To What Extent Are Free Testosterone (FT) Values Reproducible Between the Two Washingtons, and Can Calculated FT Be Used in Lieu of Expensive Direct Measurements?

Michael L. DeVan,[1] Daniel D. Bankson,[2] and Jude M. Abadie[1]

Key Words: Free testosterone; Sex hormone binding globulin; Albumin; Tandem liquid chromatography mass spectroscopy

DOI: 10.1309/6PYTC60ALVQQ59RQ

Abstract

Free testosterone (FT) measurement by equilibrium dialysis and liquid chromatography–tandem mass spectroscopy (LCMS/MS) is the "gold standard." We hypothesized that calculated FT values could substitute for measured values; compared FT results reported by Walter Reed Army Medical Center (WRAMC), Washington, DC, with results reported by the Seattle Veterans Affairs Health Care System, Seattle, WA, for 3 patient groups; and evaluated the calculated FT values by gold-standard measurements. Groups 1 and 2 included samples from 54 patients evaluated in Seattle and 94 evaluated at a primary care clinic in Alaska whose samples were analyzed in Seattle, respectively, whose care resulted in ordering an FT measurement. Group 3 included samples from 64 patients evaluated in endocrine WRAMC clinics. Calculated FT values between the 2 facilities demonstrated a strong correlation ($R^2 = 0.98$) for all 212 patients. In a comparison of calculated FT values with measured levels, group 3 had an $R^2 = 0.93$; however, samples with FT values less than 50 pg/mL had a poorer correlation ($R^2 = 0.45$). Calculated FT values may accurately reflect and be substituted in the clinical setting for gold-standard values when levels are more than 50 pg/mL.

Because diagnosing, managing, and treating patients with testosterone-related endocrinopathies rely on laboratory results, accurate assessments of testosterone levels are imperative. Even in the setting of classic clinical endocrinopathies, reliable laboratory results are essential to confirm and manage the diagnosis.[1] In men, for example, a high pretest probability of androgen deficiency in conjunction with a low testosterone level confirms a diagnosis of hypogonadism.[2] In women, testosterone values aid in managing hyperandrogenism-related disorders such as polycystic ovarian syndrome, congenital adrenal hyperplasia, androgen-secreting tumors, and androgen deficiency. Using a single analyte (ie, testosterone) to evaluate such broad clinical spectra illustrates the importance of accurate and precise testing that incorporates well-defined reference ranges.

Although testosterone is a single analyte, assays that measure different forms of testosterone can provide insight regarding different biochemical populations of the circulating hormone (ie, biologically active vs inactive). About 2% of circulating testosterone exists as a free fraction, and the remaining 98% is bound to albumin or sex hormone binding globulin (SHBG), a liver glycoprotein. Because of a strong K_a, the testosterone fraction bound to SHBG is considered not available for biologic activity.[3] Bioavailable testosterone is composed of the free testosterone (FT) and testosterone bound to albumin, which has a weak K_a.[4] Mathematical equations have been designed to calculate FT levels based on measurements of total testosterone (TT), SHBG, and albumin.[5] However, age, sex, and clinically related variables may affect these calculations and, therefore, require further investigation and validation.

Male and female children have similar SHBG levels until puberty, at which time SHBG levels decrease at a more rapid

rate in males.[6] Because of the higher estrogen/androgen ratio, SHBG levels are higher in women than in men. SHBG binds to approximately 54% and 78% of testosterone in men and women, respectively.[7] With respect to ethnicity, 1 study demonstrated significant variations in testosterone levels among European, Pakistani, and African-Caribbean people.[8] Because of the many age, sex, and ethnically related differences in testosterone levels, the decision to use calculated instead of measured FT levels should be validated by each performing laboratory.

Ours is the first study to compare and validate calculated FT results measured in 2 laboratories in patient populations from different geographic locations within the United States. These correlations may provide strong evidence for using TT, albumin, and SHBG levels to calculate FT levels when making clinical decisions for using a calculated or measured FT method. We hypothesized that calculated FT values should correlate between the Veterans Affairs (VA) Puget Sound Health Care System, Seattle, WA, and Walter Reed Army Medical Center (WRAMC), Washington, DC. In addition, we hypothesized that calculated FT levels could be used without compromising patient care in lieu of the more expensive "gold-standard" method using equilibrium dialysis and liquid tandem chromatography–mass spectroscopy (LCMS/MS). Although calculated FT levels correlated well between the 2 hospitals, calculated FT values may prove to be a better substitute in men than in women. In addition, this study demonstrated that calculated FT results may be valid for thawed and refrozen samples stored at –70°C for up to 6 months.

Material and Methods

WRAMC is a 1,280-bed facility staffed by 5,000 military and civilian personnel, including 800 physicians, and cares for about 200 inpatients a day. The VA is a 504-bed teaching hospital that treats more than 600,000 outpatient visits and about 8,300 inpatients annually. This study was approved by VA and WRAMC institutional review boards.

Human SHBG levels were determined at WRAMC and the VA by using Cobas electrochemiluminescence immunoassay kits (Hoffmann-La Roche, Basel, Switzerland) for the Roche MODULAR ANALYTICS E170 analytic platform (Hoffmann-La Roche). The measuring range of the assay is 10 to 5,764 ng/dL (0.350-200 nmol/L). TT levels were determined at WRAMC and the VA by using Elecsys testosterone monoclonal antibody test kits (Hoffmann-La Roche) for the Roche MODULAR ANALYTICS E170 analytic platform. The measuring range of the assay is 2.0 to 1,500.0 ng/dL (0.1-52.1 nmol/L). Albumin levels were determined at both facilities by using Vitros ALB slides (Ortho-Clinical Diagnostics, Rochester, NY) for the Ortho Vitros 5,1 FS (Ortho-Clinical Diagnostics) analytic platform. The assay measuring range is 1 to 6 g/dL. FT levels were measured on samples obtained from patients seen by the WRAMC endocrinology service and assayed at Quest Diagnostics Laboratory (Chantelle, VA) using equilibrium dialysis followed by LCMS/MS. Concentrations are given in picograms per milliliter. FT levels were calculated from measured SHBG, TT, and albumin levels using the equation previously described.[7] Three groups of samples were analyzed.

Group 1 included 54 samples from patients seen at the VA whose care resulted in ordering an FT that was calculated from TT, human SHBG, and albumin levels. Subsequently, all samples were frozen and transported on dry ice to WRAMC in Washington, DC. After thawing the specimens for the first time, TT, SHBG, and albumin levels were determined at WRAMC, and the FT values were calculated and compared with values calculated from analogous assay measurements performed at the VA.

Group 2 consisted of 94 samples from patients seen in a clinic in Anchorage, AK. Samples were immediately frozen and transported to the VA for FT calculation from measured TT, SHBG, and albumin levels. Subsequently, these specimens were deidentified, frozen a second time, and shipped on dry ice to WRAMC. After thawing the specimens for the second time, TT, SHBG, and albumin levels were measured and FT values calculated at WRAMC.

Group 3 consisted of 64 samples from patients seen in the WRAMC endocrinology service whose patient care resulted in ordering FT levels. During phlebotomy, patients were asked for consent to obtain an extra tube of blood for the study. One tube was sent to Quest Diagnostics for standard-of-care FT measurement via equilibrium dialysis and LCMS/MS. The second tube was frozen and later assayed at WRAMC for TT, SHBG, and albumin levels. The FT was then calculated. Subsequently, these samples were frozen a second time and shipped on dry ice to the VA for analogous testing and calculation of the FT level.

The data were analyzed using Deming regression to compare the calculated FT results obtained by the VA and WRAMC on specimens thawed once and on those thawed twice. Linear regression analysis was performed comparing FT results obtained using the measured values generated from the reference laboratory with those calculated at WRAMC.

Results

Figure 1 illustrates the calculated FT values for all study subjects whose FT values were determined by calculation by the VA and WRAMC. This entire data set of 212 samples generated a line ($y = 0.97x + 0.36$) with an R^2 value of 0.98. The calculated FT values ranged from 1 to 444 pg/mL.

DOI: 10.1309/6PYTC60ALVOO59RQ

Figure 2 illustrates results for group 3. This data set generated a line (y = 0.95x + 12.06) with an R^2 value of 0.93. When analyzed separately from this group, the 26 values less than 50 pg/mL generated a line with an R^2 value of 0.45. Of the 26 values, 12 were determined in samples from women and represent the only women in group 3. These data are not illustrated separately from group 3 data. We did not obtain demographics or clinical data from patients in groups 1 and 2.

Although the demographics and diagnoses are not given in a table or figure, the following summary describes the population of group 3. Women ranged in age from 18 to 53 years and men from 21 to 84 years. The most common diagnosis in the women was polycystic ovarian syndrome, followed by amenorrhea/infertility, overweight, adrenal mass, and osteoporosis. In men, the most common diagnosis was erectile disorder, followed by hypogonadism, prostate cancer, seminoma, pituitary neoplasm/prolactinoma, infertility/oligospermia, breast hypertrophy, empty sella syndrome, secondary testicular failure, osteoporosis, alopecia totalis, and anemia.

Table 1 shows a comparison of costs of assays required to calculate an FT between WRAMC and 3 reference laboratories. Because the individual assays used to calculate FT are automated and the calculation would be an automatic function in our laboratory information computer system, labor is not a variable. The send-out cost for measuring and reporting TT, bioavailable testosterone, and FT (measured) as determined using WRAMC's preferred pricing from Quest Diagnostics is $30.58. The analogous comparison of average reagent cost used to calculate FT values is $5.93, a difference of $24.65 per reportable test. In 2006, the WRAMC North Capital Region sent out 3,144 FT assays (2,464 from men and 680 from women). Switching to calculated FT in men only would result in an annual projected savings more than $60,000.

Discussion

This is the first study to correlate calculated FT levels in samples from different patient populations between 2 laboratories approximately 3,000 miles apart. Because the concept

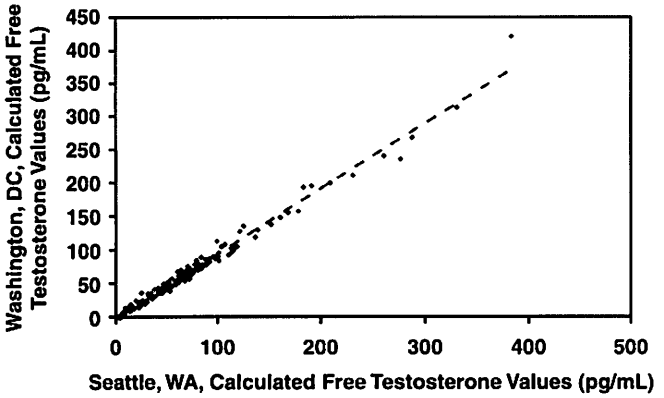

Figure 1 Free testosterone levels calculated at the Seattle, WA, Veterans Affairs and Washington, DC, Walter Reed Army Medical Center laboratories in 212 samples. y = 0.97x + 0.36. R^2 = 0.98.

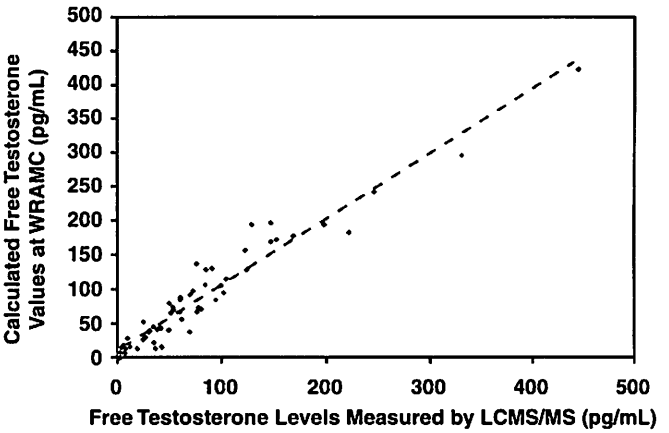

Figure 2 Comparison of measured with calculated free testosterone values in 64 samples. y = 0.95x + 12.06. R^2 = 0.93. LCMS/MS, liquid chromatography tandem mass spectroscopy; WRAMC, Walter Reed Army Medical Center.

Table 1
How Much Does Free Testosterone Cost?

Analyte	Reference Laboratory Cost ($)			Hospital Laboratory Reagent Cost ($)
	X	Y	Z	
Total testosterone*	12.01	14.65	6.75	2.23
Sex hormone binding globulin*	20.91	17.55	13.15	5.40
Albumin*	4.31	2.45	5.63	0.10
Testosterone, free and weakly bound panel†	37.23	34.65	25.53	7.73

July 2007 costs for Walter Reed Army Medical Center, Washington, DC, determined for reagents* and send-out "gold-standard"† testing at reference laboratories or calculated cost for in-house testing. Labor is not included.

DOI: 10.1309/6PYTC60ALVQQ59RQ

of FT calculation is mathematically related by mass action equilibrium formulas, an accurate and exact concentration of FT may be difficult to determine.[5] In addition, FT calculation methods may not be valid in all populations[9]; therefore, equilibrium dialysis followed by LCMS/MS has been described as the gold standard for direct FT measurements[10] and is used in many clinical settings. These factors may be the main reasons why it has been reported that some FT calculation methods demonstrate inconsistencies.[11] Furthermore, few data are available to demonstrate a low detection limit with the precision necessary to diagnose certain disorders of androgen deficiency in women[12-14] in which FT levels can be very low.

The reproducibility of calculated FT levels illustrated in Figure 1 is demonstrated throughout the entire range of data points. However, when compared with measured gold-standard values, we did not observe as strong an agreement for FT values in women as in men. Women were 12 of 26 patients whose calculated and measured FT values were lower than 50 pg/mL and demonstrated poor correlation between methods. However, calculated FT levels demonstrated good correlation throughout the range of values. We attribute this reproducible agreement directly to frozen sample stability, good clinical chemistry practice, and assay precision between similar Roche analytic platforms at the VA and WRAMC. This stability and precision remained intact for at least 6 months in samples stored at –70°C.

This study also demonstrated that calculated FT levels may clinically substitute for directly measured FT values in men but perhaps not in women. This observation may be a limitation of the equation in conjunction with a limitation of assay function comparisons at lower levels of testosterone concentration. Calculated FT accuracy is problematic in women with androgen deficiency such as hypopituitarism and is directly related to the validity of the TT and SHBG assays.[13,15] These investigators also reported that different SHBG assays yield significantly different calculated FT levels when used in the mass action equation. Our study findings support the findings of others suggesting that low testosterone level, irrespective of sex, is the main problem contributing to poor immunoassay sensitivity.[10]

According to the law of mass action, FT levels should precisely correlate with measured FT values when calculated from accurate TT and SHBG values. Although not true universally in our study, it seems that calculated FT levels are underestimated compared with measured values in women. Perhaps this observation is related to SHBG binding abnormalities resulting from testosterone competition for SHBG binding sites. Such a competition has been reported in pregnancy, in which calculated FT values were significantly lower than directly measured values.[15] Accurate FT levels are also required to diagnose androgen deficiency in women, in which the SHBG level is significantly elevated compared with the level in women without androgen deficiency. Although it has been demonstrated that accurate correlation between calculated and measured FT values is possible, we were not able to make accurate or predictable comparisons for women and for some men whose FT levels were less than 50 pg/mL in our study population.

Another reason for disagreement between calculated and measured FT values at low levels may be obvious differences in methods. Studies suggest that testosterone levels measured by immunoassay may not be reliable at lower levels.[10,12] However, values for such proposed low levels may not be well defined for many clinical decision points.

Because of the many correlation concerns, calculated FT levels should be validated based on well-established TT and SHBG assays in defined population groups. Based on our data, we demonstrate impressive assay reproducibility and suggest that calculated FT values may be clinically substituted for measured values in men when the FT levels are more than 50 pg/mL.

In 2006, 3,144 WRAMC samples were sent out for FT measurement. Of these, 743 had levels less than 50 pg/mL (in 451 women and 292 men). Therefore, if the decision were made to continue to send out FT samples from women and to use the calculated method for all FT samples from men, the projected annual savings would be slightly more than $60,000. Implementing this policy would require that all samples from men with a calculated FT value of less than 50 pg/mL be reported as "<50 pg/mL."

Irrespective of costs, studies should continue to investigate and describe relationships and possibilities for using calculated FT values to provide quality care and accurate clinical assessment in all patient populations.

From the [1]Department of Pathology and Area Laboratory Services, Core Laboratory, Walter Reed Army Medical Center, Washington, DC; and [2]Veterans Affairs Puget Sound Health Care System Pathology and Laboratory Medicine Service, Seattle, WA.

Address reprint requests to Dr Abadie: Dept of Pathology and Area Laboratory Services, Core Laboratory, Walter Reed Army Medical Center, 6900 Georgia Ave NW, Washington, DC 20011.

References

1. Matsumoto AM, Bremner BT. Serum testosterone assays: accuracy matters [editorial]. *J Clin Endocrinol Metab.* 2004;89:520-524.
2. Christina W, Catlin DH, Demers LM, et al. Measurement of total serum testosterone in adult men: comparison of current laboratory methods versus liquid chromatography-tandem mass spectrometry. *J Clin Endocrinol Metab.* 2004;89:534-543.
3. Pardridge WM, Mietul LJ, Frumar AM, et al. Effects of human sera on transport of testosterone and estradiol into rat brain. *Am J Physiol.* 1980;239:E103-E108.

DOI: 10.1309/6PYTC60ALVQQ59RQ

4. Manni A, Pardridge WM, Cefalu W, et al. Bioavailability of albumin-bound testosterone. *J Clin Endocrinol Metab.* 1985;61:705-710.
5. Vermeulen A, Verdonck L, Kaufmann JM. A critical evaluation of simple methods for the estimation of free testosterone in serum. *J Clin Endocrinol Metab.* 1999;84:3666-3672.
6. Liverman CT, Blazer DG. *Testosterone and Aging: Clinical Research Directions.* Washington, DC: Institute of Medicine, The National Academies Press; 2004.
7. Writing Group for the Women's Health Initiative Investigators. Risks and benefits of estrogen plus progestin in healthy postmenopausal women: principal results from the Women's Health Initiative randomized controlled trial. *JAMA.* 2002;288:321-333.
8. Heald AH, Ivison F, Anderson SG, et al. Significant ethnic variation in total and free testosterone concentrations. *Clin Endocrinol.* 2003;58:262-266.
9. Winters S, Kelley D, Goodpasture BB. The analog free testosterone assay: are the results in men clinically useful? *Clin Chem.* 1998;44:2178-2182.
10. Cawood ML, Field HP, Ford CG, et al. Testosterone measurement by isotope-dilution liquid chromatography–tandem mass spectrometry: validation of a method for routine clinical practice. *Clin Chem.* 2005;51:1472-1479.
11. Kapoor P, Luttrell BM, Williams D. The free androgen index is not valid for all adult males. *J Steroid Biochem Mol Biol.* 1993;45:325-326.
12. Taieb J, Mathian B, Millot F, et al. Testosterone measured by 10 immunoassays and by isotope-dilution gas chromatography–mass spectrometry in sera from 116 men, women, and children. *Clin Chem.* 2003;49:1381-1395.
13. Rosner W. An extraordinarily inaccurate assay for free testosterone is still with us [letter]. *J Clin Endocrinol Metab.* 2001;86:2903.
14. Rosner W. Errors in the measurement of plasma free testosterone [letter]. *J Clin Endocrinol Metab.* 1997;82:2014-2015.
15. Sinha-Hikim I, Arver S, Beall G, et al. The use of a sensitive equilibrium dialysis method for the measurement of free testosterone levels in healthy cycling women and in human immunodeficiency virus–infected women. *J Clin Endocrinol Metab.* 1998;83:1312-1318.

DOI: 10.1309/6PYTC60ALVQO59RQ