

## Correlation among 25-Hydroxy-Vitamin D Assays

N. Binkley, D. Krueger, D. Gemar, and M. K. Drezner

University of Wisconsin Osteoporosis Clinical Center & Research Program, University of Wisconsin, Madison, Wisconsin 53705

**Context:** Measurement of circulating 25-hydroxy-vitamin D [25(OH)D] is the accepted clinical indicator of vitamin D status. However, between-laboratory differences in measurement of this analyte exist, which may confound clinical care.

**Objectives:** We investigated the current agreement of 25(OH)D measurement in clinical laboratories and explored the possibility that simple calibration would improve between-laboratory agreement.

**Design and Participants:** Serum obtained from healthy volunteers (age 20–60 yr) and one “calibrator,” selected to have a 25(OH)D value near 30 ng/ml, were sent for 25(OH)D measurement in four clinical laboratories (laboratories A–D) using HPLC, liquid chromatography tandem mass spectroscopy, and RIA methodologies.

**Main Outcome Measures:** Serum 25(OH)D. Based upon self-report, the laboratory with the lowest interassay percent coefficient of variation was assigned as the reference to which the others were compared using linear regression and Bland-Altman analyses (Analyse-it; Analyse-it Software, Ltd., Leeds, UK).

**Results:** Good correlation was observed for 25(OH)D measurement between laboratory A and laboratories B–D ( $R^2 = 0.99, 0.81, \text{ and } 0.95$ , respectively). Modest between-laboratory variation was noted; the mean bias ranged from 2.9–5.2 ng/ml. Consistent with a systematic offset, each value in laboratory B was higher than in laboratory A, and 89% of values from laboratories B–D were higher than laboratory A. The use of a single calibrator and correction factor reduced mean between-laboratory bias for laboratories B and D.

**Conclusions:** Measurement of 25(OH)D by clinical laboratories yields similar results. The use of even a single calibrator will improve, but not resolve, between-laboratory variability. Based upon these data, in combination with reported within-individual variability, we recommend that clinicians aim for values greater than 30 ng/ml in their patients. (*J Clin Endocrinol Metab* 93: 1804–1808, 2008)

Measurement of circulating 25-hydroxy-vitamin D [25(OH)D] is accepted as the clinical indicator of vitamin D status (1). As such, 25(OH)D measurement has become widely used by clinicians with diagnoses of low vitamin D status being made, and vitamin D supplementation prescribed, based upon these values. However, in the recent past, substantial between-laboratory variability has been present, making clinical assessment of the vitamin D status of individual patients problematic (2–4). It might be assumed that the recent widespread clinical availability of liquid chromatography tandem mass spec-

troscopy (LCMSMS) and HPLC (3) technologies has improved 25(OH)D assay performance, thereby improving agreement between measurements obtained at different clinical laboratories. However, a review of recent vitamin D External Quality Assessment Scheme (DEQAS) data revealed substantial variability in individual 25(OH)D results between laboratories using LCMSMS and HPLC, despite good agreement of mean values. We hypothesized that this interlaboratory variability may reflect differences in assay calibration. The use of one or more externally supplied “calibrators,” containing an established value of

0021-972X/08/\$15.00/0

Printed in U.S.A.

Copyright © 2008 by The Endocrine Society

doi: 10.1210/jc.2007-2340 Received October 22, 2007. Accepted February 25, 2008.

First Published Online March 4, 2008

Abbreviations: CV, Coefficient of variation; DEQAS, vitamin D External Quality Assessment Scheme; LCMSMS, liquid chromatography tandem mass spectroscopy; 25(OH)D, 25-hydroxy-vitamin D.

**TABLE 1.** 25(OH)D interassay precision (percent CV)

25(OH)D concentration	"Low"		"Mid"		"High"	
	25(OH)D <sub>2</sub>	25(OH)D <sub>3</sub>	25(OH)D <sub>2</sub>	25(OH)D <sub>3</sub>	25(OH)D <sub>2</sub>	25(OH)D <sub>3</sub>
Laboratory						
A	12.6	3.9	4.9	2.6	3.2	3.3
B	11.4	11.6	8.8	8.3	8.2	7.9
C	9.7	13.5	11.6	10.7	8.8	8.4
D		5.7		7.5		

Interassay precision as self-reported by these four laboratories.

25(OH)D, could potentially be used to standardize assay performance. To evaluate the current clinical status of 25(OH)D measurement in the United States and investigate the possibility that simple assay calibration would improve between-laboratory agreement, 15 serum specimens and one "calibrator," selected to have a 25(OH)D value near 30 ng/ml, were analyzed in four clinical laboratories.

### Subjects and Methods

#### DEQAS 25(OH)D data

As part of quality control measures, over 200 laboratories worldwide participate in the DEQAS. The overall aim of this voluntary effort is to ensure analytical reliability of vitamin D assays. This is achieved through quarterly distribution of serum pools (n = 5 specimens) with subsequent reporting of these results to participants. Various methodologies are used for 25(OH)D measurement by laboratories participating in the DEQAS, including RIA, HPLC, and LCMSMS. DEQAS data are reported to participating laboratories in aggregate form with mean values reported. However, individual laboratory values are reported anonymously for HPLC and LCMSMS methods. The example data presented here are from the April, 2006, quarterly specimen distribution.

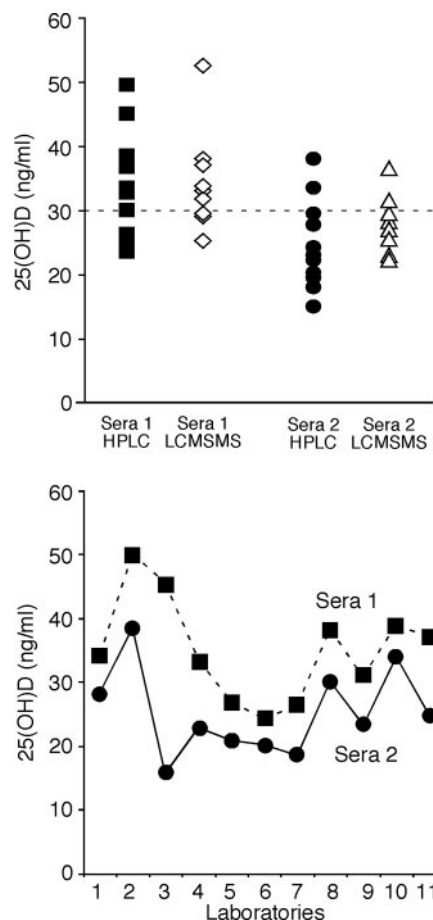
#### Study participants/clinical 25(OH)D measurement

Blood was obtained from 15 generally healthy adults (eight females, seven males; age 20–60 yr) by routine venipuncture, allowed to clot for 30 min at room temperature, then centrifuged. Study participants had no known acute or life-threatening illnesses and chemistry panels without clinically significant abnormalities. Serum aliquots were prepared and frozen at –80 C until shipped in a routine clinical manner for analysis. Aliquots of serum from one individual were obtained to serve as a "calibrator" that was known to contain approximately 30 ng/ml 25(OH)D from an independent analysis by HPLC. Aliquots of serum from all volunteers and the calibrator were sent as frozen clinical specimens to four laboratories (laboratories A–D) for 25(OH)D measurement. The methodologies used for 25(OH)D measurement included HPLC (laboratory A), LCMSMS (laboratories B and C), and RIA (DiaSorin, Inc., Stillwater, MN; laboratory D). Interassay 25(OH)D measurement precision [percent coefficient of variation (CV)] was obtained by self-report from each of the four laboratories (Table 1). This work was approved by the University of Wisconsin Health Sciences institutional review board; informed consent was obtained from all study participants.

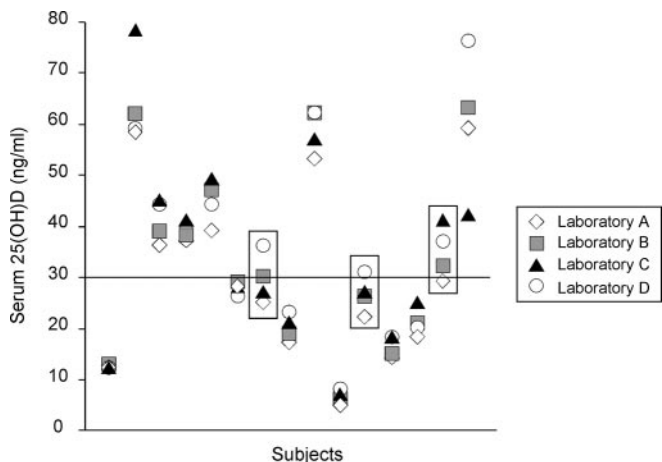
#### Data analysis

Based upon lowest self-reported percent CV, serum 25(OH)D results from laboratory A were defined for the purposes of this report as the "gold standard" to which the other laboratories were compared. Clinical vitamin D status was defined by serum 25(OH)D measurement as follows: optimal vitamin D status, more than or equal to 30 ng/ml; low vitamin D status, less than 30 ng/ml. Serum 25(OH)D measurements from the four laboratories were compared using linear regression and Bland-Altman analyses (Analyse-it; Analyse-it Software, Ltd., Leeds,

UK), an accepted methodology to compare two measurement techniques (5). Systematic bias was assessed by determination of the difference of each measurement in laboratories B–D compared with laboratory A ["delta" 25(OH)D]. For data "correction," each individual value from laboratories B–D was subsequently multiplied by the ratio of the "calibrator" value obtained in each respective laboratory compared with that obtained in laboratory A. After application of this correction factor to each individual value, 25(OH)D results from laboratories B–D were again compared with those of laboratory A using linear regression and Bland-Altman analyses as noted previously.



**FIG. 1.** DEQAS data and suggestion of systematic offset. *Top*, Despite good agreement of mean values for these two arbitrarily selected DEQAS samples (35.1 and 35.0 ng/ml by HPLC and LCMSMS for sera 1, and 25.2 and 28.9 ng/ml by HPLC and LCMSMS for sera 2), when values are considered on an individual basis, substantial between-laboratory variability is evident using either HPLC or LCMSMS. *Bottom*, Assessment of individual values from the 11 laboratories reported in the DEQAS to be using HPLC methodology suggests a systematic offset. For example, the values obtained by laboratory 8 are 25 and 28% higher for sera 1 and 2, respectively, than those obtained by laboratory 9.



**FIG. 2.** Classification of clinical vitamin D status. The clinical assessment of vitamin D status as optimal or low was identical for 80% of these 15 subjects in all four laboratories. However, three individuals (20%, noted by rectangles surrounding each individual’s values) would be variably classified as optimal or low depending on the laboratory used.

**Results**

**DEQAS 25(OH)D data**

Results are arbitrarily presented for two serum specimens (defined here as “1” and “2”). For these two sera, good agreement was reported for the mean 25(OH)D value from laboratories using HPLC (n = 9), LCMSMS (n = 8), and the DiaSorin RIA (n = 54) methodologies. Specifically, the mean (SD) values reported for these three technologies were 35.7 (8.4), 34.6 (7.3), and 31.4 (6.4) ng/ml for serum 1, and 25.5 (7.1), 28.9 (5.0), and 26.4 (5.6) ng/ml for serum 2, respectively.

The DEQAS reports individual values for those laboratories using HPLC or LCMSMS, which upon review appear to demonstrate substantial between-laboratory variability (Fig. 1A). Evaluation of 25(OH)D results for these two sera by individual laboratory using HPLC methodology identifies what may be a systematic offset that suggests the possibility that assay calibration differences might explain some of the between-laboratory differences (Fig. 1B).

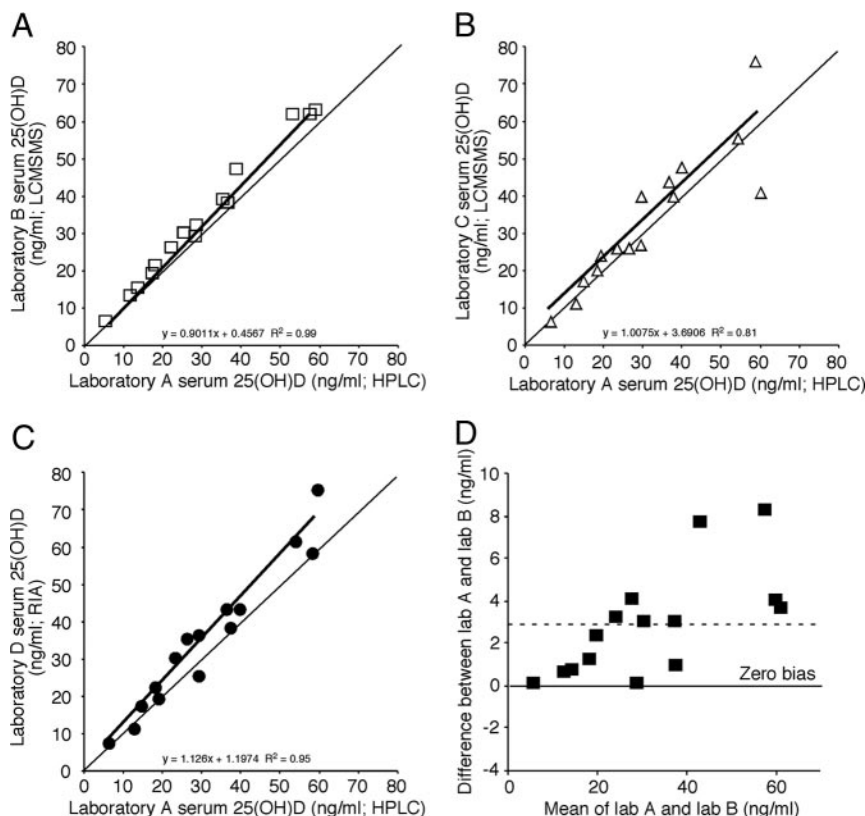
**Clinical 25(OH)D analysis**

Of these 15 specimens, 10 contained only 25(OH)D<sub>3</sub> as identified by HPLC. In the remaining five specimens, the 25(OH)D<sub>2</sub> concentration ranged from 5–19 ng/ml. Because clinical decisions are made based upon

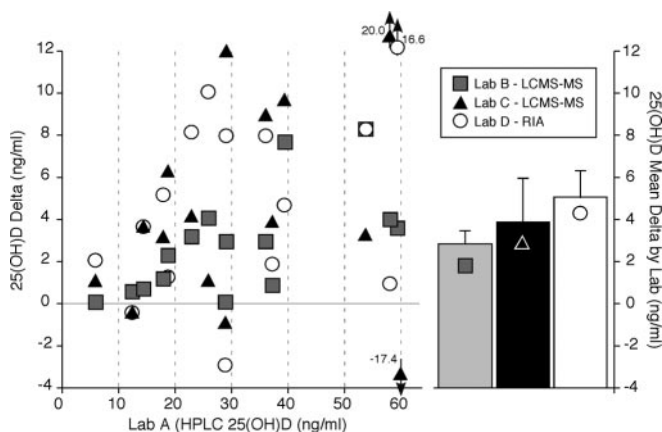
the total 25(OH)D concentration, this value was used for between-laboratory comparison. Diagnostic categorization (optimal vs. low vitamin D status) was identical in 80% (12 of 15) of subjects (Fig. 2). Good correlation was observed between laboratory A and the other three laboratories ( $R^2 = 0.99, 0.81,$  and  $0.95$ ) (Fig. 3, A–C). However, modest between-laboratory variability was noted; the mean bias of laboratories B–D ranged from +2.9 to +5.1 ng/ml when compared with laboratory A (example in Fig. 3D). This systematic positive bias leads to 89% (40 of 45) of values being higher in laboratories B–D than the corresponding laboratory A value (Fig. 4). “Correction” of the 25(OH)D values obtained by laboratories B–D reduces the mean bias (Fig. 5, A–C), e.g. after correction, laboratories A and B obtain virtually identical results (Fig. 5, A and D).

**Discussion**

Overall, there is good agreement of serum 25(OH)D determination between clinical laboratories using HPLC, LCMSMS, and a RIA methodology. However, modest variability and systematic bias are present between clinical laboratories. Based upon this variability, patients of clinicians who use a “cutpoint” diagnostic approach (e.g. 30 ng/ml) will continue to be variably identified as having low vitamin D status depending on the laboratory in which 25(OH)D measurement is performed. The presence of systematic bias between laboratories suggests that the use



**FIG. 3.** A–D, Between-laboratory agreement. Excellent correlation was observed when each individual laboratory (B–D) is compared with laboratory A (A–C). However, systematic bias was observed between laboratories. An example of this bias is shown in D in which the Bland-Altman plot demonstrates a mean bias of +2.9 ng/ml (noted by dashed line) observed when comparing laboratory A with laboratory B.



**FIG. 4.** Demonstration of systematic bias. Support for between-laboratory differences in assay calibration is gained by the demonstration that the results of the vast majority (89%) of specimens from laboratories B–D are higher than the corresponding value from laboratory A. Data presented as “Delta” 25(OH)D (the value obtained in laboratory B–D minus that obtained in laboratory A) vs. the value obtained in laboratory A. Mean “Delta” values (SEM) for laboratories B–D are presented on the right side of this figure.

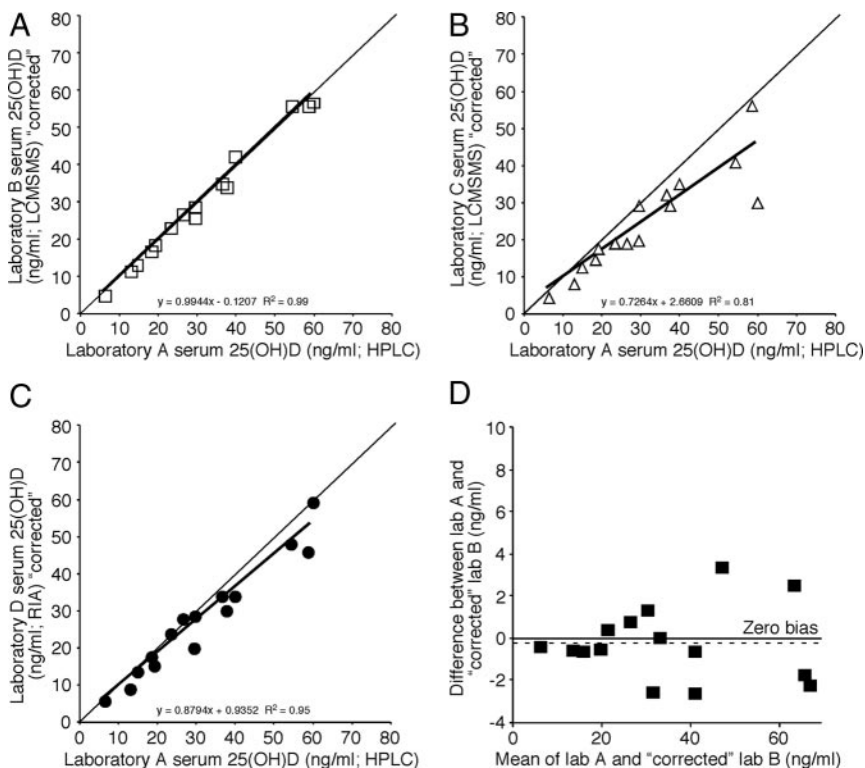
of a standardized calibrator, or calibrators, will enhance between-laboratory agreement. In addition, it is noteworthy that these data suggest better between-laboratory agreement at lower serum 25(OH)D concentrations, an observation concurrent with prior reports (3, 6).

There has been considerable discussion of what constitutes optimal vitamin D status and what cutpoint to use for a clinical diagnosis of vitamin D “deficiency” or “inadequacy.” (7–9) As could be expected, the prevalence of low D status is highly dependent upon the cutpoint selected (10). Nonetheless, there is increasing clinical consensus that values less than 30–32 ng/ml are indicative of suboptimal vitamin D status (7). It is important for clinicians to recognize that any such arbitrary cutpoint will have variability surrounding the value chosen from both analytical and biological sources. Specifically, analytical imprecision is present in all quantitative medical procedures due to human and instrument limitations. In addition, at this point in time, there is no internationally recognized primary standard, however, the National Institute of Standards and Technology is currently working to rectify this situation. Based upon these results, availability of standards should enhance between-laboratory agreement.

In addition to analytical variability, within-individual biological

variability of the parameter being measured, in this case 25(OH)D, must also be considered. Unfortunately, this aspect of 25(OH)D measurement has received only limited investigation. One study found 13–19% within-individual variability in 25(OH)D (11), which caused the authors to question the clinical use of a single 25(OH)D measurement in an individual patient. Although it is possible that this report reflects the well-known seasonal variability of vitamin D status due to variations in sunlight exposure (12–15), scrutiny of the current clinical approach of diagnosing vitamin D status based upon a single 25(OH)D measurement is clearly warranted. Furthermore, the study of within-individual variability of 25(OH)D in people on a stable vitamin D intake is clearly warranted. Moreover, for those clinicians using infrequent dosing of large amounts of vitamin D, consideration of measuring trough concentration appears to be necessary because 25(OH)D levels peak within 1 wk after dosing with ergocalciferol (D<sub>2</sub>) and subsequently decline (16). Preliminary data suggest that the measurement of trough 25(OH)D concentration may be appropriate when dosing with monthly D<sub>3</sub> as well (17). Additional study of this phenomenon is warranted.

Although further work in improving analytical 25(OH)D assay performance and investigating biological variability is indicated, clinicians are currently measuring 25(OH)D, making diagnostic assessments, and providing treatment recommendations based on clinical 25(OH)D measurement. Due to the assay variability reported here, and current (albeit limited) biological variability data, we believe it appropriate that clinicians aim to achieve 25(OH)D values above 30 ng/ml. Some clinicians may wish to consider aiming



**FIG. 5.** A–D, Clinical 25(OH)D results after use of a correction factor. A–C, Overall, use of a single calibrator improves, but does not entirely resolve, the between-laboratory bias. Specifically, before correction the mean bias of the values in comparison with laboratory A for laboratories B–D is 2.9, 3.9, and 5.1 ng/ml, respectively. After “correction” these values are –0.3, –5.7, and –2.8 ng/ml for laboratories B–D, respectively. The observation that correction produces nearly identical values for laboratories A and B is demonstrated in D.

for a 25(OH)D concentration of approximately 40 ng/ml. This suggestion is based upon the increasingly accepted recommendation that a 25(OH)D concentration of more than or equal to 30 ng/ml be considered optimal, combined with the reported approximate 20% within-individual variability (*i.e.* 6 ng/ml if one targets 30 ng/ml) and the approximate 5 ng/ml systematic bias reported here. We recognize that some individuals with a serum 25(OH)D value of 40 ng/ml could truly have a value greater than this, but given the wide margin of safety reported for vitamin D, such a recommendation seems clinically prudent to ensure that the vitamin D status of an individual is optimal (>30 ng/ml).

Limitations of this study include the small sample size, limited population studied, and the fact that not all methodologies available for 25(OH)D measurement were evaluated. In this regard, it was our intent to compare the performance of 25(OH)D assays clinically available to us; we did not seek to evaluate all available assays or results performed in research settings. However, with the interest in automated technology, we did evaluate one such instrument, the DiaSorin Liaison in a research environment. In this setting, we found good correlation ( $R^2 = 0.95$ ) with laboratory A but feel it not justifiable to include as primary data because we did not have access to this technology in a routine clinical setting. Finally, we recognize that it is possible that other technologies and the study of other populations would provide differing results than reported here.

In conclusion, there is good agreement of serum 25(OH)D measured in these four clinical laboratories. Between-laboratory agreement should be improved with the use of standard calibrators. Based upon analytical and biological variability, it seems prudent that clinicians target a serum 25(OH)D value greater than 30 ng/ml to ensure vitamin D adequacy in their patients.

## Acknowledgments

Address all correspondence and requests for reprints to: Neil Binkley, M.D., 2870 University Avenue, Suite 100, Madison, Wisconsin 53705. E-mail: nbinkley@wisc.edu.

Disclosure Statement: The authors have no conflicts of interest relevant to this work.

## References

1. Institute of Medicine (U.S.). Standing Committee on the Scientific Evaluation of Dietary Reference Intakes 1997 Dietary reference intakes for calcium, phosphorus, magnesium, vitamin D, and fluoride/Standing Committee on the Scientific Evaluation of Dietary Reference Intakes, Food, and Nutrition Board, Institute of Medicine. Washington, DC: National Academy Press
2. Lips P, Chapuy MC, Dawson-Hughes B, Pols HAP, Holick MF 1999 An international comparison of serum 25-hydroxyvitamin D measurements. *Osteoporos Int* 9:394–397
3. Lensmeyer GL, Wiebe DA, Binkley N, Drezner MK 2006 HPLC method for 25-hydroxyvitamin D measurement: comparison with contemporary assays. *Clin Chem* 52:1120–1126
4. Binkley N, Krueger D, Cowgill C, Hansen KE, Drezner MK 2003 Assay variation confounds hypovitaminosis D diagnosis: a call for standardization. *J Clin Endocrinol Metab* 89:3152–3157
5. Bland JM, Altman DG 1986 Statistical methods for assessing agreement between two methods of clinical measurement. *Lancet* 327:307–310
6. Binkley N, Novotny R, Krueger D, Kawahara TN, Daida YG, Gemar D, Lensmeyer GL, Hollis BW, Drezner MK 2007 Low vitamin D status despite abundant sun exposure. *J Clin Endocrinol Metab* 92:2130–2135
7. Dawson-Hughes B, Heaney RP, Holick MF, Lips P, Meunier PJ, Vieth R 2005 Estimates of optimal vitamin D status. *Osteoporos Int* 16:713–716
8. Lips P 2001 Vitamin D deficiency and secondary hyperparathyroidism in the elderly: consequences for bone loss and fractures and therapeutic implications. *Endocr Rev* 22:477–501
9. Heaney RP 2000 Vitamin D: how much do we need, and how much is too much. *Osteoporos Int* 11:553–555
10. Holick MF, Siris ES, Binkley N, Beard MK, Khan A, Katz J, Petruschek RA, Chen E, de Papp AE 2005 Prevalence of vitamin D inadequacy among postmenopausal North American women receiving osteoporosis therapy. *J Clin Endocrinol Metab* 90:3215–3224
11. Rejnmark L, Lauridsen AL, Brot C, Vestergaard P, Heickendorff L, Nexø E, Mosekilde L 2006 Vitamin D and its binding protein Gc: long-term variability in peri- and postmenopausal women with and without hormone replacement therapy. *Scand J Clin Lab Invest* 66:227–238
12. Kim JH, Moon SJ 2000 Time spent outdoors and seasonal variation in serum concentrations of 25-hydroxyvitamin D in Korean women. *Int J Food Sci Nutr* 51:439–451
13. Lester E, Skinner RK, Wills MR 1977 Seasonal variation in serum 25-hydroxyvitamin D in the elderly in Britain. *Lancet* 1:979–980
14. Midtby M, Magnus JH, Joakimsen RM 2001 The Tromsø study: a population-based study on the variation in bone formation markers with age, gender, anthropometry and season in both men and women. *Osteoporos Int* 12:835–843
15. Woitge HW, Scheidt-Nave C, Kissling C, Leidig-Bruckner G, Meyer K, Grauer A, Scharla SH, Ziegler R, Seibel M 1998 Seasonal variation of biochemical indexes of bone turnover: results of a population-based study. *J Clin Endocrinol Metab* 83:68–75
16. Armas LA, Hollis BW, Heaney RP 2004 Vitamin D<sub>2</sub> is much less effective than vitamin D<sub>3</sub> in humans. *J Clin Endocrinol Metab* 89:5387–5391
17. Binkley N, Gemar D, Ramamurthy R, Krueger D, Drezner MK 2007 Daily versus monthly oral vitamin D<sub>2</sub> and D<sub>3</sub>: effect on serum 25OHD concentration. *J Bone Miner Res* 22(Suppl 1):S215