National Health Statistics Reports

Number 93 ■ April 25, 2016

A Method-bridging Study for Serum 25-hydroxyvitamin D to Standardize Historical Radioimmunoassay Data to Liquid Chromatography–Tandem Mass Spectrometry

by Rosemary L. Schleicher, Ph.D., and Maya R. Sternberg, Ph.D., National Center for Environmental Health; David A. Lacher, M.D., National Center for Health Statistics; Christopher T. Sempos, Ph.D., National Institutes of Health; Anne C. Looker, Ph.D., National Center for Health Statistics; Ramon A. Durazo-Arvizu, Ph.D., and Elizabeth A. Yetley, Ph.D., National Institutes of Health; Madhulika Chaudhary-Webb, M.S., Khin L. Maw, M.S., and Christine M. Pfeiffer, Ph.D., National Center for Environmental Health; and Clifford L. Johnson, M.S.P.H., National Center for Health Statistics

Abstract

Background—Serum concentrations of 25-hydroxyvitamin D (25OHD) were measured for the National Health and Nutrition Examination Survey (NHANES) over the 1988–2006 period using a radioimmunoassay (RIA). In 2010, the Centers for Disease Control and Prevention (CDC) reissued RIA-harmonized 25OHD for NHANES 2004 and 2006, and advised users to adjust the original RIA data from 1988–1994 by using an equation. Beginning with NHANES 2007–2008, a liquid chromatography–tandem mass spectrometry (LC–MS/MS) method measured 25OHD.

Methods—A method comparison (bridging) study was designed to convert original RIA 25OHD to LC–MS/MS-equivalents. This report compares the predictive ability of a competitor regression model (Model 2) to the equations that CDC publicly released in 2015 (Model 1). The models differ by time period variable and use of transformations.

Results—The two models provided similar adjusted R^2 (Model 1: 88.9%, Model 2: 90.4%) and root mean square error of prediction (plus or minus 9 to 10 nanomoles per liter [nmol/L]). Applying these models to NHANES 1988–2006 RIA 25OHD, the Pearson correlation of LC–MS/MS-equivalent concentrations was 0.99; the median difference between models was 0 nmol/L (interquartile range: –2.8 to 1 nmol/L). In contrast to declining RIA-harmonized 25OHD, both models showed little change in LC–MS/MS-predicted 25OHD over the 1988–2006 period. For 2001–2006, both models predicted similar prevalences of 25OHD less than 30 nmol/L, which were lower than the prevalence estimates based on RIA-harmonized data. Mean weighted LC–MS/MS-equivalent concentrations based on either model were about 3 nmol/L lower for the 1988–1994 survey and about 3 nmol/L higher for the 2001–2006 surveys, effectively smoothing out temporal trends observed using the harmonized RIA data.

Conclusions—Given minimal differences between models, final selection was based on public availability of the regression data. The bridging equations provide a way to use the previous RIA results to obtain LC–MS/MS-equivalent concentrations and evaluate temporal trends in vitamin D status.

Keywords: vitamin D • regression equation • laboratory comparison • NHANES

Introduction

When important clinical methods change, laboratories perform comparison (bridging) studies to relate the old and new methods. These bridging efforts provide information useful for the continuity of care of individual patients and for evaluating trends in monitoring public health. Generally, results from the old method are converted to new-method equivalents. In the current study, the old method was a radioimmunoassay (RIA) for the vitamin D status indicator, which is serum concentrations of 25-hydroxyvitamin D (25OHD); the new method is a liquid chromatography–tandem mass spectrometry (LC–MS/MS) method traceable to Standard Reference Materials (SRM) from the National Institute of Standards and Technology (NIST).

A reformulation by the RIA manufacturer had identified a major shift in values (–12%) in kits available around 2001, and subtle drifts in values (plus or minus 5%–10%) on two occasions in 2004 and 2006, which were likely due to kit lot-to-lot variability (1). To compensate for these RIA shifts and drifts, the National Health and Nutrition Examination Survey (NHANES) had used stored specimens and long-term

U.S. DEPARTMENT OF HEALTH AND HUMAN SERVICES
Centers for Disease Control and Prevention
National Center for Health Statistics

quality control materials to harmonize the 25OHD results (1). Consequently, although the true values of serum specimens may not be known with a specified degree of assurance, specimens tested at any given time should produce approximately the same results, using adjustments, despite kit changes. This concept was endorsed by a roundtable of experts in vitamin D (1) that was convened to consider the best way to address the problem of reagent kit variability, which was confounding the interpretation of population trend data. However, these experts also agreed that an interim harmonization based on RIA results needed to be replaced as early as possible by a long-term objective to standardize the 25OHD RIA results to LC–MS/MS, allowing a more accurate interpretation of long-term trends.

To achieve the long-term objective, a subset of serum samples from NHANES III (1988–1994) and NHANES 2001–2006 that were originally tested for 25OHD using RIA were retested using an LC–MS/MS method that was traceable to SRM. These data were used to develop regression equations relating the old and new methods so that LC–MS/MS-equivalent 25OHD could be predicted for any RIA results from these surveys.

Methods

Background on NHANES

The serum specimens and 25OHD data in this study came from NHANES III and NHANES 2001–2006. NHANES is a cross-sectional survey of the civilian noninstitutionalized population of the United States, conducted by the Centers for Disease Control and Prevention's (CDC) National Center for Health Statistics (NCHS). Sampling design and data collection methods are available (2). Survey participants were interviewed at home and had blood collected at mobile examination centers. Participation rates for each survey are detailed elsewhere (3).

Laboratory measurements of 25OHD using RIA and LC–MS/MS

For the surveys conducted between 1988 and 2006, RIA (DiaSorin, Stillwater, Minn.) was used to measure total 25OHD (in duplicate) (4). A fully validated LC–MS/MS method (5), traceable to NIST reference materials, was used to measure 25-hydroxyvitamin D_2 ($25OHD_2$), 25-hydroxyvitamin D_3 ($25OHD_3$), and the C3 epimer of $25OHD_3$ for selected stored specimens for those NHANES participants during 1988–2006 having available RIA data for the bridging study (in duplicate). Once the bridging study was nearly finished, the same assay was used to measure 25OHD for all eligible participants in NHANES 2007–2010 (in singlicate). For the LC–MS/MS method, total 25OHD was defined as the sum of $25OHD_3$ and $25OHD_2$, excluding the C3 epimer of $25OHD_3$, about which less is known. The bias of the LC–MS/MS method relative to NIST SRM during the course of the bridging study was modest (3% or less) for $25OHD_3$ and $25OHD_2$ at concentrations greater than 2 nanomoles per liter (nmol/L). Furthermore, the average bias of CDC's LC–MS/MS method for a set of 50 individual donor serum samples—derived from an interlaboratory comparison study performed around the time of the bridging study—was 1.4% (95% confidence interval [CI]: 0.9%–1.8%). The serum samples were from the first interlaboratory comparison study sponsored by the Vitamin D Standardization Program (6) relative to two independent reference measurement procedures, carried out by NIST and the University of Ghent (7,8).

Quality assurance for bridging study

To achieve imprecision and bias goals for the LC–MS/MS assay (9), the 25OHD assay performance was continually monitored using quality control (QC) pools, reference materials, and external quality assurance program feedback. Calibration was performed using in-house prepared stock solutions. Value transfer from NIST SRM 2972 (ethanol matrix) was performed with each new lot of calibrators using either direct (CDC calibrators assigned using SRM 2972 calibrators) or indirect (CDC calibrators adjusted using SRM 2972 and SRM 972) value transfer procedures. Each run contained standard reference materials from NIST (10), namely, four levels of SRM 972 (serum matrix) and three levels of serum bench QC pools, which were analyzed at the beginning and end of the run. Decision rules for out-of-control runs were Westgard-type but more flexible to allow for multiple QC pools (11). For $25OHD_3$, low pools with mean concentrations of 28.5–29.7 nmol/L, medium pools with mean concentrations of 52.6–63.7 nmol/L, and high pools with mean concentrations of 86.0–92.2 nmol/L had coefficients of variation (CV) of 4% or less. For $25OHD_2$, low pools with mean concentrations of 5.0–13.9 nmol/L had CV 6%–8%, medium pools with mean concentrations of 10.8–38.9 nmol/L had CV 4%–5%, and high pools with mean concentrations of 21.6–62.7 nmol/L had CV 6%.

A number of additional quality assurance activities were conducted to confirm the validity of the bridging study results. The stability of 25OHD in long-term storage and under accelerated degradation conditions was established during the design phase of this study, because plans called for using serum that was in storage for up to 23 years for NHANES III, for which pristine (i.e., frozen once, never thawed) specimens were not available. For NHANES 2001–2006, residual serum stored at –70°C and freeze-thawed multiple times, or pristine serum stored in liquid nitrogen fumes at –130°C, were available. The 25OHD results from 35 NHANES specimens stored at –70°C for 10 years, which may have undergone multiple freeze-thaws for RIA testing, were compared with serum specimens stored at –130°C from the same 35 participants and never thawed until LC–MS/MS testing. There was 95% power for a 0.05-level equivalence test with a sample size of 35 to test that the difference of the mean responses was within equivalence bounds (–2.0 nmol/L, 2.0 nmol/L), assuming no difference between surplus and pristine samples, a correlation of 0.98, and a common

standard deviation of 16 nmol/L. No significant difference was found in $25OHD_3$ concentrations. Both storage conditions were found to be essentially the same (mean paired $25OHD_3$ difference: –0.74 [95% CI: –2.17 to 0.68] nmol/L; $p = 0.30$).

Using QC pools, 25OHD metabolites were shown to be stable in serum stored at 37°C for at least 8 days, and four freeze-thaw cycles were not associated with loss of 25OHD (12). Using three QC pools used by the laboratory during 1994–1995 and the final prediction model used to standardize NHANES data, the predicted LC–MS/MS results for these pools were shown to be within one root mean square error of prediction (rMSEP) of the LC–MS/MS results that were measured 17 years later in 2012, affirming that 25OHD is stable over long periods.

Because NHANES 1988–1994 specimens were collected in serum separator tubes, and NHANES 2001–2006 specimens were collected in red-top serum vacuum tubes, matrix equivalency needed to be established. Serum prepared from serum separator and red-top tubes from 27 blood donors were tested for 25OHD. Over 95% power was achieved for a 0.05-level equivalence test with a sample size of 27 to test that the difference of the mean responses was within equivalence bounds (–2.0 nmol/L, 2.0 nmol/L), assuming no difference between two matrices, a correlation of 0.98, and a common standard deviation of 8.5 nmol/L. No significant difference was found in $25OHD_3$ concentrations. Both matrices were found to be essentially the same (mean paired $25OHD_3$ difference: –0.01 [95% CI: –1.33 to 1.32] nmol/L; $p = 0.99$).

To determine the stability of the LC–MS/MS assay, more than 100 specimens were retested from NHANES 2005–2008, covering the period of the bridging study and the time period in which NHANES 2007–2008 samples were tested. Specifically, all specimens with adequate volume were retested from the 2005–2006 bridging study ($n = 52$; 25OHD range: 12–152 nmol/L); Bland-Altman bias of this set, retested in 2014 compared with the original testing in 2011, was not significant (difference –0.1%; 95% CI: –2.2 to 2.0%; $p = 0.92$). Specimens from NHANES 2007–2008 were also retested ($n = 73$; 25OHD range: 20–140 nmol/L); Bland-Altman bias of the retest in 2014 compared with the original testing in 2012 also was not significant (difference –0.1%; 95% CI: –1.4 to 1.3%; $p = 0.88$).

Bridging study design

The basic design of the method comparison was reviewed and evaluated by a panel of experts (1). The Clinical and Laboratory Standards Institute (Wayne, PA) method evaluation protocol, EP09–A2, was used as a guideline. The majority of the NHANES III samples (90%) were measured using the original RIA assay of 1995–1996; NHANES samples from 2000–2006 were assayed by use of the reformulated RIA assay within 12 months of collection. No specimens were collected for 25OHD measurements in 1999; because data confidentiality requirements placed participant 25OHD data from year 2000 in restricted access at the NCHS Research Data Center, the survey period 1999–2000 was not included in this analysis.

For NHANES III, the data were sorted from lowest to highest 25OHD, ranked into quartiles, and then sorted by date; overall, approximately 100–200 specimens per quartile were selected to provide a representative range of dates and concentrations. NHANES III specimens were randomized for LC–MS/MS testing. The majority of specimens were selected from 1995 and 1996 (dates of analysis), because the majority of RIA analyses were performed during these years (e.g., 1994 [19%], 1995 [35%], 1996 [44%], and 1997 [2%]). Due to low numbers, 1997 results were combined with those from 1996 for generating regression equations. Fourteen specimens per run were selected for analysis of 25OHD metabolites, but sometimes fewer (12 or 13 specimens) were tested due to insufficient quantity. NHANES III analysis was complete after 43 analytical runs.

Historical RIA data from NHANES 2001–2006 were sorted into survey year (2001, 2002, 2003, 2004, 2005, and 2006). Within each survey year, RIA data were sorted by lowest to highest 25OHD. The sorted data were divided into quartiles. Within quartiles, the data were sorted by month of laboratory measurement. For NHANES 2001–2006, 25–40 specimens per quartile per year were selected to provide a representative range of dates and concentrations; these were randomized for testing. Four specimens per year were selected for analysis of 25OHD metabolites in each run, but sometimes fewer than four specimens were tested because of insufficient specimen quantity. After 25 analytical runs, another selection of specimens was carried out, both to obtain enough specimens due to insufficient quantity and to enhance the number of specimens measured at the tails of the distribution. The method comparison data set was not limited by age and included sera from participants under age 12 years. NHANES 2001–2006 and NHANES III specimens were analyzed together. In total, 57 analytical runs were performed to complete the method comparison study.

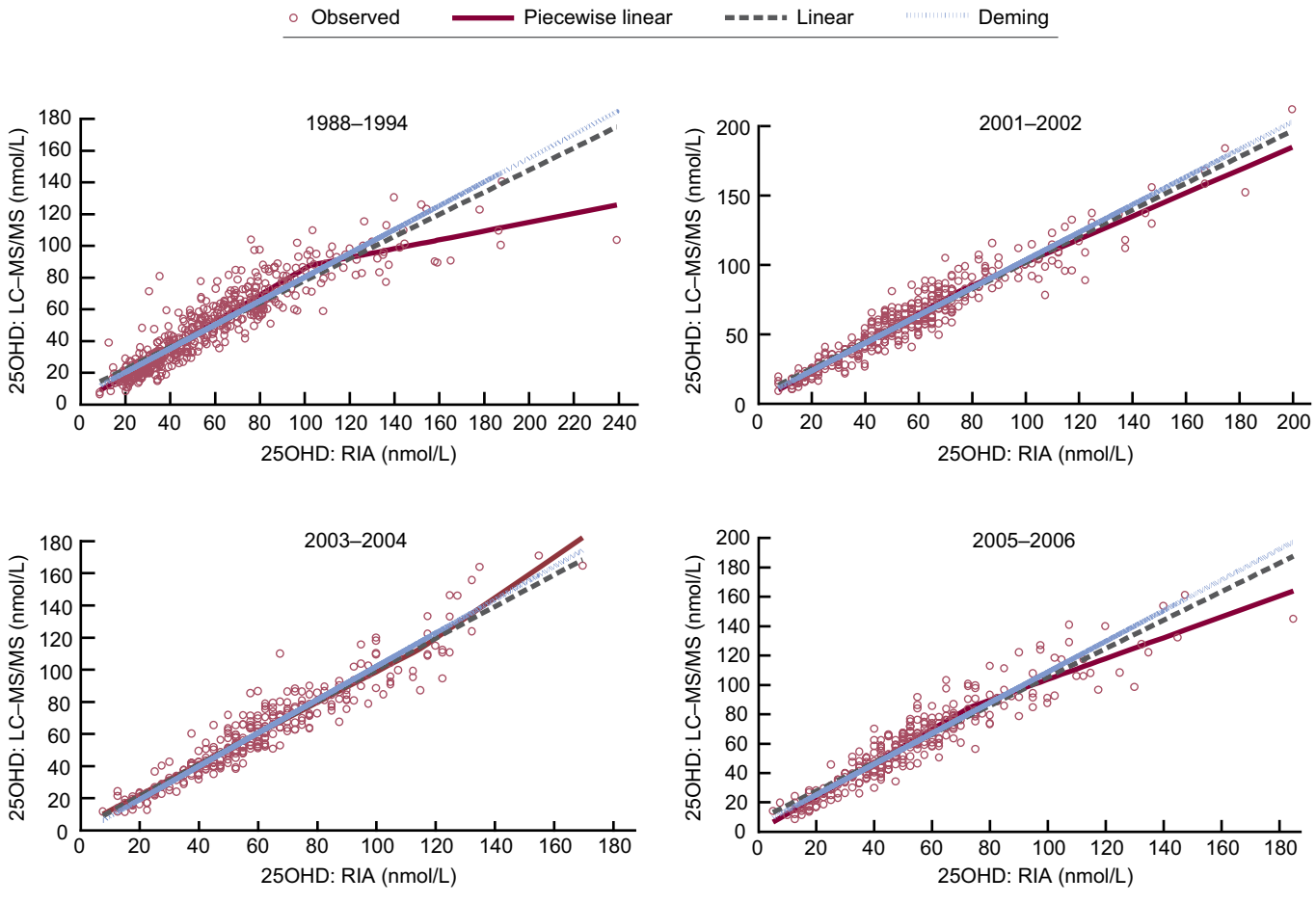
One of the major issues to address was how to model the kit-to-kit variability of RIA. Although the date of RIA measurement could be used, such exact dates and years are not part of publicly released data sets. If the proposed statistical model used measurement year, the yearly regression equations would not be publicly released due to confidentiality concerns, and NCHS would need to calculate the predicted new-method results, withhold the regression equations, and re-release the data. Therefore, the alternative of providing regression equations for each survey period was proposed. This would allow data users to access the regression equations used in the re-released data. However, because the assay fluctuations were most closely captured by using the calendar year of RIA measurement, the need for 1-year regression equations was evaluated against statistical models that used the survey period. Other features were considered, including the need for a transformation to address regression assumptions such as linearity, normality, and homogeneity; the need to address measurement error related to the RIA assay; and whether other covariates beyond RIA values and the time of measurement should be considered.

Model development

The data, based on 1,448 subjects from NHANES 1988–2006 selected as part of the crossover data set, were used to develop regression equations to predict NHANES participants' LC–MS/MS-equivalent concentrations from their previously measured RIA values. The dependent variable was the sum of $25OHD_2$ and $25OHD_3$ in nmol/L units based on LC–MS/MS. When $25OHD_2$ was less than the limit of detection (LOD), the value for $25OHD_2$ was calculated as LOD divided by the square root of 2. The primary independent variable was the original RIA result in nanograms per milliliter (ng/mL) multiplied by 2.496 to convert to nmol/L. Several statistical issues that pertain to any linear regression model such as linearity, constant variance, normality, and lack of measurement error in the covariate(s) were evaluated. Graphical diagnostics were used to assess model adequacy, as well as to identify any potentially influential data points. Potential outliers were rechecked using LC–MS/MS, but they were verified. Historically, it is known that the distribution of 25OHD is slightly skewed right with some nonconstant variance. Statistical approaches to address these types of violations of a linear regression model included the use of a square root transformation, weighted least squares, or piecewise regression. Ultimately, a square root transformation, or a combination of piecewise regression and ordinary least squares regression, was selected for the final models.

Because it was known that the manufacturer reformulated the RIA and that lot-to-lot variability in the assay kits was noted on two separate occasions during the 2001–2006 survey periods, a covariate closely aligned with the time of RIA analysis was to be used in developing the models. The choice between using survey period (publicly available) and year of RIA analysis (more closely related to the date of RIA measurement but not publicly available) to mimic the timing of any assay fluctuations as a covariate in the models was considered. Two competing models were selected for further evaluation.

Model 1 is described in an online analytical note for NHANES III (1988–1994), NHANES 2001–2006, and NHANES 2007–2010 as a data advisory



NOTES: LC–MS/MS is liquid chromatography–tandem mass spectrometry, 25OHD is 25-hydroxyvitamin D, RIA is radioimmunoassay, and nmol/L is nanomole per liter. Bridging data set consists of 1,448 stored serum samples, originally tested using RIA, from National Health and Nutrition Examination Surveys covering 1988–1994 and 2001–2006.
SOURCE: CDC/NCHS, National Health and Nutrition Examination Survey, 1988–1994 and 2001–2006.

Figure 1. Serum 25-hydroxyvitamin D concentrations measured by LC–MS/MS and predicted from Model 1 using different regression approaches for bridging study data set, stratified by survey period: National Health and Nutrition Examination Survey, 1988–1994 and 2001–2006

Table A. Individual regression models for predicting serum concentrations of LC–MS/MS-equivalent 25-hydroxyvitamin D from original radioimmunoassay 25-hydroxyvitamin D, using a bridging data set: National Health and Nutrition Examination Survey, 1988–1994 and 2001–2006

Regression model and time period	Equation
Model 1	
Survey period:	
1988–1994	If $RIA_{original} \leq 102$, then $LC–MS/MS_{equivalent} = 1.57548 + 0.8429 \bullet RIA_{original}$
	If $RIA_{original} > 102$, then $LC–MS/MS_{equivalent} = 59.2296 + 0.2788 \bullet RIA_{original}$
2001–2002	$LC–MS/MS_{equivalent} = 6.43435 + 0.95212 \bullet RIA_{original}$
2003–2004	$LC–MS/MS_{equivalent} = 1.72786 + 0.98284 \bullet RIA_{original}$
2005–2006	$LC–MS/MS_{equivalent} = 8.36753 + 0.97012 \bullet RIA_{original}$
Model 2	
Year of RIA analysis:	
1994	$SQRT(LC–MS/MS_{equivalent}) = 0.94761 + 0.80717 \bullet SQRT(RIA_{original})$
1995	$SQRT(LC–MS/MS_{equivalent}) = 1.16953 + 0.7708 \bullet SQRT(RIA_{original})$
1996	$SQRT(LC–MS/MS_{equivalent}) = 0.35182 + 0.85957 \bullet SQRT(RIA_{original})$
2001	$SQRT(LC–MS/MS_{equivalent}) = 0.6371 + 0.91432 \bullet SQRT(RIA_{original})$
2002	$SQRT(LC–MS/MS_{equivalent}) = 0.67272 + 0.96489 \bullet SQRT(RIA_{original})$
2003	$SQRT(LC–MS/MS_{equivalent}) = 0.10961 + 1.01678 \bullet SQRT(RIA_{original})$
2004	$SQRT(LC–MS/MS_{equivalent}) = 0.47863 + 0.92234 \bullet SQRT(RIA_{original})$
2005	$SQRT(LC–MS/MS_{equivalent}) = 0.17074 + 0.9854 \bullet SQRT(RIA_{original})$
2006	$SQRT(LC–MS/MS_{equivalent}) = 0.90722 + 0.97113 \bullet SQRT(RIA_{original})$

NOTES: All units are in nanomole per liter. LC–MS/MS is liquid chromatography–tandem mass spectrometry. RIA is radioimmunoassay. SQRT is square root.

SOURCE: CDC/NCHS, National Health and Nutrition Examination Survey, 1988–1994 and 2001–2006.

informing researchers that previously released nonstandardized NHANES 25OHD data have been converted to standardized 25OHD and re-released (13). A comparison of linear, Deming, and piecewise regression equations for each survey cycle (Figure 1) shows a large difference between the linear and piecewise fit for NHANES III, but negligible differences for the 2001–2006 period. Therefore, Model 1 used a piecewise linear regression for NHANES III to account for some nonlinearity, and a separate ordinary least squares regression for each subsequent survey cycle. Model 2 used a square root transformation on both the dependent variable LC–MS/MS 25OHD and the independent variable RIA 25OHD to address the mild concentration dependence and nonlinearity. Models 1 and 2 can each be formulated into single multiple linear regression (MLR) models using indicator variables and properly defined interactions. This facilitates a direct comparison of model diagnostics and adjusted R-squared. MLR Model 1 consisted of the primary independent variable (RIA 25OHD) and a set of three indicator variables (X_1–X_3) for each survey cycle (2001–2002, 2003–2004, and 2005–2006). The reference survey period was selected arbitrarily as NHANES III, so $X_1 = X_2 = X_3 = 0$. Interactions were included between each of the three indicator variables and RIA to allow for a unique slope for each survey cycle. To account for the piecewise regression for NHANES III, an indicator variable was created that was 1 when RIA 25OHD was greater than 102.2 nmol/L, and 0 otherwise; an interaction between this indicator variable and RIA 25OHD allowed for a separate slope when RIA was greater than 102.2 nmol/L during NHANES III (1988–1994). Model 1 did not use any transformation for either the dependent variable (LC–MS/MS 25OHD) or the primary independent variable (RIA 25OHD). MLR Model 2 consisted of a square root transformation for the primary independent variable (RIA 25OHD) and a set of eight indicator variables (X_1–X_8) for each year of RIA measurement (1994, 1995, 1996, 2001, 2002, 2003, 2004, and 2005). The reference year was selected arbitrarily as 2006, such that $X_1 = X_2 = \ldots X_8 = 0$. Interactions were included between each of the eight indicator variables and RIA 25OHD to allow for separate slopes for each year of RIA measurement.

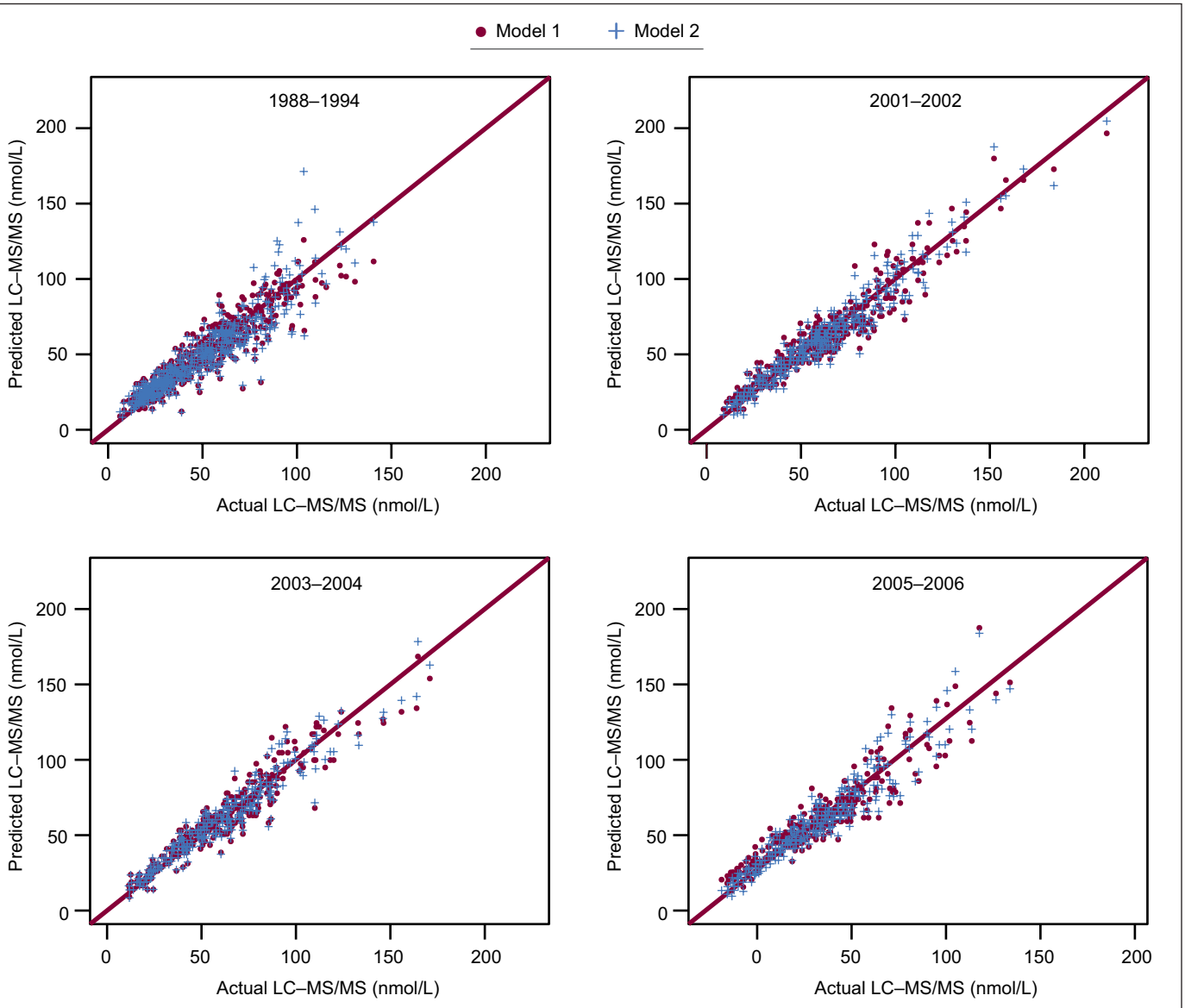
To compare the predictive ability of each model, the mean square error of prediction (MSEP) was estimated using fivefold cross validation. MSEP is the average squared difference between the observed LC–MS/MS value and the LC–MS/MS-equivalent prediction from a specified model. To estimate the MSEP using fivefold cross validation, a stratified sample of 80% of the original bridging data set was randomly selected. The strata were the year of RIA measurement to ensure 80% representation in any given year. The 80% sample was used as a training data set to estimate the model parameters, and the remaining 20% was used as a validation data set to calculate the MSEP. This was repeated five times, and the five MSEPs were averaged; the square root of the MSEP (rMSEP) was used to judge the models. The smaller the rMSEP, the better the predictive ability of a given model.

To examine the difference between the two models in terms of their impact on the assessment of temporal trends in the U.S. population, the unweighted and weighted arithmetic means, the 5th and 95th percentiles, and the prevalence of less than 30 nmol/L of LC–MS/MS-equivalent total 25OHD were calculated for each NHANES period (1988–1994, 2001–2002, 2003–2004, and 2005–2006), and results were compared with the original RIA and harmonized RIA results. NHANES examination weights were used for the weighted means and percentiles, and variance estimates based on Taylor series linearization (14) were used for the 95% CI.

Statistical analyses were performed using SAS (version 9.3, SAS Institute Inc., Cary, N.C.) and SUDAAN (version 11.0.1, RTI International, Research Triangle Park, N.C.).

Results

The individual simple linear regression equations for each survey for Model 1 and for each year of RIA testing for Model 2 are shown in Table A. Model 1 was chosen as the final model and was used to provide LC–MS/MS-equivalent 25OHD for the historical NHANES (1988–2006) that were originally tested using RIA (13). A graphical comparison of the predicted LC–MS/MS-equivalent and actual LC–MS/MS-measured 25OHD for each of the two models in

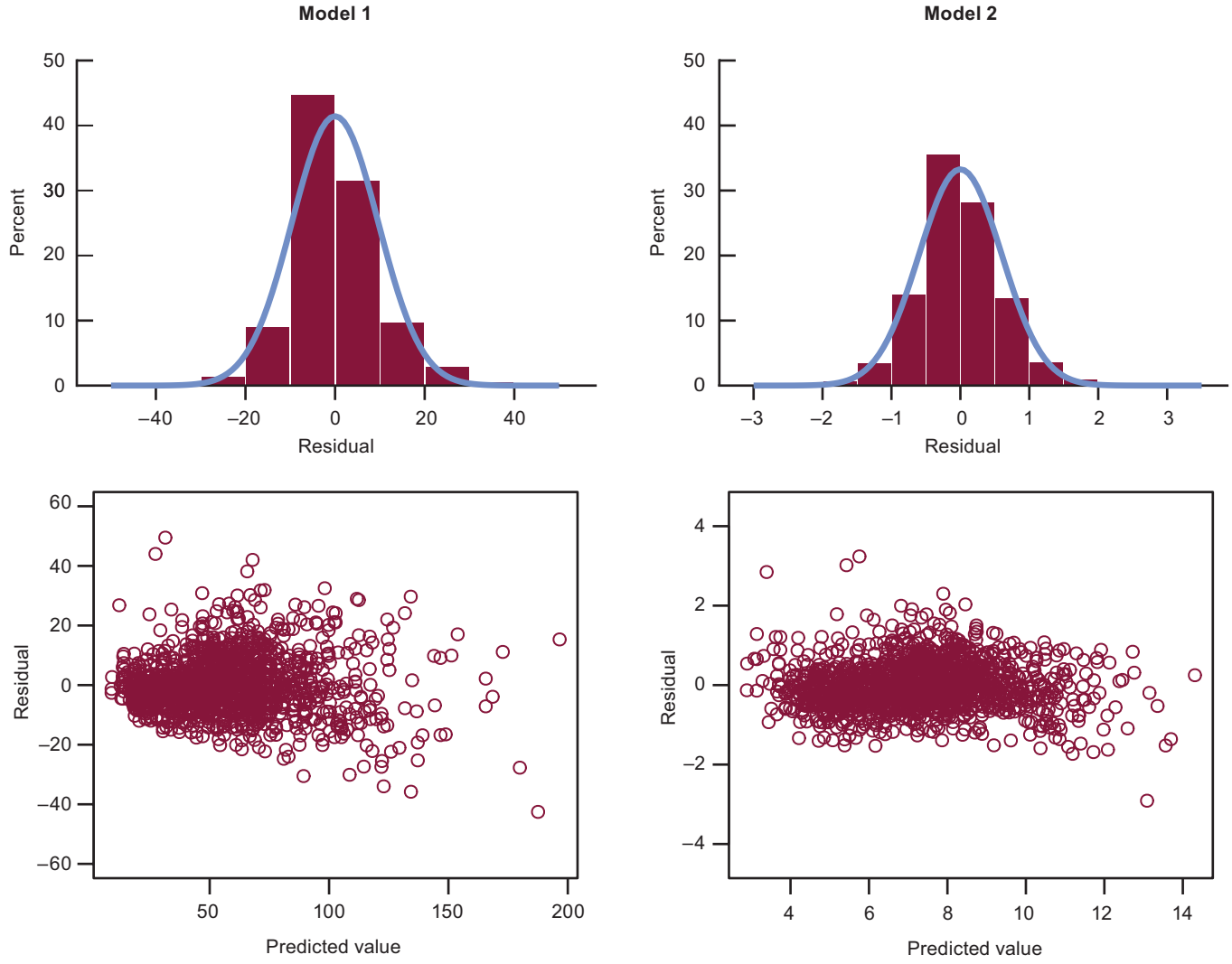


NOTES: LC–MS/MS is liquid chromatography–tandem mass spectrometry, and nmol/L is nanomole per liter. The diagonal line is the line of identity. Bridging data set consists of 1,448 stored serum samples, originally tested using radioimmunoassay (RIA), from National Health and Nutrition Examination Surveys covering 1988–1994 and 2001–2006. Prediction equations for Model 1 were prepared for each survey (6-year or 2-year cycle). For Model 2, equations were prepared for each year in which the original RIA data were generated to more fully compensate for variability in the RIA kits. Model 2 data were then sorted into survey cycles for comparison of models.
SOURCE: CDC/NCHS, National Health and Nutrition Examination Survey, 1988–1994 and 2001–2006.

Figure 2. Serum 25-hydroxyvitamin D concentrations predicted from models 1 and 2 compared with measured LC–MS/MS concentrations in bridging study data set, stratified by survey period: National Health and Nutrition Examination Survey, 1988–1994 and 2001–2006

the bridging data set showed similar predictions for each model (Figure 2). For the bridging data set, the median difference between the predicted values (Model 2 – Model 1) was less than 1 nmol/L (median = –0.77 nmol/L; interquartile range [IQR] = –3.4 to 2.3 nmol/L), and overall Pearson correlation was 0.99. A comparison of the model diagnostics, using multiple linear regression models, showed that Model 1 had slightly less preferable visual features, with the histogram of residuals being a little less normally distributed, and residuals showing a slight concentration dependence at the high end (Figure 3). The residuals from Model 2 were more closely distributed normally without concentration-dependence; the estimated R-squared from the two models were similar (Model 1, 88.9%, compared with Model 2, 90.3%).

The estimated rMSEP for the bridging study data set was plus or minus 9 to 10 nmol/L for both models, although it tended to be slightly smaller for Model 2 (Table B). However, when applied to a small independent validation data set (a set of 66 samples from NHANES 2001–2006, for which only singlicate LC–MS/MS 25OHD data were available), Model 1 had a slightly smaller rMSEP. In addition to the two models described in this report, a number of different statistical models were evaluated. These models included approaches based on weighted regression and Deming regression, with the ratio of the error



NOTES: LC–MS/MS is liquid chromatography–tandem mass spectrometry, and nmol/L is nanomole per liter. Residuals and predicted values are derived from equations shown in Table A. Units for the residuals and predicted values are nmol/L for Model 1, and square root nmol/L for Model 2. Regression Model 1 (using equations for each survey period) and Model 2 (using equations for each calendar year in which RIA data were tested) were used to generate LC–MS/MS-standardized data.
SOURCE: CDC/NCHS, National Health and Nutrition Examination Survey, 1988–1994 and 2001–2006.

Figure 3. Summary of regression diagnostics from two multiple linear regression models for predicting serum concentrations of LC–MS/MS-equivalent 25-hydroxyvitamin D from original radioimmunoassay 25-hydroxyvitamin D using a bridging data set: National Health and Nutrition Examination Survey, 1988–1994 and 2001–2006

variances either assumed to be 1 or a value estimated from CDC laboratory QC results. However, the rMSEP were much higher for these than for the two models described above and, therefore, these models are not described.

The equations from Model 1 and Model 2 were each used to predict LC–MS/MS-equivalent concentrations based on the original NHANES RIA 25OHD from each survey (1988–1994, 2001–2002, 2003–2004, and 2005–2006). The estimated distributions of the predictions made from each model were similar, except for the right tail of 1988–1994 (Figure 4). Overall, the Pearson correlation between LC–MS/MS-equivalent concentrations for models 1 and 2 across all of the surveys was 0.99, and the median difference of the predicted LC–MS/MS-equivalent concentrations between these models (Model 2 – Model 1) was 0 nmol/L (IQR = –2.8 to 1 nmol/L).

Unweighted mean 25OHD based on the original RIA and harmonized RIA concentrations, and the predictions from models 1 and 2, showed that the predicted LC–MS/MS-equivalent data from models 1 and 2 differed by less than 2 nmol/L when based on multiyear survey data (Table C). For NHANES 2001–2006, the differences between the unweighted mean predictions for models 1 and 2 were more apparent when looking at the 1-year rather than the 2-year mean predictions, ostensibly due to the effect of averaging over 2 years (Table C). As expected, the average RIA-harmonized values were lower in 1988–1994 and 2003–2004, but higher in 2005–2006, than the original RIA 25OHD values due to the harmonization efforts. Although both models 1 and 2 provided lower mean estimates for 1988–1994 than the original RIA, these standardized estimates were also lower than the harmonized RIA values. In contrast, during the entire

2001–2006 period, models 1 and 2 provided higher unweighted means than the original RIA and harmonized RIA means (Table C).

Weighted predicted means using models 1 and 2 showed small differences (2 nmol/L or less) across demographic subgroups in the different surveys (Table 1). The demographic subgroup that deviated the most when comparing the weighted mean predictions between Model 1 and Model 2 was the non-Hispanic black population. For example, in 2001–2002, the Model 1 weighted mean was 1.1 nmol/L higher for non-Hispanic black persons than Model 2, but for other demographic subgroups Model 1 was approximately 0.9 nmol/L lower than Model 2. Similarly, for 2005–2006, an approximately 0.3 nmol/L difference was noted between Model 1 and Model 2 for other demographic subgroups, but Model 1 averaged 2.9 nmol/L higher than Model 2 for non-Hispanic black persons (Table 1).

The differences in the two models' predictions of the LC–MS/MS-estimated tails of the distributions of 25OHD varied across the demographic subgroups depending on the survey cycle (Table 2). The estimated 95th percentile from each model differed the most for the 1988–1994 survey. For example, the estimated 95th percentile for Model 1 was 10 nmol/L lower for youth aged 12–19 years compared with Model 2, whereas the 95th percentile for Model 1 was 4.9 nmol/L higher for non-Hispanic black persons compared with the Model 2 estimate. The estimated 5th percentiles were relatively similar between the models across the demographic subgroups, with a less than 1 nmol/L difference from 1988 to 2004. However, for 2005–2006, the estimated 5th percentiles from Model 1 were approximately 3 nmol/L higher across all subgroups compared with Model 2.

When using each of the prediction models to generate LC–MS/MS-standardized data, neither model revealed much change in the prevalence of 25OHD less than 30 nmol/L (Table 3) over the 1988–2006 period. In contrast, based on the RIA-harmonized data, the prevalence of 25OHD less than 30 nmol/L was higher in the 2001–2006 surveys compared with the 1988–1994 survey. Compared with the RIA-harmonized data, the mean weighted LC–MS/MS-equivalent concentrations from either model were approximately 3 nmol/L lower for the 1988–1994 period and approximately 3 nmol/L higher for the 2001–2006 period; this effectively smoothed out any temporal trends in the means observed with the harmonized and the original RIA data (Figure 5).

Table B. Root mean square error of prediction from fivefold cross-validation testing for bridging study data set and independent data set

Data set	Model 1	Model 2
NHANES survey period:	nmol/L	
1988–1994	9.4	9.5
2001–2002	9.1	8.2
2003–2004	9.7	9.3
2005–2006	10.6	9.5
Independent data set	7.9	8.1

NOTES: NHANES is National Health and Nutrition Examination Survey, and nmol/L is nanomole per liter. Model 1, used for generating publicly released liquid chromatography–tandem mass spectrometry (LC–MS/MS)-equivalent data, is based on NHANES data. Model 2 is based on the year in which radioimmunoassay (RIA) data were generated (date of RIA testing is available from the NCHS Research Data Center). RIA data, collected from a set of 66 samples from NHANES 2001–2006 for which only singlicate LC–MS/MS 25-hydroxyvitamin D concentration data were available, were standardized using models 1 or 2, and the root mean square error of prediction was calculated for each model.

SOURCE: CDC/NCHS, National Health and Nutrition Examination Survey, 1988–1994 and 2001–2006.

Conclusions

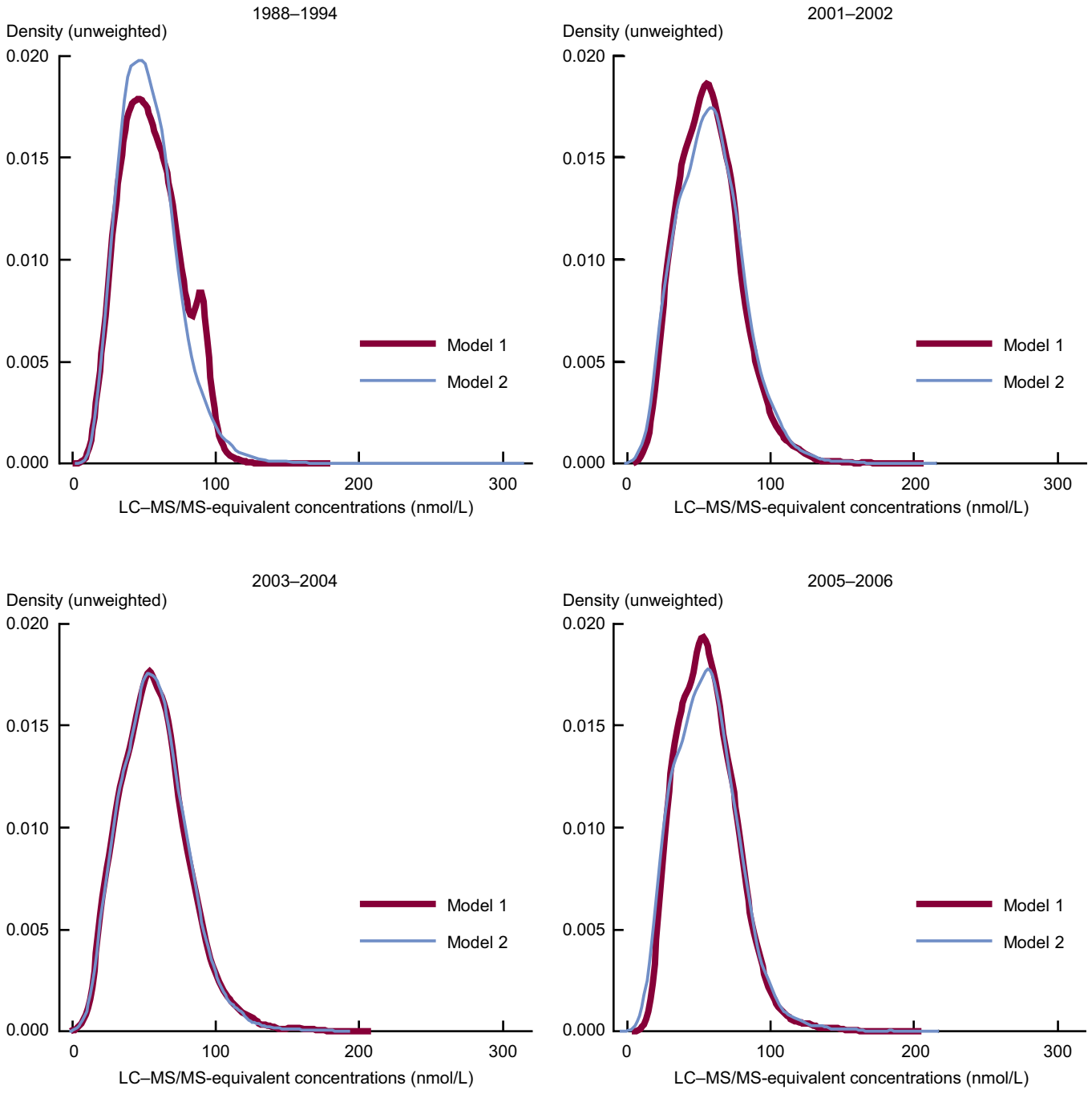
The objective of this study was to standardize previously obtained 25OHD RIA results from NHANES 1988–2006 to LC–MS/MS to allow a more accurate interpretation of long-term trends in vitamin D status in the United States over this 18-year time span. Starting with NHANES 2007–2008, the new

Table C. Serum concentrations of 25-hydroxyvitamin D measured using original radioimmunoassay, or harmonized using stored survey specimens or quality-control pool data, or standardized using regression models, for persons aged 12 years and over, by survey period, phase, or year: National Health and Nutrition Examination Survey, 1988–1994 and 2001–2006

Time period	Sample size	RIA–original	RIA–harmonized	LC–MS/MS-standardized	
				Model 1	Model 2
	n	nmol/L			
1988–1994	18,851	64.6	57.0	55.0	53.3
Phase 1: 1988–1991	8,977	65.7	58.0	55.9	54.9
Phase 2: 1991–1994	9,874	63.6	56.2	54.1	51.9
2001–2002	6,816	53.4	53.3	57.2	57.8
2003–2004	6,553	56.6	54.2	57.4	57.2
2005–2006	6,480	48.7	52.5	55.6	55.0
2001	3,507	53.9	53.9	57.7	56.9
2002	3,309	52.8	52.8	56.7	58.7
2003	3,278	52.7	51.8	53.5	55.8
2004	3,275	60.5	56.6	61.2	58.6
2005	3,241	51.0	52.4	57.9	53.8
2006	3,239	46.3	52.6	53.4	56.3

NOTES: Values are unweighted arithmetic means. RIA is radioimmunoassay; nmol/L is nanomole per liter; and LC–MS/MS is liquid chromatography–tandem mass spectrometry. Two models were used to estimate LC–MS/MS-equivalent values from RIA-original data: Model 1 uses regression equations for each survey period, and Model 2 uses regression equations for each calendar year in which RIA data were tested.

SOURCE: CDC/NCHS, National Health and Nutrition Examination Survey, 1988–1994 and 2001–2006.

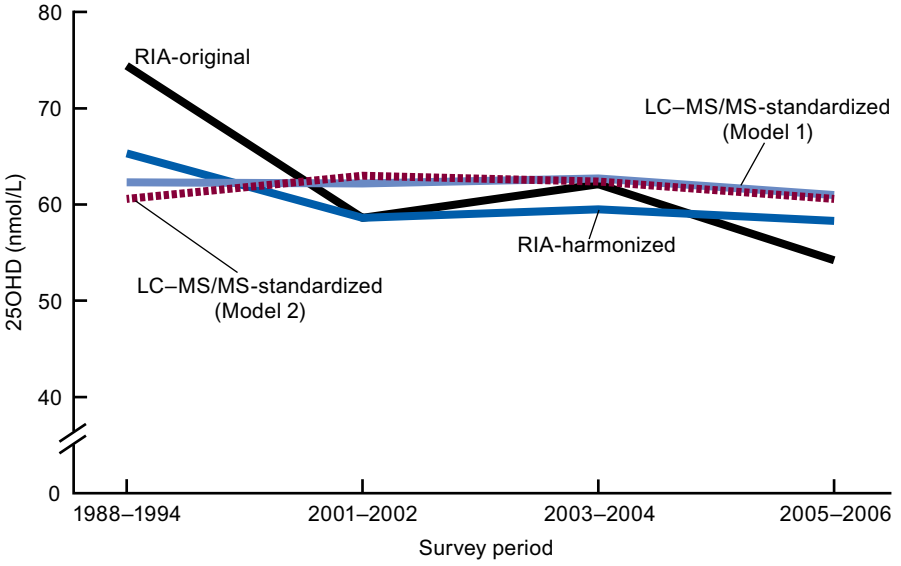


NOTES: LC–MS/MS is liquid chromatography–tandem mass spectrometry, and nmol/L is nanomole per liter. Two models were used to estimate LC–MS/MS-equivalent values from radioimmunoassay-original data: Regression Model 1 uses equations for each survey period, and Model 2 uses equations for each calendar year in which RIA data were tested.
SOURCE: CDC/NCHS, National Health and Nutrition Examination Survey, 1988–1994 and 2001–2006.

Figure 4. Estimated kernel density curves of LC–MS/MS-standardized serum 25-hydroxyvitamin D concentrations for regression models 1 and 2 for persons aged 12 years and over, stratified by survey period: National Health and Nutrition Examination Survey, 1988–1994 and 2001–2006

standardized LC–MS/MS method developed by CDC was used on a routine basis to measure 25OHD (15). Providing continuity when methods change is important for facilitating interpretation of the data. Standardized data are the best possible data to use for this purpose, because they are anchored to reference materials with specified limits on accuracy.

This report summarizes some of the technical details about how various models were evaluated for predicting LC–MS/MS-equivalent concentrations of 25OHD from RIA results. In the past, when a method was changed in NHANES, a single crossover study would typically be undertaken



NOTES: RIA is radioimmunoassay, LC–MS/MS is liquid chromatography–tandem mass spectrometry, 25OHD is 25-hydroxyvitamin D, and nmol/L is nanomole per liter. Values are weighted arithmetic means. RIA-original denotes original RIA results, and RIA-harmonized denotes results adjusted to minimize reformulation and lot-to-lot variability in the RIA kits. Regression Model 1 (using equations for each survey period) and Model 2 (using equations for each calendar year in which RIA data were tested) were used to generate LC–MS/MS-standardized data.
SOURCE: CDC/NCHS, National Health and Nutrition Examination Survey, 1988–1994 and 2001–2006.

Figure 5. Original radioimmunoassay, harmonized radioimmunoassay, and standardized LC-MS/MS serum 25-hydroxyvitamin D concentrations for persons aged 12 years and over, stratified by survey period: National Health and Nutrition Examination Survey, 1988–1994 and 2001–2006

and a regression equation would be derived to facilitate temporal trend comparisons using the old method. However, this crossover study was not as straightforward because the RIA kits that were used were not harmonized to produce comparable results over the 1988–2006 period. A number of different models were investigated, but the two models with the smallest rMSEP described in this report produced very similar results. One model used the survey periods (1988–1994, 2001–2002, 2003–2004, and 2005–2006) to derive separate equations; the other was based on the exact year of RIA measurement (1994, 1995, 1996, 2001, 2002, 2003, 2004, 2005, and 2006). Because the first model was not dependent on nonpublic-use data, this model was used to estimate vitamin D status for NHANES III and NHANES 2001–2006 data. These updated data were publicly released in 2015 with a corresponding Analytical Note (13).

Recognizing the limitations of using regression models to bridge the change in laboratory methods is important. First, these data are not actually measured LC–MS/MS values, but rather predicted from a regression model and, therefore, are inherently less variable than what would be expected had all the values actually been measured by LC–MS/MS. Secondly, the bridging data were not a random sample but were intentionally sampled across the full range of RIA concentrations and dates of RIA analyses. Therefore, the estimated relationship between LC–MS/MS and RIA using these models no longer reflects the relationship of these two variables in the underlying population; instead it resembles how laboratories typically perform method comparisons. Thirdly, regression is designed to predict the mean and is not optimal for the tails of the distribution, whereas the tails are important for estimating those at risk for deficiency or excess vitamin D intake. Fourthly, immunoassays are relatively nonspecific, often showing substantial positive and negative interferences, so the RIA data used for predicting LC–MS/MS-equivalent concentrations were less than ideal. Retesting all of the samples was not possible logistically because of the number involved (greater than 38,000).

Assay standardization is an essential process in the development of evidence-based clinical and public health guidelines for the assessment of vitamin D deficiency. Using a laboratory method that is traceable to international reference materials, a study was designed to standardize historical 25OHD RIA data from NHANES to LC–MS/MS-equivalents for 1988–2006. These predicted 25OHD data were compared with directly measured NHANES 2007–2010 LC–MS/MS 25OHD (3). Together, these data provide new insight into the long-term vitamin D status of the U.S. population.

References

1. Yetley EA, Pfeiffer CM, Schleicher RL, Phinney KW, Lacher DA, Christakos S, et al. NHANES monitoring of serum 25-hydroxyvitamin D: A roundtable summary. J Nutr 140(11):2030S–45S. 2010.
2. National Center for Health Statistics. About the National Health and Nutrition Examination Survey. Available from: http://www.cdc.gov/nchs/nhanes/about_nhanes.htm.
3. Schleicher RL, Sternberg MR, Lacher DA, Sempos CT, Looker AC, Durazo-Arvizu RA, et al. Temporal trends in the vitamin D status of the U.S. population from 1988 to 2010 using standardized serum concentrations of 25-hydroxyvitamin D. 2016 (unpublished report).
4. Pfeiffer C. Laboratory procedure manual: 25-hydroxyvitamin D in serum, NHANES 2003–2004. Atlanta, GA: National Center for Environmental Health. Available from: http://www.cdc.gov/nchs/data/nhanes/nhanes_03_04/l06vid_c_met_Vitamin_D.pdf.
5. Schleicher RL, Encisco SE, Chaudhary-Webb M, Paliakov E, McCoy LF, Pfeiffer CM. Isotope dilution ultra performance liquid chromatography–tandem mass spectrometry method for simultaneous measurement of 25-hydroxyvitamin D_2, 25-hydroxyvitamin D_3 and 3-epi-25-hydroxyvitamin D_3 in human serum. Clin Chim Acta 412(17–18):1594–9. 2011.
6. 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 243 Suppl:32–40. 2012.

7. Tai SS, Bedner M, Phinney KW. Development of a candidate reference measurement procedure for the determination of 25-hydroxyvitamin D_3 and 25-hydroxyvitamin D_2 in human serum using isotope-dilution liquid chromatography–tandem mass spectrometry. Anal Chem 82(5):1942–8. 2010.
8. Stepman HC, Vanderroost A, Van Uytfanghe K, Thienpont LM. Candidate reference measurement procedures for serum 25-hydroxyvitamin D_3 and 25-hydroxyvitamin D_2 by using isotope-dilution liquid chromatography–tandem mass spectrometry. Clin Chem 57(3):441–8. 2011.
9. Stöckl D, Sluss PM, Thienpont LM. Specifications for trueness and precision of a reference measurement system for serum/plasma 25-hydroxyvitamin D analysis. Clin Chim Acta 408(1–2):8–13. 2009.
10. Phinney KW, Bedner M, Tai SS, Vamathevan VV, Sander LC, Sharpless KE, et al. Development and certification of a standard reference material for vitamin D metabolites in human serum. Anal Chem 84(2):956–62. 2012.
11. Caudill SP, Schleicher RL, Pirkle JL. Multi-rule quality control for the age-related eye disease study. Stat Med 27(20):4094–106. 2008.
12. Encisco SE, Sternberg M, Chaudhary-Webb M, Paliakov E, Schleicher RL. Effects of sample matrices, extreme storage temperatures, and repeated freeze/thaw cycles on integrity of 25-hydroxyvitamin D. Clin Chem 57(S10):A91. 2011.
13. National Center for Health Statistics. National Health and Nutrition Examination Survey: Analytical note for 25-hydroxyvitamin D data analysis using NHANES III (1988–1994), NHANES 2001–2006, and NHANES 2007–2010 (October 2015). Available from: http://wwwn.cdc.gov/Nchs/Nhanes/VitaminD/AnalyticalNote.aspx?h=http://wwwn.cdc.gov/Nchs/Nhanes/2009-2010/VID_F.htm&t=VID_F%20Doc.
14. Cochran WG. Sampling techniques. 3rd ed. Hoboken, NJ: Wiley. 1977.
15. Schleicher RL, Looker AC, Yetley EA, Lacher DA, Sempos CT, Taylor CL, et al. Nationally representative estimates of serum 25-hydroxyvitamin D metabolite concentrations measured by liquid chromatography–tandem mass spectrometry in the U.S. population: National Health and Nutrition Examination Survey 2007–2010. J Nutr. In press 2016.

Table 1. Original radioimmunoassay, harmonized radioimmunoassay, and standardized LC-MS/MS serum 25-hydroxyvitamin D concentrations for persons aged 12 years and over, by demographic subgroup and survey period: National Health and Nutrition Examination Survey, 1988–1994 and 2001–2006

Group	1988–1994	2001–2002	2003–2004	2005–2006
Total[1]	nmol/L (confidence interval)			
RIA-original	74.4 (72.7–76.2)	58.6 (56.7–60.6)	62.1 (58.5–65.6)	54.2 (51.8–56.7)
RIA-harmonized	65.3 (63.8–66.8)	58.6 (56.7–60.6)	59.5 (56.7–62.2)	58.3 (55.4–61.1)
LC–MS/MS-standardized (Model 1)	62.3 (61.1–63.5)	62.2 (60.4–64.1)	62.7 (59.3–66.2)	61.0 (58.6–63.4)
LC–MS/MS-standardized (Model 2)	60.6 (59.3–61.9)	63.0 (60.6–65.4)	62.4 (59.9–65.0)	60.6 (57.6–63.7)
Age (years)				
12–19:				
RIA-original	80.0 (76.6–83.5)	59.4 (57.1–61.7)	63.3 (58.7–67.8)	55.2 (51.6–58.8)
RIA-harmonized	70.0 (67.1–73.0)	59.4 (57.1–61.7)	60.7 (56.7–64.7)	58.9 (55.1–62.8)
LC–MS/MS-standardized (Model 1)	66.2 (64.1–68.4)	63.0 (60.8–65.2)	63.9 (59.4–68.4)	61.9 (58.5–65.4)
LC–MS/MS-standardized (Model 2)	64.8 (62.2–67.4)	63.8 (61.1–66.6)	63.7 (59.8–67.7)	61.2 (57.1–65.3)
20–39:				
RIA-original	77.6 (75.2–80.1)	59.2 (56.9–61.4)	62.3 (58.3–66.3)	55.8 (52.7–59.0)
RIA-harmonized	68.0 (66.0–70.1)	59.2 (56.9–61.4)	59.6 (56.3–62.8)	59.8 (56.3–63.3)
LC–MS/MS-standardized (Model 1)	64.4 (62.8–66.0)	62.8 (60.6–64.9)	62.9 (59.0–66.9)	62.5 (59.5–65.6)
LC–MS/MS-standardized (Model 2)	63.0 (61.2–64.7)	63.4 (60.6–66.2)	62.7 (59.6–65.8)	62.1 (58.4–65.8)
40–59:				
RIA-original	71.1 (69.2–73.0)	58.8 (56.2–61.3)	61.5 (57.2–65.8)	53.3 (50.8–55.9)
RIA-harmonized	62.5 (60.9–64.1)	58.8 (56.2–61.3)	59.0 (55.7–62.3)	57.2 (54.2–60.2)
LC–MS/MS-standardized (Model 1)	60.1 (58.7–61.5)	62.4 (59.9–64.8)	62.2 (58.0–66.4)	60.1 (57.7–62.6)
LC–MS/MS-standardized (Model 2)	58.2 (56.8–59.6)	63.2 (60.4–66.0)	61.9 (58.7–65.0)	59.6 (56.3–62.9)
60 and over:				
RIA-original	68.6 (67.3–70.0)	56.7 (54.1–59.3)	61.8 (59.3–64.4)	52.6 (50.1–55.2)
RIA-harmonized	60.4 (59.3–61.5)	56.7 (54.1–59.3)	59.4 (57.3–61.5)	57.1 (54.4–59.8)
LC–MS/MS-standardized (Model 1)	58.4 (57.4–59.5)	60.4 (58.0–62.9)	62.5 (60.0–65.0)	59.4 (57.0–61.9)
LC–MS/MS-standardized (Model 2)	56.3 (55.3–57.3)	61.2 (58.2–64.2)	62.1 (60.4–63.8)	59.8 (56.8–62.7)
Sex				
Male:				
RIA-original	78.7 (76.7–80.6)	59.6 (57.6–61.7)	62.4 (58.7–66.1)	54.3 (52.1–56.5)
RIA-harmonized	68.9 (67.2–70.6)	59.6 (57.6–61.7)	59.8 (56.8–62.8)	58.3 (55.6–61.0)
LC–MS/MS-standardized (Model 1)	65.6 (64.3–66.9)	63.2 (61.3–65.1)	63.1 (59.4–66.7)	61.1 (58.9–63.2)
LC–MS/MS-standardized (Model 2)	63.8 (62.3–65.2)	64.1 (61.6–66.6)	62.8 (60.0–65.6)	60.7 (57.7–63.8)
Female:				
RIA-original	70.5 (68.6–72.4)	57.7 (55.3–60.0)	61.7 (58.3–65.1)	54.2 (51.3–57.1)
RIA-harmonized	62.0 (60.4–63.6)	57.7 (55.3–60.0)	59.2 (56.6–61.8)	58.2 (55.0–61.3)
LC–MS/MS-standardized (Model 1)	59.2 (57.9–60.6)	61.3 (59.1–63.5)	62.4 (59.0–65.8)	60.9 (58.2–63.7)
LC–MS/MS-standardized (Model 2)	57.7 (56.3–59.1)	62.0 (59.4–64.7)	62.1 (59.7–64.5)	60.5 (57.2–63.9)
Race and Hispanic origin				
Mexican American:				
RIA-original	63.7 (61.8–65.6)	51.1 (47.8–54.4)	53.4 (49.9–57.0)	44.1 (40.0–48.2)
RIA-harmonized	56.3 (54.7–57.8)	51.1 (47.8–54.4)	50.8 (47.9–53.8)	47.9 (43.3–52.6)
LC–MS/MS-standardized (Model 1)	54.7 (53.3–56.2)	55.0 (51.9–58.2)	54.2 (50.7–57.7)	51.2 (47.2–55.1)
LC–MS/MS-standardized (Model 2)	52.8 (51.5–54.1)	56.6 (52.9–60.3)	54.3 (51.0–57.6)	50.9 (46.1–55.8)
Non-Hispanic black:				
RIA-original	49.2 (47.1–51.3)	34.6 (33.3–35.8)	39.9 (36.6–43.1)	34.3 (32.1–36.6)
RIA-harmonized	44.0 (42.3–45.8)	34.6 (33.3–35.8)	38.8 (35.7–41.8)	36.4 (34.0–38.8)
LC–MS/MS-standardized (Model 1)	42.8 (41.1–44.6)	39.3 (38.2–40.5)	40.9 (37.7–44.1)	41.7 (39.5–43.9)
LC–MS/MS-standardized (Model 2)	41.8 (40.0–43.5)	38.2 (36.2–40.2)	41.2 (38.3–44.0)	38.8 (35.9–41.8)
Non-Hispanic white:				
RIA-original	80.3 (78.3–82.2)	64.0 (61.7–66.2)	67.8 (64.3–71.4)	59.6 (57.5–61.8)
RIA-harmonized	70.2 (68.6–71.9)	64.0 (61.7–66.2)	65.0 (62.7–67.3)	64.0 (61.9–66.1)
LC–MS/MS-standardized (Model 1)	66.7 (65.4–68.0)	67.3 (65.2–69.4)	68.4 (64.9–71.9)	66.2 (64.1–68.3)
LC–MS/MS-standardized (Model 2)	64.9 (63.5–66.4)	68.3 (65.8–70.7)	67.9 (65.6–70.3)	66.3 (63.9–68.7)

[1]"Other" race and Hispanic-origin group not shown but included in total estimates.

NOTES: Values are weighted arithmetic means (95% confidence interval) that are not age-standardized. LC–MS/MS is liquid chromatography–tandem mass spectrometry; nmol/L is nanomole per liter; and RIA is radioimmunoassay. RIA-original is original RIA results; RIA-harmonized is results adjusted to minimize reformulation and lot-to-lot variability in RIA kits. Regression Model 1 (using equations for each survey) and Model 2 (using equations for each calendar year in which RIA data were tested) were used to generate LC–MS/MS-standardized data.

Table 2. LC–MS/MS-equivalent tail percentiles for serum 25-hydroxyvitamin D for persons aged 12 years and over, by demographic subgroup and survey period: National Health and Nutrition Examination Survey, 1988–1994 and 2001–2006

Group	5th percentile				95th percentile			
	1988–1994	2001–2002	2003–2004	2005–2006	1988–1994	2001–2002	2003–2004	2005–2006
Total[1]	nmol/L (confidence interval)							
RIA-original	32.0 (30.8–33.1)	21.9 (20.4–23.6)	23.5 (20.2–26.2)	19.8 (18.4–21.2)	127 (124–133)	96.3 (92.7–103)	103 (97.7–112)	89.2 (86.5–93.0)
RIA-harmonized	29.6 (28.6–30.5)	21.9 (20.4–23.6)	23.5 (20.1–25.9)	21.3 (19.5–23.0)	110 (107–114)	96.3 (92.7–103)	97.2 (92.9–104)	95.3 (91.8–99.3)
LC–MS/MS-standardized (Model 1)	28.6 (27.6–29.5)	27.4 (25.8–28.9)	24.8 (21.6–27.5)	27.6 (26.2–29.0)	94.8 (93.6–96.2)	98.1 (94.7–104)	103 (97.8–112)	94.9 (92.3–98.6)
LC–MS/MS-standardized (Model 2)	28.7 (27.3–30.0)	27.1 (24.6–28.3)	26.1 (23.6–28.5)	25.0 (22.2–27.6)	99.5 (96.7–104)	101 (98.0–107)	102 (98.0–107)	97.4 (94.0–102)
Age (years)								
12–19:								
RIA-original	35.6 (32.2–39.3)	24.5 (20.8–27.2)	24.9 (21.2–28.0)	19.7 (17.4–21.8)	137 (129–152)	95.9 (91.7–102)	108 (98.5–123)	91.4 (84.8–102)
RIA-harmonized	32.6 (29.7–35.7)	24.5 (20.8–27.2)	24.8 (21.0–27.3)	21.2 (17.6–23.9)	118 (112–131)	95.9 (91.7–102)	101 (93.9–112)	96.1 (88.5–111)
LC–MS/MS-standardized (Model 1)	31.6 (28.7–34.7)	29.7 (26.3–32.3)	26.2 (22.5–29.2)	27.5 (25.2–29.5)	97.3 (95.3–102)	97.7 (93.8–104)	108 (98.5–123)	97.1 (90.6–107)
LC–MS/MS-standardized (Model 2)	31.7 (28.9–34.1)	28.3 (24.6–32.4)	26.9 (23.7–30.1)	25.0 (19.0–28.5)	107 (102–118)	101 (96.3–109)	104 (98.4–116)	97.8 (91.2–112)
20–39:								
RIA–original	32.4 (30.1–33.7)	21.2 (19.4–23.1)	22.6 (19.2–25.6)	19.9 (18.4–21.5)	135 (126–140)	102 (94.6–110)	109 (99.5–119)	94.9 (91.4–100)
RIA–harmonized	29.9 (28.0–31.0)	21.2 (19.4–23.1)	22.5 (19.0–25.4)	21.1 (19.0–23.3)	116 (109–121)	102 (94.6–110)	104 (95.7–110)	99.0 (95.3–112)
LC–MS/MS-standardized (Model 1)	28.9 (27.0–30.0)	26.7 (24.9–28.4)	23.9 (20.6–26.9)	27.7 (26.2–29.2)	96.8 (94.4–98.3)	103 (96.5–111)	109 (99.5–118)	100 (97–106)
LC–MS/MS-standardized (Model 2)	28.8 (26.8–30.5)	24.9 (22.3–27.8)	24.5 (21.4–27.6)	24.5 (20.2–28.3)	105 (99.3–109)	105 (101–113)	107 (101–115)	102 (97.4–113)
40–59:								
RIA–original	31.3 (29.5–32.6)	21.4 (18.7–24.4)	23.7 (18.8–27.4)	19.3 (17.1–21.2)	123 (117–126)	94.7 (90.4–104)	99.4 (92.3–115)	87.1 (84.9–91.3)
RIA–harmonized	28.9 (27.4–30.0)	21.4 (18.7–24.4)	23.6 (18.7–26.9)	20.8 (18.8–22.8)	106 (101–109)	94.7 (90.4–104)	94.3 (88.9–102)	93.9 (88.8–99.8)
LC–MS/MS-standardized (Model 1)	27.9 (26.4–29.0)	26.9 (24.2–29.7)	25.0 (20.2–28.7)	27.1 (25.0–29.0)	93.4 (91.9–94.5)	96.6 (92.5–105)	99.5 (92.4–115)	92.9 (90.7–96.9)
LC–MS/MS-standardized (Model 2)	27.9 (26.9–29.1)	27.0 (22.2–28.8)	26.2 (21.3–29.5)	24.5 (22.5–27.1)	95.1 (92.7–99.2)	101 (95.2–107)	98.2 (93.2–107)	96.3 (91.1–102)
60 and over:								
RIA–original	31.3 (29.8–32.6)	23.1 (21.1–24.4)	24.2 (21.8–26.1)	20.6 (18.6–22.7)	115 (113–119)	91.7 (88.3–94.8)	100 (95.1–106)	83.6 (79.9–89.3)
RIA–harmonized	29.0 (27.7–30.1)	23.1 (21.1–24.4)	24.2 (21.8–25.8)	22.3 (20.4–24.3)	99.3 (97.5–103)	91.7 (88.3–94.8)	93.9 (90.0–102)	90.3 (87.4–93.7)
LC–MS/MS-standardized (Model 1)	28.0 (26.7–29.1)	28.4 (26.5–29.6)	25.5 (23.1–27.4)	28.4 (26.4–30.4)	91.2 (90.6–92.3)	93.7 (90.5–96.7)	100 (95.2–106)	89.4 (85.9–95.0)
LC–MS/MS-standardized (Model 2)	28.2 (26.7–29.4)	27.1 (24.6–29.8)	26.4 (23.8–28.6)	26.7 (23.0–29.7)	90.8 (89.1–92.8)	95.9 (92.3–101)	98.0 (94.7–105)	92.5 (89.3–96.1)
Sex								
Male:								
RIA–original	36.1 (34.6–37.9)	25.2 (22.8–26.6)	27.2 (22.8–29.9)	21.5 (19.4–23.4)	130 (126–136)	95.2 (91.1–102)	96.8 (92.4–104)	86.3 (83.0–90.8)
RIA–harmonized	33.0 (31.8–34.5)	25.2 (22.8–26.6)	26.5 (22.6–29.0)	23.2 (20.8–25.2)	112 (108–118)	95.2 (91.1–102)	91.7 (88.5–97.5)	92.6 (87.6–98.5)
LC–MS/MS-standardized (Model 1)	32.0 (30.8–33.5)	30.4 (28.2–31.7)	28.5 (24.1–31.1)	29.2 (27.2–31.0)	95.5 (94.2–97.3)	97.1 (93.1–104)	96.9 (92.5–104)	92.1 (88.8–96.5)
LC–MS/MS-standardized (Model 2)	32.1 (30.7–33.4)	29.3 (27.3–31.5)	29.4 (25.6–32.0)	27.4 (24.2–29.2)	101 (98.3–106)	101 (95.8–108)	96.0 (93.0–101)	94.7 (89.6–101)
Female:								
RIA-original	29.4 (27.8–31.0)	20.0 (17.9–21.8)	20.8 (18.9–23.0)	18.7 (17.1–20.2)	124 (121–130)	98.3 (93.1–106)	110 (102–119)	92.4 (87.8–96.6)
RIA-harmonized	27.3 (26.0–28.7)	20.0 (17.9–21.8)	20.7 (18.8–23.0)	20.1 (18.2–21.9)	107 (105–112)	98.3 (93.1–106)	104 (97.5–110)	96.8 (92.9–107)
LC–MS/MS-standardized (Model 1)	26.3 (25.0–27.7)	25.5 (23.5–27.2)	22.1 (20.3–24.3)	26.5 (25.0–28.0)	93.9 (93.0–95.5)	100 (95.1–108)	110 (102–118)	98.0 (93.6–102)
LC–MS/MS-standardized (Model 2)	26.6 (25.1–27.9)	24.7 (21.5–27.2)	23.7 (21.2–25.3)	23.7 (20.7–26.7)	97.6 (94.4–102)	102 (98.3–111)	107 (103–114)	99.0 (94.7–109)

See footnotes at end of table.

Table 2. LC–MS/MS-equivalent tail percentiles for serum 25-hydroxyvitamin D for persons aged 12 years and over, by demographic subgroup and survey period: National Health and Nutrition Examination Survey, 1988–1994 and 2001–2006—Con.

Group	5th percentile				95th percentile			
	1988–1994	2001–2002	2003–2004	2005–2006	1988–1994	2001–2002	2003–2004	2005–2006
Race and Hispanic origin	nmol/L (confidence interval)							
Mexican American:								
RIA-original	30.3 (28.6–32.0)	22.2 (21.1–23.3)	21.8 (19.8–23.8)	17.9 (14.8–20.2)	104 (102–108)	81.8 (77.6–91.1)	85.0 (81.8–89.2)	70.4 (66.2–78.4)
RIA-harmonized	28.1 (26.7–29.6)	22.2 (21.1–23.3)	21.8 (19.8–23.8)	19.4 (17.3–21.1)	90.6 (88.4–93.8)	81.8 (77.6–91.1)	80.2 (76.7–83.8)	76.3 (71.4–84.6)
LC–MS/MS-standardized (Model 1)	27.1 (25.7–28.6)	27.6 (26.5–28.6)	23.1 (21.2–25.1)	25.7 (22.7–28.0)	88.3 (87.4–89.4)	84.3 (80.3–93.1)	85.3 (82.1–89.4)	76.7 (72.6–84.4)
LC–MS/MS–standardized (Model 2)	27.4 (25.2–28.9)	27.4 (24.8–29.8)	24.0 (21.6–26.1)	23.1 (18.7–26.2)	83.6 (81.0–86.1)	88.3 (83.8–97.8)	84.9 (81.7–88.7)	78.8 (73.4–113)
Non-Hispanic black:								
RIA-original	21.6 (20.1–22.7)	12.9 (9.28–14.0)	14.1 (12.6–15.4)	12.7 (11.6–13.3)	87.7 (85.5–90.9)	64.9 (62.3–67.5)	71.1 (65.5–79.0)	62.1 (58.9–69.6)
RIA-harmonized	20.8 (19.5–21.7)	12.9 (9.28–14.0)	13.9 (11.5–15.3)	13.7 (12.8–14.5)	76.5 (74.7–79.2)	64.9 (62.3–67.5)	67.4 (63.6–73.7)	65.4 (62.3–70.4)
LC–MS/MS-standardized (Model 1)	19.8 (18.5–20.7)	18.7 (15.3–19.8)	15.6 (14.1–16.8)	20.7 (19.6–21.3)	75.5 (73.7–78.2)	68.2 (65.7–70.7)	71.6 (66.1–79.4)	68.6 (65.5–75.9)
LC–MS/MS-standardized (Model 2)	20.2 (18.8–21.2)	16.8 (14.1–19.1)	16.2 (14.4–17.8)	15.9 (15.8–16.5)	70.6 (69.0–72.9)	68.8 (64.0–73.8)	71.7 (68.9–76.5)	68.5 (65.4–72.6)
Non-Hispanic white:								
RIA-original	38.1 (36.5–39.4)	29.3 (27.6–30.7)	30.4 (26.8–33.9)	27.0 (24.8–29.0)	133 (129–138)	101 (95.9–108)	110 (102–118)	93.5 (90.3–96.7)
RIA-harmonized	34.7 (33.3–35.8)	29.3 (27.6–30.7)	29.6 (27.5–32.0)	30.5 (26.7–31.8)	115 (111–119)	101 (95.9–108)	103 (97.3–109)	98.4 (95.7–106)
LC–MS/MS-standardized (Model 1)	33.7 (32.3–34.8)	34.4 (32.7–35.7)	31.6 (28.1–35.0)	34.5 (32.4–36.5)	96.3 (95.1–97.7)	103 (97.7–109)	109 (102–117)	99.1 (96.0–102)
LC–MS/MS-standardized (Model 2)	33.5 (32.1–34.9)	33.5 (31.6–35.4)	32.4 (29.0–36.1)	33.3 (31.6–35.0)	104 (100–108)	106 (101–111)	107 (102–115)	101 (97.6–107)

[1]"Other" race and Hispanic origin group not shown but included in total estimates.

NOTES: LC–MS/MS is liquid chromatography–tandem mass spectrometry; nmol/L is nanomole per liter; and RIA is radioimmunoassay. Values are weighted percentiles; data were standardized to LC–MS/MS-equivalents by using Model 1, for which regression equations were generated for each survey period (1988–1994, 2001–2002, 2003–2004, and 2005–2006), or Model 2, for which regression equations were generated for each calendar year in which RIA data were tested (1994, 1995, 1996, 2001, 2002, 2003, 2004, 2005, and 2006).

SOURCE: CDC/NCHS, National Health and Nutrition Examination Survey, 1988–1994 and 2001–2006.

Table 3. Prevalence of serum 25-hydroxyvitamin D concentrations less than 30 nmol/L from original radioimmunoassay, harmonized radioimmunoassay, and standardized LC-MS/MS concentrations for persons aged 12 years and over, by demographic subgroup and survey period: National Health and Nutrition Examination Survey, 1988–1994 and 2001–2006

Group	1988–1994	2001–2002	2003–2004	2005–2006
Total[1]		Percent (confidence interval)		
RIA-original	3.8 (3.2–4.5)	9.0 (7.3–11)	7.5 (5.3–10)	12 (9.2–15)
RIA-harmonized	5.2 (4.5–6.1)	9.0 (7.3–11)	8.1 (5.9–11)	10 (7.8–13)
LC–MS/MS-standardized (Model 1)	6.0 (5.2–6.9)	5.4 (4.1–7.0)	7.5 (5.3–10)	5.2 (3.8–6.9)
LC–MS/MS-standardized (Model 2)	5.9 (4.9–7.1)	7.0 (5.2–9.4)	7.5 (5.3–10)	7.5 (5.4–10)
Age (years)				
12–19:				
RIA-original	2.8 (2.0–4.0)	7.0 (4.4–11)	6.5 (4.2–9.8)	12 (8.1–17)
RIA-harmonized	3.9 (2.9–5.2)	7.0 (4.4–11)	7.0 (4.6–11)	10 (6.8–14)
LC–MS/MS-standardized (Model 1)	4.2 (3.2–5.4)	4.1 (2.6–6.4)	6.5 (4.2–9.8)	5.3 (3.3–8.2)
LC–MS/MS-standardized (Model 2)	4.2 (3.1–5.6)	5.7 (3.2–9.7)	6.5 (4.2–9.8)	7.8 (4.7–13)
20–39:				
RIA-original	3.7 (2.8–4.8)	9.5 (7.7–12)	8.2 (5.7–12)	12 (8.7–16)
RIA-harmonized	5.0 (3.9–6.4)	9.5 (7.7–12)	9.2 (6.6–13)	9.6 (6.9–13)
LC–MS/MS-standardized (Model 1)	6.0 (5.0–7.2)	6.0 (4.6–7.7)	8.2 (5.7–12)	5.1 (3.7–7.1)
LC–MS/MS-standardized (Model 2)	5.7 (4.3–7.4)	7.6 (5.6–10)	8.2 (5.7–12)	7.5 (5.0–11)
40–59:				
RIA-original	4.2 (3.4–5.1)	9.4 (7.2–12)	7.1 (4.5–11)	12 (9.2–16)
RIA-harmonized	5.8 (4.9–6.8)	9.4 (7.2–12)	7.4 (4.8–11)	11 (8.3–14)
LC–MS/MS-standardized (Model 1)	6.6 (5.6–7.9)	5.7 (4.0–8.1)	7.1 (4.5–11)	5.5 (3.9–7.8)
LC–MS/MS-standardized (Model 2)	6.5 (5.4–7.8)	7.3 (5.3–10)	7.1 (4.5–11)	8.2 (6.0–11)
60 and over:				
RIA-original	4.1 (3.3–5.0)	8.8 (6.7–11)	7.6 (5.6–10)	11 (8.6–14)
RIA-harmonized	5.7 (4.8–6.7)	8.8 (6.7–11)	8.3 (6.1–11)	9.3 (7.2–12)
LC–MS/MS-standardized (Model 1)	6.4 (5.5–7.5)	4.4 (3.3–5.9)	7.6 (5.6–10)	4.6 (3.3–6.4)
LC–MS/MS-standardized (Model 2)	6.6 (5.4–8.1)	6.2 (4.2–9.1)	7.6 (5.6–10)	6.3 (4.5–8.8)
Sex				
Male:				
RIA-original	2.3 (1.9–2.7)	6.8 (5.4–8.5)	5.1 (3.2–8.1)	10 (7.5–13)
RIA-harmonized	3.2 (2.7–3.8)	6.8 (5.4–8.5)	5.7 (3.8–8.7)	8.6 (6.5–11)
LC–MS/MS-standardized (Model 1)	3.7 (3.1–4.3)	3.8 (3.0–4.7)	5.1 (3.2–8.1)	4.1 (2.9–5.6)
LC–MS/MS-standardized (Model 2)	3.6 (2.9–4.5)	5.1 (3.7–7.1)	5.1 (3.2–8.1)	6.2 (4.3–8.8)
Female:				
RIA-original	5.2 (4.2–6.3)	11 (8.8–14)	9.7 (7.2–13)	13 (10–17)
RIA-harmonized	7.1 (6.0–8.4)	11 (8.8–14)	10.4 (7.8–14)	12 (8.7–15)
LC–MS/MS-standardized (Model 1)	8.2 (7.1–9.5)	6.9 (4.9–9.5)	9.7 (7.2–13)	6.2 (4.5–8.5)
LC–MS/MS-standardized (Model 2)	8.0 (6.6–9.7)	8.8 (6.4–12)	9.7 (7.2–13)	8.8 (6.2–12)
Race and Hispanic origin				
Mexican American:				
RIA-original	4.6 (3.7–5.8)	10 (7.7–13)	9.3 (6.5–13)	20 (14–27)
RIA-harmonized	6.4 (5.2–7.9)	10 (7.7–13)	11 (8.1–15)	17 (11–24)
LC–MS/MS-standardized (Model 1)	7.3 (6.1–8.8)	5.3 (3.8–7.2)	9.3 (6.5–13)	7.7 (4.3–13)
LC–MS/MS-standardized (Model 2)	7.5 (5.8–9.6)	6.2 (4.0–9.5)	9.3 (6.5–13)	11 (6.8–18)
Non-Hispanic black:				
RIA-original	17 (14–20)	42 (39–46)	30 (23–37)	43 (37–49)
RIA-harmonized	22 (19–26)	42 (39–46)	31 (24–38)	40 (35–46)
LC–MS/MS-standardized (Model 1)	24 (21–28)	28 (25–33)	30 (23–37)	22 (18–27)
LC–MS/MS-standardized (Model 2)	24 (20–29)	36 (30–42)	30 (23–37)	34 (27–42)
Non-Hispanic white:				
RIA-original	1.7 (1.3–2.1)	3.9 (3.0–5.1)	3.5 (2.5–4.9)	5.3 (3.8–7.3)
RIA-harmonized	2.4 (1.9–3.0)	3.9 (3.0–5.1)	3.8 (2.8–5.2)	4.3 (3.3–5.5)
LC–MS/MS-standardized (Model 1)	2.8 (2.4–3.5)	2.2 (1.6–2.9)	3.5 (2.5–4.9)	2.1 (1.6–2.8)
LC–MS/MS-standardized (Model 2)	2.8 (2.2–3.6)	2.8 (2.1–3.8)	3.5 (2.5–4.9)	2.8 (2.1–3.6)

[1]"Other" race and Hispanic-origin group not shown but included in total estimates.

NOTES: LC–MS/MS is liquid chromatography–tandem mass spectrometry; nmol/L is nanomole per liter; and RIA is radioimmunoassay. Values are weighted prevalences (95% confidence interval) that were not age-standardized. RIA-original is original RIA results; RIA-harmonized is results adjusted to minimize reformulation and lot-to-lot variability in RIA kits. Regression Model 1 (based on survey period) and Model 2 (based on the year in which the original RIA data were generated) were used to generate LC–MS/MS-standardized data.

SOURCE: CDC/NCHS, National Health and Nutrition Examination Survey, 1988–1994 and 2001–2006.

U.S. DEPARTMENT OF HEALTH & HUMAN SERVICES

Centers for Disease Control and Prevention
National Center for Health Statistics
3311 Toledo Road, Room 5419
Hyattsville, MD 20782–2064

For more NCHS NHSRs, visit:
http://www.cdc.gov/nchs/products/nhsr.htm.

National Health Statistics Reports ■ Number 93 ■ April 25, 2016

Acknowledgments

The authors gratefully acknowledge the contribution of laboratory members Sara Encisco and Paul Yoon of CDC's National Center for Environmental Health. We also thank Dr. Paul Coates, NIH Office of Dietary Supplements, for his valuable scientific input and support.

Suggested citation

Schleicher RL, Sternberg MR, Lacher DA, et. al. A method-bridging study for serum 25-hydroxyvitamin D to standardize historical radioimmunoassay data to liquid chromatography–tandem mass spectrometry. National health statistics reports; no 93. Hyattsville, MD: National Center for Health Statistics. 2016.

For e-mail updates on NCHS publication releases, subscribe online at: http://www.cdc.gov/nchs/govdelivery.htm.
For questions or general information about NCHS: Tel: 1–800–CDC–INFO (1–800–232–4636) • TTY: 1–888–232–6348
Internet: http://www.cdc.gov/nchs • Online request form: http://www.cdc.gov/info
DHHS Publication No. 2016–1250 • CS263882