Published in final edited form as:

Ann Clin Biochem. 2010 May ; 47(Pt 3): 233–241. doi:10.1258/acb.2010.009112.

Validity of estimating non-SHBG bound testosterone and estradiol from total hormone measurements in boys and girls

Brian L. Egleston¹, D. Walt Chandler², and Joanne F. Dorgan¹

Brian L. Egleston: Brian.Egleston@fccc.edu

¹ Fox Chase Cancer Center, 333 Cottman Ave., Philadelphia, PA 19130, phone: 215-214-3917, fax: 215-728-2553

² Endocrine Sciences/Esoterix, Division of LabCorp, 4301 Lost Hills Road, Calabasas Hills, CA 91301

Abstract

Background—Circulating levels of bioavailable estradiol and testosterone are often desirable for clinical practice or investigational studies of children. However, assays to measure circulating hormone levels might not always be accessible. We sought to validate the empirical calculation of circulating bioavailable testosterone and estradiol in children.

Methods—663 eight to ten year olds were recruited to the Dietary Intervention Study in Children (DISC). DISC was a randomized clinical trial designed to test efficacy of a dietary intervention to reduce serum cholesterol (LDL-C) in children with elevated cholesterol. Assay measures of estradiol, testosterone, sex hormone-binding globulin concentration (SHBG), and albumin concentration in girls as well as dihydrotestosterone in boys were measured for up to 10 years. We calculated measures of circulating non-SHBG bound estradiol and testosterone from total hormone levels using the law of mass action. We compared proportional differences in assay measured minus calculated non-SHBG bound hormone levels versus their averages using GEE-estimated linear regressions.

Results—On average, calculated values overestimated assay measured values (−11.7% for non-SHBG bound estradiol in girls and −2.6% for non-SHBG bound testosterone in boys). The intercept and slope of the regression for non-SHBG bound estradiol in girls were −0.13 (95% CI −0.14 to −0.12) and 0.005 (95% CI 0.003 to 0.007), respectively. The intercept and slope for non-SHBG bound testosterone in boys were −0.16 (95% CI −0.17 to −0.14) and 0.0006 (95% CI 0.0005–0.0006).

Conclusion—While calculated values might be useful for research purposes, they are generally not close enough for clinical purposes.

Introduction

The bioavailable fractions of estradiol and testosterone include hormones that are freely circulating and bound to albumin but not bound to sex hormone binding globulin (SHBG) and are collectively the non-SHBG bound fraction. Circulating levels of non-SHBG bound estradiol and testosterone are desirable for clinical practice or investigational studies. However,

Correspondence to: Brian L. Egleston, Brian.Egleston@fccc.edu.

Competing interests: At the time the assays were conducted, DWC held stock in Esoterix Endocrinology, Inc. (Calabasas Hills, CA).

Ethical approval: This work is covered by Institutional Review Board protocol #09-841 of the Fox Chase Cancer Center.

Guarantor: BLE

Contributorship: BLE: Study design, data analysis, manuscript writing. DWC: Data acquisition, assay validation, manuscript writing. JFD: Study design, data acquisition, data analysis, manuscript writing.

assays to measure circulating levels of these hormones are often not done because the tests are not always readily accessible, might require more serum than is available, or might be considered too expensive. In addition, some methods of direct measurement are reported to be too inaccurate to be effective.^{1,2} As a result, a number of authors have published or validated formulae for estimating bioavailable testosterone and bioavailable estradiol using total hormone levels in blood.^{3,4,5} DeVan and colleagues⁶ estimated that using calculated rather than directly measured bioavailable testosterone in one medical center would save in excess of \$60,000 per year if such calculations are accurate. It must be noted that calculated estimates using formulas are only as accurate as the measured values input into the formula making accurate and specific measures of testosterone, estradiol and SHBG an absolute requirement.

We know of no study that has validated the empirical calculation of circulating bioavailable testosterone and estradiol in children. In this paper, we investigate the validity of estimating non-SHBG bound estradiol and testosterone in children using formulas based on the law of mass action.

Subjects and Methods

Data came from the Dietary Intervention Study in Children (DISC). DISC was a multicenter randomized controlled clinical trial sponsored by the National Heart, Lung, and Blood Institute (NHLBI) to test safety and efficacy of a dietary intervention to reduce serum low-density lipoprotein cholesterol (LDL-C) in children with elevated LDL-C. Design and results of DISC have been described.^{7, 8, 9} Briefly, between 1988 and 1990, 663 eight to ten year olds with elevated LDL-C were randomly assigned to a dietary intervention to reduce fat intake or to usual care at one of six DISC clinical centers. The initial DISC protocol was designed for 3 years and was subsequently extended with planned intervention and follow-up of all participants until 18 years of age. The timing of the last blood draw relative to study initiation was variable due to early termination of the study in 1997 for lack of a treatment effect.

DISC recruited 362 boys and 301 girls through schools, health maintenance organizations, and pediatric practices. Boys were eligible if they were 8.6–10.8 years old and girls were eligible if they were 7.8–10.1 years old. Eligibility requirements stipulated that children have serum LDL-C level in the 80th to 98th percentiles, had no major illness, were not taking medications that affect lipid levels or growth, were at least in the 5th percentile for height and in the 5th–90th percentiles for weight for height, were Tanner stage 1 for genital and pubic hair development, and had normal psychosocial and cognitive development.

Assent was obtained from DISC participants and written informed consent was obtained from their parents or guardians prior to randomization and again when the study was extended. The DISC protocol and Hormone Ancillary Study were approved by Institutional Review Boards at all participating centers, and a National Heart, Lung, and Blood Institute-appointed independent data and safety monitoring committee provided oversight.

Hormone Assays

Blood was collected at a visit through a single blood sample by venipuncture in the morning after an overnight fast. Serum was separated by centrifugation after the blood sample was kept at room temperature for at least 45 minutes to allow complete clotting. Serum was then aliquoted and stored in glass vials at -80°C .

Dorgan and colleagues^{10,11} provide extensive details of the assays used for this study. We summarize some of their characteristics here. Hormone assays were performed by Esoterix Endocrinology, Inc. (Calabasas Hills, CA) using standard procedures. Estradiol (E_2) was measured using a modification of the procedure developed by Wu and Lundy.¹² Serum samples

were extracted with hexane: ethyl acetate, 80:20 (vol/vol). The extract was then washed with dilute base, concentrated and chromatographed on Sephadex LH20 micro columns (Sigma, St. Louis, MO). E₂ was specifically eluted using benzene: methanol, 85:15 (vol/vol). E₂ was quantified by RIA in duplicate using antiserum raised to an Estradiol-6-oxime-BSA conjugate.

Testosterone (T) was measured using a modification of the procedure developed by Furuyama et al.¹³ Samples were extracted with hexane: ethyl acetate, 90:10 (vol/vol), and the extracts were applied to aluminum oxide micro columns. The columns were washed with hexane containing 0.55% ethanol, and T was specifically eluted using hexane containing 1.4% ethanol. T in eluates was quantified in duplicate by RIA using antiserum raised to a testosterone-3-oxime-BSA conjugate.

To measure DHT, serum samples were first extracted with seven volumes of hexane: ethyl acetate. Extracts were then evaporated to dryness and redissolved in potassium permanganate to oxidize steroids containing conjugated ketones. DHT was then selectively re-extracted. Duplicate aliquots of each purified sample were measured by RIA using antiserum raised to a DHT-3-oxime-BSA conjugate.

SHBG was measured by a radioimmunoassay. The serum sample and an SHBG monoclonal antibody labeled with ¹²⁵I were incubated with plastic beads coated with a different SHBG monoclonal antibody. The beads were washed to remove unbound label and the bound radioactivity was measured.

The percent non-SHBG bound estradiol and the percent non-SHBG bound testosterone were determined by ammonium sulfate precipitation as described by Mayes and Nugent.¹⁴ The concentration of non-SHBG bound steroid was then calculated as the product of its total concentration and the percent that was non-SHBG bound. In the paper, we refer to these as assay measured values. Assay measured non-SHBG bound T was collected for boys only, while assay measured non-SHBG bound E₂ was collected for girls only.

Serum albumin was not measured in all subjects at all visits; however the subjects were healthy, allowing for its approximation when missing using the average measured albumin level.

For external quality assurance, three sex-specific external quality controls indistinguishable from participant samples were included in each hormone batch. Lab personnel were blinded to which samples were participant samples and which were quality control samples. As reported by Dorgan,^{10,11} coefficients of variation (CV) of these assays, as estimated from the external quality control samples, were 6%–30% for E₂ for girls, 6%–26% for T for both sexes, 13%–24% for DHT for boys, and 15% for SHBG for girls. Higher CVs were generally related to lower mean hormone levels. The limits of detection for T, E₂, and DHT, were 3.0 ng/dL, 0.5 ng/dL, and 2.0 ng/dL, respectively.

RIA and mass spectrometry estimates of E₂, DHT, and T at levels commonly seen in children showed good agreement as described by Dorgan and colleagues.¹¹ This was similar to unpublished adult data from Esoterix Endocrinology that found a regression slope of MS on RIA of 1.14 for E₂ (coefficient of determination, R²=0.98) and a slope of RIA on MS of 1.09 for T (R²= 0.99), indicating some bias.

Calculating non-SHBG bound E₂ and T

Assay measures of total estradiol (E₂), testosterone (T), sex hormone-binding globulin concentration (C_{SHBG}), and albumin concentration (C_a) used to calculate the bioavailable fractions of estradiol including free and non-SHBG bound levels of estradiol in girls were collected at years 1, 3, 5, 7 and the last follow-up visit of the study. Assay measures of total

E_2 , T, C_{SHBG} , C_a and dihydrotestosterone (DHT) used to calculate free and non-SHBG bound levels of testosterone in boys were collected at years 3, 5, 7 and the last follow-up visit of the study. Because few children completed both a 7 year follow-up and a last visit due to early study termination, we do not present the 7-year follow-up data from children with both visits in tables. However, these values were used in plots described below.

We used the methods detailed in Rinaldi et al.³ to first estimate the amount of free testosterone (fT) and free estradiol (fE₂) for boys and girls using this data. We then used these results to estimate the amount of non-SHBG bound T and non-SHBG bound E₂ for comparison with assay measured values. As cited in Rinaldi et al.³, we used the single equation models⁴, and the multi-equation models (3 equations for boys, 2 equations for girls)⁵ to estimate fT and fE₂. We could not use the three-equation model in girls since DHT data were not collected for them. However, girls' DHT concentrations remain low throughout puberty.

The equations are based on the mass action law. T, DHT and E₂ circulate in serum free or are bound to albumin or SHBG. The equations use the following affinity constants in liters/mol for albumin as cited in Rinaldi and colleagues³: $K_aT=4.06 \times 10^4$ (testosterone), $K_aE_2=4.21 \times 10^4$ (estradiol), $K_aDHT=3.5 \times 10^4$ (DHT), and the following respective constants in liters/mol for C_{SHBG} : $K_sT=1 \times 10^9$, $K_sE_2=3.14 \times 10^8$, $K_sDHT=3 \times 10^9$. Specifically, we calculated fT and fE₂ in girls by solving for fT and fE₂ using the following set of equations adapted from Rinaldi.³ In the equations, fT and fE₂ are the only two unknowns.

$$\begin{aligned} E_2 - fE_2 \times \left(1 + \frac{K_a E_2 \times C_a}{1 + K_a E_2 \times fE_2 + K_a T \times fT} + \frac{K_s E_2 \times C_{SHBG}}{1 + K_s E_2 \times fE_2 + K_s T \times fT} \right) &= 0 \\ T - fT \times \left(1 + \frac{K_a T \times C_a}{1 + K_a E_2 \times fE_2 + K_a T \times fT} + \frac{K_s T \times C_{SHBG}}{1 + K_s E_2 \times fE_2 + K_s T \times fT} \right) &= 0 \end{aligned}$$

We compared this to results from the single equation method reproduced from Rinaldi³ that does not use testosterone data:

$$E_2 + (E_2 - K_s E_2 \times C_{SHBG} - K_a E_2 \times C_a - 1) \times fE_2 - (K_a E_2 \times C_a + 1) \times fE_2^2 = 0$$

Similarly, we calculated fT, fE₂, and fDHT in boys using the following set of equations also reproduced from Rinaldi.³ In the equations, fT, fE₂, and fDHT are the only unknown quantities.

$$\begin{aligned} E_2 - fE_2 \times \left(1 + \frac{K_a E_2 \times C_a}{1 + K_a E_2 \times fE_2 + K_a T \times fT + K_a DHT \times fDHT} + \frac{K_s E_2 \times C_{SHBG}}{1 + K_s E_2 \times fE_2 + K_s T \times fT + K_s DHT \times fDHT} \right) &= 0 \\ T - fT \times \left(1 + \frac{K_a T \times C_a}{1 + K_a E_2 \times fE_2 + K_a T \times fT + K_a DHT \times fDHT} + \frac{K_s T \times C_{SHBG}}{1 + K_s E_2 \times fE_2 + K_s T \times fT + K_s DHT \times fDHT} \right) &= 0 \\ DHT - fDHT \times \left(1 + \frac{K_a DHT \times C_a}{1 + K_a E_2 \times fE_2 + K_a T \times fT + K_a DHT \times fDHT} + \frac{K_s DHT \times C_{SHBG}}{1 + K_s E_2 \times fE_2 + K_s T \times fT + K_s DHT \times fDHT} \right) &= 0 \end{aligned}$$

We compared this to results from the single equation method reproduced from Rinaldi³ that does not use estradiol or DHT data:

$$T + (T - K_s T \times C_{SHBG} - K_a T \times C_a - 1) \times fT - (K_a T \times C_a + 1) \times fT^2 = 0$$

The single equation methods can be calculated analytically using the quadratic formula. We solved for unknown values in the two and three equation methods using the optimization package "optim" macro in R (The R Foundation for Statistical Computing, Vienna, Austria).

To validate our findings, we compared calculated serum non-SHBG bound E_2 and T with assay measured results. We used calculated fT , calculated fE_2 , and assay measured C_a to estimate the amount of serum non-SHBG bound E_2 and the amount of non-SHBG bound T with the following set of equations: non-SHBG bound E_2 in moles/liter is $K_a E_2 \times C_a \times fE_2 + fE_2$, and non-SHBG bound T in moles/liter is $K_a T \times C_a \times fT + fT$. For ease of interpretation, we report these as ng/dL. We provide S.I. conversion factors in the tables.

We substituted the observed mean value of C_a , 4.5 g/dL, for individuals with missing assay measured albumin levels. The observed mean was the same for boys and girls.

We examined Bland-Altman¹⁵ plots to investigate agreement between calculated and observed values. For the y-axis of the Bland-Altman plots, we plotted the proportion of the difference with respect to the average of the two values. Negative proportions indicate that calculated values were larger than assay measured values. For example, for the Bland-Altman plots of non-SHBG bound hormones, the y-axis consists of the points created by the following formula:

$$\frac{\text{Assay measured value} - \text{Calculated value}}{(\text{Assay measured value} + \text{Calculated value})/2}.$$

We report the quartiles of these proportional differences in the tables.

We used simple linear regressions fit on the data used to generate the Bland-Altman plots to estimate the relationship of the calculated with assay measured values. To more evenly spread the data, we depict the x-axis on the log scale in the Bland-Altman plots. However, we did not use log transformations for the regression analyses. We examined multiple linear regressions in which we included the average of calculated and assay measured hormone values, age, and the interaction between age and average values to investigate if the relationship between assay measured and calculated values differed by age. We also examined regressions in which we included average and average squared to test for nonlinear trends. We estimated all regressions using generalized estimating equations (GEE) assuming unstructured correlation matrices and robust standard errors to account for the correlation of multiple measurements within the same child.¹⁶

As a sensitivity analysis to the robustness of our results, we compared the GEE estimated simple linear regressions with Deming¹⁷ regressions of the observed regressed on the calculated values. The Deming regressions consider both model response and covariate to be measured with error, but do not account for the correlation of measurements over time. We assumed equal measurement error variances for the Deming regressions since we did not have duplicate measurements to estimate the variances. We estimated all regressions using the “xtgee” and “deming” commands in STATA version 10 (StataCorp, College Station, Texas).

To convert E_2 from ng/dL to pmol/L, multiply by 36.76. To convert T in ng/dL to pmol/L, multiply by 34.72.

Results

Estradiol in girls

There were 301 girls enrolled in DISC. Of these, 291 girls had sufficient data to compare assay measured non-SHBG bound E_2 with calculated values in this analysis, with an average of 3.4 measurements per girl between the first and last assessment of E_2 . Of the 987 observations used to compare calculated with assay measured values, 568 had missing C_a measurements replaced with the observed mean of 4.5 g/dL (SD 0.2). In analyses, 97 (9.8%) of the total E_2 measurements and 55 (5.7%) of the total T measurements that were below the limit of detection were set to the lower limit (0.5 and 3.0 ng/dL respectively).

The mean age of the girls at the first measurement of E_2 was 10.3 (SD 0.6, range 9.1–11.8 years), while the mean age of the girls at the last measurement of E_2 was 16.7 (SD 0.9, range 14.6–19.1 years). Over the course of the study, the great majority of girls had assay measured E_2 in the measurable range of the assays employed. The average total value of assay measured E_2 in girls over the course of the study was 6.13 ng/dL (SD=6.66, range=0.50–52.0). The average amount of assay measured non-SHBG bound E_2 was 3.17 ng/dL (SD=3.59, range 0.1–29.4).

In figure 1a we compare fE_2 calculated using the single equation and two equation methods. We see that the proportional discrepancies between the two methods are very small (a maximum of 0.01 ng/dL when the average of the two values is equal to 0.86, and a relative difference of less than 2.5% overall). The reason that the estimates are close is that the denominators in the fractions of the two equation method are both close to 1, while fE_2^2 and $fE_2 \times E_2$ are both small relative to the other terms in the single equation formula. This results in the two methods producing almost identical estimates. Because the two methods produced calculations that were so close, we used the single equation method to calculate non-SHBG bound E_2 since this method is easier for researchers to implement. The rest of the paper describes estimates obtained using the single equation method. Using the single equation method, the average percentage of calculated fE_2 was 1.94% (SD 0.41%, range 0.78%–3.11%).

In figure 2, we present the Bland-Altman¹⁵ plots of the proportional difference between assay measured and calculated non-SHBG bound E_2 . In figure 2a and figure 2b, we present the data from those with observed albumin so that we can compare the figures when using observed albumin and when substituting the mean albumin value for the observed value. The general spread of the figure does not change much when using observed values and mean substitution. In figure 2c, we present a plot with all of the data. In general, the absolute difference between observed and calculated values was less than four ng/dL in figure 2c. The average bias was –11.7% (indicating general overestimation of assay measured values by calculated values).

In table 1, the last column suggests that in percentage terms, the differences between assay measured and calculated values were sizable. However, many of the large percentage differences occurred with very low levels of non-SHBG bound E_2 . For example, the largest percentage difference occurred when observed non-SHBG bound E_2 was equal to 0.10 ng/dL and calculated non-SHBG bound E_2 was equal to 0.22 ng/dL. While large in percentage terms, it is unclear whether the absolute magnitude of such a difference would be clinically relevant.

The intercept and slope of the GEE-estimated simple linear regression of the proportional difference of assay measured minus calculated non-SHBG bound E_2 regressed on the average of the two values (figure 2c) were –0.13 (95% confidence interval, CI, –0.14 to –0.12) and 0.005 (95% CI 0.003 to 0.007). An intercept and slope equal to zero would indicate perfect prediction on average of assay measured non-SHBG bound E_2 from calculated non-SHBG bound E_2 . In table 1, we see that assay measured non-SHBG bound E_2 was slightly less on average than calculated non-SHBG bound E_2 . Hence, there is evidence of overestimation on average of assay measured non-SHBG bound E_2 when using the calculated values, but the overestimation decreases as hormone levels increase. The correlation of the proportional differences in measurements within individuals over time was not of consistent magnitude with the correlations between time points ranging from –0.08 to 0.44, as estimated by the working correlation matrix from the GEE fit regression.

The intercept and slope from a GEE-fit model estimated only using observations with assay measured albumin levels (figure 2a) were –0.15 (95% CI –0.16 to –.13) and 0.003 (95% CI –0.002 to 0.007), respectively. A quadratic term added to the model was not statistically significant (slope=0.00005 for quadratic term, $p=0.705$) suggesting a simple linear

relationship. The intercept and slope from a simple Deming regression of observed values regressed on calculated values were -0.16 (95% CI -0.25 to -0.07) and 0.96 (95% CI 0.92 to 0.99).

The interaction term in the multiple linear regression on the Bland-Altman data assessing whether the predictive value of calculated non-SHBG bound E_2 differed by age was statistically significant (coefficient $= -0.002$, $p = 0.010$). The main effects term for the average effect in the interaction model was 0.032 ($p = 0.002$). This suggests that the relationship of expected bias with level of non-SHBG bound E_2 was decreased for older girls: effect of average on proportional difference $= (0.032 - 0.002 * \text{age}) * \text{average}$.

Testosterone in boys

There were 362 boys enrolled in DISC. Of these, 346 had sufficient data to compare assay measured non-SHBG bound T with calculated values, with an average of 2.7 measurements per boy between the first and last assessment of T. Of the 937 observations used to compare calculated with assay measured values, 677 had missing C_a measurements replaced with the observed mean of 4.5 g/dL (SD 0.2). In analyses, 168 (17.9%) of the total E_2 measurements and 3 (0.3%) of the total T measurements that were below the limit of detection were set to the lower limit (0.5 and 3.0 ng/dL respectively).

The mean age of the boys at the first visit for which we had complete data was 13.0 (SD 0.7 , range 11.7 – 15.0 years), while the mean age of the boys at the last measurement of T was 17.3 (SD 0.8 , range 15.6 – 19.4 years). The average value of assay measured total T in boys over the course of the study was 493 ng/dL (SD $= 258$, range $= 3$ – 1186). The average amount of assay measured non-SHBG bound T was 237 ng/dL (SD $= 151$, range 0.3 – 795).

In figure 1b we present fT calculated with the three equation method compared to the single equation method. We see that the single equation method significantly underestimates fT compared to the three equation method. The mean fT in boys using the single equation method was 6.8 compared with 8.2 using the three equation method. Because of the discrepancy, we used the three equation method for inferences below. The average percentage of calculated fT using the three equation method was 1.56% (SD 0.53% , range 0.34% – 2.96%).

In figure 3 we present the Bland-Altman¹⁵ plots of the proportional difference between assay measured and calculated non-SHBG bound T. As with girls, the general spread of the data was similar when comparing proportional differences derived using observed albumin compared with mean substitution of the same observed values (figures 3a and 3b). In figure 3c, we present the proportional differences between observed and calculated values for the entire sample. The proportional differences between observed and calculated values for many observations were sizable. The average bias was -2.6% (indicating general overestimation of assay measured values by calculated values).

The intercept and slope of the GEE-estimated simple linear regression of the proportional difference in assay measured non-SHBG bound T minus calculated non-SHBG bound T regressed on the average of the two values (figure 3c) were -0.16 (-0.17 to -0.14) and 0.0006 (95% CI 0.0005 – 0.0006), respectively. Hence, there is evidence of average overestimation of assay measured non-SHBG bound T from the calculated values when levels are low, but the overestimation decreases as hormone levels increase. The correlation of the proportional differences in measurements within individuals over time was inconsistent, with the correlations ranging from 0.02 to 0.35 , as estimated by the working correlation matrix from the GEE fit regression. The GEE-fit intercept and slope from a model estimated only using observations with assay measured albumin levels (figure 3a) were -0.20 (95% CI -0.23 to -0.18) and 0.0009 (95% CI 0.0008 – 0.0011). A quadratic term of the average added to the

model was statistically significant (slope = -0.000001 for quadratic term, $p < 0.001$) suggesting a non-linear relationship. The intercept and slope from a simple Deming regression of observed values regressed on calculated values were -13.2 (95% CI -16.0 to -10.4) and 1.08 (95% CI 1.07 – 1.10), respectively.

There was one outlying observation excluded from figure 3 for ease of presentation in which the calculated non-SHBG bound T is quite different from the assay measured non-SHBG bound T. This represented data from one boy whose assay measured total T level was 901 ng/dL and calculated non-SHBG bound T was 395 (percentage difference compared with average of two of -164%), but whose assay measured non-SHBG bound T was 39 ng/dL. This contrasts with mean assay measured total T of 210 ng/dL (range 93 – 410) among the 14 other boys with assay measured non-SHBG bound T between 36 and 44 ng/dL, inclusive. The inconsistency suggests that there might have been a measurement or recording error in this boy's assay measured non-SHBG bound T. The coefficients from a GEE-estimated simple linear regression without this boy's data were the same as those reported above.

In table 2, we present means of assay measured and calculated non-SHBG bound T at each visit. Except for the 3 year visit, all of the assay measured values were higher on average. The percentage discrepancies were large, particularly when the average of calculated and assay measured non-SHBG bound T was small.

The interaction term in the multiple linear regression on the Bland-Altman data assessing whether the predictive value of calculated non-SHBG bound T differed by age was statistically significant (coefficient = -0.0001 , $p < 0.001$). The main effects term for the average effect in the interaction model was 0.003 ($p < 0.001$). This suggests that relationship of amount of bias with level of non-SHBG bound T was decreased for older boys: effect of average on proportional difference = $(0.003 - 0.0001 * \text{age}) * \text{average}$.

Discussion

This is the first study to investigate the relationship of calculated non-SHBG bound E_2 in girls and non-SHBG bound T in boys with assay measured values of the same hormones. While we have found good correlation of the measurements, as expected, there were some significant differences between measurements within individuals. Since calculated non-SHBG bound E_2 and non-SHBG bound T were constructed using calculated fE_2 and fT , it is likely that our findings are generalizable to fE_2 and fT values. Although there were many large percentage differences between observed and calculated hormone values, many of the largest percentage discrepancies occurred at lower observed hormone levels. Hence, it is unclear whether the small absolute magnitudes of the differences would be clinically relevant even though the percentage differences are large.

Among the boys, there was one outlier in which the calculated non-SHBG bound T did a poor job of estimating the assay measured non-SHBG bound T. This might have been an assay measurement or reporting error since the assay measured non-SHBG bound T was so much smaller than the assay measured total T relative to other boys' measurements.

The Bland-Altman plots demonstrated that the individual differences between assay measured and calculated non-SHBG bound hormone levels were proportionally sizable. This suggests that using the calculated values might only be appropriate for research purposes in which population level inferences are more of interest than subject-specific inferences. The use of the calculated values would be inappropriate in clinical settings in which greater precision is required.

Our findings that there might be discrepancies between assay measured and calculated values of non-SHBG bound T are consistent with Sartorius and colleagues' findings that there was some bias and variability between calculated and assay measured fT.¹⁸ However, Sartorius and colleagues¹⁸ found that calculated values overestimated fT, while we found that calculated values underestimated non-SHBG bound T for larger assay measured levels.

A limitation of this work is that we did not have assay measurements of fT and fE₂ in the children. Still, the percent fT and fE₂ values that we observed were consistent with what one might expect in children.

Also, our measurements of T and E₂ were made using RIA while gas chromatography-mass spectrometry coupled with gas chromatography (GC/MS) or high pressure liquid chromatography (HPLC/MS) is currently the recommended method. In previous work, Dorgan et al.¹¹ used regression analysis to demonstrate that there was good agreement between T and E₂ measured in children using the RIA used in this study and GC/MS. The accuracy of our RIA measurements contributed to the good agreement we observed between calculated and measured bioavailable E₂ and T. Inaccurate measurement of any of the hormones used in the calculations would generally result in invalid estimates of the bioavailable fractions.

In conclusion, we have found average bias of -2.6% in calculations of non-SHBG bound testosterone in boys and average bias of -11.7% of non-SHBG bound estradiol in girls using total levels of hormones measured using RIA after extraction and chromatography. These findings provide some evidence that calculated free levels of testosterone and estradiol would have similar bias. Good correlation of RIA after extraction and chromatography with levels found using HPLC and mass spectrometry suggest that this technique will also provide similar calculated results.¹⁹

Acknowledgments

Funding: Research reported here was supported in part by NIH grant P30 CA 06927 and an appropriation from the Commonwealth of Pennsylvania. The Dietary Intervention Study in Children (DISC) Hormone Ancillary Study was supported by cooperative agreements U01-HL37947, U01-HL37948, U01-HL37954, U01-HL37962, U01-37966, U01-HL37975, and U01-HL38110 from the National Heart, Lung, and Blood Institute supplemented by the National Cancer Institute.

The investigators express appreciation to the boys and girls who participated in DISC and their families. We also would like to thank Dr. William Rosner, Dr. Samuel Litwin, and Dr. Hua Min for their useful comments regarding the paper.

References

1. Rosner W. An extraordinarily inaccurate assay for free T is still with us. *J Clin Epidemiol* 2001;86(6): 2903.
2. Fritz KS, McKean AJS, Nelson JC, Wilcox RB. Analog-based free T test results linked to total T concentrations, not free T concentrations. *Clin Chem* 2008;54:512-516. [PubMed: 18171714]
3. Rinaldi S, Geay A, Dechaud H, Biessy C, Zeleniuch-Jacquotte A, Akhmedkhanov A, Shore RE, Riboli E, Toniolo P, Kaaks R. Validity of free T and free estradiol determinations in serum samples from postmenopausal women by theoretical calculations. *Cancer Epidemiol Biomarkers Prev* 2002;11:1065-1071. [PubMed: 12376508]
4. Vermeulen A, Verdonck L, Kaufman JM. A critical evaluation of simple methods for the estimation of free T in serum. *J Clin Endocrinol Metab* 1999;84(10):3666-3672. [PubMed: 10523012]
5. Sodergard R, Backstrom T, Shanbhag V, Carstensen H. Calculation of free and bound fractions of T and estradiol-17beta to human plasma proteins at body temperature. *J Steroid Biochem* 1982;16:801-810. [PubMed: 7202083]

6. DeVan ML, Bankson DD, Abadie JM. To what extent are free T (FT) values reproducible between the two Washingtons, and can calculated FT be used in lieu of expensive direct measurements. *Am J Clin Pathol* 2008;129:459–563. [PubMed: 18285270]
7. DISC Collaborative Research Group. Dietary intervention study in children (DISC) with elevated low-density-lipoprotein cholesterol. Design and baseline characteristics. *Ann Epidemiol* 1993;3:393–402. [PubMed: 8275216]
8. The Writing Group for the DISC Collaborative Research Group. Efficacy and safety of lowering dietary intake of fat and cholesterol in children with elevated low-density lipoprotein cholesterol. The Dietary Intervention Study in Children (DISC). *JAMA* 1995;273:1429–1435. [PubMed: 7723156]
9. Obarzanek E, Kimm SY, Barton BA, Van Horn LL, Kwiterovich PO Jr, Simons-Morton DG, Hunsberger SA, Lasser NL, Robson AM, Franklin FA Jr, Lauer RM, Stevens VJ, Friedman LA, Dorgan JF, Greenlick MR. Long-term safety and efficacy of a cholesterol-lowering diet in children with elevated low-density lipoprotein cholesterol: seven-year results of the Dietary Intervention Study in Children (DISC). *Pediatrics* 2001;107:256–264. [PubMed: 11158455]
10. Dorgan JF, Baer DJ, Albert PS, Judd JT, Brown ED, Corle DK, Campbell WS, Hartman TJ, Tejpar AA, Clevidence BA, Giffen CA, Chandler DW, Stanczyk FZ, Taylor PR. Serum hormones and the alcohol-breast cancer association in postmenopausal women. *J Natl Cancer Inst* 2001;93:710–715. [PubMed: 11333294]
11. Dorgan JF, Fears TR, McMahon RP, Aronson Friedman L, Patterson BH, Greenhut SF. Measurement of steroid sex hormones in serum: a comparison of radioimmunoassay and mass spectrometry. *Steroids* 2002;67:151–158. [PubMed: 11856538]
12. Wu CH, Lundy LE. Radioimmunoassay of plasma estrogens. *Steroids* 1971;18:91–111. [PubMed: 5107186]
13. Furuyama S, Mayes DM, Nugent CA. A radioimmunoassay for plasma testosterone. *Steroids* 1970;16:415–428. [PubMed: 5533947]
14. Mayes DM, Nugent CA. Determination of plasma testosterone by the use of competitive protein binding. *J Clin Endocrinol Metab* 1968;28:1169–1176. [PubMed: 4877626]
15. Bland JM, Altman DG. Statistical methods for assessing agreement between two methods of clinical measurement. *Lancet* 1986;1(8476):307–310. [PubMed: 2868172]
16. Zeger SL, Liang KY, Albert PS. Models for longitudinal data: A generalized estimating equation approach. *Biometrics* 1988;44:1049–1060. [PubMed: 3233245]
17. Cornbleet PJ, Gochman N. Incorrect least-squares regression coefficients in method-comparison analysis. *Clin Chem* 1979;25(3):432–438. [PubMed: 262186]
18. Sartorius G, Ly LP, Sikaris K, McLachlan R, Handelsman DJ. Predictive accuracy of variability in calculated free testosterone estimates. *Ann Clin Biochem* 2009;46:137–143. [PubMed: 19225026]
19. Rosner W, Auchus RJ, Azziz R, Sluss PM, Raff H. Position statement: Utility, limitations, and pitfalls in measuring testosterone: an Endocrine Society position statement. *J Clin Endocrinol Metab* 2007;92(2):405–13. [PubMed: 17090633]

Figure 1.

Plots comparing fE₂ and fT (ng/dL) calculated using multi-equation and single equation methods. The y-axis depicts the proportional difference in values with respect to the average. Note that the range of the y-axis is much larger for fT than for fE₂.

Figure 2.

Plots comparing observed and calculated non-SHBG bound E_2 in girls. The y-axis depicts the proportional difference in values with respect to the average. The dashed line indicates the mean proportional difference while the dotted lines indicate 2 standard deviations above and below the mean. Observed albumin data were used in figure 2a, while the mean albumin values were substituted for observed values in figure 2b. All of the data were used in figure 2c.

Figure 3.

Plots comparing observed and calculated non-SHBG bound T in boys. The y-axis depicts the proportional difference in values with respect to the average. The dashed line indicates the mean proportional difference while the dotted lines indicate 2 standard deviations above and below the mean. Observed albumin data were used in figure 3a, while the mean albumin values were substituted for the observed values in figure 3b. All data were used for figure 3c. One outlier in figure 3c was excluded (difference=-1.64).

Table 1

Estimates of total E₂ and non-SHBG bound E₂ (ng/dL) among girls and changes in girls with complete data. To convert to pmol/L, multiply by 36.76.

Visit	N	Age Mean (SD)	Total E ₂ Mean (SD)	Assay non-SHBG bound E ₂ Mean (SD)	Calculated non-SHBG bound E ₂ Mean (SD)	25th/50th/75th percentile of assay minus calculated differences as a percent of average of the two values
1 year	189	10.3 (0.6)	1.26 (1.81)	0.53 (0.68)	0.64 (0.89)	-26%, -15%, -4%
3 Years	273	12.2 (0.6)	4.94 (2.47)	2.47 (2.41)	2.81 (2.66)	-20%, -14%, -7%
Change 1 to 3 years	183		3.61 (3.93)	1.93 (2.02)	2.16 (2.23)	-25%, -13%, -4%
5 years	226	14.3 (0.6)	8.96 (8.58)	4.68 (4.69)	4.76 (4.47)	-9%, -4%, 1%
Change 3 to 5 years	218		4.16 (8.49)	2.26 (4.71)	2.02 (4.57)	-14%, -3%, 10%
Last (variable)	229	16.7 (0.9)	8.14 (6.90)	4.16 (3.54)	4.89 (4.11)	-26%, -14%, -6%
Change 5 years to last	194		- .42 (10.31)	-0.27 (5.72)	0.39 (5.96)	-34%, -7%, 11%

Table 2

Estimates of total T and non-SHBG bound T (ng/dL) among boys and changes in boys with complete data. Non-SHBG bound testosterone was not measured on boys at the one-year visit. To convert to pmol/L, multiply by 34.72.

Visit	N	Age Mean (SD)	Total T Mean (SD)	Assay non-SHBG bound T Mean (SD)	Calculated non-SHBG bound T Mean (SD)	25th/50th/75th percentile of assay minus calculated differences as a percent of average of the two values
3 Years	294	13.0 (0.7)	279 (248)	96 (103)	97 (97)	-25%, -9%, 2%
5 years	275	15.0 (0.7)	539 (196)	249 (117)	236 (101)	-4%, 4%, 10%
Change 3 to 5 years	231		262 (199)	151 (97)	138 (87)	-5%, 6%, 19%
Last (variable)	290	17.3 (0.8)	630 (184)	337 (110)	332 (95)	-6%, 2%, 8%
Change 5 years to last	245		91 (173)	88 (110)	96 (101)	-20%, 3%, 27%