Editorial

25-Hydroxyvitamin D assay standardisation and vitamin D guidelines paralysis

Overview: vitamin D guidelines paralysis

Vitamin D guidelines development is in a state of paralysis. There is an outward sign of a growing consensus between the two groups of nations: (1) the UK and the Netherlands with (2) Australia–New Zealand, European Union (EU) and the USA in how to define the lower limit of vitamin D adequacy based on serum total 25-hydroxyvitamin D (25(OH)D) concentration (Fig. 1), that is, the sum of serum 25(OH)D$_{2}$ and 25(OH)D$_{3}$ concentrations. But the two groups do not agree if there are data to support defining additional vitamin D physiological states, e.g. insufficiency, sufficiency and toxicity based on 25(OH)D concentrations. Moreover, those national recommendations conflict with a third set of recommendations from non-governmental medical societies and organisations (1), for example, the Endocrine Society, with no prospect for resolving the differences among the three approaches (1,2). Given the wealth of vitamin D research, it could be expected that the controversy would be resolved and consensus reached. Why has this not occurred?

To understand this situation, we need to understand the requirements for making and revising guidelines and policies that result from those guidelines. Understanding those requirements leads to a rather straightforward conclusion: Making and revising guidelines requires 25(OH)D data from rigorously conducted research and nationally representative surveys (3,4). In addition, nationally representative survey data are required to accurately and precisely assess the current state of vitamin D status in the population by person, season and geographical location, and to monitor changes over time in order to develop rational guidelines and associated policies and revise them over time (5,6). But variability of 25(OH)D assays thwarts attempts to resolve the controversy (7).

Variability of 25(OH)D assays is widely recognised (7–10). As such, only standardised 25(OH)D data provide the necessary level of accuracy and precision essential to the process of developing evidence-based vitamin D guidelines and policies (5–6). In the last 10 years, enormous progress has been made in collecting nationally representative survey data that meet those requirements (11–16), but despite this progress, we remain woefully behind in generating the vitamin D research data necessary to break out of the paralysis.

Thus, we have at a minimum three basic approaches for defining different states of vitamin D status, and no way to determine which is most appropriate. The goals of this Commentary are to describe the origins of the problem and to propose a set of recommendations based on work of the Vitamin D Standardization Program (VDSP) that may provide a way forward in developing rational vitamin D status guidelines.

Origin of vitamin D guidelines, 25(OH)D assay variability and policy paralysis

The UK appears to be the first country to have adopted serum 25(OH)D cut-points for defining vitamin D status. In 1991, the UK Committee on Medical Aspects of Food and Nutrition Policy (COMA) Dietary Reference Values (DRV) report stated that ‘Plasma 25-OHD concentrations in rickets range from not detectable to about 8 ng/ml’ (17). That appears to be the first recommended cut-off for 25(OH)D. The quoted source was a 1976 paper by Arnaud et al. (18) reporting serum 25(OH)D concentrations ranging between approximately 8 and 20 ng/ml in nine cases of rickets (Fig. 2). As a result, it appears that the UK DRV 1991 Committee acted conservatively in picking an 8 ng/ml (20 nmol/l) 25(OH)D cut-point. In the 1998 UK COMA report, the 25(OH)D cut-point was raised to 10 ng/ml (25 nmol/l) (19). In this case, the quoted source was a 1986 paper by Grindulis et al. (20). It is not clear why the value was increased to 10 ng/ml (25 nmol/l), but it may have been as simple as 10 is a round number easy for physicians to remember. The recent 2016 UK Scientific Advisory Committee on Nutrition (SACN) retained the 1998 COMA 10 ng/ml (25 nmol/l) concentration to define the lower limit of adequacy, ‘based on evidence suggesting risk of rickets and osteomalacia is increased at concentrations below this level’ (21). Notably, variability within and among 25(OH)D assays was cited as an important limitation for interpretation of studies that have examined the relationship between serum 25(OH)D concentration and health outcomes (21).

Concern about 25(OH)D assay variability is not a new problem. The problem was first reported in 1983–1984 (22–24) and in many publications since that time (25,26).
As a result, the Vitamin D External Quality Assessment Scheme (DEQAS) was introduced in 1989 to improve the reliability of 25(OH)D assays. In 2013, DEQAS became an accuracy-based external quality assessment scheme (EQAS) with Joint Committee for Traceability in Laboratory Medicine (JCTLM)-approved reference measurement procedures from the National Institute for Standards and Technology (NIST) and the Centers for Disease Control and Prevention (CDC) being used to provide the target values, that is, true 25(OH)D concentration, for serum samples used in DEQAS.

Currently, there are primarily two basic types of assays in use: (1) automated immunoassays, which are commercially developed and marketed; and (2) chromatography-based assays, which are primarily laboratory-developed using HPLC or LC-MS/MS. Trends in the performance of assays used by participating laboratories have been tracked since the initiation of DEQAS.

Recent DEQAS results (January 2019) document that 25(OH)D assay variability persists among the ten assays most commonly used by participating laboratories. Clearly, mean bias, defined as the percentage difference from the true concentration, varies by sample within an assay, and there is enormous variability around the mean for all the assays. Moreover, there is a great deal of variability among the different assays. HPLC and especially LC-MS/MS assays are often assumed to be the gold standard; however, results in Table 1 demonstrate that a great deal of variability exists among the laboratories using them. Similar results were found in the recent analysis of data from the College of American Pathologists (CAP) Accuracy-Based Vitamin D (ABVD) survey. Thus, using a chromatography-based assay and participating in an external quality assessment programme does not assure that just because a laboratory uses an HPLC or LC-MS/MS assay the results are either accurate or precise.
requirements of approximately 98% of the population, that is, 20 ng/ml (50 nmol/l). 25(OH)D cut-points were based on relationships to markers of bone health, and there was no single quoted source for the suggested cut-points.

Thus, the IOM 2011 guidelines were consistent with the UK guidelines and IOM 1997 in defining the lower limit of adequacy, but it was more expansive in setting 25(OH)D levels to define inadequacy 12–20 ng/ml (30–50 nmol/l), sufficiency 20–30 ng/ml (50–75 nmol/l), no added benefit 30–50 ng/ml (75–125 nmol/l) and possible harm >50 ng/ml (>125 nmol/l) such that we can speak of overlapping but non-congruent sets of national guidelines (1) (Fig. 1). Moreover, the authors of the report commented that ‘Currently, different assays for the determination of serum 25OHD levels are in use, and they provide disparate results. In turn, reported measures are confounded by the need to understand the assay used and research reports contain results that are not easily compared. The role of standard reference materials and interlaboratory collaboration is an important aspect of overcoming the challenges that the assay methodologies present’ (55). In short, this report acknowledged that without assay standardisation, results from different studies are not comparable and, ideally, they should not be pooled to develop consensus results.

National committees like those in the UK and USA appropriately tend to be very conservative. Once guidelines have been established, they are re-evaluated only when there is sufficient new data (30) and updated only when that data are irrefutable and there is consensus supporting change. No government agency wants to face a situation of having to retract one set of guidelines and replace it with a new set. With such a change, there is a loss of confidence in the entire process by the public at large, clinicians and scientists, all of which can result in the loss of necessary political support, that is, money, to develop and revise guidelines. Moreover, government laws, regulations, rules and programmes may be based on the guidelines, and changing and then unravelling existing guidance would not only result in lost confidence but an enormous cost in money and resources. As a result, once deficiency was defined in IOM 1997, it became the fulcrum around which future guidelines would revolve as seen in IOM 2011. An important issue that was hinted at in both IOM 2011 and SACN 2016 reports is that without assay standardisation further revision of 25(OH)D cut-points would be difficult, that is, paralysis.

Following the release of IOM 2011 guidelines, several other countries and medical societies released vitamin D guidelines (Fig. 1, Table 2). National agencies, more often than not, look to the guidelines of other countries when setting their own. The tendency is for national agencies to adopt the guidelines of other countries. In this case, the Nordic countries (37), the Swiss Federal Commission on Nutrition (38) along with the EFSA (39) adopted IOM 2011 guidelines or IOM 2011-like guidelines for interpreting 25(OH)D concentrations, while the Netherlands, being a bit more conservative, adopted the UK 1998 COMA recommendations to define a 25(OH)D concentration of 10 ng/ml (25 nmol/l) as the level at which risk of rickets and osteomalacia increases for persons aged 0–70 years (40). All these groups have adopted guidelines consistent with the conservative spirit of the UK DRV 1991 and IOM 2011 guidelines. That is, they define a serum 25(OH)D concentration of 25–30 nmol/l (10–12 ng/ml) as the lower limit of adequacy – an indicator of high risk of vitamin D deficiency (Fig. 1, Table 2). In addition, several organisations (41–51), including the Global Consensus Recommendations on Prevention and Management of Nutritional Rickets (51) (Table 2), have adopted guidelines consistent with IOM 2011 that define cut-points for insufficiency, sufficiency and possible harm.

Table 1 Mean bias from the ‘true’ sample concentration for the ten most commonly used assay platforms of the laboratories participating in the Vitamin D External Quality Assessment Scheme (DEQAS)*

<table>
<thead>
<tr>
<th>Assay</th>
<th>DEQAS sample 550</th>
<th>DEQAS sample 546</th>
<th>DEQAS sample 549</th>
<th>DEQAS sample 547</th>
<th>DEQAS sample 548</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Labs</td>
<td>Mean bias (%)</td>
<td>Mean bias (%)</td>
<td>Mean bias (%)</td>
<td>Mean bias (%)</td>
</tr>
<tr>
<td></td>
<td>(n)</td>
<td>95% CI</td>
<td>95% CI</td>
<td>95% CI</td>
<td>95% CI</td>
</tr>
<tr>
<td>Abbott Architect New</td>
<td>64</td>
<td>-7</td>
<td>-24, 9</td>
<td>-4</td>
<td>-15, 6</td>
</tr>
<tr>
<td>Bechman Unicel</td>
<td>35</td>
<td>6</td>
<td>-38, 50</td>
<td>-8</td>
<td>-31, 15</td>
</tr>
<tr>
<td>DiaSorin Liaison Total</td>
<td>185</td>
<td>18</td>
<td>-4, 41</td>
<td>7</td>
<td>-10, 25</td>
</tr>
<tr>
<td>HPLC</td>
<td>17</td>
<td>7</td>
<td>-52, 66</td>
<td>-1</td>
<td>-41, 40</td>
</tr>
<tr>
<td>IDS-SYS</td>
<td>31</td>
<td>0</td>
<td>-44, 44</td>
<td>5</td>
<td>-23, 33</td>
</tr>
<tr>
<td>IDS-SYS New</td>
<td>17</td>
<td>4</td>
<td>-25, 33</td>
<td>6</td>
<td>-25, 37</td>
</tr>
<tr>
<td>LC-MS/MS</td>
<td>143</td>
<td>-1</td>
<td>-23, 21</td>
<td>0</td>
<td>-20, 20</td>
</tr>
<tr>
<td>Roche Total 25OHD</td>
<td>76</td>
<td>5</td>
<td>-24, 35</td>
<td>4</td>
<td>-16, 24</td>
</tr>
<tr>
<td>Roche Vitamin D Total II</td>
<td>62</td>
<td>6</td>
<td>-21, 33</td>
<td>-5</td>
<td>-19, 9</td>
</tr>
<tr>
<td>Siemens ADVIA</td>
<td>48</td>
<td>34</td>
<td>-15, 83</td>
<td>12</td>
<td>-16, 40</td>
</tr>
<tr>
<td>Centaur</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

*January 2019 distribution for samples 546–550. Results are displayed by lowest to highest serum 25-hydroxyvitamin D concentrations of DEQAS samples. Data source: DEQAS Laboratory Report for January 2019 distribution (http://www.deqas.org/). [CDC reference measurement procedure true concentration target value; Mineva et al. (20).]
Table 2  Selected* recommendations for interpreting serum total 25-hydroxyvitamin D concentrations by type of committee, year of publication and consistency with UK DRV 1991, IOM2011 or Endocrine Society 2012 recommendations

<table>
<thead>
<tr>
<th>Type of committee</th>
<th>Year of publication</th>
<th>Consistent with</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Governmental</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>UK DRV(17)</td>
<td>1991</td>
<td>UK DRV 1991</td>
</tr>
<tr>
<td>US IOM(33)</td>
<td>1997</td>
<td>UK DRV 1991</td>
</tr>
<tr>
<td>UK COMA(19)</td>
<td>1998</td>
<td>UK DRV 1991</td>
</tr>
<tr>
<td>The Netherlands(40)</td>
<td>2012</td>
<td>UK DRV 1991†</td>
</tr>
<tr>
<td>UK SAGN(51)</td>
<td>2011</td>
<td>US IOM 2011</td>
</tr>
<tr>
<td>US IOM(50)</td>
<td>2011</td>
<td>US IOM 2011</td>
</tr>
<tr>
<td>NORDIC(37)</td>
<td>2012</td>
<td>US IOM 2011</td>
</tr>
<tr>
<td>European Food Safety Authority(39)</td>
<td>2016</td>
<td>US IOM 2011</td>
</tr>
<tr>
<td><strong>Non-governmental/medical societies</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Lawson Wilkins Pediatric Endocrine Society(41)</td>
<td>2008</td>
<td>US IOM 2011</td>
</tr>
<tr>
<td>German, Austrian, &amp; Swiss Nutrition Societies (DACH)(42)</td>
<td>2012</td>
<td>US IOM 2011</td>
</tr>
<tr>
<td>Australia/New Zealand(43,44)</td>
<td>2012/2013</td>
<td>US IOM 2011</td>
</tr>
<tr>
<td>British Paediatric(45)</td>
<td>2012</td>
<td>US IOM 2011</td>
</tr>
<tr>
<td>Spanish Association of Pediatrics(47)</td>
<td>2012</td>
<td>US IOM 2011</td>
</tr>
<tr>
<td>European Society for Paediatric Gastroenterology, Hepatology and Nutrition(48)</td>
<td>2013</td>
<td>US IOM 2011</td>
</tr>
<tr>
<td>Global Consensus Recommendations(51)</td>
<td>2016</td>
<td>US IOM 2011</td>
</tr>
<tr>
<td>Canadian Paediatric Society(52)</td>
<td>2007</td>
<td>Endocrine Society</td>
</tr>
<tr>
<td>International Osteoporosis Foundation(54)</td>
<td>2010</td>
<td>Endocrine Society</td>
</tr>
<tr>
<td>Scientific Advisory Council of Osteoporosis Canada(55)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Endocrine Society(56)</td>
<td>2011</td>
<td>Endocrine Society</td>
</tr>
<tr>
<td>Central Europe(56)</td>
<td>2013</td>
<td>Endocrine Society</td>
</tr>
<tr>
<td>Society for Adolescent Health and Medicine(57)</td>
<td>2013</td>
<td>Endocrine Society</td>
</tr>
<tr>
<td>American Geriatrics Society Workshop(58)</td>
<td>2014</td>
<td>Endocrine Society</td>
</tr>
<tr>
<td>National Osteoporosis Foundation(59)</td>
<td>2014</td>
<td>Endocrine Society</td>
</tr>
<tr>
<td>Vitamin D Council(60)</td>
<td>2015</td>
<td>Endocrine Society</td>
</tr>
<tr>
<td>United Arab Emirates(61)</td>
<td>2016</td>
<td>Endocrine Society</td>
</tr>
<tr>
<td>Japanese Medical Societies(62)†</td>
<td>2017</td>
<td>Endocrine Society†</td>
</tr>
<tr>
<td>Italian Pediatric Society(63)</td>
<td>2018</td>
<td>Endocrine Society</td>
</tr>
</tbody>
</table>

*Adapted, in part, from Bouillon(1) and Saggese et al(63).
†Expert panel supported by the research programme of intractable diseases; Ministry of Health, Labour and Welfare, Japan; the Japanese Society for Bone and Mineral Research; and the Japan Endocrine Society. It is not clear if this was a ‘governmental’ set of guidelines.
‡Recommendations for: ages 0–70; serum 25(OH)D level >12 ng/ml (30 nmol/l); and age >70: risk of bone fractures and serum 25(OH)D > 20 ng/ml (50 nmol/l).

Does this indicate a developing consensus among health agencies? Possibly, but we believe it likely indicates conformity rather than evidence-based consensus.

In 2011 a third distinct set of guidelines – Evaluation, Treatment, and Prevention of Vitamin D Deficiency: An Endocrine Society Clinical Practice Guideline (emphasis added) – was released by the Endocrine Society(52), and while they were quite similar to the 2007 guidelines released by the Canadian Paediatric Society(53) (Table 2), they have come to be the focus of opposition to the UK DRV 1991 and IOM 2011 guidelines. Several medical societies and non-governmental organisations have adopted the Endocrine Society guidelines(54–63). The Endocrine Society set the 25(OH)D concentration to define deficiency, insufficiency, sufficiency and possible harm at <20 ng/ml (50 nmol/l), 21–29 ng/ml (52.5–72.5), 30–100 ng/ml (75–250 nmol/l) and >100 ng/ml (>250 nmol/l), respectively based on the 2007 paper by Holick(64).

The Endocrine Society guidelines were quite different from anything proposed by the UK DRV 1991 or IOM 2011 guidelines (Table 3). They set off a firestorm of debate between the authors of the Endocrine Society and the authors of IOM 2011 guidelines that has continued ever since.

But there is even more confusion. The Endocrine Society stated that their guidelines were designed for clinical practice, while the IOM 2011 and later UK SAGN 2016 demurred by stating that their guidelines were not for use in clinical practice but were public health guidelines for the general healthy non-diseased population(41,55,52). We find it difficult to understand how a government agency can define interpretive guidelines for 25(OH)D – especially to define vitamin D deficiency – and not have them be clinically relevant. This emphasises the point that public health and clinical guidelines both need to revolve around identical sets of interpretive cut-points; otherwise there will be chaos and confusion among government agencies, physicians and the general public – as currently exists.

Given the sharp distinctions between the IOM 2011 and Endocrine Society guidelines, how can we go about determining the ‘best’ or ‘most appropriate’ cut-points given our current state of knowledge (Table 3)? Will meta-analyses solve the problem? There are an ever-increasing number published every year (Fig. 3), and yet we do not appear to be any closer to a resolution to this dilemma.

The fundamental reason why meta-analyses of the currently available data will not resolve the problem is that,
without assay standardisation, meta-analyses based on 25(OH)D concentrations are quite simply uninterpretable as it is impossible to pool the disparate results from different studies in any reasonable fashion. Currently, to our knowledge, there are only three meta-analyses based on standardised 25(OH)D levels(65–67), and all were conducted as part of the ODIN project(68). It was suggested recently that unstandardised results might end up providing the same answer as standardised data (69). However, given the amount of assay variability that exists, without standardisation, we will never know if the results are correct.

In today’s evidence-based world, guidelines and policy simply cannot be made based on serendipity. In summary, without accurate and precise data, countries with current guidelines committees are stuck in place, and other agencies/medical societies wanting to develop guidelines appear to be selecting one set of guidelines or the other without the data needed to resolve the differences. That means that without accurate and precise data, there is an inadequate basis to advance the vitamin D field and establish if 25(OH)D increases or decreases the risk of non-skeletal diseases. Thus, the paralysis that has ensued expands and worsens.

**VDSP: a ready solution**

The National Institutes of Health (NIH) Office of Dietary Supplements (ODS) established the VDSP in 2010 and coordinated its efforts until 2018. Since 2018 it has been coordinated as an independent agency. From the beginning, VDSP has been an international collaborative effort to standardise the laboratory measurement of serum total 25(OH)D and other potential markers of vitamin D status in order to improve clinical and public health practice(66,70).

A standardised laboratory measurement is defined as one that provides the ‘true’ total 25(OH)D concentration as measured by the three JCTLM-recognised reference measurement procedures(71). Serum 25(OH)D measurements can be ‘prospectively’ standardised using a standardised assay, or they can be ‘retrospectively’ standardised, after the fact, using methods developed by the VDSP(14,72,73).

Two examples from representative national surveys highlight the importance of having standardised 25(OH)D data to evaluate current levels, trends and seasonal differences in both representative national health surveys and vitamin D research. The first example is from the US National Health and Nutrition Examination Surveys (NHANES). In four surveys from 1988 to 2006, there appeared to be a dramatic drop in mean serum 25(OH)D levels for all persons aged >12 based on the original assay measurements (Fig. 4). DiaSorin Radioimmunoassay was used originally in all four surveys. At the time it could not be determined if the trend was real or not. However, when VDSP methods were used to retrospectively standardise the results for all four surveys, it became clear that the ‘decline’ was an assay artefact(74).

The second example is from the 2011–2012 Australian Health Survey (AHS) where it was found that the prevalence of vitamin D deficiency, that is, <20 ng/ml or

---

**Table 3** Comparison of Institute of Medicine(35) v. Endocrine Society(52) cut-points for serum total 25-hydroxyvitamin D (nmol/l)

<table>
<thead>
<tr>
<th>Interpretation</th>
<th>IOM</th>
<th>Endocrine Society</th>
</tr>
</thead>
<tbody>
<tr>
<td>Deficient</td>
<td>&lt;12</td>
<td>&lt;20</td>
</tr>
<tr>
<td>Insufficient</td>
<td>12–20</td>
<td>21–29</td>
</tr>
<tr>
<td>Sufficient</td>
<td>20–30</td>
<td>30–100</td>
</tr>
<tr>
<td>No added benefit</td>
<td>30–50</td>
<td></td>
</tr>
<tr>
<td>Possible harm</td>
<td>&gt;50</td>
<td>&gt;100</td>
</tr>
</tbody>
</table>

*nmol/l = ng/ml x 2.5.
50 nmol/l, increased dramatically in winter months compared to summer months. Countries without representative standardised 25(OH)D data can only guess in the development of vitamin D policy, programmes and their costs, which can further lead to a loss of political support for national surveys.

Two important questions are then: (1) how to select an assay; and (2) what are the key phases in VDSP.

VDSP criteria for selecting an assay are the following:

1. Fit for use;
2. Certified by the CDC Vitamin D Standardization Certification Program as being standardised and having an appropriate measurement range or be a documented standardised laboratory-developed HPLC or LC-MS/MS assay with an appropriate measurement range;
3. Appropriate level of assay precision and accuracy; and
4. Meets VDSP assay standardisation criteria in your ‘hands’ or laboratory.

‘Fit for use’ is a new criterion and has been added with the realisation that some immunoassays do not function appropriately in all patient populations. Because serum 25(OH)D was measured with a prospective VDSP-standardised LC-MS/MS assay, it is now possible for national and state governments to develop policy and cost estimates for programmes to ameliorate the situation based on accurate and precise estimates of current 25(OH)D levels and the prevalence of deficiency by region, age, sex, ethnic group and other potential risk factors.


Fig. 5 Prevalence of vitamin D deficiency in Australian summer months (serum 25(OH)D concentration <20 ng/ml (50 nmol/l); survey weighting factors were used to calculate representative prevalence figures for each state or territory). Australian Health Survey non-indigenous participants by state and territory, Australia, 2011–2012. Source: Australian Bureau of Statistics (75) (10 %; 10 to <20 %; 20 to <30 %; 30 to <40 %; 40 to <50 %)
accurately, or you may need one that measures total 25(OH)D in specific conditions, for example, pregnancy or in different pathophysiological states. Importantly, some immunoassays do not function well in pregnant women or in people with different diseases(79). Additionally, data suggest that some immunoassays do not measure serum 25(OH)D well(79,80). This can be a problem when ergocalciferol (vitamin D2) is used to treat vitamin D deficiency, as is the case in the USA or is used by vegetarians. It is apparent that if an assay cannot be verified to be fit for use, it should not be utilised.

If you are planning to select an immunoassay, we suggest that you first see which ones are currently, or have been in the past, certified by the CDC as meeting VDSP performance criteria of having a total CV \(\leq 10\%\) and a mean bias with the range of –5 to \(+5\\%\)(72,74). The list of current and past CDC-certified assays is available from the CDC website: https://www.cdc.gov/labstandards/hs.html.

The CDC list also gives information on the assays’ stated measurement range. VDSP recommends using an assay that does have an appropriate measurement range for the population it will be used in; for example, it should be able to measure 25(OH)D in persons who are deficient.

Representative national nutrition surveys need to have the very highest level of accuracy and precision. Therefore, it has been recommended that a standardised LC-MS/MS assay be selected(81).

There is another reason for national health surveys and researchers to consider using an LC-MS/MS assay. At present, the vitamin D field is in a great deal of flux where additional vitamin D metabolites, for example, 3-epi-25(OH) D3, 24R,25(OH)2D3 and vitamin D-binding protein (VDBP), may turn out to be essential to assessing vitamin D status(82,83). For national health surveys to maintain political support, they need to be flexible enough to respond to the needs that were not anticipated at the time the survey was designed. As a result, we suggest that, if possible, it would be prudent to measure 25(OH)D2, 25(OH)D3, 3-epi-25(OH)D3, 24R,25(OH)2D3 and possibly VDBP as well in those surveys. Given that researchers around the world are generally working with very limited budgets, we suggest that research grant applications include measurement of those compounds where it fits in with the hypotheses being tested and, otherwise, request funds to collect and appropriately store serum samples for potential future analyses.

It is possible to standardise the measurements for those analytes and 25(OH)D to 24R,25(OH)2D3 ratio and VDBP as well given that there are reference methods and reference materials available(27,84–86). NIST SRMs 972a, 1949 and
2973 along with selected DEQAS samples provide target values for 25(OH)D2, 25(OH)D3, 3-epi-25(OH)D3 and 24R,25(OH)2D3 (14,87,88). NIST SRM 1949 includes target values for VDBP (Table 4) (88). However, at present, VDSP statistical criteria to define standardisation/traceability for those compounds have not been defined. Currently, the VDSP and the International Federation of Clinical Chemistry and Laboratory Medicine are collaborating to define those statistical criteria.

To accomplish assay standardisation, we recommend the following two-phase approach:

- Phase 1: verification of fit for use
- Phase 2: calibration of assay to meet VDSP criteria, that is, total CV < 10% and mean bias ± 5% (14,72).

Verification of fit for use can be accomplished by testing the assay against a VDSP standardised LC-MS/MS assay. On the other hand, NIST SRM 1949 includes reference measurement target values in sera from pregnant women in the first, second and third trimesters and should be used to verify that the intended assay is ‘fit for use’ in pregnant women (Table 4) (88).

An essential point that needs repeating is that laboratories cannot assume that just because an immunoassay assay is CDC-certified it will function appropriately in their hands (89,90). We recommend a testing period in order to verify that an immunoassay is standardised especially since there is generally very little an individual laboratory can do to ‘calibrate’ an immunoassay.

If we started today to conduct all vitamin D research using only assays that were both ‘fit for use’ and VDSP-standardised, would we be able to decide between IOM 2011 and Endocrine Society guidelines (Table 3)? The sad fact is that we could not, until substantial amounts of new, standardised data are available. Since the discovery of serum 25(OH)D3 in 1968 (91), approximately 80 000 vitamin D papers have been published (Fig. 7). Virtually all of those, where serum total 25(OH)D was measured, reported only

Table 4 Reference values for Standard Reference Material® 1949: Frozen Human Prenatal Serum*

<table>
<thead>
<tr>
<th>Constituent†</th>
<th>Non-pregnant</th>
<th>First trimester</th>
<th>Second trimester</th>
<th>Third trimester</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Mean</td>
<td>Expanded uncertainty</td>
<td>Mean</td>
<td>Expanded uncertainty</td>
</tr>
<tr>
<td>25(OH)D3 (ng/ml)</td>
<td>24·98</td>
<td>0·28</td>
<td>26·01</td>
<td>0·22</td>
</tr>
<tr>
<td>25(OH)D2 (ng/ml)</td>
<td>N/A</td>
<td>N/A</td>
<td>N/A</td>
<td>N/A</td>
</tr>
<tr>
<td>3-epi-25(OH)D3 (ng/ml)</td>
<td>1·32</td>
<td>0·06</td>
<td>1·20</td>
<td>0·05</td>
</tr>
<tr>
<td>VDBP (μg/ml)</td>
<td>211·5</td>
<td>2·8</td>
<td>286·7</td>
<td>3·8</td>
</tr>
<tr>
<td>VDBP (μmol/kg)</td>
<td>4·01</td>
<td>0·05</td>
<td>5·43</td>
<td>0·06</td>
</tr>
</tbody>
</table>

3-epi-25(OH)D3, 3-epi-25-hydroxyvitamin D3; VDBP, vitamin D-binding protein.
*A additional target values given for measures of thyroid function, as well as for copper, selenium and zinc. For additional details, see National Institute for Standards and Technology. Certificate of Analysis Standard Reference Material® 1949: Frozen Human Prenatal Serum.
†Equations to convert from ng/ml to nmol/l: (1) 25(OH)D3: 2·496 × ng/ml; (2) 25(OH)D2: 2·423 × ng/ml; (3) 3-epi-25(OH)D3: 2·496 × ng/ml.
‡No concentration value provided in the Certificate of Analysis.

Fig. 7 Vitamin D papers published since 1922. Source: PubMed: https://www.ncbi.nlm.nih.gov/pubmed/ (accessed September 2019)
unstandardised data. As a result, we recommend that the vitamin D field needs to work with researchers, journal editors and research funding agencies around the world to identify and promote keys studies for retrospective standardisation using methods and procedures developed by the VDSP(75,92). At that point, the retrospectively standardised data could be re-analysed and published in meta-analyses based solely on standardised results(93).

Importantly, retrospective standardisation can be accomplished relatively inexpensively as shown in the re-analyses of the Canadian Health Measures and HunMen study data(12,50). Such approaches could be utilised in virtually all vitamin D research. Journal editors are encouraged to require either ‘prospectively’ standardised or ‘retrospectively’ standardised 25(OH)D data as a condition for publication.

We recognise that all science evolves. Similarly, there are weaknesses in the current VDSP performance criteria that will, hopefully, be improved as the vitamin D field progresses. One worthy of brief mention is the VDSP performance criteria for mean bias(54). Nonetheless, we believe it is past time to implement the recommendations laid out here.

Conclusions

It is past time for the vitamin D research field to embrace reporting only standardised 25(OH)D data. The assays utilised must be: (1) fit for use in the population studied; and (2) standardised (either prospectively or retrospectively) across the appropriate measurement range. Moreover, funding agencies, for example, the US NIH, and journals must make this essential for study funding and publication. Failure to do so will perpetuate the current paralysis and preclude moving the vitamin D field forward.

Acknowledgements

Acknowledgements: The authors would like to thank Dr. Mona S. Calvo, Dr. Peter J. Gergen and Mr. Lawrence Kessenich for reading and very kindly providing thoughtful comments and suggestions to an early draft. Figures 1 and 2 are being used with the permission of Nature Reviews Endocrinology and Pediatrics, respectively. Conflict of interest: C.T.S.: The NIH/ODS established the VDSP in December 2010 and supported it through April 2018. Dr. C.T.S. coordinated the VDSP throughout that entire time period as an NIH employee and then, after his retirement, as a consultant to NIH/ODS. Since April 2018, Dr. C.T.S. has continued to coordinate the VDSP privately in order to fulfil its goal of promoting standardised measurement of vitamin D status in order to improve clinical and public health practice worldwide. N.B. has no conflicts of interest. Financial support: This research received no specific grant from any funding agency in the public, commercial or not-for-profit sectors. Authorship: Formulating commentary outline: C.T.S. and N.B. Data analyses: C.T.S. Writing article: C.T.S. and N.B. Reviewing article: C.T.S. and N.B.

CT Sempo and N Binkley

1Vitamin D Standardization Program LLC, Havre de Grace, MD, USA
Email sempoche@gmail.com

2Osteoporosis Clinical Research Program and Institute on Aging, Department of Medicine, University of Wisconsin-Madison, Madison, WI, USA

References


