



HHS Public Access

Author manuscript

J Urol. Author manuscript; available in PMC 2023 July 25.

Published in final edited form as:

J Urol. 2012 April ; 187(4): 1369–1373. doi:10.1016/j.juro.2011.11.095.

Performance of Total Testosterone Measurement to Predict Free Testosterone for the Biochemical Evaluation of Male Hypogonadism

Bradley D. Anawalt*,

Department of Medicine, Seattle, Washington

James M. Hotaling,

Department of Urology, Seattle, Washington

Thomas J. Walsh,

Department of Urology, Seattle, Washington

Veterans Administration Puget Sound Health Care System, Seattle, Washington

Alvin M. Matsumoto

Department of Medicine, Seattle, Washington

University of Washington and Geriatric Research, Education and Clinical Center, Seattle, Washington

Veterans Administration Puget Sound Health Care System, Seattle, Washington

Abstract

Purpose: Guidelines recommend serum total testosterone measurement as the initial test to evaluate male hypogonadism, reserving free testosterone assessment for men with suspected sex hormone-binding globulin abnormalities or total testosterone near the lower limit of normal. We determined the performance of total testosterone measurement as a test to identify men with normal vs low free testosterone.

Materials and Methods: We examined the electronic medical records of all 3,672 men evaluated for hypogonadism by a serum testosterone panel, including total testosterone, sex hormone-binding globulin, albumin and calculated free testosterone, from January 1, 1997 through December 31, 2007 in a network that serves veterans in Washington.

Results: The sensitivity and specificity of low total testosterone (less than 280 ng/dl) to rule out and predict low calculated free testosterone was 91.0% and 73.7%, respectively. At thresholds of less than 350 and less than 400 ng/dl the sensitivity of total testosterone for low calculated free testosterone increased to 96.8% and 98.2%, and at thresholds of less than 150 and less than 200 ng/dl specificity increased to 98.9% and 92.6%, respectively.

Conclusions: Total testosterone between 280 and 350 ng/dl is not sensitive enough to reliably exclude hypogonadism. Total testosterone must exceed 350 to 400 ng/dl to reliably predict normal

*Correspondence and requests for reprints: 1959 Northeast Pacific St., Seattle, Washington 98195 (telephone: 206-897-1409; FAX: 206-616-4663; banawalt@medicine.washington.edu).

free testosterone. Except when levels are less than 150 ng/dl total testosterone measurement has low specificity for the biochemical diagnosis of hypogonadism.

Keywords

testis; hypogonadism; testosterone; sex hormone-binding globulin; diagnosis

MALE hypogonadism is a common disorder that affects an estimated 5% to 10% of men older than 30 years in the United States and the prevalence increases to 20% to 40% in men older than 70 years.¹⁻³ Evaluation for hypogonadism may be indicated in men with weakness, low libido, infertility and osteoporosis.⁴ Also, clinicians often measure serum testosterone to assess possible hypogonadism in various clinical settings, including diabetes mellitus, obesity, exogenous corticosteroid and opiate use, depression and sarcopenia.

To our knowledge the best laboratory test to use for initial screening for hypogonadism is unknown. TT assays are easy to perform, inexpensive, widely available and generally accurate. However, many common conditions, eg obesity, diabetes mellitus and aging, affect circulating SHBG and, therefore, affect circulating TT without necessarily affecting free (unbound) and weakly bound (bound to albumin) testosterone. Serum FT and weakly bound testosterone represent the bioavailable forms of circulating testosterone.⁵

FT measured by equilibrium dialysis is generally considered the clinical gold standard for the biochemical diagnosis of hypogonadism. This method of measuring FT is costly, time-consuming and not often used in clinical practice. FT can also be calculated by measuring serum TT, SHBG and albumin, and using one of the published, validated formulas.^{6,7} cFT corresponds well with FT measured by equilibrium dialysis.^{3,6,7}

The original and the recently updated Endocrine Society guidelines to evaluate hypogonadism recommend TT measurement as the initial screening test with FT or bioavailable testosterone measured only in men in whom testosterone is “near the lower limit of normal range and in whom alterations of SHBG are suspected.”^{3,8} To our knowledge it has not yet been determined in a large population of men how well TT measurement performs as a predictor of FT, which is the benchmark for biochemical confirmation of hypogonadism.

We determined how well TT performed as a predictor of low cFT using a cFT assay methodology that correlates well with FT measurements by equilibrium dialysis.^{9,10} We examined data on a large cohort of male veterans evaluated for hypogonadism from 1997 through 2007. We hypothesized that 1) low TT less than 280 ng/dl (multiply by 0.0347 to convert to nmol/l) is sensitive to rule out but not specific to predict low FT and 2) TT less than 150 ng/dl, a level associated with severe hypogonadism,³ is specific to predict low FT.

MATERIALS AND METHODS

To determine the performance of low TT to predict low serum FT we examined the electronic medical records of all men seen at the Veterans Administration Puget Sound Health Care System from January 1, 1997 through December 31, 2007. This institution

includes a large network of outpatient clinics and a 504 bed teaching hospital that serves veterans from Washington, Idaho and Alaska.

We identified all men evaluated for hypogonadism during this period with a serum testosterone panel, including TT, SHBG, albumin, cFT and calculated bioavailable testosterone. If a patient had multiple panels done during the study, we used the first panel for analysis. Men prescribed androgen therapy (testosterone or gonadotropin treatment) at initial testing were excluded from study.

The 3,672 men included in analysis tended to be middle-aged or older (mean \pm SEM age 59.7 ± 2.0 , range 20 to 98), white and obese (table 1). We abstracted the electronic medical record for ICD-9-CM inpatient and outpatient diagnostic codes, pharmacy and laboratory data for conditions such as diabetes mellitus, glucocorticoids and prescription opiates (United States Food and Drug Administration schedule II only) that commonly perturb SHBG. About a third of the men had diabetes mellitus, defined as a recorded ICD-9 code, hemoglobin A1c greater than 6.5%, or use of insulin or another glucose lowering drug. Almost half of them were obese, defined as a body mass index of greater than 30 kg/m^2 (table 1). Within 90 days of the testosterone panel, about half of the men were prescribed potent opioids and about a tenth were prescribed a systemic glucocorticoid.

All data were abstracted and aggregated by a single research coordinator. All information was de-identified to expunge any potential relationship between data and specific individuals. The Veterans Administration Puget Sound Research and Development Committee, and the human subjects review committee approved the study. An informed consent waiver was authorized.

Hormone Assays

TT was determined using the Elecsys[®] monoclonal antibody test kit. The measuring range for the TT assay was 2.0 to 1,500.0 ng/dl (normal range 280 to 800). SHBG was determined using the Cobas[®] electrochemiluminescence immunoassay kit. The measuring range of the SHBG assay was 0.35 to 200 nmol/l (normal range 10 to 80). Albumin was determined using the VITROS[®] 5,1 analytic platform assay. The measuring range of the albumin assay was 1 to 6 gm/dl (normal range 3.5 to 2). FT was calculated using the Vermeulen formula and the normal range was 34 to 194 pg/ml.⁶ All TT and cFT levels were determined at a central laboratory.

Statistical Analysis

We used standard descriptive statistics to characterize the patients. We generated ROC curves to determine the performance of TT measurements to predict low cFT (less than 34, 35, 36, 37, 38, 39 and 40 pg/ml, respectively). Using the ROC for TT as the continuous variable vs cFT less than 34 pg/ml, which was the lower limit of normal for this assay, we determined the sensitivity, specificity, and negative and positive LR for the TT thresholds 100, 150, 200, 250, 280, 300, 350, 400, 450 and 500 ng/dl.

cFT served as the benchmark for the biochemical confirmation of hypogonadism. We determined the discrimination of low TT for low cFT using a nonparametric method

computation of AUROC. As the primary analysis, we used cFT less than 34 pg/ml to define hypogonadism. We performed additional AUROC analyses with Stata[®] 11.0 for Macintosh[®] using higher cFT thresholds of less than 35, 36, 37, 38, 39 and 40 pg/ml, respectively.

RESULTS

ROC Curves

Overall TT performed well as a predictor of cFT. The ROC with TT as the continuous variable vs cFT less than 34 pg/ml (lower limit of normal for our assay) had an AUROC of 0.93, indicating that TT was generally a good predictor of cFT. The ROC for TT vs cFT less than 35, 36, 37, 38, 39 and 40 pg/ml was similar at 0.93, 0.93, 0.92, 0.92, 0.92 and 0.92, respectively.

Sensitivity and Specificity

Although the ROC indicated that TT generally predicted cFT, 61.7% of all patients with low TT had normal cFT (fig.1). The prevalence of low cFT was 15.2% and 2.1% of the men had low cFT but normal TT.

We used the ROC to determine the sensitivity and specificity of clinically useful thresholds of TT to exclude and predict low cFT, respectively (table 2). Using the assay lower limit of normal (less than 280 ng/dl) the sensitivity of TT to identify all patients with low cFT was only 91%. The sensitivity of low TT did not exceed 95% until threshold TT was defined as less than 350 ng/dl.

The specificity of TT measurement for low cFT was generally low with only 73.7% specificity at the lower limit of normal of the TT assay (less than 280 ng/dl). Low cFT specificity did not exceed 95% until threshold TT was defined as less than 150 ng/dl.

Likelihood Ratios

Table 3 shows the positive and negative LRs of low cFT at a specific TT threshold. The positive LR for biochemical hypogonadism, defined as low cFT, exceeded 10 only when TT was less than 200 ng/dl. TT below the lower limit of normal (less than 280 ng/dl) had a positive LR of only 3.8. On the other hand, the negative LR of TT above the lower limit of normal (280 ng/dl or greater) predicted normal cFT fairly strongly (negative LR 0.12 for low cFT). TT 400 ng/dl or greater had a robust negative LR of less than 0.04, indicating a low likelihood of low cFT.

DISCUSSION

In a large cohort of men evaluated for hypogonadism we assessed the performance of TT measurement to predict low cFT. Based on ROC curves TT and cFT generally correlated well. Also, as defined by low cFT, the sensitivity of low TT to rule out biochemical hypogonadism was fair at 91%. Because hypogonadism is a syndrome with clinically important sequelae, it is necessary to use a threshold TT of 350 or 400 ng/dl when the sensitivity of the test to exclude hypogonadism is 97% to 98%.

Although TT measurement using a threshold of 350 to 400 ng/dl appears to be a sensitive initial screening test to rule out hypogonadism, it has poor specificity to predict biochemical hypogonadism, as defined by low cFT. Hypogonadism is a syndrome that is usually treated for years with androgen replacement therapy. Before subjecting a man to the inconvenience, expense and potential risks of androgen replacement therapy, it is critical to use an accurate test with high specificity to confirm biochemical evidence of hypoandrogenemia in a man with clinical symptoms and signs of hypogonadism. In our cohort of unselected men in whom a clinician suspected clinical hypogonadism the specificity of low TT (less than 280 ng/dl) to rule in low cFT was only 74%. Lowering the TT threshold to 150 or 200 ng/dl increased specificity to 93% and 99% but significantly decreased sensitivity to 59% and 77%, respectively.

The first clinical implication of our study is that TT measurement is an appropriate initial test in a man who might have hypogonadism. TT that exceeds the lower limit of normal (greater than 280 to 300 ng/dl) has a negative LR of 0.1 to 0.12 for low cFT and it decreases the pretest probability of hypogonadism by 30% to 45%, enough to exclude hypogonadism in most men.^{10,11} TT greater than 350 or 400 ng/dl has a negative LR of 0.04 to 0.06 and excludes the diagnosis of hypogonadism in virtually all men.

The second clinical implication of our study is that in populations of men with a high prevalence of obesity, diabetes mellitus and conditions that affect SHBG, low TT does not reliably predict biochemical hypogonadism. Thus, TT cannot be used to make or confirm the definitive diagnosis of hypogonadism in many men. Because low TT confers a positive LR of 3.5, increasing the pretest probability by about 15% for low cFT, it has modest clinical value. However, it is not until the TT threshold of less than 200 ng/dl that LR exceeds 10, increasing the pretest probability by about 45%.^{11,12} Therefore, in a man with a high probability of hypogonadism, eg one with small prepubertal testes and eunuchoid proportions, low TT would suffice to make the diagnosis of hypogonadism. However, in the more common clinical scenario of a man with one or more nonspecific symptoms or signs of hypogonadism, eg decreased libido and osteoporosis, TT between 200 and 300 ng/dl does not reliably predict low FT and is insufficient to confirm the diagnosis of hypogonadism.

The strengths of our study include the large number of men evaluated for possible hypogonadism, their wide age range, measurement of all samples with a single, highly accurate methodology and using a central laboratory to determine all testosterone and cFT levels. In a similar patient population our specific method of cFT measurement correlated well with the gold standard for FT, which is equilibrium dialysis followed by tandem mass spectroscopy.^{9,10}

A weakness of our study is the absence of data on the true prevalence of hypogonadism in our cohort. Without detailed information on symptoms and signs of hypogonadism in our cohort, we do not know how well TT measurement performed to predict hypogonadism, defined as clinical evidence of hypogonadism plus low androgen. Nevertheless, our study reflects the practical reality that clinicians often diagnose a syndrome such as hypogonadism based on some clinical suspicion of the syndrome plus an abnormal result of a laboratory test thought to be specific for the condition.

Another study limitation is that all serum hormones were not measured in the early morning, which is the optimal time for testing.^{4,5} This limitation does not undermine the principal finding of our study that TT measurement has low specificity to predict low cFT.

An additional limitation is that testosterone assays are notoriously variable.⁵ This variability could limit the extrapolation of our results to men in the general population, in which a variety of TT assays is used. However, if a clinician uses a TT assay that conforms to current guidelines and has a lower limit of normal of 280 to 300 ng/dl, our conclusions are broadly applicable.

Finally, there is controversy about whether FT is the gold standard to assess male gonadal status.¹³ The FT hypothesis holds that testosterone bound to SHBG is not bioavailable to tissue. This hypothesis is based on sparse clinical evidence but it is widely accepted as the benchmark for biochemical evidence of hypogonadism.¹³ Assuming that the FT hypothesis is correct, our data indicate that low TT has low specificity for diagnosing hypogonadism.

Our findings indicate that accurate, precise measurement of FT is critical to confirm the diagnosis of hypogonadism in men with modestly decreased TT. Since validated methods to ascertain cFT are not offered at many clinical laboratories, it is critical that accurate methodologies become widespread.¹⁴

CONCLUSIONS

If interpreted correctly, TT measurement is a good initial test to evaluate possible male hypogonadism. Although normal TT (greater than 280 to 300 ng/dl on an accurate, precise assay) significantly lowers the likelihood of hypogonadism, a level between 280 and 350 ng/dl is not sensitive enough to reliably exclude hypogonadism. TT must exceed 350 to 400 ng/dl to reliably predict normal FT.

TT measurement is a poor predictor of low FT. Only when TT is low (less than 150 ng/dl) are specificity and LR high enough to confirm the diagnosis of hypogonadism in a symptomatic man. Although serum TT measurement is a good initial test to assess for possible male hypogonadism, accurate assessment of serum FT is necessary in most men to biochemically confirm the diagnosis of male hypogonadism.

Acknowledgments

Study received approval from the Veterans Administration Puget Sound Research and Development, and Human Subjects Review committees.

Supported by the National Institute of Child Health through Cooperative Agreements U54 HD-12629 and U54-HD-42454 (BDA), and the Department of Veterans Administration (TJW, AMM).

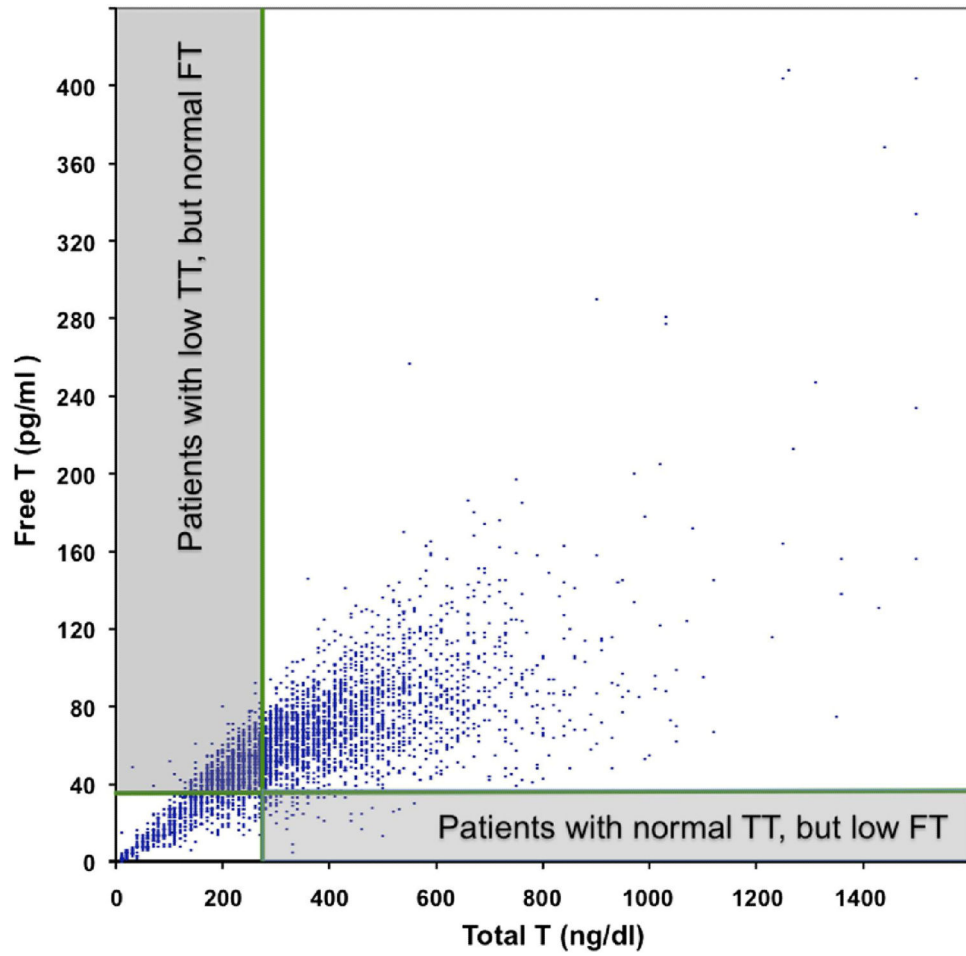
Abbreviations and Acronyms

AUROC	area under receiver operating characteristic
cFT	calculated FT
FT	free testosterone

LR	likelihood ratio
SHBG	sex hormone-binding globulin
TT	total testosterone

REFERENCES

1. Araujo AB, Esche GR, Kupelian V et al. : Prevalence of symptomatic androgen deficiency in men. *J Clin Endocrinol Metab* 2007; 92: 4241. [PubMed: 17698901]
2. Araujo AB, O'Donnell AB, Brambilla DJ et al. : Prevalence and incidence of androgen deficiency in middle-aged and older men: estimates from the Massachusetts Male Aging Study. *J Clin Endocrinol Metab* 2004; 89: 5920. [PubMed: 15579737]
3. Harman SM, Metter EJ, Tobin JD et al. : Longitudinal effects of aging on serum total and free testosterone levels in healthy men. Baltimore Longitudinal Study of Aging. *J Clin Endocrinol Metab* 2001; 86: 724. [PubMed: 11158037]
4. Bhasin S, Cunningham GR, Hayes FJ et al. : Testosterone therapy in adult men with androgen deficiency syndromes: an Endocrine Society Clinical Practice Guideline. *J Clin Endocrinol Metab* 2010; 95: 2536. [PubMed: 20525905]
5. Rosner W, Auchus RJ, Azziz R et al. : Utility, limitations and pitfalls in measuring testosterone: an Endocrine Society Position Statement. *J Clin Endocrinol Metab* 2007; 92: 405. [PubMed: 17090633]
6. Vermeulen A, Verdonck L and Kaufman JM: Acritical evaluation of simple methods for the estimation of free testosterone in serum. *J Clin Endocrinol Metab* 1999; 84: 3666. [PubMed: 10523012]
7. Sartorius G, Ly LP, Sikaris K et al. : Predictive accuracy and sources of variability in calculated serum free testosterone estimates. *Ann Clin Biochem* 2009; 46: 137. [PubMed: 19225026]
8. Bhasin S, Cunningham GR, Hayes FJ et al. : Testosterone therapy in adult men with androgen deficiency syndromes: an Endocrine Society Clinical Practice Guideline. *J Clin Endocrinol Metab* 2006; 91: 1995. [PubMed: 16720669]
9. DeVan ML, Bankson DD and Abadie JM: 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? *Am J Clin Pathol* 2008; 129: 459. [PubMed: 18285270]
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. [PubMed: 15961553]
11. McGee S: Simplifying likelihood ratios. *J Gen Intern Med* 2002; 17: 647.
12. Grimes DA and Schulz KF: Refining clinical diagnosis with likelihood ratios. *Lancet* 2005; 365: 1500. [PubMed: 15850636]
13. McLachlan RI: Certainly more guidelines than rules. *J Clin Endocrinol Metab* 2010; 95: 2610. [PubMed: 20525908]
14. Vesper HW and Botelho JC: Standardization of testosterone measurements in humans. *J Steroid Biochem Mol Biol* 2010; 121: 513. [PubMed: 20302935]



Of men with low serum TT 61.7% had normal serum cFT but only 2.1% with normal TT had low cFT. Only 38.3% of men with low serum TT had low cFT but 97.9% with normal TT had normal cFT. T, testosterone.

Table 1.

Patient characteristics

	% Pts
Race:	
White	59.4
Black	17.1
Asian	10.1
Native American	3.2
Unknown	10.2
Diabetes mellitus	32.6
Obesity	45.0
Prescription medication:	
Opioids	48.4
Glucocorticoids	9.4

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

Table 2.

Sensitivity and specificity of low TT for low cFT

Threshold TT (ng/dl)	% Sensitivity	% Specificity
Less than 100	33.9	99.9
Less than 150	58.8	98.9
Less than 200	76.5	92.6
Less than 250	87.6	81.9
Less than 280*	91.0	73.7
Less than 300	93.0	67.9
Less than 350	96.8	53.3
Less than 400	98.2	41.0
Less than 450	98.8	31.3
Less than 500	99.3	23.1

* Lower limit of normal.

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

Table 3.

TT positive LRs to predict and negative LRs to exclude low cFT

Threshold TT (ng/dl)	LR*
Pos (less than):	
100	339
150	53.5
200	10.3
250	4.8
280 [†]	3.5
300	2.9
350	2.1
400	1.7
450	1.4
500	1.3
Neg (or greater):	
100	0.66
150	0.42
200	0.25
250	0.15
280 [†]	0.12
300	0.10
350	0.06
400	0.04
450	0.04
500	0.03

Multiply testosterone concentration by 0.0347 to convert to nmol/l.

* Higher vs lower positive LR indicates greater vs decreased likelihood of low calculated cFT.

[†] TT lower limit of normal was 280 ng/dl.