_{Total}$

This plot compares participant results for sample VitDQAP-III ("Vial A" on x-axis) with SRM 909c ("Vial B" on y-axis) from the Winter 2014 Comparability Study (Exercise 8). There are two blue consensus boxes, one for results from IA methods (- - - - -) and one for results from LC methods (- - - - -). Participant results (depicted as circles) that are within the consensus range for both study materials are within the blue consensus boxes for their technique. Conversely, results that fall outside of (or on the edge of) either of the consensus boxes are not included in the consensus ranges and are highlighted with their code numbers. The NIST values for the materials are denoted with a red diamond symbol (◆), and the red line (——) represents the Youden 45 °line centered on the NIST values. Participant values spread along the line indicate differences in calibration. Values at right-angles far from the line represent methods that respond differently to the two samples. The result for Participant 110 for SRM 909c is off the y-axis (79.6 ng/mL).

*7.4 Reporting*

The tabular and graphical results for each exercise of the VitDQAP were assembled into formal reports that summarized the exercise and provided a detailed discussion of the data, results, and significant conclusions. The reports were distributed to study participants soon after the closing date for each exercise. Participants were afforded the opportunity to correct any errors and have any concerns addressed prior to publishing the exercise results as a NISTIR (Table 1). The participant results were also summarized and reported in two separate peer-reviewed manuscripts, one summarizing Exercises 1 to 6 [14], and the second summarizing Exercises 7 to 12 [1].

## 8. RESULTS AND DISCUSSION

### *8.1 Results for Controls*

For the first six exercises of the program, the ethanol based SRM 2972 (25-Hydroxyvitamin $D_2$ and $D_3$ Calibration Solutions) was distributed as the control material. As this SRM was intended for calibration of analytical methods, it was postulated that this material would be a good mimic of participant solutions and would inform them of any potential biases in their calibration prior to analyzing the study materials. However, we learned from participants using IA methods that the ethanolic solutions were not compatible with their assays, which generally use matrix-matched (i.e., serum-based) materials as calibrants. Over the course of the first 6 exercises, the data returned for SRM 2972 decreased due to the incompatibility of the ethanol matrix with most IA methods and in exercise 7, the control was switched to the human serum material SRM 968d-L1 to be more applicable to all methods.

Prior to establishing SRM 968d-L1 as the measurement control material in Exercises 7 through 12, the material was evaluated as a blinded study sample in Exercises 2 through 6. Therefore, participant performance for this material could be evaluated across the duration of the VitDQAP except for the pilot study, Exercise 1. Figure 10 provides a graphical representation of the results for $25(OH)D_{Total}$ in SRM 968d-L1 over time. This box-and-whisker plot depicts the participant performance summarized by consensus boxes representing results for all methods (LC and IA), IA methods only, and the LC methods only. Figure 10 further provides the NIST value range for this material to facilitate comparison to the true value.

Several key trends in the participant performance can be observed in Figure 10. The box and whisker heights indicate that the IA results exhibit consistent, relatively large variability over time, whereas the LC results show a decrease in variability during the VitDQAP. Additionally, the IA results are consistently biased high relative to both the LC results and to the NIST value. While the LC results are also consistently biased somewhat high relative to the NIST value, the agreement was improved from Exercise 7 to Exercise 12. While the reasons for the improvement in the variability and accuracy of the LC methods are not completely known, the conversion of SRM 968d-L1 to a control material in Exercise 7 is one potential contributing factor. When used as a control, participants were provided with the NIST value at the onset of each exercise and were encouraged to repeat their measurements until they achieved a result that agreed with the NIST value. The change in instructions to participants could partially explain the better alignment of the reported LC results with the NIST value in the latter exercises. In contrast, the IA methods rely on kits from the manufacturer and are not generally ‘adaptable,’ hence the IA performance was irrespective of whether it was a blinded study sample or a known control.

A summary of a cohort of laboratories that participated across most exercises of the VitDQAP, and their results for SRM 968d-L1 are provided in a table and box-and whisker plot, respectively, in Appendix C. The cohort results for SRM 968d-L1 suggest similar trends in community measurement accuracy and precision as seen in Figure 10 for all methods (IA and LC) and for LC methods only.

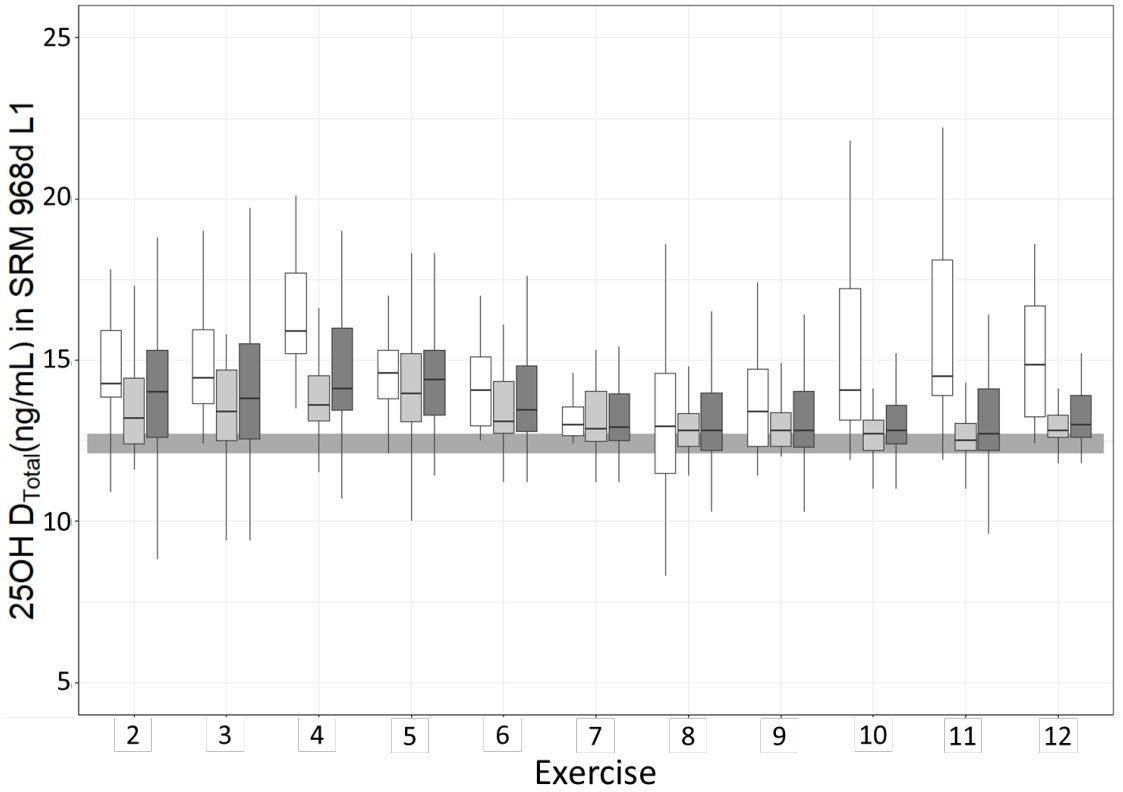



**Figure 10**: Box-and-Whisker Plot for 25(OH)$D_{Total}$ for SRM 968d-L1

This plot summarizes the results reported for SRM 968d-L1 when used as a blinded sample (exercises 2 to 6) and as the control serum (exercises 7 to 12) throughout the VitDQAP. Each box represents the 25 % to 75 % quartile range, and the horizontal line contained within representing the median value (50 % quartile). The results for IA (☐), LC (▒), and IA and LC (■) are depicted separately. The error bars (whiskers) represent the empirically determined 95 % range. The gray-shaded bar represents the range bound by the NIST value and its $U_{95}$ uncertainty.

### *8.2 Results for 25(OH)$D_{Total}$ in Study Samples*

The results for SRM 968d-L1 are representative of the results for most study materials evaluated in the VitDQAP, which also contained 25(OH)$D_3$ as the predominant metabolite comprising 25(OH)$D_{Total}$. For these materials, the interlaboratory CVs were relatively large and in the range from ≈7 % to ≈20 %, and the results from both IA and LC techniques were consistently biased somewhat high, indicating accuracy is a continuing issue in the community. The source of the high bias can stem from different factors. For example, a high bias could arise if the purity of the standards used to calibrate the participant methods was not appropriately considered. The 25-hydroxyvitamin D standards are notably hygroscopic, and the purity of the primary standards was rigorously determined and accounted for in the NIST values for the control and study materials that provide the accuracy base in the VitDQAP. Another probable contributor of high bias is measurement interference from other vitamin D metabolites in human serum. A discussion of two potential interferents leading to a high bias follows.

3-epi-25(OH)$D_3$. This metabolite is not generally considered to be a source of bias for the major IA methods, but it can present a bias to LC methods that do not chromatographically separate it

from $25(OH)D_3$. Since 3-epi-$25(OH)D_3$ is detected at the same mass or mass transition as $25(OH)D_3$, it cannot be separated using $MS^n$. The 3-epi-$25(OH)D_3$ generally correlates with the level of $25(OH)D_3$ in the sample. Therefore, for many patient samples, 3-epi-$25(OH)D_3$ will be a consistent source of bias for LC methods that do not fully separate the metabolites.

In the VitDQAP, this measurement interference was evaluated in a single study material, SRM 972a-L4, which was fortified with 3-epi-$25(OH)D_3$. The program results for the evaluation of this material are presented in the report for Exercise 7 [8], in Wise et.al. [1], and in Figure 11 in this current report. The results for $25(OH)D_{Total}$ are bimodal, with 21 out of 30 LC-$MS^n$ values and all 3 LC-UV values biased high relative to both the NIST value and the IA median value, with the other 9 LC-$MS^n$ values clustered near the NIST value. The LC values clustered around the NIST value likely used methods that chromatographically separate the 3-epi-$25(OH)D_3$ from $25(OH)D_3$, but most of the LC methods used in the VitDQAP did not separate the metabolites. LC participants in the VitDQAP were encouraged to modify their methods as the metabolites are readily separable using methods with some cyanopropyl or pentafluorphenylpropyl analytical columns, but modifying approved methods can be challenging for some clinical laboratories.

Most of the IA results agreed with the NIST value for SRM 972a-L4, indicating 3-epi-$25(OH)D_3$ does not pose a significant bias for IA methods. One caveat is that the 3-epi-$25(OH)D_3$ was spiked in the material and not bound the same way as the $25(OH)D_3$, which could impact its detection in some immunoassays.

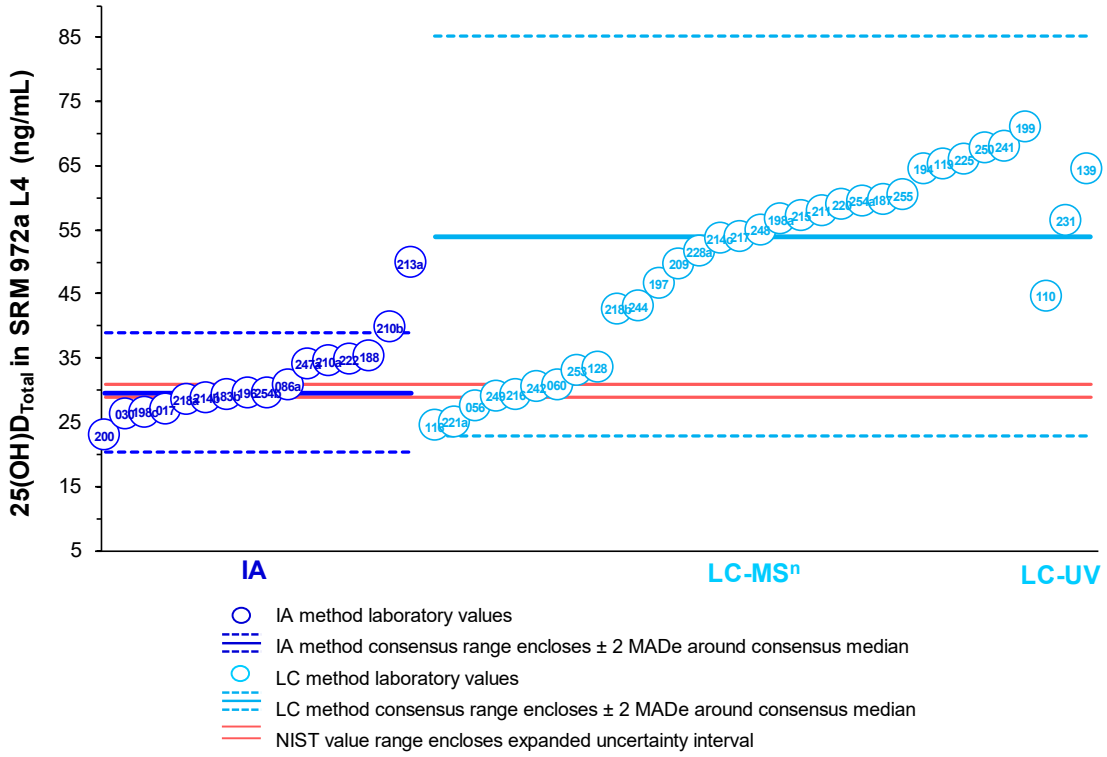



**Figure 11**: Consensus Data Plot for $25(OH)D_{Total}$ in SRM 972a-L4, Exercise 7

The description of the lines and color-coding is the same as in the caption for Figure 8.

24R,25-dihydroxyvitamin $D_3$. This metabolite is not a significant source of bias for LC techniques but is a potential source of bias for IA methods where it can cross-react with the binding assay and yield high results for $25(OH)D_{Total}$. Like 3-epi-$25(OH)D_3$, 24R,25-dihydroxyvitamin $D_3$ is generally expected to correlate with $25(OH)D_3$ levels in human serum samples and could pose a consistent source of bias for some IA methods.

Measurement (and potential interference with IA methods) of the 24R,25-dihydroxyvitamin $D_3$ metabolite was not explored in the VitDQAP. However, it is increasingly studied as a potential analyte for clinical diagnostics. For the last three exercises of the VitDQAP, two participants voluntarily reported values for this metabolite in at least one of the studies; values are reported in NISTIRs 8142, 8134, and 8169 [11, 12, 13]. Just prior to the end of the VitDQAP, NIST also developed and published a RMP for 24R,25-dihydroxyvitamin $D_3$ [24], which was used to assign values in the SRMs specifically designed for analysis of vitamin D metabolites in human serum, SRM 972a and SRM 2973.

*8.3 Study Materials with Appreciable 25(OH)$D_2$*

While the general trend observed in the VitDQAP was for the participant results to be biased high relative to the NIST value, a low bias was also observed for most of the IA results for materials containing appreciable amounts of the $25(OH)D_2$ metabolite. In the VitDQAP, measurement interference from the $25(OH)D_2$ metabolite was explored with three study materials: SRM 972-L3, which was fortified with $25(OH)D_2$; SRM 972a-L3, which contained a relatively high endogenous concentration of $25(OH)D_2$; and VitDQAP-III, which contained a measurable 'intermediate' amount of endogenous $25(OH)D_2$ (see Table 2).

The program results for SRM 972-L3 and SRM 972a-L3 were reported in Exercise 3 [4] and Exercise 6 [7], respectively, and collectively in Bedner et al. [14] as well as Figure 12 in this report. Both SRM 972-L3 and SRM 972a-L3 contain significant levels of $25(OH)D_2$, with NIST values of 26.4 ± 2.0 ng/mL and 13.3 ± 0.3 ng/mL, respectively. For SRM 972-L3, which contained augmented $25(OH)D_2$, all IA results for $25(OH)D_{Total}$ were biased low relative to the NIST value, which is likely attributable to two factors. IA methods that do not utilize an extraction step tend to under-recover the spiked $25(OH)D_2$, which is not bound to the matrix like endogenous $25(OH)D_2$. Bias also likely arises from the non-equivalent, lower response to $25(OH)D_2$ versus the $25(OH)D_3$ metabolite, both of which comprise $25(OH)D_{Total}$. For SRM 972a-L3, which contains only endogenous $25(OH)D_2$, 14 out of 18 results IA results were biased low relative to the NIST value. The overall difference in performance between the IA methods and the LC methods for SRM 972a-L3 was not as dramatic as it was for SRM 972-L3.

The VitDQAP-III material was evaluated both in Exercise 8 (Winter 2014) and Exercise 10 (Winter 2015), and it was specifically obtained to represent an 'intermediate' level of $25(OH)D_2$ with a NIST value of 6.5 ± 0.2 ng/mL. The program results for VitDQAP-III from Exercise 8 are presented in Figure 8 in this report. Unlike the results for SRM 972-L3 and SRM 972a-L3, the IA method results for VitDQAP-III overlap almost completely with the LC results (Figure 8). However, like the results for SRM 972-L3 and SRM 972a-L3, the median IA result is biased lower that the NIST and LC median values. For VitDQAP-III, the level of $25(OH)D_2$ is likely not high enough to reveal major differences that are discernable from the overall variability of the results for IA and LC.

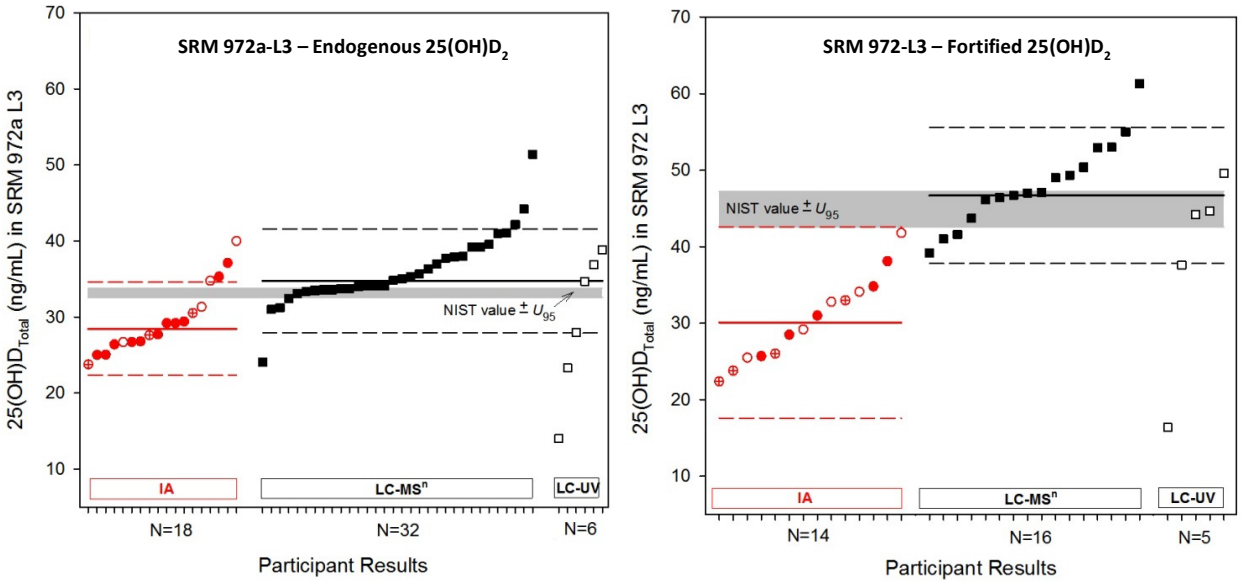



**Figure 12**: Consensus Data Plots for 25(OH)$D_{Total}$ in SRM 972a-L3 and SRM 972-L3

These plots compare results reported for SRM 972-L3 in the Winter 2011 Comparability Study (Exercise 3) and SRM 972a-L3 in the Summer 2012 Comparability Study (Exercise 6). The results from the individual methods are displayed with different symbols, including: CLIA (●), EIA, (⊕), RIA (○), LC-$MS^n$ (■), and LC-UV (□). For each of the techniques within both graphs, IA and LC, the solid lines (——) and (——) represent the consensus median and the dashed lines (- - - - -) and (- - - - -) represent approximate 95 % confidence intervals (2 × $MAD_e$). The grey-shaded bars represent the ranges bound by the NIST values with its $U_{95}$ uncertainty.

While the difference in results between IA and LC for materials with high levels of 25(OH)$D_2$ was one of the most significant observations in the VitDQAP, 25(OH)$D_2$ is likely to become less of a concern in serum patient samples over time. Since vitamin $D_2$ is not the human form of vitamin D, its only source is through dietary intake of certain foods and supplements. Many vitamin D supplements in the US now contain vitamin $D_3$, whereas they used to contain vitamin $D_2$. The decreased prevalence of vitamin $D_2$ in human diets correlates with decreased production of its metabolite 25(OH)$D_2$, and hence decreased impact to IA method performance for 25(OH)$D_{Total}$.

## 9. ACKNOWLEDGEMENTS

Financial support for the VitDQAP was provided by the National Institutes of Health, Office of Dietary Supplements (NIH-ODS). We thank our NIH-ODS colleagues Paul Coates, Joseph Betz, Chris Sempos, Stephen Wise, and Adam Kuszak for their many years of support for the program.

We thank our NIST colleagues David Duewer, Karen Phinney, and Benjamin Place for their editorial assistance. The time and effort of the analysts and management of the participating laboratories of the VitDQAP are also gratefully acknowledged.

## 10. REFERENCES


[1] 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 (2017) 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. *Journal of AOAC International* 100:1260-1276. https://doi.org/10.5740/jaoacint.17-0305

[2] Bedner M, Lippa KA (2012) NIST/NIH Vitamin D Metabolites Quality Assurance Program Report of Participant Results: Winter 2010 Comparability Study (Exercise 1). NISTIR 7890. https://doi.org/10.6028/NIST.IR.7890

[3] Lippa KA, Bedner M, Tai SSC (2012) NIST/NIH Vitamin D Metabolites Quality Assurance Program Report of Participant Results: Summer 2010 Comparability Study (Exercise 2). NISTIR 7891. https://doi.org/10.6028/NIST.IR.7891

[4] Bedner M, Lippa KA (2012) NIST/NIH Vitamin D Metabolites Quality Assurance Program Report of Participant Results: Winter 2011 Comparability Study (Exercise 3). NISTIR 7892. https://doi.org/10.6028/NIST.IR.7892

[5] Lippa KA, Bedner M, Tai SSC (2012) NIST/NIH Vitamin D Metabolites Quality Assurance Program Report of Participant Results: Summer 2011 Comparability Study (Exercise 4). NISTIR 7893. https://doi.org/10.6028/NIST.IR.7893

[6] Bedner M, Lippa KA, Tai SSC (2013) NIST/NIH Vitamin D Metabolites Quality Assurance Program Report of Participant Results: Winter 2012 Comparability Study (Exercise 5). NISTIR 7894. https://doi.org/10.6028/NIST.IR.7894

[7] Lippa KA, Bedner M (2013) NIST/NIH Vitamin D Metabolites Quality Assurance Program Report of Participant Results: Summer 2012 Comparability Study (Exercise 6). NISTIR 7895. https://doi.org/10.6028/NIST.IR.7895

[8] Bedner M, Lippa KA, Tai SSC (2015) NIST/NIH Vitamin D Metabolites Quality Assurance Program Report of Participant Results: Summer 2013 Comparability Study (Exercise 7). NISTIR 8000. https://doi.org/10.6028/NIST.IR.8000

[9] Bedner M, Tai SSC (2016) NIST/NIH Vitamin D Metabolites Quality Assurance Program Report of Participant Results: Winter 2014 Comparability Study (Exercise 8). NISTIR 8133. https://doi.org/10.6028/NIST.IR.8133

[10] Bedner M, Tai SSC (2016) NIST/NIH Vitamin D Metabolites Quality Assurance Program Report of Participant Results: Summer 2014 Comparability Study (Exercise 9). NISTIR 8141. https://doi.org/10.6028/NIST.IR.8141

[11] Bedner M (2016) NIST/NIH Vitamin D Metabolites Quality Assurance Program Report of Participant Results: Winter 2015 Comparability Study (Exercise 10). NISTIR 8142. https://doi.org/10.6028/NIST.IR.8142

[12] Bedner M (2016) NIST/NIH Vitamin D Metabolites Quality Assurance Program Report of Participant Results: Summer 2015 Comparability Study (Exercise 11). NISTIR 8143. https://doi.org/10.6028/NIST.IR.8143

[13] Bedner M, Burdette CQ (2017) NIST/NIH Vitamin D Metabolites Quality Assurance Program Report of Participant Results: Summer 2016 Comparability Study (Exercise 12). NISTIR 8169. https://doi.org/10.6028/NIST.IR.8169

[14] Bedner M, Lippa KA, Tai SSC (2013) An assessment of 25-hydroxyvitamin D measurements in comparability studies conducted by the Vitamin D Metabolites Quality Assurance Program. *Clinica Chimica Acta* 426:6-11. https://doi.org/10.1016/j.cca.2013.08.012

[15] Bedner M, Phinney KW (2012) Development and Comparison of Three Liquid Chromatography-Atmospheric Pressure Chemical Ionization/Mass Spectrometry Methods for Determining Vitamin D Metabolites in Human Serum. *Journal of Chromatography. A* 1240: 132-139. https://doi.org/10.1016/j.chroma.2012.03.091

[16] Duewer DL, Thomas JB (2013) NIST Micronutrients Measurement Quality Assurance Program Winter 2008 Comparability Studies: Results for Round Robin LXIII Fat-Soluble Vitamins and Carotenoids in Human Serum and Round Robin 28 Ascorbic Acid in Human Serum. NISTIR 7880-9. https://dx.doi.org/10.6028/NIST.IR.7880-9

[17] Tai SS, Bedner M, Phinney KW (2010) Development of a candidate reference measurement procedure for the determination of 25-hydroxyvitamin $D_3$ and 25-hydroxyvitamin $D_2$ in human serum using isotope-dilution liquid chromatography-tandem mass spectrometry. *Analytical Chemistry* 82:1942-1948. https://doi.org/10.1021/ac9026862

[18] Certificate of Analysis, SRM 1950 Metabolites in Frozen Human Plasma. https://www-s.nist.gov/srmors/certificates/1950.pdf

[19] Certificate of Analysis, SRM 972a Vitamin D Metabolites in Frozen Human Serum. https://www-s.nist.gov/srmors/certificates/972A.pdf

[20] 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 (2017) Development of an Improved Standard Reference Material for Vitamin D Metabolites in Human Serum. *Analytical Chemistry* 89:4907-4913. https://doi.org/10.1021/acs.analchem.6b05168

[21] Certificate of Analysis, SRM 968f Fat-Soluble Vitamins in Frozen Human Serum. https://www-s.nist.gov/srmors/certificates/968f.pdf

[22] Certificate of Analysis, SRM 2973 Vitamin D Metabolites in Frozen Human Serum (High Level). https://www-s.nist.gov/srmors/certificates/2973.pdf

[23] Tai SSC, Nelson MA, Bedner M, Lang BE, Phinney KW, Sander LC, Yen JH, Betz JM, Sempos CT, Wise SA (2017) Development of Standard Reference Material (SRM) 2973 Vitamin D Metabolites in Frozen Human Serum (High Level). *Journal of AOAC International* 100:1294-1303. https://doi.org/10.5740/jaoacint.17-0182

[24] Tai SS, Nelson MA (2015) Candidate Reference Measurement Procedure for the determination of (24R),25-Dihydroxyvitamin $D_3$ in human serum using isotope-dilution liquid chromatography-tandem mass spectrometry. *Analytical Chemistry* 87:7964-7970. https://doi.org/10.1021/acs.analchem.5b01861

# 11. APPENDIX A: LIST OF PARTICIPATING ORGANIZATIONS

## INDUSTRY or NON-PROFIT

Acibadem Labmed Clinical Laboratories
Acquity Labs
Affiliated Medical Services Laboratory Inc.
Alfred Pathology Service
ARUP Laboratories, Inc.
Ascend Clinical
Beijing Lawke Health Lab
Biolab Medical Unit
BioReference Laboratories
BML
Canisius-Wilhelmina Ziekenhuis
Care S.r.l.
Centro Laboratuvaries
Craft Technologies, Inc.
DIA Source ImmunoAssays
Diasorin
DSM Nutritional Products, Ltd
ESA - A Dionex Company
Fred Hutchinson Cancer Research Center
Fujirebio Inc.
Global Diagnostic Laboratory
Green Cross Reference Laboratory
Immunodiagnostic Systems Inc
Instituto de Investigaciones Metabolicas
Mater Pathology
Med Fusion
Metabolon, Inc.
Metametrix Clinical Laboratory
MuirLab
NMS LABS
Novilytic
Nutritional Biochemistry Lab, CVS
Pathology Associates Medical Laboratories
PerkinElmer
PPD
Quantimetrix Corporation
Randox Laboratories Ltd.
Rheumatic Disease Center
Seegene Medical Foundation
Seoul Clinical Laboratories (SCL)
Shenzhen New Industries Biomedical Engineering Co. Ltd. (SNIBE Co. Ltd.)
South Bend Medical Foundation
Thermo Fisher Scientific
TPMG Kaiser Regional Laboratories
Zivak Technologies

### HOSPITAL or CLINIC

Cheil General Hospital & Women's Healthcare Center
Children’s Hospital National Medical Center
Children’s Mercy Hospital
Harborview Medical Center
Jeffrey A. Alper, M.D. Laboratory
Laboratoire de Biochimie Hôpital Bichat
Laboratoire de Biochimie Spécialisée
Laboratorio Analisi - Azienda Ospedaliera di Desio e Vimercate
Lahey Clinic
Massachusetts General Hospital
Mayo Clinic
Providence Regional Medical Center
Royal Wolverhampton NHS Trust, New Cross Hospital
Seattle Childrens Regional Medical Center
St. Joseph Medical Center
St. Paul's Hospital
The Children’s Hospital (Colorado)
The Children's Hospital of Philadelphia
University Hospital of North Staffordshire NHS Trust

### GOVERNMENT

Centers for Disease Control and Prevention (CDC)
Medical Research Council MRC), Human Nutrition Research (HNR)
Kenya Medical Research Institute (KEMRI)
NASA, Nutritional Biochemistry Lab
Pathology Queensland-Central
TÜBİTAK UME
U.S. Food and Drug Administration
USDA-ARS, Western Human Nutrition Center

### UNIVERSITY HOSPITAL or CLINIC

Centre Hospitalier Universitaire de Grenoble
Hospital of the University of Pennsylvania
Hospital Universitario Puerta de Hierro Majadahondra, Edificio Laboratorios (Peine 7) - Planta 1a
Marmara Universitesi Pendik Egitim Arastirma Hastanesi
Ramathibodi Hospital
Seoul National University Bundang Hospital
Tufts Medical Center
UCLA Center for Human Nutrition
Umass Memorial Medical Center
University of California San Diego Health, Moores Cancer Center
University of California San Diego, Center for Advanced Laboratory Medicine
University of North Carolina (Medical Center)

University of Pittsburgh Medical Center
University of Wisconsin Hospital & Clinics
VCU Medical Center Clinical Chemistry Laboratory
Winthrop University Hospital

### ACADEMIC

Albert Einstein College of Medicine
CUHK Li Ka Shing Medical Sciences
Emory University School of Medicine
Harvard T.H. Chan School of Public Health
Medical College of Georgia
Muhimbili University of Health and Allied Sciences
Oregon Health and Science University
Tufts University
University of Alabama Birmingham
University of Hawai‘i at Manoa
University of Wisconsin-Madison, Wisconsin Primate Research Center/ACTR Core Lab

## 12. APPENDIX B: FULL PARTICIPANT TIMELINE


| Code | |
|---|---|
| CLIA | LC-MS/MS |
| EIA | LC-UV |
| RIA | LC-ECD |
| CLEIA | |

| | | Exercise | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| **Participant** | **Method** | **1** | **2** | **3** | **4** | **5** | **6** | **7** | **8** | **9** | **10** | **11** | **12** |
| 213b | EIA | | | | | | | | | | | | |
| 214a | RIA | | | | | | | | | | | | |
| 214b | CLIA | | | | | | | | | | | | |
| 214c | LC-MS/MS | | | | | | | | | | | | |
| 215 | LC-MS/MS | | | | | | | | | | | | |
| 216 | LC-MS/MS | | | | | | | | | | | | |
| 217 | LC-MS/MS | | | | | | | | | | | | |
| 218a | CLIA | | | | | | | | | | | | |
| 218b | LC-MS/MS | | | | | | | | | | | | |
| 219 | LC-MS/MS | | | | | | | | | | | | |
| 220a | LC-MS/MS | | | | | | | | | | | | |
| 221a | LC-MS/MS | | | | | | | | | | | | |
| 221b | LC-UV | | | | | | | | | | | | |
| 221c | LC-MS | | | | | | | | | | | | |
| 222 | CLIA | | | | | | | | | | | | |
| 223 | LC-MS/MS | | | | | | | | | | | | |
| 225 | LC-MS/MS | | | | | | | | | | | | |
| 228a | LC-MS/MS | | | | | | | | | | | | |
| 228b | CLIA | | | | | | | | | | | | |
| 231a | LC-UV | | | | | | | | | | | | |
| 231b | CLIA | | | | | | | | | | | | |
| 234 | LC-MS/MS | | | | | | | | | | | | |
| 236 | CLIA | | | | | | | | | | | | |
| 241 | LC-MS/MS | | | | | | | | | | | | |
| 242 | LC-MS/MS | | | | | | | | | | | | |
| 243a | LC-UV | | | | | | | | | | | | |
| 243b | LC-MS/MS | | | | | | | | | | | | |
| 244 | LC-MS/MS | | | | | | | | | | | | |
| 245 | LC-UV | | | | | | | | | | | | |
| 247a | CLIA | | | | | | | | | | | | |
| 247b | EIA | | | | | | | | | | | | |
| 248 | LC-MS/MS | | | | | | | | | | | | |
| 249 | LC-MS/MS | | | | | | | | | | | | |
| 250 | LC-MS/MS | | | | | | | | | | | | |
| 251 | LC-MS/MS | | | | | | | | | | | | |
| 253 | LC-MS/MS | | | | | | | | | | | | |
| 254a | LC-MS/MS | | | | | | | | | | | | |
| 254b | CLIA | | | | | | | | | | | | |
| 255 | LC-MS/MS | | | | | | | | | | | | |
| 256 | CLIA | | | | | | | | | | | | |
| 257 | CLIA | | | | | | | | | | | | |
| 258 | CLIA | | | | | | | | | | | | |
| 259 | LC-MS/MS | | | | | | | | | | | | |
| 260 | EIA | | | | | | | | | | | | |
| 261 | CLIA | | | | | | | | | | | | |
| 262 | CLIA | | | | | | | | | | | | |
| 263 | CLIA | | | | | | | | | | | | |
| 264 | LC-MS/MS | | | | | | | | | | | | |
| 265 | LC-MS/MS | | | | | | | | | | | | |
| 266a | LC-UV | | | | | | | | | | | | |
| 266b | EIA | | | | | | | | | | | | |
| 267 | CLEIA | | | | | | | | | | | | |
| 268a | RIA | | | | | | | | | | | | |
| 268b | EIA | | | | | | | | | | | | |
| 269 | LC-MS/MS | | | | | | | | | | | | |
| 270 | LC-MS/MS | | | | | | | | | | | | |
| 271 | LC-MS/MS | | | | | | | | | | | | |
| 272 | LC-MS/MS | | | | | | | | | | | | |
| 273 | EIA | | | | | | | | | | | | |
| 274 | CLIA | | | | | | | | | | | | |
| **Unique participants** | | 16 | 36 | 29 | 38 | 50 | 47 | 41 | 57 | 51 | 47 | 44 | 31 |
| **Total reported results** | | 17 | 39 | 35 | 45 | 57 | 56 | 49 | 71 | 63 | 58 | 55 | 38 |

| Code | |
|---|---|
| CLIA | LC-MS/MS |
| EIA | LC-UV |
| RIA | LC-ECD |
| CLEIA | |

## 13. APPENDIX C: SUMMARY AND RESULTS FOR A COHORT OF 10 LABORATORIES THAT PARTICIPATED ACROSS MOST EXERCISES

| 10 Lab Cohort | | |
|---|---|---|
| **Participant** | **Method** | **Exercises** |
| 056a | LC-MS/MS | 2 to 12 |
| 110 | LC-UV | 2 to 12 |
| 116 | LC-MS/MS | 2 to 12 |
| 188 | CLIA | 2 to 12 |
| 194 | LC-MS/MS | 2, 3, 5-10, 12 |
| 196 | CLIA | 2 to 12 |
| 197 | LC-MS/MS | 2 to 12 |
| 199 | LC-MS/MS | 2 to 12 |
| 209 | LC-MS/MS | 3 to 12 |
| 211 | LC-MS/MS | 3 to 12 |