Limitations of Platform Assays to Measure Serum 25OHD Level Impact on Guidelines and Practice Decision Making

M Rahme[1], **Shaar L Al**[1], **R Singh**[3], **R Baddoura**[2], **G Halaby**[2], **A Arabi**[1], **RH Habib**[1], **R Daher**[1], **D Bassil**[1], **K El-Ferkh**[1], **M Hoteit**[1], and **Fuleihan G El-Hajj**[1]

[1.] American University of Beirut Medical Center

[2.] St Joseph University, Beirut Lebanon,

[3.] Mayo Clinic Laboratories, Mayo Clinic, USA.

Abstract

Context: Liquid Chromatography Mass Spectroscopy (LC-MS/MS) is the preferred method to measure 25 hydroxyvitamin D (25OHD) levels, but laboratories are increasingly adopting automated platform assays.

Objective: We assessed the performance of commonly used automated immunoassays, with that of LC-MS/MS, and the National Institute of Standards and Technology (NIST) reference values, to measure 25OHD levels.

Methods/Setting: We compared serum 25OHD levels obtained from 219 elderly subjects, enrolled in a vitamin D trial, using the Diasorin Liaison platform assay, and the tandem LC-MS/MS method. We also assessed the performance of the Diasorin and Roche automated assays, expressed as mean % bias from the NIST standards, based on the vitamin D External Quality Assessment Scheme (DEQAS) reports, from 2013-2017.

Results: Serum 25OHD levels were significantly lower in the Diasorin compared to LC-MS/MS assay at baseline, 18.5 ± 7.8 vs 20.5 ± 7.6 ng/ml ($p<0.001$), and all other time points. Diasorin (25OHD) = 0.76x LC-MS/MS (25OHD) + 4.3, $R^2=0.596$. The absolute bias was independent of 25OHD values, and the pattern unfit for any cross-calibration. The proportion of subjects considered for vitamin D treatment based on pre-set cut-offs differed significantly between the 2

Address queries to: Ghada El-Hajj Fuleihan, MD, MPH Calcium Metabolism and Osteoporosis Program WHO Collaborating Center for Metabolic Bone Disorders American University of Beirut Medical Center, Beirut-Lebanon P.O. Box: 113-6044/C8 Telephone: 961-1-737868 gf01@aub.edu.lb

Department of Internal Medicine, Division of Endocrinology, Calcium Metabolism and Osteoporosis Program, WHO Collaborating Center for Metabolic Bone Disorders, (GEHF, AA,MR,DB, KEF, MH, LAS); Department of Internal Medicine (RH); and Department of Pathology & Laboratory Medicine (RD), American University of Beirut Medical Center, Beirut- Lebanon

Department of Endocrinology (GH) and Rheumatology (RB), Hotel Dieu de France, Beirut-Lebanon

Division of Clinical Biochemistry and Immunology, Mayo Clinic, Rochester, Minnesota, USA (RS)

All authors state that they have no conflict of interest

assays. There also was wide variability in the performance of both automated assays, compared to NIST reference values.

Conclusion: The performance of most widely used automated assays is sub-optimal. Our findings underscore the pressing need to re-consider current practices with regards to 25OHD measurements, interpretation of results from research studies, meta-analyses, the development of vitamin D guidelines, and their relevance to optimizing health.

Introduction

Hypovitaminosis D is quite prevalent worldwide, and when chronic, may have negative, all-be-it subtle, repercussions on skeletal health across the life course (1-3). The effects of this condition on a variety of acute (asthma, infection) and chronic, non-communicable diseases (cancer, diabetes, and cardiovascular) are more debatable (1-3).

Serum 25-hydroxyvitamin D (25OHD) reflects vitamin D nutritional status (4). With the recognition of the high prevalence of vitamin D insufficiency in the general population, testing for 25OHD levels has become an increasingly common practice. The Vitamin D External Quality Assessment Scheme (DEQAS), reported that in the Fall 2014, 26 different assay methods were used in over 1100 laboratories from 53 countries worldwide (5). The most common types of assays prototypes used by various manufacturers today are the antibody-based methods (kit-based), the increasingly popular rapid high-output automated methods (platform- based), and the high-performance liquid chromatography (HPLC)-based methods with either UV or mass spectrometric (MS) detection (5). The increasing high-test volume requires a simple, reproducible, high-throughput method, which accurately measures total circulating 25OHD ($25OHD_2$ and $25OHD_3$), thus the increasing popularity of platform assays in many commercial or hospital-based laboratories. In contrast, many academic research centers, and clinical centers of excellence are favoring the more established "gold standard" of LC-MS/MS.

Data from the National Health And Nutritional Epidemiologic Survey (NHANES) reveals that the mean serum 25OHD levels increased by 5-6 nmol/L (2-2.4ng/ml) between the periods of 1998-2006 and 2007-2010, findings that may be in part explained by the use of different assay methods, and possibly increasing vitamin D supplementation (6). Our group has reported a similar temporal trend in Lebanese patients (7). Potential source of variability between assays include differences in vitamin D assay methods, vitamin D metabolite extraction from vitamin D binding protein (DBP); differences in the detection of $25OHD_2$, cross-reactivity with 3-epi- $25OHD_3$ and other vitamin D metabolites, and matrix interferences, to name a few (8-11). The differences in measured serum 25OHD levels, within and also between assays, results in underestimation or overestimation of the actual 25OHD level depending on the specific assay used (5, 9, 10, 12), and on the type of supplement used by the patient (Vitamin D2 or vitamin D3) (13, 14). As part of our quality assurance program when transitioning from a radio-immunoassay to an automated platform assay, we had compared results from sera run in parallel on both assays on 494 samples, and showed significant differences between the radio-immuno-assay and the platform assay,

differences of −38 to +19 ng/mL, a bias that was independent of the serum 25OHD level (12).

The overall objective of this study is to investigate the variability in the performance of vitamin D assays, and its impact on the assessment of vitamin D status of subjects tested, and on interventions based on pre-set vitamin D guideline based cut-offs.

Methods

Study design

We systematically assessed the performance of 2 of the most commonly used automated platform immunoassays to-date, to measure serum 25OHD, using two approaches.

First, we compared the performance of DiaSorin LIAISON® 25OHD total assay (DiaSorin, Stillwater, MN, USA) against the established liquid chromatography-tandem mass spectrometry (LC-MS/MS) assay, based on samples obtained from elderly subjects enrolled in a vitamin D randomized trial, run in parallel using both assays.

Second, we assessed the performance of both the DiaSorin LIAISON® and the Roche Elecsys COBAS® (Roche Diagnostics International, Rotkreuz, Switzerland) platform assays, based on the calculated mean % bias obtained with both assays for the highest and the lowest NIST reference standard value, using the quarterly reports obtained from DEQAS, from 2013-2017.

Study population and trial design

Elderly, overweight, ambulatory subjects with a serum 25OHD between 10 and 30 ng/ml (using the Diasorin platform assay) at screening, were recruited from the greater Beirut Area. In total 257 elderly subjects were randomly assigned to receive high dose (3,750 IU/day) or low dose (600 IU/day) of vitamin D, and all subjects received 1000 mg of elemental calcium citrate (Europharm); 35 subjects (14%) did not complete the study with no vitamin D data available post study discontinuation (NCT01315366) (15). Briefly, 219 subjects had baseline 25OHD levels measured in parallel using both assays at study entry. Subjects had a mean age (range) of 70 (67-74) years, a mean BMI of 29.2 (26.8-32.4) kg/m^2, and 55% were females. Overall, 77% had one or more co-morbidity, and 69% met the definition of pre-diabetes but were not on oral hypoglycemic drugs (15).

Serum 25OHD measurements and assay methodology

Samples were measured in parallel using Diasorin Liaison total rapid automated assay chemiluminescent (Diasorin, Stillwater, MN) at American University of Beirut (AUB), and the LC-MS/MS method at the Mayo Clinic laboratories, Mayo Clinic, Rochester, Minnesota, USA. The LC-MS/MS at the Mayo Clinic Laboratories is the reference method traceable to the National Institute of Standards and Technology (NIST). Additional parallel measurements using both methods were done in subsets of patients at 6 months (N=171) and 12 months (N=126). Hormonal assays were run in batches and samples for levels drawn serially on the same patient (0,6,12 months) were included within the same assay run, as singlets, both for Diasorin and LC- MS/MS assays. Serum 25OHD values are reported in

nanograms per milliliter, and rounded to one decimal, (to convert to nanomoles per liter (SI units), multiply by 2.496).

Both institutions are participants in DEQAS for several years (www.deqas.org).The Clinical Chemistry Laboratory at AUB also receives quality assurance, evaluation, and accreditation from the College of American Pathologists (CAP) (www.cap.org).

Diasorin liaison platform assay

Diasorin Liaison uses a chemiluminescent immunoassay where serum 25OHD concentration is inversely proportional to the obtained chemiluminescent signal. The manufacturer reports an analytical range of 4 −150 ng/ml, an intra-assay precision (variability) of 7.7-12.7% for values between 5.8 and 35 ng/ml, and an inter-assay precision of 11.6 −25% for values between 5.8 and 35 ng/ml. Cross reactivity is reported by the manufacturer to be 104% to 25 OHD_2 and 100% to $25OHD_3$ (www.diasorin.com). However, according to the College of American Pathologists (CAP) Accuracy Based Vitamin D survey 2011, it was estimated to be 90% for $25OHD_3$ and 67% for $25OHD_2$ in our laboratory. At AUB, the intra-assay precision (CV) for duplicate analysis was $4.61 \pm 3.2\%$, and the inter-assay precision (CV) was $9.68 \pm 5.35\%$, (n=30 patients); and using manufacturer's quality control samples the inter-assay precision was 7.1% and 5.2% for mean 25OHD concentrations of 16.6 and 55.4 ng/ml, respectively (n = 55).

Liquid chromatography mass spectroscopy (LC-MS/MS)

Liquid Chromatography-Tandem Mass Spectrometry (LC-MS/MS) (ThermoFisher Scientific, Franklin, Massachusetts 02038 and Applied Biosystems-MDS Sciex, Foster City, CA 94404) assays were performed at the Mayo Clinic Laboratories. Intra-assay CVs were 3.8%, 2.4%, and 4.7% and inter-assay CVs were 6.4%, 6.8%, and 5.0%, at 24, 52, and 140 ng/mL, respectively.

Assessment of method specific assay bias based on quarterly DEQAS reports

Each quarter, DEQAS sends 5 unknown 25OHD samples to over 1,100 laboratories partaking in its quality assurance program. These unknown samples are run by each laboratory based on their specific method and the results are mailed back to DEQAS. Once all results are received, DEQAS generates a general report with details on method specific mean 25OHD values obtained for each method separately and specifies the NIST target value for each unknown sample. We used the quarterly reports received from DEQAS to calculate the *method mean % bias* [calculated as (Method specific mean total 25(OH)D level - NIST target 25(OH)D value/NIST 25(OH)D target value) X 100], for the 2 most commonly used automated platform assays to-date, Diasorin Liaison and Roche Elecsys, and for LC-MS/MS methods, for the period 2013-2017. We also further compared the performance of the Roche assay calculating the laboratory *individual % bias,* obtained when DEQAS quarterly samples were also assayed, independently and in parallel, in 2 laboratories using this same methodology over the same period at our institution.

Statistical analysis

Serum 25OHD levels were rounded to first decimal and mean or percent differences were rounded to nearest integer. Continuous variables were summarized using mean and standard deviation (SD), or median and range, as applicable based on normality of data assessed by the Shapiro-Wilk test. Difference in the mean 25OHD levels was tested using paired t test or Wilcoxon signed rank test. Measurements of these assays were regressed against each other, and a linear regression model fit was generated accordingly. Subjects were also categorized according to their 25OHD levels [Low (<10 ng/ml), intermediate (10-19.99 ng/ml), desirable (20 ng/ml)] based on each assay, and agreement of patient grouping based on the comparison assays was tested using McNemar Test. Bland-Altman plots of the difference (or percentage difference) in 25OHD levels between both assays versus the LC-MS/MS value, considered the best method, were generated for all subjects at baseline, 6, 12 months, and for all time points combined. Analyses were done using SigmaPlot 11.0 (Systat Software Inc., San Jose, CA) and IBM SPSS software version 24.0 (SPSS, Chicago, Illinois). Two tailed p-value<0.05 was considered significant.

Results

Comparison between LC-MS/MS and Diasorin Liaison assays

Baseline: The mean serum 25OHD levels obtained on 219 subjects at baseline using Diasorin Liaison were systematically and significantly lower 18.5 ± 7.8 ng/ml (mean $\pm$ SD), from the corresponding levels obtained with LC-MS/MS, 20.5 ± 7.6 ng/ml, (P< 0.0001). The absolute bias assessed with Bland Altman graphs, averaged 2 ± 6 ng/ml, [95% CI= −10, 14] and was independent of the LC-MS/MS 25OHD value (Figure 1A), and the percent bias averaged 6.9 ± 29 % (Figure 1B). Comparable results were obtained with serum 25OHD levels at 6 or 12 months, and in the complete dataset (Table, and Appendix Figures).

Results of linear regression analyses of serum 25OHD levels by Diasorin as a function of levels obtained by LC-MS/MS, at various time points, were as follows:

Complete data set (N=516): 25OHD Liaison=0.756 LC-MS/MS+4.26, R^2=0.596

Baseline data set (N=219): 25OHD Liaison=0.693 LC-MS/MS+4.36, R^2=0.460

6 months data set (N=171): 25OHD Liaison=0.722 LC-MS/MS+6.16, R^2=0.528

12 months data set (N=126): 25OHD Liaison=0.653 LC-MS/MS+8.36, R^2=0.490

Impact of assay methodology (LC-MS/MS and Diasorin Liaison) on patient classification

There was a 10-15 % difference in the classification of study subjects as having less than desirable levels of serum 25OHD levels whether one considered a cut-off of 20 ng/ml Institute of Medicine (IOM) (16), or 30 ng/ml (Endocrine society guidelines) (17).

Baseline 25OHD values: selecting a cut-off at 20 ng/ml, 64 % of subjects would then be classified as having less than desirable 25OHD levels using the Diasorin Liaison, compared to 50% using the LC-MS/MS assay (P<0.001). For a cut-off at 30 ng/ml, 91% of

subjects would be considered for treatment using Diasorin Liaison, compared to 89 % with LC-MS/MS (p=0.44).

Complete dataset: Using a desirable 25 OHD level of   20 ng/ml, 38% of subjects would then be classified as having less than desirable 25OHD levels using Diasorin Liaison, compared to 27% using the LC-MS/MS assay (P<0.001) (Figure 2A). Similarly, for a desirable 25OHD level of   30 ng/ml, 74% of subjects would be considered for treatment using Diasorin Liaison, compared to 68% according to results obtained by LC-MS/MS (P=0.001) (Figure 2A).

The distribution curve of 25OHD levels at baseline by LC-MS/MS had a more normal distribution, and was shifted to the right for low values, and to the left for high values, as compared to the one obtained using the Diasorin assay (Figure 2B).

Concordance rates between the LC-MS/MS and Diasorin Liaison assay results

Concordance of proportions defined depends on the selected category approach. The overall concordance rate for classification between the two assays was 80%, and for values below 10 ng/ml it was 67% (Figure 3). The weighted k coefficients at baseline, 6, 12 months, and for all data points were 0.52, 0.43, 0.48, and 0.60 respectively, p<0.04 for all coefficients.

To further evaluate differences between assays, the difference in 25OHD levels obtained with both assays was calculated for each subject: the proportions of subjects that varied by less than 20% was 55%, between 20-40% it was 34%, and by more than 40% it was 11% (Figure 4).

Performance of the Roche Elecsys platform assay at AUBMC 2013-2017.

The mean method specific % bias for the 2 most commonly used automated platform assays to- date, compared to NIST target values varied widely overt the 5 year period. Indeed, the mean % bias varied from −18% to 11% for the Diasorin-Laison assay, for the lowest NIST mean target value of 33.6 ± 8.1 nmol/L (13.5 ± 3.2 ng/ml), and from −17% to 8% for the Diasorin assay, for the highest NIST mean target value of 100.1 ± 16.5 nmol/L (40 ± 6.6 ng/ml) (Figure 5, Panels A and B). The variations were wider for the Roche Elecsys assay, between - 27% to 24% for the lowest NIST mean target value, and from −3% to 35% for the highest NIST mean target value. Conversely, the accuracy of LC-MS/MS was superior, and although it had a consistent positive bias compared to target NIST values, it was, nevertheless, much closer to the desirable range of ± 5% bias, set by the vitamin D Standardization Program (VDSP) (9, 10, 18). The individual % bias from NIST reference values obtained from running the unknown DEQAS samples in 2 laboratories at our institution that use the same Roche Elecsys platform assay, also varied widely, from −30% to 11% and −37% to 35% for the mean lowest NIST target value, and between −9% to 35% and −4% to 22% for the mean highest NIST target value, over the same study period (Figure 5 Panels C and D).

Discussion

We demonstrate a substantial difference in serum 25OHD levels obtained with a commonly used platform automated assay, Diasorin Laison, as compared to LC-MS/MS, at baseline and post-intervention, in a large sample of elderly overweight ambulatory individuals who completed a one-year vitamin D trial. Compared to NIST reference values, both the LC-MS/MS based assay, and the automated Diasorin Liaison and Roche Elecsys platform assays showed variations in assay performance, that were however larger and more unpredictable for the Roche Elecsys platform. Roche has recently released the Elecsys Total II 25OHD assay, with an anticipated superior performance, and that is certified by the CDC VDSP (19).

In our study, the platform Diasorin assay resulted on average in lower values than LC-MS/MS, with a concordance rate of only 67% with LC-MS/MS at values below 10 ng/ml values, and of 80% at higher values. Considering the whole sample, almost ½ of the paired samples had a mean difference exceeding 20%. The distribution of the absolute difference for individual paired values also spanned a wide range, between −30 to 19 ng/ml for baseline values, and −27 to 28 ng/ml, at 12 months post-supplementation. The difference was random, thus ruling out a systematic difference in calibration, and making any attempt to correction to the more established LC-MS/MS values impossible.

There is a multitude of methods to measure serum 25(OH) assays, and the wide variations noted between methods is not new (12, 20-26). Collection methods with additives (gel and clot activator) and presence of 25OHD2 and 3-epi 25OHD3 all affected results obtained depending on assay type (25). Cavalier et al. demonstrated that re-standardization of the various commercial immunoassays will reduce variations noted, but that substantial differences remain if sera with a different matrix (pregnant women, or patients with renal failure) are used (26). It is nevertheless increasingly worrisome as this situation persists despite major efforts for standardization, such as launched with the VDSP effort almost 8 years ago (11, 27). Janssen et al also compared the performance of various assays to the accepted LC-MS/MS, and showed a similar modest concordance rate between the Diasorin and LC-MS/MS of 70% (14). A study from Germany based on testing a total of 300 surplus single EDTA plasma samples revealed regression equations between the LC-MS/MS and Diasorin assays that were similar to those we describe in our study (22).

The positive bias of LC-MS/MS assay in our study has been noted by others and seems to be well recognized. (28, 29) The DEQAS program affords further insights into the performance over time of various assays including LC-MS/MS, comparing values obtained each quarter for unknown samples received by each laboratory to their corresponding NIST target values, in a large sample of laboratories using a spectrum of technologies worldwide (29). Based on data obtained from quarterly cycles from 2013-2017, we demonstrate that the mean bias values for LC-MS/MS compared to the NIST standards are consistently positive, but on average closest to the target 5% mean bias set by the VDSP. However, mean bias values are much more variable and unpredictable for the platform assays. Indeed, they are most consistently negative for the Diasorin platform assay, as had been previously reported, but nevertheless with wide variations, which are even larger and less predictable with the Roche

automated assay. Others have also reported on the large variations expressed as mean % bias from target NIST at any cycle, and that also vary widely across cycles for the same assay method, including the Diasorin platform assay (5, 11).

DEQAS has allowed further systematic insight into sources of 25OHD assay variations as being, in part, due to interference from other circulating vitamin D metabolites, such as $24,25 (OH)_2D$ metabolites (R or S isomers), and the 3-epimer 25(OH)D (30). Indeed, while crossreactivity with 24 hydroxylated metabolites is a major limitation with ligand-binding assays, it is not with LC-MS/MS (30). Conversely, whereas the 3-epimer shows cross-reactivity of 56% with competitive protein binding assays, it is not resolved by all LC-MS/MS assays, and is not detected in any of the antibody based methods (30).

Our study has some limitations. It did not systematically evaluate pre-analytical, analytical (matrix), or post-analytical factors that could account for observed differences in 25OHD between assays. But some of these limitations are inherent to vitamin D measurements per-se (for example skipping the extraction step in automated assays) and do not reflect on the quality of work performed in our laboratories, both of which are DEQAS proficient. Although trial samples were not assayed within the same time frame at both institutions, blood samples were processed, sera separated, and immediately frozen at −70 degrees, until assay performance at AUB and were shipped on dry ice to the Mayo Clinic laboratories. Vitamin D is a steroid that is very stable as opposed to peptide hormones. We consider the LC-MS/MS assay to be the "gold standard" in our comparisons of samples obtained from the clinical trial, and although LC- MS/MS is the best option available today, it is a heterogeneous group of methods with their own limitations. They nevertheless still do perform better considering the % mean bias data compared with NIST standards as compared to other technologies.

Our study has many important contributions. It is based on a large sample of specimens, spanning a wide spectrum of vitamin D levels, and the results obtained are therefore quite robust. We capture and compare assay performance based on trial specimens over a protracted time period (over 4 years, 2011-2015), thus confirming the on-going nature of the problem of assay variation, in spite of the launch and continued efforts of the VDSP initiative. Based on our assessment of % bias from NIST standards across methods and manufacturers, our study also confirms that the problem of assay variation is widespread across methodologies, and also on-going (2013-2017). Finally, we also demonstrate wide variation in % bias from NIST reference values with the use of an identical methodology at 2 DEQAS certified laboratories, within the same institution.

In contrast to older days when Diasorin RIA was almost universally used in research and practice and a single transposable cut-off was used worldwide, the landscape is quite different and most challenging today. The increase in the flurry of assays over the last 2 decades, and the above described assay variations have major implications for clinical practice, in terms of diagnosis, intervention thresholds, and target levels reached post-supplementation (12, 18). Indeed, in our study classification using the same IOM recommended 25OHD cut-off of 20 ng/ml varied by 10-15%, and this variation was lower using the Endocrine society desirable level of 30 ng/ml. The overall concordance rate

between paired samples was only 80% and as low as 67% for 25OHD values below 10 ng/ml, resulting in projected large differences in number of individuals who would be recommended supplementation at the public health level. Suh variations will continue to constitute a major obstacle to harmonization of international vitamin D guidelines (11, 12). These assay limitations have also probably contributed to the increasing number of reference and research laboratories that have switched to LC-MS/MS assays over time (31). Although labor intensive, they still provide a better option if available, as we do not anticipate substantial further improvement in platform assays over the nearby future (9, 10, 12, 27, 31, 32).

In conclusion, the variability in 25 OHD assay performance that plagues most platform automated technologies incurs major limitations to the quality of the evidence obtained from vitamin D randomized trials and meta-analyses, formulation of vitamin D guidelines, and thus ultimately affects interventions to be implemented at the public health level or and clinical practice. Considering the above, it would be most advisable to abide by international recommendations, and refrain from measuring serum 25OHD levels in apparently healthy ambulatory individuals (16, 17, 33). Serum 25 OHD measurements should be performed only in high risk individuals (metabolic bone disorders, liver failure, chronic kidney disease, certain medications, etc,) and in the setting of clinical trials. Measurements are ideally performed at centers that consistently partake and fulfill the most rigorous vitamin D proficiency programs, such as VDSP (34, 35). Such approach would pave the way forward to optimize the quality of the evidence from research studies, allow the determination of robust vitamin D intervention cut-offs, and harmonize public health and clinical practice strategies worldwide.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgements

References

1. Holick MF. Vitamin D deficiency. N Engl J Med. 2007;357(3):266–81. [PubMed: 17634462]

2. Lips P, van Schoor NM. The effect of vitamin D on bone and osteoporosis. Best Pract Res Clin Endocrinol Metab. 2011;25(4):585–91. [PubMed: 21872800]

3. Bouillon R, Van Schoor NM, Gielen E, Boonen S, Mathieu C, Vanderschueren D, et al. Optimal vitamin D status: a critical analysis on the basis of evidence-based medicine. J Clin Endocrinol Metab. 2013;98(8):E1283–304. [PubMed: 23922354]

4. El-Hajj Fuleihan G, Bouillon R, Clarke B, Chakhtoura M, Cooper C, McClung M, et al. Serum 25-Hydroxyvitamin D Levels: Variability, Knowledge Gaps, and the Concept of a Desirable Range. J Bone Miner Res. 2015;30(7):1119–33. [PubMed: 25952470]

5. Carter GD, Berry J, Durazo-Arvizu R, Gunter E, Jones G, Jones J, et al. Quality assessment of vitamin D metabolite assays used by clinical and research laboratories. J Steroid Biochem Mol Biol. 2017;173:100–104. [PubMed: 28315391]

6. Schleicher RL, Sternberg MR, Lacher DA, Sempos CT, Looker AC, Durazo-Arvizu RA, et al. A Method-bridging Study for Serum 25-hydroxyvitamin D to Standardize Historical Radioimmunoassay Data to Liquid Chromatography-Tandem Mass Spectrometry. Natl Health Stat Report. 2016(93):1–16.

7. Hoteit M, Al-Shaar L, Yazbeck C, Bou Sleiman M, Ghalayini T, Fuleihan Gel H. Hypovitaminosis D in a sunny country: time trends, predictors, and implications for practice guidelines. Metabolism. 2014;63(7):968–78. [PubMed: 24874590]

8. Phinney KW, Tai SS, Bedner M, Camara JE, Chia RRC, Sander LC, et al. Development of an Improved Standard Reference Material for Vitamin D Metabolites in Human Serum. Anal Chem. 2017;89(9):4907–13. [PubMed: 28375002]

9. Binkley N, Sempos CT. Standardizing vitamin D assays: the way forward. J Bone Miner Res. 2014;29(8):1709–14. [PubMed: 24737265]

10. Sempos CT, Betz JM, Camara JE, Carter GD, Cavalier E, Clarke MW, et al. General Steps to Standardize the Laboratory Measurement of Serum Total 25-Hydroxyvitamin D. Journal of AOAC International. 2017;100(5):1230–3. [PubMed: 28766476]

11. Binkley N, Carter GD. Toward Clarity in Clinical Vitamin D Status Assessment: 25(OH)D Assay Standardization. Endocrinol Metab Clin North Am. 2017;46(4):885–99. [PubMed: 29080641]

12. Barake M, Daher RT, Salti I, Cortas NK, Al-Shaar L, Habib RH, et al. 25- hydroxyvitamin D assay variations and impact on clinical decision making. J Clin Endocrinol Metab. 2012;97(3):835–43. [PubMed: 22238386]

13. Ong L, Saw S, Sahabdeen NB, Tey KT, Ho CS, Sethi SK. Current 25-hydroxyvitamin D assays: do they pass the test? Clin Chim Acta. 2012;413(13–14):1127–34. [PubMed: 22465235]

14. Janssen MJ, Wielders JP, Bekker CC, Boesten LS, Buij s MM, Heijboer AC, et al. Multicenter comparison study of current methods to measure 25-hydroxyvitamin D in serum. Steroids. 2012;77(13):1366–72. [PubMed: 22925701]

15. El-Hajj Fuleihan G, Baddoura R, Habib RH, Halaby G, Arabi A, Rahme M, et al. Effect of vitamin D replacement on indexes of insulin resistance in overweight elderly individuals: a randomized controlled trial. Am J Clin Nutr. 2016;104(2):315–23. [PubMed: 27413130]

16. Ross AC, Manson JE, Abrams SA, Aloia JF, Brannon PM, Clinton SK, et al. The 2011 report on dietary reference intakes for calcium and vitamin D from the Institute of Medicine: what clinicians need to know. J Clin Endocrinol Metab. 2011;96(1):53–8. [PubMed: 21118827]

17. Holick MF, Binkley NC, Bischoff-Ferrari HA, Gordon CM, Hanley DA, Heaney RP, et al. Evaluation, treatment, and prevention of vitamin D deficiency: an Endocrine Society clinical practice guideline. J Clin Endocrinol Metab. 2011; 96(7):1911–30. [PubMed: 21646368]

18. Thienpont LM, Stepman HC, Vesper HW. Standardization of measurements of 25- hydroxyvitamin D3 and D2. Scand J Clin Lab Invest Suppl. 2012;243:41–9. [PubMed: 22536761]

19. Vitamin D Standardization-Certification Program (CDC VDSCP) Certified Total 25-hydroxyvitamin D Procedures. https://www.cdc.gov/labstandards/pdf/hs/CDC_Certified_Vitamin_D_Procedures-508.pdf (Last Accessed August 2018)

20. Lai JK, Lucas RM, Banks E, Ponsonby AL. Variability in vitamin D assays impairs clinical assessment of vitamin D status. Intern Med J. 2012;42(1):43–50. [PubMed: 21395958]

21. Glendenning P, Taranto M, Noble JM, Musk AA, Hammond C, Goldswain PR, et al. Current assays overestimate 25-hydroxyvitamin D3 and underestimate 25-hydroxyvitamin D2 compared with HPLC: need for assay-specific decision limits and metabolite-specific assays. Annals of Clin Biochem. 2006;43:23–30.

22. Roth HJ, Schmidt-Gayk H, Weber H, Niederau C. Accuracy and clinical implications of seven 25-hydroxyvitamin D methods compared with liquid chromatography-tandem mass spectrometry as a reference. Annals of Clin Biochem. 2008;45:153–9.

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

24. De Koning L, Al-Turkmani MR, Berg AH, Shkreta A, Law T, Kellogg MD. Variation in clinical vitamin D status by DiaSorin Liaison and LC-MS/MS in the presence of elevated 25-OH vitamin D2. Clin Chim Acta. 2013;415:54–8. [PubMed: 22975531]

25. Yu S, Cheng X, Fang H, Zhang R, Han J, Qin X, et al. 25OHD analogues and vacuum blood collection tubes dramatically affect the accuracy of automated immunoassays. Sci Rep. 2015;5:14636. [PubMed: 26420221]

26. Cavalier E, Lukas P, Crine Y, Peeters S, Carlisi A, Le Goff C, et al. Evaluation of automated immunoassays for 25(OH)-vitamin D determination in different critical populations before and after standardization of the assays. Clin Chim Acta. 2014;431:60–5. [PubMed: 24508999]

27. Sempos CT, Vesper HW, Phinney KW, Thienpont LM, Coates PM. Vitamin D status as an international issue: national surveys and the problem of standardization. Scand J Clin Lab Invest Suppl. 2012;243:32–40. [PubMed: 22536760]

28. Berry DJ, Dutton J, Fraser WD, Jarvelin MR, Hypponen E. Harmonization Study Between LC-MS/MS and Diasorin RIA for Measurement of 25-Hydroxyvitamin D Concentrations in a Large Population Survey. J Clin Lab Anal. 2017;31(3).

29. DEQAS (Vitamin D External Quality Assessment Scheme) website http://www.deqas.org/ (Last Accessed August 2018).

30. Carter GD, Jones JC, Shannon J, Williams EL, Jones G, Kaufmann M, et al. 25- Hydroxyvitamin D assays: Potential interference from other circulating vitamin D metabolites. J Steroid Biochem Mol Biol. 2016;164:134–8. [PubMed: 26718874]

31. Carter GD, Berry J, Durazo-Arvizu R, Gunter E, Jones G, Jones J, et al. Hydroxyvitamin D assays: An historical perspective from DEQAS. J Steroid Biochem Mol Biol. 2018;177:30–5. [PubMed: 28734989]

32. Binkley N, Dawson-Hughes B, Durazo-Arvizu R, Thamm M, Tian L, Merkel JM, et al. Vitamin D measurement standardization: The way out of the chaos. J Steroid Biochem Mol Biol. 2017;173:117–21. [PubMed: 27979577]

33. Hanley DA, Cranney A, Jones G, et al. Vitamin D in adult health and disease: a review and guideline statement from Osteoporosis Canada. CMAJ. 2010;182(12):E610–8. [PubMed: 20624868]

34. Sempos CT, Vesper HW, Phinney KW, Thienpont LM, Coates PM. Vitamin D status as an international issue: national surveys and the problem of standardization. Scand J of Clin and Lab Invest. 2012;243:32–40.

35. Cashman KD, Kiely M, Kinsella M, Durazo-Arvizu RA, Tian L, Zhang Y, et al. 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(6):1235–42. [PubMed: 23615829]

Highlights

- Substantial variations between Vitamin D assays remain

- This is particularly relevant to the commonly used, automated platform assays

- Such variations limit the formulation of desirable Vitamin D levels and guidelines

- Participation in Vitamin D assay standardization by major laboratories is a priority
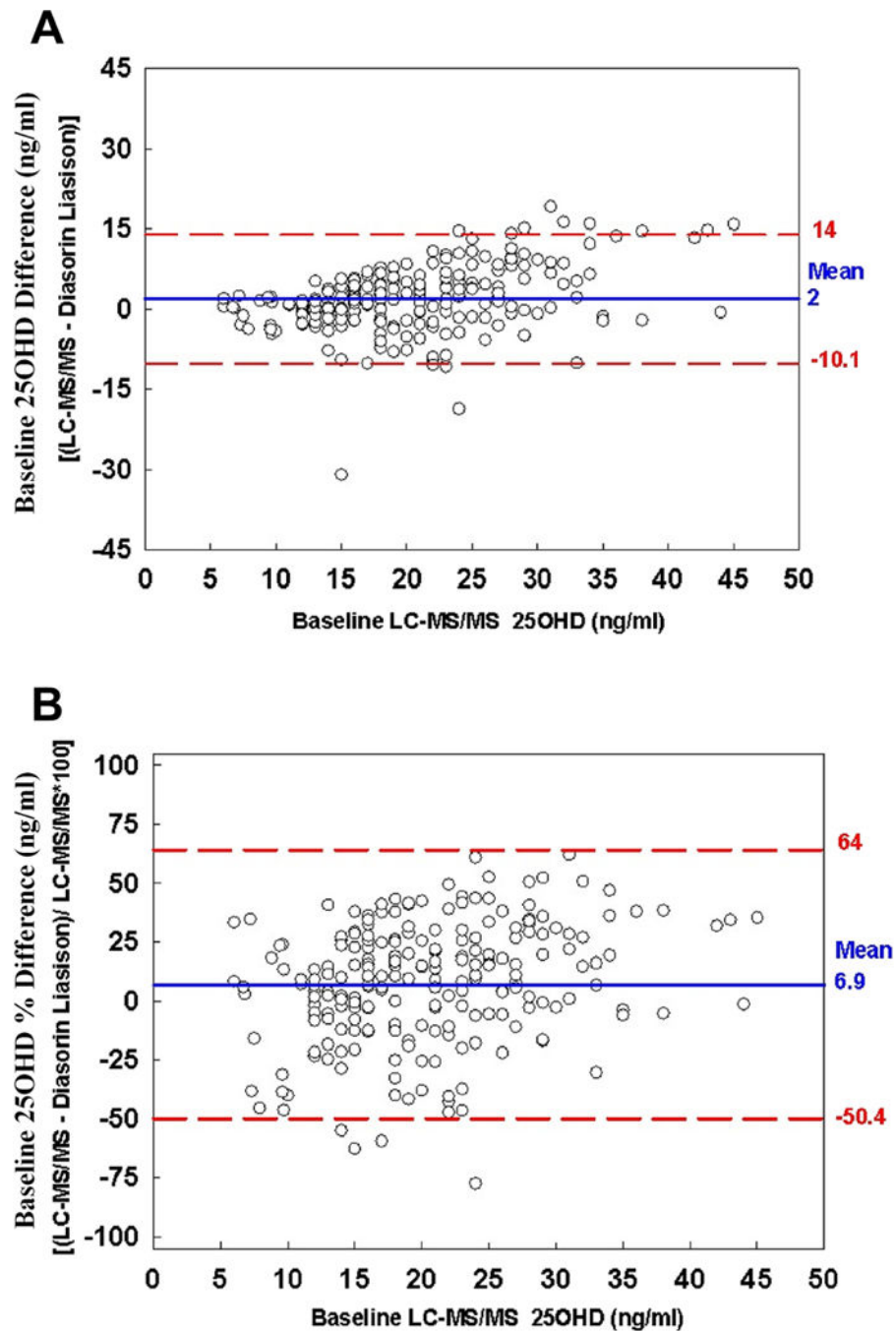
Figure 1.
Panel A: Bland-Altman plot of the baseline difference between Diasorin Liaison and LC-MS/MS 25-OHD levels vs. baseline LC-MS/MS 25OHD in 219 subjects (symbols). Solid horizontal line, mean difference between the two assays (or bias); dashed horizontal lines, 95% confidence interval.
Panel B. Bland-Altman plot of the baseline percentage difference [(LC-MS/MS - Diasorin Liaison)/LC-MS/MS*100]. Solid horizontal line, mean difference between the two assays (or bias); dashed horizontal lines, 95% confidence interval.
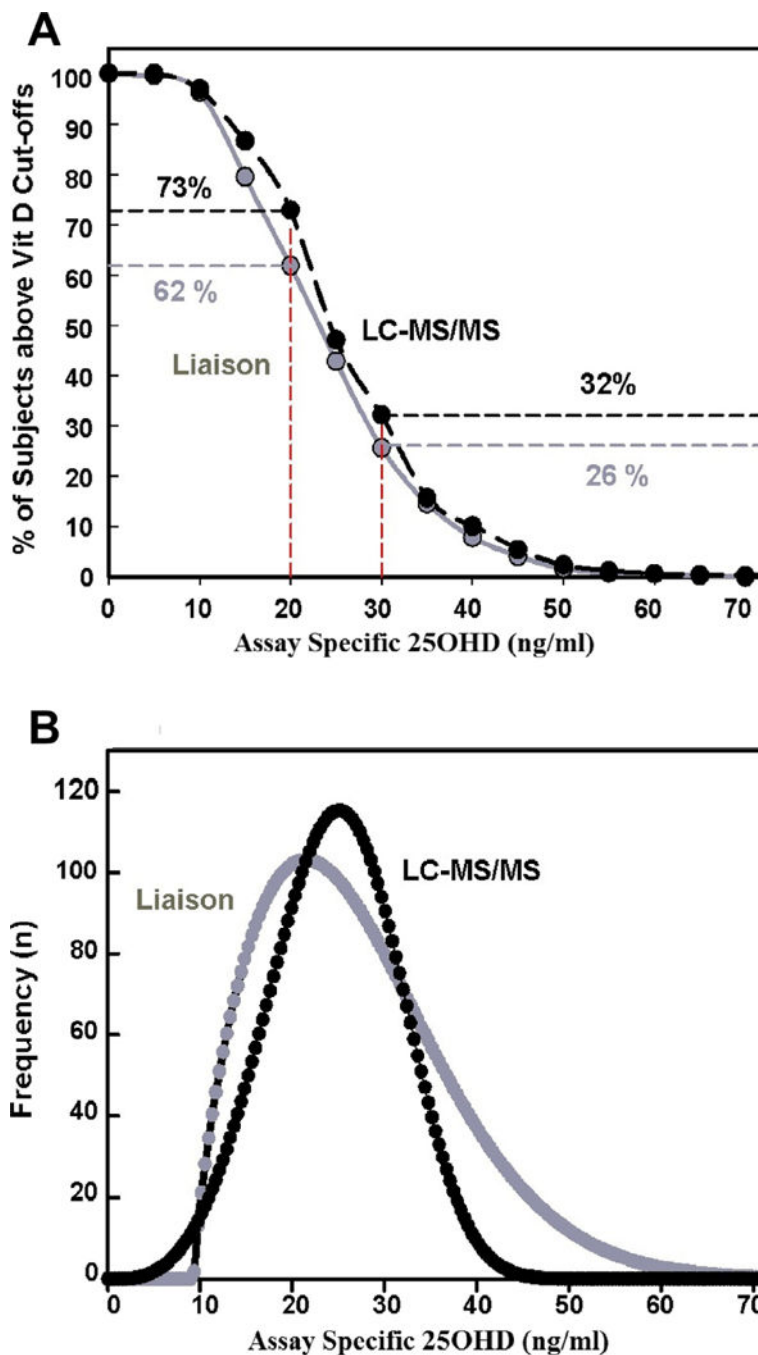
Figure 2. Proportion (Panel A) or frequency (Panel B) of subjects with 25OHD levels in the total cohort (n 516) above or at indicated assay specific 25OHD level.
Diasorin Liaison (gray) and the LC-MS/MS (black) assays.

Figure 3. Comparison of Diasorin Liaison vs. LC-MS/MS 25OHD levels in the overall data set (n 516) subjects.

[Liaison] = 4.3 + 0.76 [LC-MS/MS]
$r^2 = 0.60$; N=516
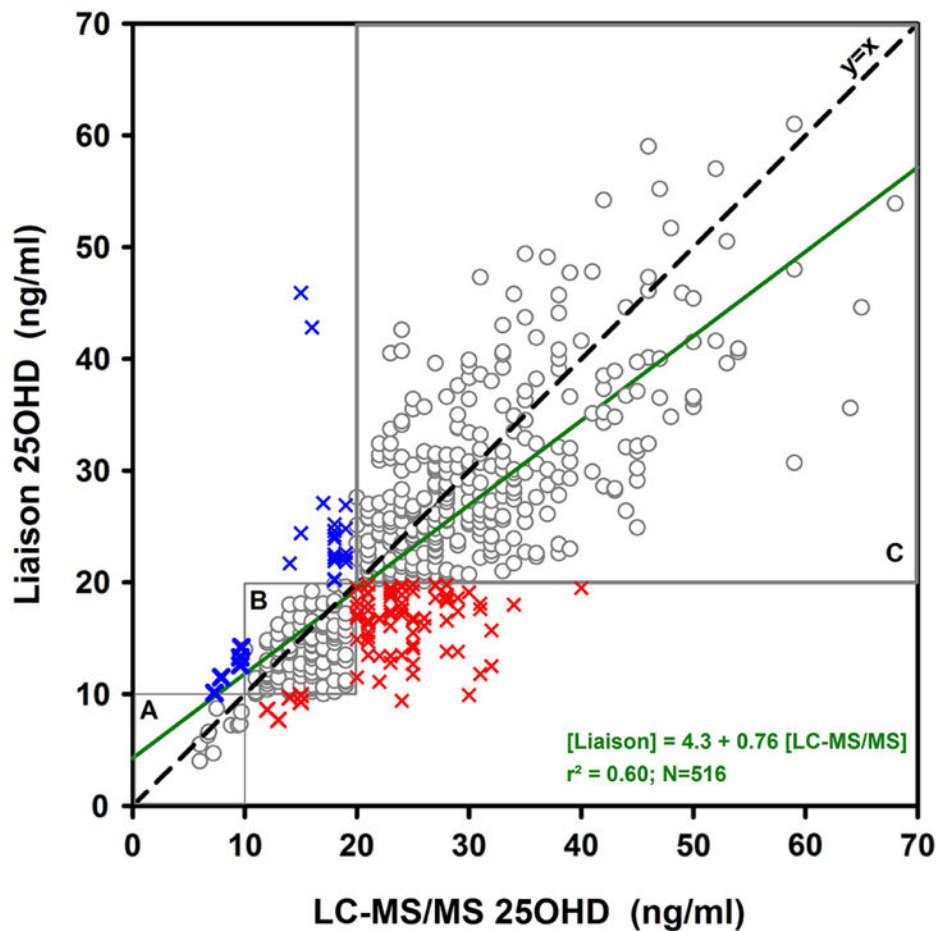
○ Concordant
✗ Discordant - Liaison > LC-MS/MS
✗ Discordant - Liaison < LC-MS/MS

			Liaison Method	
Group (25OHD)	25OHD	N*	Concordant	Discordant
Very Low	<10 ng/ml	15	10 (67%)	5 (33%)
Low	10-19.9 ng/ml	125	101 (81%)	24 (19%)
Normal	≥20 ng/ml	376	300 (80%)	76 (20%)
Overall		516	411 (80%)	105 (20%)

* counts based on LC-MS/MS Method (Gold Standard)

Figure 3. Comparison of Diasorin Liaison vs. LC-MS/MS 25OHD levels in the overall data set (n 516) subjects.

Symbols represent individual data points using both assays. Gray light circles represent concordant points at increasing 25OHD cut-offs [A (<10 ng/ml), B (10–20ng/ml), C (>20 ng/ml). Dashed line, Line of identity; solid line, linear regression model fit with its equation.
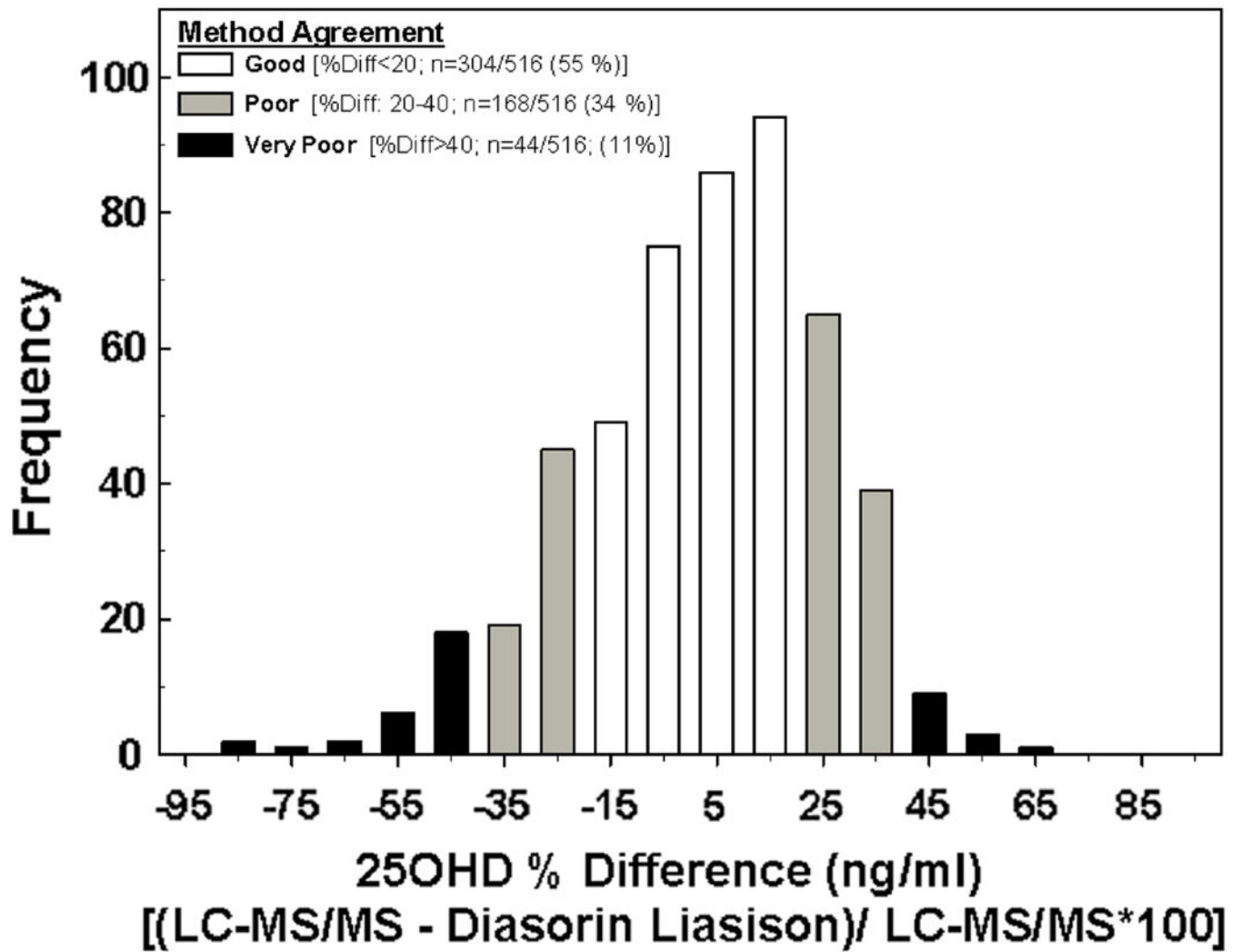
Figure 4:
Frequency distribution of the percentage difference in 25-OHD levels in the overall data set 516 subjects.
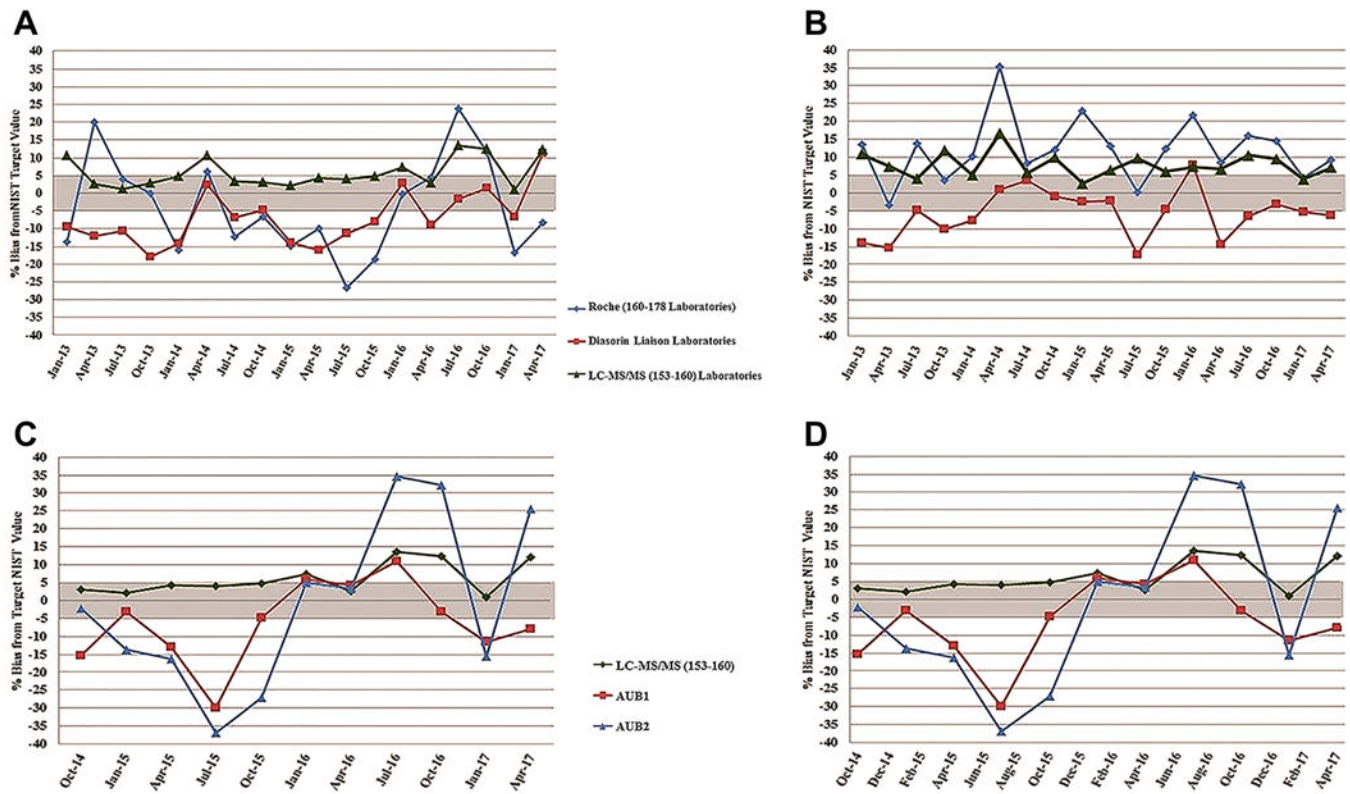
Figure 5.

(A and B): Method specific mean % bias from NIST lowest (panel A) and highest (panel B) 25OHD) target values, for the 2 most commonly used platform assay methods/manufacturers and for LC-MS/MS methods. The number of laboratories varied by DEQAS report cycle, as shown in legend, depending on the cycle.

Figure 5 (C and D): Method specific % bias from NIST lowest (Panel C) and highest (Panel D) 25OHD target values, at 2 AUB laboratories.

The mean lowest 25(OH)D NIST target value (panels A and C) obtained by deriving the mean using the lowest 25(OH)D NIST target value for each cycle, over the DEQAS report period 2013–2017: 33.6 ± 8.1 nmol/L (13.4±3.2 ng/ml). Mean highest 25(OH)D NIST target value (panels B and D) obtained by deriving the mean using the highest 25(OH)D NIST target value for each cycle, over the report period 2013–2017: 100.07 ± 16.5 nmol/L (40.0±6.6 ng/ml).

Table:

Summary of 25OHD Levels Obtained in Elderly RCT using LC-MS/MS and Diasorin Liaison 25OHD Assays

	Test	N	Mean ± SD	Median	Minimum	Maximum	p-value
Baseline	AUB (Diasorin Liaison) (ng/ml)	219	18.5 ± 7.8	16.6	4	45.9	<0.0001
	Mayo (LC-MS/MS) (ng/ml)	219	20.5 ± 7.6	19	6	45	
	Difference (ng/ml) *	219	2 ± 6.2	1.7	−30.9	19.2	
	Difference (%) ‡	219	6.9 ± 29.3	10	−206	61.9	
6 months	AUB (Diasorin Liaison) (ng/ml)	171	28.1 ± 9.9	26.8	8.6	61	<0.0001
	Mayo (LC-MS/MS) (ng/ml)	171	30.4 ± 10	29	11	68	
	Difference (ng/ml) *	171	2.3 ± 7.4	2.7	−17.5	28.4	
	Difference (%) ‡	171	5.8 ± 23.2	10	−76.1	67	
12 months	AUB (Diasorin Liaison) (ng/ml)	126	28.5 ± 9.5	26.5	9.7	54.2	0.001
	Mayo (LC-MS/MS) (ng/ml)	126	30.9 ± 10.2	28	14	65	
	Difference (ng/ml) *	126	2.4 ± 7.6	2.7	−26.8	28.3	
	Difference (%) ‡	126	5.1 ± 26.4	9.5	−167.5	60.9	
All	AUB (Diasorin Liaison) (ng/ml)	516	24.1 ± 10.2	23.3	4	61	<0.0001
	Mayo (LC-MS/MS) (ng/ml)	516	26.3 ± 10.4	25	6	68	
	Difference (ng/ml) *	516	2.2 ± 6.9	2.2	−30.9	28.4	
	Difference (%) ‡	516	6.1 ± 26.7	9.8	−206	67	

*
Difference= 25OHD (LC-MS/MS) - (Diasorin Liaison)

‡
%Difference= (LC-MS/MS -Diasorin Liaison)/ LC-MS/MS *100