How Accurate Are Assays for 25-Hydroxyvitamin D? Data from the International Vitamin D External Quality Assessment Scheme, *Graham D. Carter,[1] Richard Carter,[2] Julia Jones,[1*] and Jacqueline Berry[3]* ([1] Endocrine Laboratory, Clinical Chemistry Department, Charing Cross Hospital, London, United Kingdom; [2] Sovereign Software Services, Haywards Heath, West Sussex, United Kingdom; [3] Vitamin D Research Group, Medicine, Manchester Royal Infirmary, Manchester, United Kingdom.: * address correspondence to this author at: Endocrine Laboratory, Clinical Chemistry Department, Charing Cross Hospital, Fulham Palace Road, London, W6 8RF, United Kingdom; fax 44-208-846-7007, e-mail julia.jones@imperial.ac.uk)

A recent report highlighted the interlaboratory variability in serum 25-hydroxyvitamin D (25OHD) results *(1)*. Concerns have also been raised about the apparent inability of some 25OHD assays to reliably measure 25-hydroxyvitamin D_2 ($25OHD_2$) *(2–4)*. The accurate measurement of this metabolite is essential for the monitoring of vitamin-D-deficient patients receiving ergocalciferol, which is the only supplement used in the United States *(5)* and is widely elsewhere.

The international Vitamin D Quality Assessment Scheme (DEQAS) has been monitoring the performance of 25OHD assays since 1989 and now has >100 registered participants in 18 countries. In essence, DEQAS is an ongoing, multicenter trial of the methods used by its participants and provides a unique opportunity to assess the accuracy and specificity of 25OHD methods as well as the analytical performance of a large number of their users.

The organization of DEQAS has been described elsewhere *(6)*, and details are available on the DEQAS website (www.deqas.org). In brief, five samples of normal human sera are sent out at 3-month intervals. Participants are asked to measure the total 25OHD concentration in each and return their results within 6 weeks. After statistical analysis of results *(7)*, participants receive a report giving an All-Laboratory Trimmed Mean (ALTM) and Standard Deviation (SD) for each sample. A small study conducted in 1997 *(8)* showed that the ALTM was a good surrogate for the "true" (target) value produced by gas chromatography–mass spectrometry. The accuracy of each result is defined by its percentage bias from the ALTM. Results for each sample are also grouped by method, and a method mean (MM) is produced. The overall accuracy of each method can be assessed from the percentage bias of the method mean from the ALTM: $\{[(MM - ALTM)/ALTM] \times 100\}$.

The percentage bias of method means was calculated for each sample (n = 88) sent out between January 2000 and January 2004, except for the Nichols automated assay (n = 49) and the IDS enzyme immunoassay (EIA; n = 23); results for these methods were available only from 2001 and 2002, respectively. Samples known to contain $25OHD_2$ were excluded. The ALTMs (nmol/L) for the samples ranged from 12.1 to 118.5 (median, 39.3). The mean percentage bias over this period was calculated for each major method group. The results are presented in Fig. 1.

With the exception of the Nichols automated procedure, the mean bias of 25OHD methods over this period was within 7% of the ALTM. Plots of percentage bias vs ALTM (not shown) revealed that in the IDS EIA, bias increased with concentration, whereas bias was decreased slightly in the Nichols assay at higher concentrations. Since its inclusion in the scheme, the Nichols assay has, for most samples, shown a marked positive bias, averaging ~31%. Detection of the small amounts (<10 nmol/L) of 24,25-dihydroxyvitamin D found in human serum can lead to high results in nonchromatographic methods for 25OHD. However, this is unlikely to cause the positive bias in the Nichols assay because the stated cross-reactivity (100%) is the same as that quoted for the two other methods (DiaSorin and IDS RIA) used by the majority of DEQAS participants. The positive intercept (8.7 nmol/L) of the regression line method mean vs ALTM (see below) suggests that a sample matrix effect is a contributory factor. This would not be surprising. Early attempts to simplify competitive protein binding assays for 25OHD, by abandoning the preliminary chromatographic step, failed because interfering substances in the sample matrix led to high results *(9, 10)*. The possibility that sodium azide, formerly used as a preservative in DEQAS samples, could have interfered in the chemiluminescence endpoint of the Nichols assay was considered but discounted because the bias persisted after the use of azide was abandoned in 2002. The continued use of the ALTM as the target value for 25OHD assays has also been questioned because the statistics could be distorted by the dominance of two methods (the DiaSorin and IDS RIAs), which currently account for >60% of the submitted results. However, the more rigorous chromatographic protein-binding assay and HPLC both continue to give results

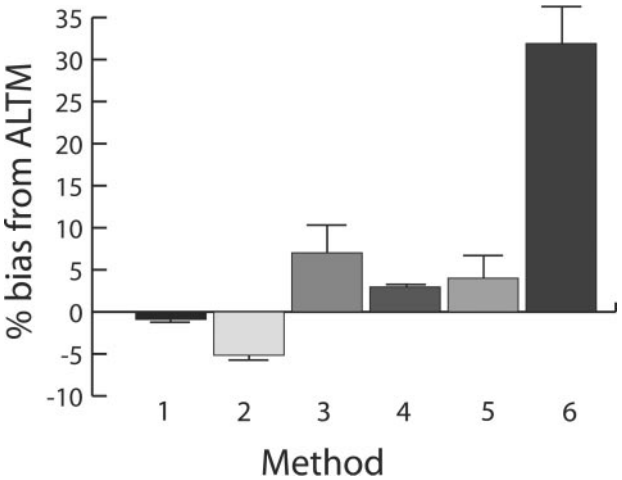

Fig. 1. Accuracy of 25OHD methods used by DEQAS participants.

Each *column* shows the mean deviation (*% bias*) and SE from the target value (ALTM) over the period January 2000 to January 2004 unless stated otherwise. Method means of samples known to contain $25OHD_2$ were excluded. *Method 1*, DiaSorin RIA; *method 2*, IDS RIA; *method 3*, IDS EIA (October 2002 to January 2004); *method 4*, competitive protein-binding assay; *method 5*, HPLC; *method 6*, Nichols automated chemiluminescence assay (July 2001 to January 2004).

close to the ALTM. This supports the use of the ALTM as a target value, although its validity is under constant review.

Linear regression analysis was performed to define the relationship between method means and the ALTM (x); the results were as follows (correlation coefficients in parentheses): DiaSorin = $0.97x + 0.64$ nmol/L ($r = 0.99$); IDS RIA = $0.98x - 1.13$ nmol/L ($r = 0.99$); IDS EIA (all results) = $1.28x - 9.02$ nmol/L ($r = 0.95$); IDS EIA (results <80 nmol/L) = $1.06x - 1.82$ nmol/L ($r = 0.96$); HPLC = $0.98x + 2.15$ nmol/L ($r = 0.96$); Nichols = $1.1x + 8.7$ nmol/L ($r = 0.90$). P was <0.001 for all correlation coefficients.

Among the five samples distributed in January 2004 were two (samples 4 and 5) containing endogenous $25OHD_2$. A summary of the results is presented in Table 1.

The wide range of results submitted by users of the same method illustrates the degree to which 25OHD assays are operator-dependent. The one automated procedure (Nichols) produced more consistent values, but for the samples (1, 2, and 3) containing only $25OHD_3$, the results were again higher than those of other methods.

For samples 4 and 5, in which $25OHD_2$ is the predominant metabolite, method means for the IDS RIA (IDS RIA) and the Nichols automated assay were lower than those of other methods. The inference is that the IDS RIA and the Nichols automated procedure underestimate $25OHD_2$, the latter by a considerable margin. Indeed, the presence in these samples of $25OHD_3$, albeit in smaller amounts than $25OHD_2$, disguises the true extent of the problem. IDS acknowledge that their RIA has only a 75% cross-reactivity with $25OHD_2$ *(11)*, but Nichols claims a 100% cross-reactivity and has presented data to support this *(12)*. The reason for the disparity between the manufacturer's claims and the DEQAS findings is unclear. Low recovery of $25OHD_2$ by the Nichols automated assay has also been observed in clinical samples (Dr. Carol Wagner, ???, personal communication). Vitamin-D-deficient neonates failed to show an increase in 25OHD, despite receiving large doses of ergocalciferol. When the samples were reanalyzed by the DiaSorin assay, the expected increase in 25OHD was observed. This suggests that the problem is not confined to DEQAS samples, an explanation originally proposed by the manufacturer to explain the overrecovery of $25OHD_3$. In the IDS RIA, the underrecovery of $25OHD_2$ can be attributed to differences in antibody specificity for the two forms of the metabolite. Interestingly, the same antibody is used in the nonisotopic version of the IDS assay, which despite the published cross-reactivity data of 75%, did not appear to underestimate $25OHD_2$ in the DEQAS samples.

In summary, an international quality assessment scheme has demonstrated that, for samples containing only $25OHD_3$, most commercial 25OHD methods are

Table 1. Method means (nmol/L) and CVs for total 25OHD results submitted for the DEQAS samples distributed in January 2004.[a]

	Sample				
Method	**1**	**2**	**3**	**4**	**5**
CPBA[b] (n = 2)					
Method mean, nmol/L (CV, %)	24.9 (17)	57.2 (4.5)	101.6 (27)	101.2 (12)	72.5 (12)
Median (range), nmol/L	24.9 (21.3; 28.4)	57.2 (55.0; 59.3)	101.6 (78.8; 124.5)	101.2 (91.3; 111.0)	72.5 (65.0; 80.0)
DIAS (n = 36)					
Method mean, nmol/L (CV, %)	31.5 (20)	59.7 (18)	87.7 (16)	106.5 (19)	68.8 (18)
Median (range), nmol/L	31.5 (17.3–42.4)	59.0 (41.8–79.2)	84.2 (65.5–117.0)	105.5 (80.0–209.0)	67.5 (47.0–106.0)
HPLC (n = 5)					
Method mean, nmol/L (CV, %)	22.9 (17)	55.1 (18)	82.0 (20)	80.8 (28)	55.0 (40)
Median (range), nmol/L	25.3 (17.0–26.4)	52.0 (46.6–77.0)	79.5 (62.5–109.0)	90.8 (38.0–101.8)	56.8 (12.0–82.0)
IDS EIA (n = 9)					
Method mean, nmol/L (CV, %)	28.6 (15)	58.2 (15)	79.1 (14)	110.1 (24)	67.7 (14)
Median (range), nmol/L	29.5 (21.0–34.0)	59.0 (46.0–69.0)	78.0 (62.5–93.7)	117.7 (68.0–139.7)	70.0 (47.0–81.0)
IDS RIA (n = 23)					
Method mean, nmol/L (CV, %)	30.6 (15)	57.5 (17)	90.3 (16)	65.6 (14)	48.4 (13)
Median (range), nmol/L	30.3 (24.0–45.0)	60.0 (42.7–71.0)	89.8 (69.9–117.2)	67.0 (52.8–82.4)	47.8 (38.8–61.3)
NICH ADV (n = 14)					
Method mean, nmol/L (CV, %)	37.3 (13)	72.7 (10)	109.5 (8.7)	36.6 (22)	38.0 (19)
Median (range), nmol/L	38.3 (28.0–46.5)	70.5 (65.0–85.0)	108.5 (97.0–145.5)	34.3 (22.8–49.0)	37.6 (29.0–60.3)
ALTM					
Method mean, nmol/L (CV, %)	31.1 (21)	60.2 (19)	90.4 (19)		
n	91	92	92		

[a] Samples 4 and 5 contained $25OHD_2$ and 25-hydroxyvitamin D_3; HPLC results indicated that $25OHD_2$ comprised ~80% (sample 4) and 66% (sample 5) of the total 25OHD. The ALTM, which includes results from minor methods not individually listed, is given for samples 1, 2, and 3 only. The ALTM is an inappropriate target value for samples containing $25OHD_2$.

[b] CPBA, competitive protein-binding assay; DIAS, DiaSorin RIA; NICH ADV, Nichols Advantage automated chemiluminescence assay.

capable of giving results close to the target value, but the results are highly operator dependent. The Nichols automated assay gives more consistent results but generally produces higher values than other methods. For samples containing predominantly $25OHD_2$, the Nichols procedure and, to a lesser extent, the IDS RIA, gave considerably lower results than other methods. The underrecovery of $25OHD_2$ by the Nichols assay, which has also been observed in clinical samples, occurred despite the manufacturer's claim that the method is equally specific for both forms of 25OHD. The treatment of vitamin-D-deficient patients could be severely compromised by the use of assays that underestimate $25OHD_2$.

The validity of 25OHD results will, justifiably, continue to be questioned. The only way for laboratories to demonstrate the accuracy of their results is to participate in an external quality assessment scheme and to make details of their performance available to clinical colleagues.

Since this article was submitted, Nichols Institute Diagnostics has issued a technical note in which they acknowledge that some samples containing substantial quantities of $25OHD_2$ give low results in the Advantage automated assay.

References

1. Binkley N, Krueger D, Cowgill L, Plum L, Lake E, Hansen KE, et al. Assay variation confounds the diagnosis of hypovitaminosis D: a call for standardization. J Clin Endocrinol Metab 2004;89:3152–7.
2. Hollis BW. Comparison of commercially available ^{125}I-based methods for the determination of circulating 25-hydroxyvitamin D. Clin Chem 2000;46: 1657–61.
3. Berry JL, Martin J, Mawer EB. 25-Hydroxyvitamin D assay kits: speed at the expense of accuracy? In: Norman AW, Bouillon R, Thomasett M, eds. Vitamin D endocrine system; structural, biological, genetic and clinical aspects. Riverside, CA: University of California, 2000:797–800.
4. Glendenning P, Noble JM, Taranto M, Musk AA, McGuiness M, Goldswain P, et al. Issues of methodology and metabolite recognition for 25-hydroxyvitamin D when comparing the DiaSorin radioimmunoassay and the Nichols Advantage automated chemiluminescence protein-binding assay in hip fracture cases. Ann Clin Biochem 2003;40:546–51.
5. Hollis B. Editorial: the determination of circulating 25-hydroxyvitamin D: no easy task. J Clin Endocrinol Metab 2004;89:3149–51.
6. Carter GD, Carter CR, Gunter E, Jones J, Jones G, Makin HJL, et al. Measurement of vitamin D metabolites: an international perspective on methodology and clinical interpretation. J Steroid Biochem Mol Biol 2004; 89–90:467–71.
7. Healy MJR. Outliers in clinical chemistry quality-control schemes. Clin Chem 1979;25:675–7.
8. Carter GD, Nolan J, Trafford DJ, Makin HJL. Gas chromatography-mass spectrometry (GC-MS) target values in the international quality assessment scheme (EQAS) for 25 hydroxyvitamin D (25OHD). In: Norman AW, Bouillon R, Thomasett M, eds. Vitamin D; chemistry, biology and clinical applications of the steroid hormone. Riverside, CA: University of California, 1997:737–8.
9. Belsey RE, DeLuca HF, Potts JT. A rapid assay for 25-OH-vitamin D_3 without preparative chromatography. J Clin Endocrinol Metab 1974;38:1046–51.
10. Dorantes LM, Arnaud SB, Arnaud CD, Kilgust KA. Importance of the isolation of 25-hydroxyvitamin D before assay. J Lab Clin Med 91:791–6.
11. IDS Ltd. IDS 25OHD radioimmunoassay [Product Insert]. Tyne and Wear, UK: IDS Ltd.
12. Nichols Institute Diagnostics. Nichols Advantage 25 OHD assay [Product Insert]. San Juan Capistrano, CA: Nichols Institute Diagnostics.

DOI: 10.1373/clinchem.2004.040683