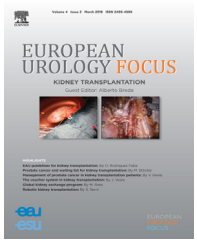
available at www.sciencedirect.com
journal homepage: www.europeanurology.com/eufocus



Grey Zone – Andrology

Current Diagnostic Criteria for Testosterone Deficiency Are Inadequate

Abraham Morgentaler *

Men's Health Boston, Beth Israel Deaconess Medical Center, Harvard Medical School, MA, USA

1. Introduction

It is remarkable that 80 yr since testosterone (T) therapy became available and 20 yr into the modern era of T therapy [1], there is still great uncertainty regarding diagnosis of T deficiency (TD; also known as hypogonadism).

This past week, I saw in my office, a 37-yr-old otherwise healthy man who had noticed loss of muscle mass over 2 yr, accompanied by decreased libido, unexplained fatigue, and depressed mood. His total T was 315 ng/dl (10.9 nmol/l). An endocrinologist told him that his T level was normal, and no treatment with testosterone was justified. Note that this value is considered normal by American Urological Association (AUA) guidelines (normal >300 ng/dl or 10.4 nmol/l) [2] and low by European Association of Urology (EAU) guidelines (normal >348 ng/dl or 12.1 nmol/l) [3]. Does he or does he not have a low T level? A second patient had characteristic symptoms of TD with total T of 290 ng/dl (10.1 nmol/l). Although this value was below the diagnostic threshold for nearly all published guidelines, it was above the normal reference range of 270 ng/dl (9.4 nmol/l) provided by the laboratory. He too was advised by his physician that his T level was normal, and he was thus not a candidate for treatment.

These cases highlight a common clinical problem. If a man's total T concentration is above a diagnostic threshold, does that mean there is no possibility of response to treatment? Should health care providers trust their clinical acumen or should they rather trust a blood test result? The truthful answer is that we do not know what the likelihood is that he might respond well to treatment because the studies necessary to evaluate this have not been performed. Our diagnostic criteria for TD are inadequate.

It is important to recognize that there is no clear threshold for total T that reliably separates men who have the condition from men who do not [4], or that predicts who will or will not respond to treatment. The consequence of this grey zone within andrology is that we deprive some men of treatment who may benefit while exposing other men to unnecessary risks of treatment. Moreover, we confuse clinicians and frustrate patients when there is lack of clarity on such a fundamental issue.

2. The current status of diagnostic criteria

For several decades after T therapy became available in the 1930s, blood testing for hormones was laborious and not readily available [1]. Treated cases had obvious deficiencies of T based on clinical grounds, such as low or absent libido, decreased body hair or beard, erectile dysfunction, "soft" bodies with poorly defined musculature, and small testes. These cases were rare and primarily consisted of men with major abnormalities, such as pituitary tumors, anorchia or atrophic testes, or genetic issues such as Klinefelter syndrome. These diagnoses were made almost exclusively on clinical presentation and not on laboratory results. In contrast, the vast majority of cases of TD currently are diagnosed in men without major medical conditions based mainly on blood tests for testosterone. There is an urgent need to define what is a normal level is and what constitutes a low level to accurately diagnose TD.

This has been a challenge. In the United States (US), the Food and Drug Administration applied a threshold of

* Corresponding author. Department of Urology, Beth Israel Deaconess Medical Center, Harvard Medical School, 200 Boylston St, Suite A309, Massachusetts, Chestnut Hill, MA 02467, USA. Tel.: +617 277 5000; Fax: +617 277 5444.
E-mail addresses: amorgent@bidmc.harvard.edu, amorgent@yahoo.com, dr.morgentaler@menshealthboston.com.

https://doi.org/10.1016/j.euf.2018.08.014

Please cite this article in press as: Morgentaler A, Current Diagnostic Criteria for Testosterone Deficiency Are Inadequate. Eur Urol Focus (2018), https://doi.org/10.1016/j.euf.2018.08.014

300 ng/dl (10.4 nmol/l) for evaluating new testosterone formulations indicated for the treatment of hypogonadism, and this was the value recommended by the first set of Endocrine Society guidelines in 2006 [4] and the new AUA guidelines [2]. The EAU and other international groups recommend 348 ng/dl (12.1 nmol/l) [3]. Some experts use 400 ng/dl (13.9 nmol/l) based on clinical experience. Which threshold is correct or most accurate? We do not know. Perhaps, we are not even asking the right question. Rather than attempting to determine whether a man has a normal or abnormal T concentration, it seems more reasonable to ask: What is the likelihood that a symptomatic man with a particular T value will respond to treatment?

One attempt to provide a scientific basis for a diagnostic threshold was to use young healthy men as a reference population. Data from the Framingham Heart Study in non-obese men aged ≤40 yr revealed the 2.5th percentile for total T to be 348 ng/dl (12.1 nmol/l), coincidentally identical to the value recommended by the EAU [5]. The 2018 version of the Endocrine Society guidelines proposed using "harmonized reference ranges" with a 2.5th percentile value of 264 ng/dl (9.2 nmol/l) [6]. This harmonized reference range required mutual adjustment of four datasets further calibrated downward by standards at the Center for Disease Control in the US [7]. The clinical utility of any of these reference ranges is unknown.

Wu et al [8] used responses to clinical questionnaires and blood test results from the European Male Aging Study (EMAS) to help define the condition. The best statistical fit was obtained with the combination of reduced erection quality, libido, and nocturnal erections, combined with a serum T concentration of 11 nmol/L (316 ng/dl) and calculated free T (cFT) less than 220 pmol/l (64 pg/ml).

One of the most confusing sources for determination of what is a normal level is laboratory-provided reference ranges due to enormous variation from one laboratory to another. In one survey of 25 laboratories in the US, 17 provided different reference ranges, with 300% variation from the lowest level categorized as "normal" to the highest [9]. This means that the same test result will be categorized as low by one laboratory and within the normal range by another.

3. Why has it been so hard to establish a threshold?

Any clinician with more than passing experience with T therapy has observed that the characteristic symptoms of TD occur over a wide range of serum T concentrations in men. There are several possible factors:

1. There is great inter-individual variation in biology between men with T, so that one man with a total T concentration of 10 nmol/l is markedly symptomatic and another with the same level is completely asymptomatic.
2. It has been postulated that the magnitude of the decline of T concentrations in a given individual matters more than the absolute value. This is an interesting concept, but entirely speculative. To assess this possibility, we would require knowledge of serum T concentrations in men from 20 to 30 yr earlier.
3. There is evidence that the androgen receptor has varying degrees of sensitivity determined by the number of CAG repeats in its gene [10].
4. Total T is the wrong test

4. The problem with total testosterone

Testosterone is a lipophilic molecule; in its free form, it is able to pass easily through the lipid bilayer membrane of all cells. However, in humans, 40–60% of circulating testosterone is tightly bound to its carrier molecule, sex hormone-binding globulin (SHBG). This tight binding renders this testosterone fraction unavailable to cells. Higher SHBG concentrations result in higher total T concentrations, even though FT or bioavailable T may be low. In a study of 1000 consecutive male patients seen at a men's health center, the range of concentrations varied nearly 20-fold, and 5.6% had elevated levels [11]. Similar results were seen for men aged <55 yr. Using an online calculator for FT, for any given total T value, the cFT will drop by one-half by raising SHBG concentration from the low end of normal (20 nmol/l) to the upper end (60 nmol/l). This makes total T an unreliable test of androgen status.

Antonio et al [12] investigated how well total T and cFT performed with regard to their association with clinical symptoms of TD based on data from EMAS. Men with low cFT but normal total T demonstrated a high rate of characteristic symptoms of TD. In contrast, men with normal cFT and low total T had almost no symptoms. Moreover, symptomatic men with total T above the threshold appear to respond symptomatically as well to T therapy as men with total T below the threshold, as long as free T values were low [13]. Simply, symptoms and response to treatment correspond to free T concentrations and not to total T.

Yet, nearly all clinical guidelines recommend total T thresholds as the primary guide to the biochemical diagnosis of TD. The EAU guidelines [3] state: "In cases with discrepancy between T levels and symptoms, FT levels should be analyzed." The Endocrine Society recommends that FT should be measured primarily in cases when there is suspected abnormality of SHBG [6]. The AUA fails to mention FT at all [2].

Why do these various guidelines continue to recommend total T despite its insensitivity? Because this is how it has been done for a long time. We need a new approach.

5. Other diagnostic rules

Not only are total T thresholds arbitrary but there are also many rules about measurements: test results should be repeated on different days; some guidelines recommend fasting; the specimen must be obtained in the early morning even for men aged >40 yr for whom there is minimal diurnal variation [14]. These might be all reasonable steps

Please cite this article in press as: Morgentaler A, Current Diagnostic Criteria for Testosterone Deficiency Are Inadequate. Eur Urol Focus (2018), https://doi.org/10.1016/j.euf.2018.08.014

if shown to improve accuracy. However, there is no evidence that any of these steps influence the reliability of diagnosis and the likelihood of symptomatic response to treatment.

6. What studies are needed?

The earliest bioassay in medicine was the rooster's coxcomb , used to help with the isolation of testosterone in the 1930s since it grew in the capon (castrated rooster) upon injections of various testicular extracts [1]. What remains to be determined is whether we can identify a human biomarker that can serve as an indicator of androgen status. It is detrimental to the health of our patients for an arbitrary level of total T to remain as a rigid threshold. Indeed, symptoms may be more reliable than total T.

Based on 40 yr of research and clinical experience with T [15], I believe that FT is the most accurate indicator of a man's androgen status, and I will offer treatment to a symptomatic man with low FT, regardless of his total T value. In nearly all cases, a discrepancy between total T and FT values is explained by a generous SHBG concentration. Future research is required to investigate this more thoroughly and prospectively.

Calculated FT is the most practical measurement of FT for most clinicians, requiring only total T and SHBG concentrations, and access to an online calculator. Albumin is usually included in the equation; however, it has little influence on cFT results, and most calculators provide a constant value for convenience. Direct measurement of FT, often called the analog assay, has an excellent correlation with both equilibrium dialysis (EqD) and cFT [16]; however, it requires a different scale than cFT or EqD due to lower numerical values.

A key study to be performed is to take a population of men with a defined symptom or set of symptoms characteristic of TD and expose them to treatment, regardless of baseline T levels. The goal of such a study is to determine the likelihood of symptomatic response based on baseline concentrations of androgen tests, which at a minimum should include total T, FT, and bioavailable T. The most predictive of these will be the most clinically useful. Clinicians would do well to think about T levels providing information as to the likelihood of symptomatic response rather than "normal" versus "abnormal".

There will be great value for the world of andrology if the diagnosis of TD came to be based on solid evidence, rather than rigid adherence to arbitrary thresholds for unreliable tests.

Conflicts of interest: The authors have nothing to disclose.

References

[1] Morgentaler A, Traish A. The history of testosterone and the evolution of its therapeutic potential. Sex Med Rev 2018;13:. http://dx.doi.org/10.1016/j.sxmr.2018.03.002, pii: S2050-0521(18)30040-4.
[2] Mulhall JP, Trost LW, Brannigan RE, et al. Evaluation and management of testosterone deficiency: AUA Guideline. J Urol 2018;200:423–32.
[3] Dohle GR, Arver S, Bettocchi C, Jones TH, Kliesch S. European Association of Urology guidelines: Male Hypogonadism. uroweb.org/guideline/male-hypogonadism/. Accessed July 23, 2018.
[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 2006;91:1995–2010.
[5] Bhasin S, Pencina M, Jasuja GK, et al. Reference ranges for testosterone in men generated using liquid chromatography tandem mass spectrometry in a community-based sample of healthy nonobese young men in the Framingham Heart Study and applied to three geographically distinct cohorts. J Clin Endocrinol Metab 2011;96:2430–9.
[6] Bhasin S, Brito JP, Cunningham GR, et al. Testosterone therapy in men with hypogonadism: an endocrine society clinical practice guideline. J Clin Endocrinol Metab 2018;103:1715–44.
[7] Travison TG, Vesper HW, Orwoll E. Harmonized reference ranges for circulating testosterone levels in men of four cohort studies in the United States and Europe. J Clin Endocrinol Metab 2017;102:1161–73.
[8] Wu FC, Tajar A, Beynon JM, et al. Identification of late-onset hypogonadism in middle-aged and elderly men. N Engl J Med 2010;363:123–35.
[9] Lazarou S, Reyes-Vallejo L, Morgentaler A. Wide variability in laboratory reference values for serum testosterone. J Sex Med 2011;3:1085–9.
[10] Zitzmann M. The role of the CAG repeat androgen receptor polymorphism in andrology. Front Horm Res 2009;37:52–61.
[11] Krakowsky Y, Conners W, Morgentaler A. Serum concentrations of sex hormone-binding globulin vary widely in younger and older men: clinical data from a men's health practice. Eur Urol Focus 2017;3:. http://dx.doi.org/10.1016/j.euf.2017.05.007, pii: S2405-4569(17)30126-8.
[12] Antonio L, Wu FC, O'Neill TW, et al. Low free testosterone is associated with hypogonadal signs and symptoms in men with normal total testosterone. J Clin Endocrinol Metab 2016;101:2647–57.
[13] Reyes-Vallejo L, Lazarou S, Morgentaler A. Subjective sexual response to testosterone replacement therapy based on initial serum levels of total testosterone. J Sex Med 2007;4:1757–62.
[14] Crawford ED, Barqawi AB, O'Donnell C, Morgentaler A. The association of time of day and serum testosterone concentration in a large screening population. BJU Int 2007;100:509–13.
[15] Morgentaler A. Controversies and advances with testosterone therapy: a 40-year perspective. Urology 2016;89:27–32.
[16] Kacker R, Hornstein A, Morgentaler A. Free testosterone by direct and calculated measurement versus equilibrium dialysis in a clinical population. Aging Male 2013;16:164–8.

Please cite this article in press as: Morgentaler A, Current Diagnostic Criteria for Testosterone Deficiency Are Inadequate. Eur Urol Focus (2018), https://doi.org/10.1016/j.euf.2018.08.014