



Published in final edited form as:

Epidemiology. 2018 May ; 29(3): 448–452. doi:10.1097/EDE.0000000000000810.

The Effects of Long-Term Storage on Commonly Measured Serum Analyte Levels

Cynthia Kleeberger^a, David Shore^b, Elaine Gunter^c, Dale P. Sandler^d, and Clarice R. Weinberg^e

^aSocial & Scientific Systems, Inc., Durham, NC

^bWestat, Durham, NC

^cSpecimen Solutions, LLC, Tucker, GA

^eEpidemiology Branch, National Institute of Environmental Health Sciences, Research Triangle Park, NC

^eBiostatistics and Computational Biology Branch, National Institute of Environmental Health Sciences, Research Triangle Park, NC

Abstract

Background—Cohort studies typically bank biospecimens for many years prior to assay and investigators do not know whether levels of analytes have degraded.

Methods—We collected control samples from 22 non-study participants using the same enrollment criteria and specimen collection, processing, and storage protocols as The Sister Study. Serum samples were assayed for 21 analytes at collection and 6 years later. For each sample, the difference between the result at baseline and at 6 years was calculated for each analyte.

Results—Some of the analytes experienced a marked decrease in concentration after six years of frozen storage in liquid nitrogen vapor, compared to their baseline value. The confidence interval for the mean paired difference excluded 0 for 8 of the 21 analytes tested (aspartate transaminase, total cholesterol, estradiol, glucose, HDL cholesterol, luteinizing hormone, protein, and triglycerides). Two analytes, lactate dehydrogenase and sex hormone binding globulin, increased substantially in concentration over time (confidence interval excluded 0). For compounds substantially affected by storage time, the internal laboratory control variance was greater than the estimated mean percent change for HDL cholesterol and luteinizing hormone, indicating that extent of degradation for these analytes did not exceed technical variation.

Conclusions—Despite evidence for systematic changes over long-term storage, correlations between baseline and later measures were high with little relation between size of the correlation and estimated mean difference across time points. QC experiments to assess the impact of long-term storage on anticipated analytes of interest are important in planning cohort studies with banked samples.

Correspondence: Cynthia Kleeberger, 1007 Slater Road, Durham, NC 27703. ckleeberger@s-3.com.

Conflicts of Interest

The authors report no conflicts of interest.

Keywords

blood serum; cohort analysis; validation studies

Introduction

Long-term cohort studies often bank specimens for many years prior to assay, but effects of long-term storage on serum analytes are largely unknown. Even when temperatures are optimal, degradation can produce measurement errors.¹ Such effects could bias estimation and produce empirical confounding with other factors under study that change with follow-up time and are associated with storage time.²

We assessed long-term effects of storage time on selected analytes, using sampling and storage protocols designed for The Sister Study, a nationwide cohort study underway at the National Institute of Environmental Health Sciences (NIEHS). At time of enrollment, between 2003 and 2009, more than 50,000 participants provided blood and urine specimens, which were aliquoted into separate containers and frozen in vapor-phase liquid nitrogen for long-term storage and later analysis.

In a parallel quality control (QC) effort, we collected and similarly stored samples from 22 women who were not in The Sister Study. We wanted to assess impact of storage and handling on assayed levels of analytes to identify those for which the assay is sufficiently reproducible and stable over time. That way, differences in levels among individuals in the cohort would be detectable and meaningful.

Materials and Methods

Specimen Collection Methods

Mimicking Sister Study specimen collection and processing, 22 non-pregnant adult women without breast cancer each donated four 10.0-mL red top Vacutainer® (B-D Life Sciences, Inc.) tubes of whole blood, which were allowed to clot for 30 minutes and then centrifuged in the laboratory for 15 min at $1500 \times g$ to isolate serum in a consistent manner to minimize pre-analytic variability. Resulting serum was aliquotted into one 4.5-mL Nunc® cryovial and approximately sixteen 0.5-mL CryoBioSystem® (CBS) straws per donor.

Cryovials and straws were stored frozen overnight at -80°C . The next morning, the cryovials were thawed rapidly in a 37°C water bath with slight agitation. They were removed from the water bath when a small amount of ice crystals remained and shipped by FedEx Overnight to Quest Diagnostics, Inc. at refrigerated temperatures using frozen cold packs. Specimens were analyzed for designated analytes on the day of arrival.

This sample collection was designed to analyze degradation over time for clinical analytes that may be important in epidemiologic analyses. We selected the following 21 analytes: total cholesterol, HDL cholesterol, triglycerides, creatinine, alanine transaminase (ALT), aspartate transaminase (AST), sodium, potassium, chloride, blood urea nitrogen, glucose, phosphorus, protein, albumin, lactate dehydrogenase (LDH), calcium, luteinizing hormone

(LH), C-reactive protein (CRP), cortisol, estradiol, and sex hormone binding globulin (SHBG). Baseline assays were conducted on all 21 analytes after a single freeze/thaw cycle, i.e. upon arrival at Quest Diagnostics.

Straws were stored in vapor phase liquid nitrogen tanks (-180°C) for 6 years and then thawed in a 37°C water bath as above. Storage temperatures were continuously monitored over the 6-year period using two thermocouples and one resistance temperature detector sensor. There was no indication of temperatures drifting out of range in the liquid nitrogen tanks where these straws were stored. The straws were stored in two different tanks and seven different goblet containers. All samples were stored well below the opening of the tank. The same serum volume per participant as baseline (4.5 mL) was shipped with cold packs to the same lab and analyzed using the same laboratory assay protocols used six years earlier on baseline samples.

To evaluate laboratory analytical performance, we formulated two serum QC pool materials for insertion into each batch. Each pool was formed by mixing equal aliquots of serum from over 20 anonymous donors to produce a standard serum, providing for comparison of results across batches and estimation of inter-batch and intra-batch variability. Separate pools were created for pre-menopausal (21 donors) and post-menopausal women (23 donors), aliquoted into 0.5-mL CBS straws and stored in VP-LN₂ along with other Sister Study samples. Prior to shipping to Quest Diagnostics for testing, straws were rapidly thawed in a 37°C water bath with slight agitation, as above. Contents of the straws were decanted into cryovials for insertion into each of the six testing batches (Figure). Each batch contained two pre-menopausal and two post-menopausal pooled serum samples. As at baseline, samples were shipped overnight to Quest Diagnostics using frozen cold packs.

Analyte testing was run on four different platforms/instruments at Quest Diagnostics. There was no difference in instrumentation between baseline and six-year post-baseline testing. However, given that time-points were six years apart, reagent lots used for calibration were necessarily different. Equipment used was as follows:

- Beckman Coulter AU5400®: Lipid panel (cholesterol, triglycerides, HDL), albumin calcium, chloride, creatinine, glucose, LDH, phosphorus, potassium, total protein, sodium, AST, ALT, urea nitrogen.
- Siemens Centaur XP®: LH, cortisol, estradiol
- Dade Behring BNII Nephelometer®: CRP
- Immunlite 2000 Immunoassay®: SHBG

This study was approved by the NIEHS IRB and Copernicus Group IRB. Verbal informed consent was obtained from all subjects.

Statistical Methods

When pooled QC replicates were used to assess batch effects, graphs suggested that batch four was quite dissimilar to the other five batches. T statistics comparing batch four to other batches combined showed a marked difference in means for 17 of the 21 analytes. Furthermore, using the pooled replicates for each analyte, we took each batch in turn and

compared it to the other five batches combined. We calculated a z-score as $Z = (\text{TestBatchMean} - \text{CombinedBatchesMean}) / \text{CombinedBatchesSE}$. The batch 4 z-score was typically an order of magnitude larger than the other z-scores. This provided evidence that the exclusion of batch 4 was warranted and, therefore, was omitted from the remaining analyses. The Z score test results have been included as supplemental digital content (eTable).

When levels were below the limit of detection (LOD), the value $\text{LOD}/\text{SQRT}(2)$ was substituted. In baseline testing, there were six determinations below the LOD for CRP (<1 mg/L) and one that was below the LOD for both estradiol (<7 pg/ml) and CRP. In the 6-years-later testing, two measurements fell below the LOD for CRP LOD (<1 mg/L), four below LOD for estradiol (<15 pg/ml) and five below LOD for both.

For individual samples, we calculated the difference between the 6-year result and the baseline result for each analyte. The difference, $Y_{ij} = \text{result}_{6\text{ year}} - \text{result}_{\text{baseline}}$, was modeled using a mixed-effects model, with random effects for batch (b_i) and a fixed effect for the overall mean difference: $Y_{ij} = u + b_i + e_{ij}$, where, $b_i \sim N(0, s^2_{\text{batch}})$, $i=1,2,3,5,6$.

This model allowed estimation of the mean difference, i.e. u , and associated 95% confidence interval (CI) based on equating $u^{\text{hat}} / \text{se}(u^{\text{hat}})$ to the upper 97.5% percentile and the lower 2.5% percentile for $t_{4\text{df}}$ where u^{hat} is the estimate of u . Note that negative values of the difference measure suggest degradation over storage time. We used SAS® Proc Mixed Type 3 estimators, $\text{se}(u^{\text{hat}}) = [\text{MS}(\text{batch})/18]^{1/2}$. We considered including a fixed effect for menopause status but found it had negligible impact on assessment of the mean difference.

Results

Results from 18 individuals are presented in the table below for each analyte and include estimate, standard error, 95% CI for mean difference (later minus baseline), baseline mean, percent change of the means, internal lab coefficient of variance data (% CV), and Spearman Rho correlation. The column for % CV is based on internal laboratory cumulative control data collected from multiple runs for each specific analyte using the same instrumentation and during the same period as the analyses of our specimens. Control data are available separately for baseline and follow-up but have been aggregated as an average of the two for purpose of reporting in the table.

Results indicate that some analytes experienced a marked decrease in concentration after six years of frozen storage, compared to their baseline value (Table). The estimated mean difference was found to have a 95% CI that excluded zero for 9 of 21 analytes tested (AST, total cholesterol, estradiol, glucose, HDL cholesterol, LH, protein, sodium, and triglycerides.) However, the decrease in sodium may be due to assay drift at Quest Diagnostics where evidence strongly suggested that Beckman Coulter sodium testing under-recovered National Institute of Standards and Technology (NIST) standards by 2.5 mmol/L (JM Konopka, written communication, December 2016).

Two analytes, LDH and SHBG, increased substantially in concentration over time and had a 95% CI that excluded zero. Ten analytes showed little change.

For compounds that were substantially affected by storage time, the internal laboratory control variance was greater than the estimated change in the analyte for HDL cholesterol and LH, indicating that degradation estimated from the model for these analytes was within assay variability. Furthermore, percent CV was similar to percent change for protein and triglycerides. We conclude that the strongest evidence of systematic degradation was found in four analytes - AST, total cholesterol, estradiol, and glucose (with possibly systematic differences in protein and triglycerides). However, correlations between baseline and time-lagged measures were largely quite good (Table) and there was little correspondence between size of the correlation and the standardized difference in means for the two timepoints.

Discussion

We conducted this study to investigate effects of long-term storage on serum samples that are analyzed long after collection. This information is important to The Sister Study and other similar epidemiological studies that store samples for many years prior to analysis.³

There is minimal literature quantifying analyte degradation after very long periods of storage. Even when literature is available, results are not always consistent. We know of no directly comparable studies that evaluated the effect of storing serum at -180°C for as long as six years. Literature is especially sparse for long storage in VP-LN₂. Consequently, the current analysis provides useful supplemental data for cohort studies storing samples for long-term periods. High correlations between time periods, together with the fact that there was little relation between size of the correlation and difference in means, suggest that for analytes found to degrade over time in this study, the extent of degradation as a function of storage time could be adjusted for in epidemiological models, to correct for some of the resulting measurement error.

Similar results were found in other studies for cholesterol, triglycerides, and HDL where serum levels decreased after storage at -70°C for up to 7 years.⁴ The change in HDL was not large in our data but consistent with findings from another study.⁵ Little comparable literature was found for estradiol, but what was found indicated no degradation after three years at -80°C , in contrast with the large decline we saw after six years.⁶ AST was found to be stable for 1 year in another study⁷, whereas we observed a loss of almost 7% after six years. In this study, we found a small reduction in levels of potassium. Results for potassium stability are inconsistent across the literature.^{7, 8} Our results indicate little loss of CRP, which is consistent with another study.⁹

We found a decrease in glucose after six years, which may be due to the fact that our bloods were not drawn into a collection tube with preservative, causing degradation after thawing for the second measurement.¹⁰ Although the CI for the adjusted difference in glucose between baseline and six years excluded zero, percent loss over time was low and correlation with baseline level was high.

The increase found in LDH and SHBG may possibly be related to molecular binding alterations. For example, with LDH, rupture of a molecular bond releasing the enzyme from a bound inactive form may be possible.^{11, 12}

Although some readers may find the high correlations between baseline levels and later-measured levels to be reassuring, we caution that systematic reductions or elevations in levels can nevertheless disturb the relative rankings of values and can also cause misclassification when using percentile-based categories, such as quartiles. Also, if the analyte is dichotomized according to clinical standards based on normal range, as is often done for cholesterol or fasting glucose, these assignments can become storage-time-dependent. Finally, if a case-cohort analysis is planned, and analytes for the random subcohort are analyzed at baseline, while incident cases are analyzed later, serious bias can result.

Many pre-analytical factors influence sample degradation, not just storage time. Specimen collection and handling, serum separation from cells¹³, time to freezing, storage temperature, and thawing technique can all play critical roles.^{14,15} For analytes that were found to degrade, we are unable to determine when the degradation occurred and whether or not those effects are linear over time. Additional time points would be valuable for future studies of this kind. Although the testing laboratory provided QC variance data, we acknowledge that the laboratory did not have long-term in-house QC pools, nor did we have data from standard reference materials such as NIST samples in our batches, to establish accuracy or analytical comparability over time.

We conclude that there are differences in assay results after long-term frozen storage although for most analytes correlation coefficients were high and percent change small. Our results are relevant to biobanks or cohort studies that bank samples and for which built-in QC experiments could assess impact of long-term storage on anticipated analytes of interest. QC materials can be from non-study participants, as in this analysis, or from study participants. For cohorts and samples that exist for many years, it is often difficult to predict analytes that will become relevant. Results of such assessments would be useful to others for evaluating potential impact of measurement error from storage times on study findings.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

Source of Funding

Supported by the Intramural Research Program of the National Institutes of Health, National Institute of Environmental Health Sciences.

The authors thank Quest Diagnostics, Inc. for testing the samples, providing the results, and providing quality control variance data. We thank Kristie Dantzer for laboratory coordination of sample collection, storage, transfer, and returned results.

References

1. Hubel A, Spindler R, Skubitz APN. Storage of human biospecimens: selection of the optimal storage temperature. *Biopreservation and Biobanking*. 2014; 12:165–175. [PubMed: 24918763]
2. Kugler K, Hackl W, Mueller L, Fiegl H, Graber A, Pfeiffer R. The impact of sample storage time on estimates of association in biomarker discovery studies. *J Clin Bioinform*. 2011; 1:1–8.
3. Baker M. Biorepositories: Building better biobanks. *Nature*. 2012; 486:141–146. [PubMed: 22678297]
4. Shih WJ, Bachorik PS, Haga JA, Myers GL, Stein EA. Estimating the long-term effects of storage at –70C on cholesterol, triglyceride, and HDL-cholesterol measurements in stored sera. *Clinical Chemistry*. 2008; 46(3):351–364.
5. Bausserman LL, Saritelli AL, Milosavljevic D. High-density lipoprotein subfractions measured in stored serum. *Clinical Chemistry*. 1994; 40(9):1713–1716. [PubMed: 8070080]
6. Bolelli G, Muti P, Micheli A, Sciajno R, Franceschetti F, Krogh V, Pisani P, Berrino F. Validity for epidemiological studies of long-term cryoconservation of steroid and protein hormones in serum and plasma. *Cancer Epidemiology Biomarkers and Prevention*. 1995; 4(5):509–513.
7. Brinc D, Chan MK, Venner A, et al. Long-term stability of biochemical markers in pediatric serum specimens stored at –80 C: A CALIPER Substudy. *Clinical Biochemistry*. 2012; 45(10):816–826. [PubMed: 22510430]
8. Dimagno EP, Corle D, O'Brien JF, Masnyk IJ, Go VLW, Aamodt R. Effect of long-term freezer storage, thawing, and refreezing on selected constituents of serum. *Mayo Clinic Proceedings*. 1989; 64(10):1226–1234. [PubMed: 2593713]
9. Doumatey AP, Zhou J, Adeyemo A, Rotimi C. High sensitivity C-reactive protein (Hs-CRP) remains highly stable in long-term archived serum. *Clinical Biochemistry*. 2014; 47(4):315–318. [PubMed: 24373927]
10. Bruns D, Knowler W. Stabilization of glucose in blood samples, why it matters. *Clin Chem*. 2009; 55(5):850–852. [PubMed: 19282352]
11. Geisler P, Iossifides I, Eichman M. 1964 Preservation of the lactic dehydrogenase activity of platelets by freezing in dimethylsulfoxide and plasma. *Blood*. 1964; 24(6):761–764. [PubMed: 14235358]
12. Holl K, Lundin E, Kaasila M, et al. Effect of long-term storage on hormone measurements in samples from pregnant women: The experience of the Finnish Maternity Cohort. *Acta Oncol*. 2008; 47(3):406–412. [PubMed: 17891670]
13. Boyanton BL Jr, Blick KE. Stability studies of twenty-four analytes in human plasma and serum. *Clinical Chemistry*. 2002; 48(12):2242–2247. [PubMed: 12446483]
14. Hubel A, Aksan A, Skubitz AP, Wendt C, Zhong X. State of the art in preservation of fluid biospecimens. *Biopreserv Biobank*. 2011; 9:237–244. [PubMed: 24850337]
15. Betsou F, Gunter E, Clements J, et al. Identification of evidence-based biospecimen quality-control tools: A report of the International Society for Biological and Environmental Repositories (ISBER) Biospecimen Science Working Group. *J Mol Diagn*. 2013; 15:3–16. [PubMed: 23195791]

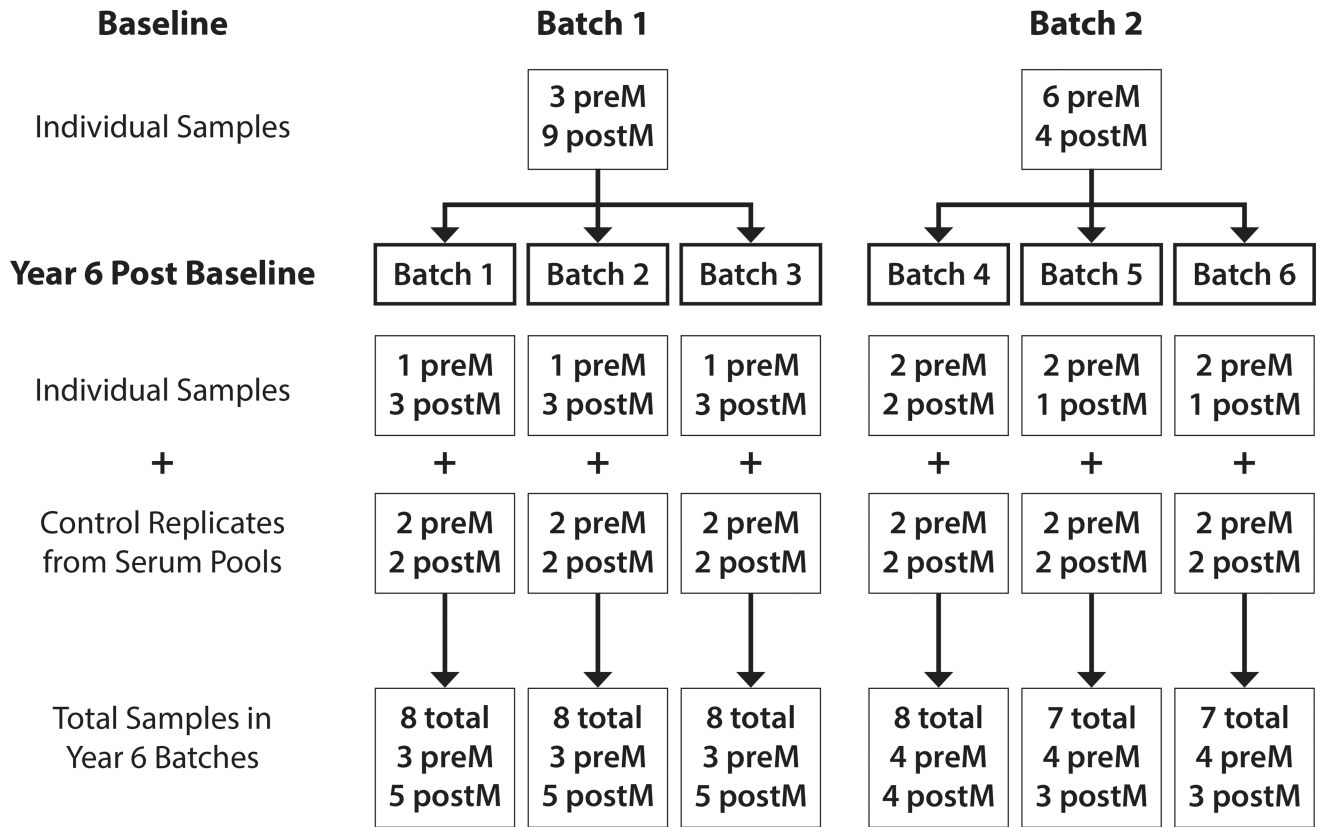


FIGURE 1. Design of Batches

Schematic showing the study design, including 22 specimens from 22 women, 13 of whom were post-menopausal, assayed in two batches at baseline and in 6 batches 6 years later, each of which included 2 replicates of the pre-menopausal pool and 2 of the post-menopausal pool. The control pools were used to assess batch effects.

Change in serum analytes over time (N=18 individuals per analyte)

TABLE

Serum Analyte	Units	Estimate	SE	95% CI	Baseline Mean	% Change	Lab Variance (% CV)	Spearman Rho Correlation
ALBUMIN	g/dL	0.06	0.05	(-0.09, 0.20)	4.34	1.0	1.9	0.93
ALT	U/L	-0.36	0.27	(-1.09, 0.38)	20.06	-1.7	4.9	0.98
AST	U/L	-1.46	0.41	(-2.59, -0.33)	22.28	-6.9	3.9	0.92
C-REACTIVE PROTEIN	mg/L	-0.39	0.16	(-0.83, 0.05)	3.66	-11.9	8.7	0.99
CALCIUM	mg/dL	0.07	0.08	(-0.14, 0.28)	9.26	0.9	1.5	0.84
CHLORIDE	mmol/L	-1.56	0.76	(-3.66, 0.54)	104.33	-1.6	1.0	0.65
CHOLESTEROL	mg/dL	-6.38	2.00	(-11.9, -0.84)	226.89	-3	2.0	0.98
CORTISOL, TOTAL	mcg/dL	-0.93	0.51	(-2.35, 0.49)	13.79	-7.6	5.5	0.95
CREATININE	mg/dL	-0.02	0.02	(-0.08, 0.04)	0.77	-2.6	3.3	0.82
ESTRADIOL	pg/mL	-23.67	6.66	(-42.1, -5.20)	83.33	-38.8	8.7	0.89
GLUCOSE	mg/dL	-2.93	0.70	(-4.88, -0.99)	80.89	-3.8	2.0	0.92
HDL CHOLESTEROL	mg/dL	-2.17	0.58	(-3.78, -0.57)	62.22	-3.6	4.6	0.99
LDH	U/L	17.54	2.29	(11.2, 23.9)	161.22	9.8	3.4	0.87
LH	mIU/mL	-0.75	0.25	(-1.44, -0.07)	24.76	-3.2	4.9	0.99
PHOSPHATE	mg/dL	-0.06	0.05	(-0.19, 0.07)	3.39	-1.8	2.3	0.97
POTASSIUM	mmol/L	-0.09	0.03	(-0.17, -0.00)	4.16	-2.2	1.3	0.96
PROTEIN, TOTAL	g/dL	-0.15	0.05	(-0.28, -0.02)	7.17	-2.2	1.8	0.98
SHBG	nmol/L	5.56	1.06	(2.63, 8.50)	65.72	7.8	5.6	0.99
SODIUM	mmol/L	-3.45	0.83	(-5.76, -1.14) ^a	141.72	-2.5 ^a	0.9	0.66
TRIGLYCERIDE	mg/dL	-2.53	0.66	(-4.37, -0.69)	105.33	-2.5	1.9	0.99
UREA NITROGEN	mg/dL	-0.06	0.13	(-0.41, 0.30)	14.17	-0.4	2.7	0.99

ALT (alanine transaminase), AST (aspartate transaminase), HDL (high-density lipoprotein), LDH (lactate dehydrogenase), LH (luteinizing hormone), SHBG (sex hormone binding globulin), SE (standard error), CI (confidence interval), CV (coefficient of variation)

^aDecrease in sodium may be due to assay drift at Quest Diagnostics