Clinical Chemistry 55:7
1300–1302 (2009)

Opinions

25-Hydroxyvitamin D Assays: The Quest For Accuracy

Graham D. Carter[1*]

The recently reported difficulties experienced by Quest Laboratories *(1)* have once again highlighted concerns about the reliability of 25-hydroxyvitamin D (25-OH-D)[2] results, particularly those generated by liquid chromatography–tandem mass spectrometry (LC-MS/MS), the method used by Quest. In the *New York Times* article of January 8, 2009 *(1)*, a spokesman for Quest admitted that some erroneous results were reported because of problems with calibration and that some of their laboratories "did not always follow proper procedures." One suspects these shortcomings were influenced by the eye-watering number of 25-OH-D requests received by Quest: 500 000 per month according to John Cannell, quoted in a pathologist's newsletter *(2)*. This unfortunate episode dramatically emphasizes the need for a rigorous internal quality-assurance system. Such a system must include a role for a quality manager (preferably independent) who should be responsible for the monitoring of internal quality controls and the results of external proficiency-testing schemes. Some of us learned of the Quest problem through John Cannell's *Vitamin D Council* newsletter sent out in July 2008 *(3)*. In the same issue, he was promoting, without apparent irony, a commercial kit designed to "accurately" measure 25-OH-D in patient-generated blood spots, a technique that cannot easily be monitored by external proficiency-testing schemes.

Those of us who work in clinical laboratories know that "stuff happens." At some stage in our careers, many of us will probably have to contact clinical colleagues to admit reporting errors of some sort, although almost certainly not on the scale reported in the *New York Times*. It is to Quest's credit that they admitted the problem and offered to repeat these analyses free of charge.

It is important that the publicity given to Quest's problems not lead to LC-MS/MS being regarded as an inherently unreliable technique. The introduction of LC-MS/MS represents a hugely important development in the evolution of 25-OH-D testing, as it has for many other clinical analytes. It is a welcome antidote to the onward march of ever-simpler methods that have sacrificed analytical rigor on the altar of expediency. It is only a matter of time before the development of a commercial, Food and Drug Administration–approved LC-MS/MS "kit" that will presumably satisfy those who rather dismissively refer to the technique as a home-brew assay. It is worth remembering that the widely used DiaSorin RIA kit was once the home-brewed assay of a distinguished vitamin D research scientist *(4)*.

The suggestion that only the DiaSorin RIA kit gives "accurate" results *(5)* and that by inference other methods, including LC-MS/MS, are inaccurate is not based on evidence. The international Vitamin D External Quality Assessment Scheme (DEQAS) has been monitoring 25-OH-D assays for 2 decades, with individual results and method means compared with a consensus mean [All-Laboratory Trimmed Mean (ALTM)]. We originally showed this approach to be a good surrogate for values obtained by a rigorous GC-MS method *(6)*. Now, however, the scheme has many more participants who use a wider variety of methods, and the ALTM can no longer be regarded as the accurate or "true" value. This point was eloquently made by Lensmeyer et al. *(7)* in response to a suggestion that the DiaSorin RIA method gave accurate results because the method mean was close to the ALTM *(5)*.

Intuitively, the specificity afforded by mass spectrometry might be regarded as imbuing LC-MS/MS with greater accuracy than immunoassays, but as with all higher-technology methods, it is a technique requiring the skills of an experienced analyst. The use of inexperienced analysts might contribute to the rather high interlaboratory imprecision of LC-MS/MS–generated results, although the LC-MS/MS results submitted for the October 2008 distribution of DEQAS (Table 1) show interlaboratory CVs to be no worse than for most immunoassays.

Calibration errors are clearly a potential problem, and 2 recent studies *(8, 9)* have demonstrated the im-

[1] DEQAS Organizer, Imperial College Healthcare NHS Trust, London, UK.
* Address correspondence to the author at: Imperial College Healthcare NHS Trust, Oncology/Endocrine Laboratory, Charing Cross Hospital, Fulham Palace Rd., London W6 8RF, UK. Fax 44-(0)20-8846-7007; e-mail b.carter1@which.net.
Received February 23, 2009; accepted April 14, 2009.
Previously published online at DOI: 10.1373/clinchem.2009.125906
[2] Nonstandard abbreviations: 25-OH-D, 25-hydroxyvitamin D; LC-MS/MS, liquid chromatography–tandem mass spectrometry; DEQAS, Vitamin D External Quality Assessment Scheme; ALTM, All-Laboratory Trimmed Mean; RMP, Reference Measurement Procedure; IDS, Immunodiagnostic Systems; NHANES, National Health and Nutrition Examination Survey.

Table 1. Method means and ALTM for the 5 samples distributed by DEQAS in October 2008.

	Mean 25-OH-D concentration, nmol/L (CV)				
Sample no.	341	342	343	344	345
Method					
DiaSorin RIA (n = 40)	52.0 (18%)	80.4 (18%)	106.5 (16%)	37.8 (21%)	62.3 (17%)
DiaSorin Liaison T[a] (n = 144)	52.2 (15%)	75.8 (15%)	103.7 (14%)	31.0 (17%)	58.6 (15%)
IDS RIA (n = 29)	56.9 (15%)	89.6 (15%)	128.8 (15%)	39.1 (14%)	67.3 (14%)
IDS EIA (n = 84)	55.1 (13%)	83.6 (15%)	111.4 (17%)	34.8 (13%)	62.4 (16%)
IDS EIA, automated (n = 32)	56.8 (12%)	85.9 (15%)	113.6 (16%)	36.4 (13%)	63.1 (12%)
Roche 25-OH-D_3 (n = 26)	53.0 (14%)	75.0 (11%)	94.3 (12%)	43.4 (15%)	57.2 (15%)
HPLC (n = 16)	58.7 (25%)	94.3 (20%)	124.4 (21%)	38.9 (17%)	66.6 (18%)
LC-MS/MS (n = 39)	56.5 (15%)	94.5 (13%)	125.4 (13%)	39.6 (14%)	66.3 (18%)
ALTM (n = 437)	54.2 (15%)	81.8 (17%)	109.9 (17%)	35.2 (20%)	61.6 (16%)

[a] DiaSorin Liaison Total; EIA, enzyme immunoassay.

provement in interlaboratory precision that can be achieved when laboratories use the same calibrator for LC-MS/MS assays. Unfortunately, existing commercial suppliers of calibrators and controls are unwilling to divulge details of their manufacture; however, the NIST is producing reference sera (SRM 972) with values assigned by LC-MS/MS *(10)*. This action should be helpful in assessing the accuracy of 25-OH-D methods. NIST is also producing solvent-based calibrators (SRM 2972), an approach that potentially offers a way of standardizing to an accurate primary reference the calibration of HPLC and LC-MS/MS assays. The reference materials should be available during 2009.

One perceived disadvantage of LC-MS/MS assays is that they were not used for the clinical studies on which current 25-OH-D reference data are based. For this reason, LC-MS/MS users have been advised to calibrate their assays against the DiaSorin RIA, a method that has been widely used in epidemiologic surveys *(11)*. This approach is unsatisfactory for at least 3 reasons.

First, in practice, calibration to an RIA would mean applying an average correction factor to each result. Bland–Altman difference plots often reveal a large spread of results when one or both of the 25-OH-D methods are immunoassays *(12, 13)*. In one study that compared LC-MS/MS to the DiaSorin RIA, the differences between LC-MS/MS results and the mean of both methods spanned a broad range [−40 nmol/L to 40 nmol/L (−16 μg/L to 16 μg/L)] *(12)*.

Second, the near future will see the arrival of the long-awaited Reference Measurement Procedure (RMP) for 25-OH-D, which will almost certainly be based on mass spectrometry. If the present differences between immunoassays and LC-MS/MS results are confirmed, immunoassay manufacturers will surely have no option but to recalibrate their methods against the RMP. The vitamin D community will then have to accept that some 25-OH-D results obtained in the past by the methods available at the time are simply wrong.

Third, immunoassay kits are subject to performance changes over time, perhaps due to the reformulation of standards or reagents (a change of antibody being an extreme example of the latter). Manufacturers implicitly acknowledge such changes when they "recalibrate" their kits, presumably by allocating different values to their standards, to bring sample results into line with another method. Immunodiagnostic Systems (IDS) recalibrated their 25-OH-D enzyme immunoassay in 2006 after both DEQAS and internal QC data revealed a significant positive bias from their RIA method; DiaSorin recently modified their automated Liaison assay, a method used by an increasing number of DiaSorin customers, including LabCorp, a major competitor of Quest Diagnostics. DEQAS results revealed the modified assay to have a mean negative bias against the DiaSorin RIA of approximately 5% during the April 2007 to January 2008 distribution cycle; the bias is still evident in the results for October 2008 (Table 1). Such "methodologic drift" caused by assay reformulations complicates data interpretation in long-term surveys such as the National Health and Nutrition Examination Survey (NHANES) *(14)*. Nonimmunoassay methods such as LC-MS/MS should be less vulnerable to long-term performance change, problems with calibration notwithstanding.

Until a new body of data is collected with assays calibrated against an accepted RMP, 25-OH-D results should be interpreted with due regard to methodologic differences and without recourse to "fudge factors,"

which are inherently unreliable. Adjusting LC-MS/MS results to match those of any immunoassay, however well established, would no doubt be described as pragmatic, but I believe that most clinical chemists would share my belief that it is more akin to "cooking the books" (albeit with the best of intentions). Given the interlaboratory imprecision of most 25-OH-D methods (Table 1), the problem of method-related differences may have been overstated.

My experience in organizing DEQAS has convinced me that no analyst or method has a monopoly on virtue in the field of 25-OH-D measurements. The Quest affair has engendered a distinct whiff of schadenfreude in some quarters. Before the whiff turns into a stench, perhaps we should remind ourselves that only God knows whether a result (for any analyte) is the true value, an aphorism users of the late and unlamented Nichols Advantage method will probably endorse. We must hope that the advent of an RMP for 25-OH-D will bring us a little closer to analytical nirvana.

Author Contributions: *All authors confirmed they have contributed to the intellectual content of this paper and have met the following 3 requirements: (a) significant contributions to the conception and design, acquisition of data, or analysis and interpretation of data; (b) drafting or revising the article for intellectual content; and (c) final approval of the published article.*

Authors' Disclosures of Potential Conflicts of Interest: *No authors declared any potential conflicts of interest.*

Role of Sponsor: The funding organizations played no role in the design of study, choice of enrolled patients, review and interpretation of data, or preparation or approval of manuscript.

References

1. New York Times. Pollack A. Quest acknowledges errors in vitamin D tests. http://www.nytimes.com/2009/01/08/business/08labtest.html?_r=1&emc=eta1 (Accessed May 2009).
2. Doctor notices different vitamin D results over time. The Dark Report 2008;15:6–9.
3. Vitamin D Council. Cannell JJ. Supplementing with vitamin D. The Vitamin D Newsletter. 2008 Jul. http://www.vitamindcouncil.org/newsletter/2008-july.shtml (Accessed May 2009).
4. Hollis BW, Napoli JL. Improved radioimmunoassay for vitamin D and its use in assessing vitamin D status. Clin Chem 1985;31:1815–9.
5. Schmidt JA. Measurement of 25-hydroxyvitamin D revisited. Clin Chem 2006;52:2304–5.
6. Carter GD, Nolan J, Trafford DJ, Makin HJL. Gas chromatography-mass spectrometry (GC-MS) target values in the international external 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: proceedings of the Tenth Workshop on Vitamin D, Strasbourg, France, May 24–29, 1997. Riverside (CA): Printing and Reprographics, University of California; 1997. p 737–8.
7. Lensmeyer G, Wiebe D, Binkley N, Drezner M. Reply. Clin Chem 2006;52:2305–6.
8. Yates AM, Bowron A, Calton L, Heynes J, Field H, Rainbow S, Keevil B. Interlaboratory variation in 25-hydroxyvitamin D_2 and 25-hydroxyvitamin D_3 is significantly improved if common calibration material is used. Clin Chem 2008;54:2082–4.
9. Carter GD, Jones JC. Use of a common standard improves the performance of liquid chromatography-tandem mass spectrometry methods for serum 25-hydroxyvitamin-D. Ann Clin Biochem 2009;46:79–81.
10. Phinney KW. Development of a standard reference material for vitamin D in serum. Am J Clin Nutr 2008;88:511S–512S.
11. Hollis BW. Assessment of vitamin D status and definition of a normal circulating range of 25-hydroxyvitamin D. Curr Opin Endocrinol Diabetes Obes 2008;15:489–94.
12. Maunsell Z, Wright DJ, Rainbow SJ. Routine isotope-dilution liquid chromatography–tandem mass spectrometry assay for simultaneous measurement of the 25-hydroxy metabolites of vitamin D. Clin Chem 2005;51:1683–90.
13. 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. Ann Clin Biochem 2008;45:153–9.
14. Looker AC, Pfeiffer CM, Lacher DA, Schleicher RL, Picciano MF, Yetley EA. Serum 25-hydroxyvitamin D status of the US population: 1988–1994 compared with 2000–2004. Am J Clin Nutr 2008;88:1519–27.