Washington State Health Care Authority

Health Technology Assessment

Testosterone Testing

Final Evidence Report

February 6, 2015

Health Technology Assessment Program (HTA)
Washington State Health Care Authority
PO Box 42712
Olympia, WA 98504-2712
(360) 725-5126
hca.wa.gov/hta
shtap@hca.wa.gov

Testosterone Testing

A Health Technology Assessment
Prepared for Washington State Health Care Authority

Final REPORT

February 6, 2015

Acknowledgement
This report was prepared by:

Hayes, Inc.
157 S. Broad Street Suite 200
Lansdale, PA 19446
P: 215.855.0615 F: 215.855.5218

Table of Contents

List of Tables

EVIDENCE SUMMARY

The **EVIDENCE SUMMARY** summarizes background information, the methods and search results for this report, findings with respect to the Key Questions, and payer policies and practice guidelines. The **EVIDENCE SUMMARY** also includes conclusions and an assessment of the quality of the evidence for each Key Question. In general, references are not cited in the **EVIDENCE SUMMARY**. The **EVIDENCE SUMMARY** ends with an **Overall Summary and Discussion**. The **TECHNICAL REPORT** provides additional detail, with full citation, regarding background information, study results, and payer policies and guidelines.

Summary of Clinical Background

Low Testosterone and Hypogonadism

A recent analysis of the *general* population of men in the United States reported the following estimates of the prevalence of low testosterone regardless of symptoms that might signal clinical syndrome: 9.0% in men aged 45 to 54 years, 16.5% in men aged 55 to 64 years, and 18.3% in men aged 65 to 74 years. These estimates were derived from the National Health and Nutrition Examination Survey III (NHANESIII), which defined low testosterone levels as < 300 nanograms per deciliter (ng/dL) (10.4 nanomoles per liter [nmol/L]).

Hypogonadism is defined as a clinical syndrome resulting from a failure of the testis to produce physiological levels of testosterone and/or a normal number of spermatozoa. The causes of hypogonadism represent disruption of 1 or more levels of the hypothalamic-pituitary-testicular axis.

Primary hypogonadism is caused by abnormalities at the testicular level, whereas secondary hypogonadism is caused by dual defects in both the testis and the pituitary. **Table 1** lists possible causes of primary and secondary hypogonadism. The correct classification of hypogonadism as primary or secondary is relevant since impairment of spermatogenesis (infertility) can be corrected in patients with secondary hypogonadism but not in most patients with primary hypogonadism. Furthermore, secondary hypogonadism can indicate a pituitary tumor, other disorders related to the pituitary gland, genetic disorders, or systemic illness.

Age-Related Hypogonadism (Symptomatic Androgen Deficiency)

Low serum testosterone alone does not constitute a diagnosis of androgen deficiency or clinical hypogonadism. Diagnosis of a clinical condition requires the presence of certain characteristic symptoms as well as abnormally low serum testosterone. The literature generally distinguishes between hypogonadism diagnosed on the basis of signs and symptoms associated with aging (plus low serum testosterone levels) and hypogonadism due to a disorder (congenital or acquired) of the hypothalamus, pituitary, or testis. The latter category has been referred to by some experts as “organic” hypogonadism. *Symptomatic androgen deficiency*, or simply *androgen deficiency*, are other terms often used in place of age-related hypogonadism, which is a form of secondary hypogonadism.

Whether age-related hypogonadism represents a pathological condition per se is a matter of controversy. The prevalence of symptomatic androgen deficiency depends not only on age, but also on which symptoms are considered for the diagnosis and the cutoff value assumed for defining normal serum testosterone levels. There is no definitive cutoff value for normal testosterone. Serum levels for men of all ages are typically compared with reference ranges for young men, which suggest cutoff values of 280 to 300 ng/dL (9.7 to 10.4 nmol/L). There is also no consensus on a standard symptom profile. **Table 2** lists some terms that are sometimes used to describe age-related decline in serum testosterone, along with the inadequacies of those terms.

Characteristic symptoms of age-related hypogonadism are not highly specific and may reflect any of variety of clinical factors. The latest guideline on testosterone therapy from the Endocrine Society includes a consensus- and experience-based list of the symptoms and signs suggestive of androgen deficiency (see **Table 3**). These include specific symptoms and signs such as reduced libido, very small or shrinking testes, and signs of osteoporosis (height loss, low trauma fracture). The list also includes less specific symptoms and signs such as decreased energy, depressed mood, and increased body fat or body mass index (BMI).

Epidemiological studies have demonstrated that testosterone levels decline with age at an estimated rate of 1% to 2% per year. A substantial proportion of older men have levels below the lower limit of the normal range for young, healthy men. By one estimate, 30% of men in their seventies have total testosterone values in the abnormally low range and 50% of men in their seventies have free testosterone values in the abnormally low range, irrespective of symptoms of androgen deficiency, when cutoff points are based on normal reference ranges for young men. Population studies have estimated the prevalence of *symptomatic* androgen deficiency in the general population of middle-aged to elderly men to be 2% to 6% (see **Table 4**).

Screening Tools for Diagnosing Androgen Deficiency Without Testosterone Testing

Self-report case detection instruments and structured interview formats may be used to identify individuals who could be classified as having androgen deficiency on the basis of symptoms alone. These tools are not highly sensitive or specific. The details of these tools are presented in **Appendix IV.**

Selecting Patients for Testosterone Testing

Guidelines recommend against screening in general populations. The purpose of testing for and treating low serum testosterone may be to improve symptoms associated with low testosterone per se (e.g., erectile dysfunction or fatigue) or to reduce the adverse consequences of the medical condition associated with low testosterone (e.g., to reduce cardiovascular events in men with heart failure or to improve glucose control in men with diabetes). For both purposes, the utility of testing depends on the ability to identify clinical populations in which testing would have high diagnostic yield. Thus, investigators have attempted to identify symptoms, signs, and medical conditions that characterize men most likely to have low levels of testosterone.

Evidence for Associations Between Low Testosterone Levels and Signs or Symptoms of Hypogonadism

The strength of association between the signs and symptoms described in **Table 3** and low serum testosterone, and the causal nature of those relationships, are uncertain. **Appendix III** presents the

results of observational studies designed to explore specific associations. The clearest association between low serum testosterone and symptoms is with symptoms of sexual dysfunction. At least 2 observational studies have demonstrated such associations. Four observational studies reported inconclusive evidence regarding poor health status, poor physical performance, or increased psychological symptoms as indicators of low levels of testosterone.

Given the uncertainty regarding the contribution of low serum testosterone to the symptoms associated with age-related hypogonadism, the Endocrine Society guidelines make only weak recommendations to test in presence of symptoms that are thought to be associated with hypogonadism. In many cases, testosterone testing should be preceded by other forms of testing such as a comprehensive metabolic panel, complete blood count, and thyroid-stimulating hormone (TSH) level. In severely obese patients complaining of fatigue, evaluation of obstructive sleep apnea may be warranted before testing for low testosterone.

Medical Conditions Raising Suspicion of Low Serum Testosterone

According to the Endocrine Society's guideline on testosterone therapy and other expert sources, testosterone measurement may be warranted in the following situations (Bhasin et al., 2010; McGill et al., 2012; Pantalone and Faiman, 2012):

- A mass in, radiation of, or disease of sellar region (a depression in the upper surface of the sphenoid bone in which the pituitary gland sits).*
- Use of medications that affect testosterone production or metabolism (e.g., glucocorticoids, anabolic steroids, or opioids).*
- Human immunodeficiency virus (HIV)-associated weight loss.*
- Osteoporosis or low trauma fracture (especially in a young man).*
- Type 2 diabetes.
- End-stage renal disease and maintenance hemodialysis.
- Moderate to severe chronic obstructive pulmonary disease (COPD).

* According to the Endocrine Society, these conditions warrant testing for low testosterone even in the absence of characteristic symptoms. Consideration of testing in men with the other conditions should be based on the presence of concomitant characteristic symptoms (see **Table 3**).

Pathology associated with the pituitary, which produces the hormones (luteinizing hormone [LH] and follicle-stimulating hormone [FSH]) that regulate testosterone production, can be expected to affect testosterone levels. Certain medications are known to suppress the hypothalamic-pituitary gonadal axis and small studies have suggested an association between androgen deficiency and opioid or glucocorticoid use. A systematic review identified 4 studies measuring the association between opioid use and testosterone levels; 3 studies showed an inverse relationship (higher opioid dose was associated with lower testosterone levels) and 1 study showed no relationship. No systematic reviews investigating the effects of glucocorticoids or other medications were identified.

The literature reviewed for the present report did not provide a biological rationale for the link between low testosterone and most of the chronic diseases in the Endocrine Society list of medical conditions

that may prompt testosterone testing. However, there is some evidence in support of certain links. Prevalence estimates suggest an association between androgen deficiency and chronic kidney disease, chronic obstructive COPD, and the effects of HIV infection or HIV treatment. Systematic reviews have shown low testosterone to be associated with osteoporosis, metabolic syndrome, type 2 diabetes, cardiovascular disease, and cardiovascular mortality. The most recent systematic reviews evaluating a link between endogenous testosterone and health conditions are summarized in **Appendix V**.

The Clinical Significance of Low Testosterone: Association of Low Serum Levels of Testosterone with Health Outcomes

The long-term *consequences* of low testosterone levels are largely unknown for the largest subpopulations in which low testosterone levels are common: older men and men with chronic illness. Some associations reflected in **Appendix V** suggest testosterone deficiency could be a predisposing factor to poorer health, but there is no consensus on the extent to which low serum testosterone is a true health problem. A meta-analysis of 11 studies computed a relative risk (RR) of 1.35 (95% CI, 1.13 to 1.62) for the association between lower versus higher testosterone levels and all-cause mortality. The size of the RR does not suggest a strong association (RR ≥ 2 or RR ≤ 0.5) and the pooled estimate had high statistical heterogeneity. Although low serum testosterone has been shown to be associated with some chronic diseases and symptom complexes, strong associations have not been demonstrated, except for the relationship with metabolic syndrome and sexual dysfunction. Some health conditions, such as obesity and type 2 diabetes, may be the *cause* rather than the result of low testosterone levels (see also the following section on **Environmental and Modifiable Lifestyle Influences on Testosterone Levels**). Some systematic review authors surmised that low testosterone may simply be a marker for concurrent, even occult, disease rather than a causative agent, even though a temporal relationship is suggested by some of the analyses. Other narrative and systematic review authors have acknowledged that current knowledge does not allow a conclusion as to whether improving testosterone levels could improve a comorbid chronic condition or improving the comorbidity could raise testosterone levels. For other indications such as sexual dysfunction, this ambiguity does not exist. Lastly, systematic reviews investigating the link between low testosterone and medical conditions have not included an adjustment for health behavior, an issue discussed in the next section.

Environmental and Modifiable Lifestyle Influences on Testosterone Levels

Epidemiological studies have demonstrated a decline in average circulating testosterone levels over recent decades, independent of age. Other evidence suggests several possible contributors to this secular trend: increases in BMI (adiposity is linked with diminished testosterone levels), smoking cessation, certain health behaviors, traumatic life events, environmental toxins, and changes in assay technology. Increasing polypharmacy is also a conceivable contributor. Studies that typify the evidence of a link (or lack thereof) between environmental or lifestyle factors and low testosterone levels are described in the **TECHNICAL REPORT**. One study of 3453 men aged 65 to 83 years found that a poorer lifestyle score, based on 8 health-related behaviors, was associated with lower levels of testosterone. It is noteworthy that 3 studies found a link between low testosterone and BMI, but the studies did not establish the direction of effect. The authors of the listed studies, as well as review articles, suggested promotion of healthy behaviors as a potential treatment for age-related declines in serum testosterone prior to consideration of medical treatment of low testosterone. Nutritional deficiencies, eating disorders, and excessive exercise can also contribute to low levels of testosterone.

Medical Treatment of Low Testosterone Levels

Testosterone replacement therapy is considered the medical treatment of choice when fertility is not required. Since serum testosterone down-regulates the hypothalamus-pituitary axis, exogenous testosterone actually inhibits sperm production. Reversible conditions should be addressed as possible causes of low serum testosterone and corrected where possible before initiating testosterone therapy. Testosterone therapy comes in various formulations: intramuscular, subcutaneous, transdermal, buccal, and oral. Different forms have different advantages and disadvantages.

Guidelines suggest that testosterone therapy can be offered for maintenance of secondary sex characteristics, improvement of sexual function, promotion of a sense of well-being, or correction of the effects of HIV treatment or high doses of glucocorticoids. The evidence of efficacy varies considerably by indication. Clinicians are advised to also consider erectile dysfunction as the reason for low libido and to consider established treatments (treatment other than testosterone therapy) for erectile dysfunction before offering testosterone therapy, especially in older men because of the risk of adverse treatment effects.

Efficacy of Testosterone Therapy

The Endocrine Society guideline considered the evidence of testosterone treatment efficacy for all indications to be of low quality. There is no evidence that the effectiveness of testosterone therapy varies by formulation. The benefits and safety of testosterone therapy in asymptomatic men have not been demonstrated.

See **Appendix VI** for a summary of the most recent systematic reviews of the efficacy of exogenous testosterone. NOTE: Some systematic reviews had not been published at the time the Endocrine Society guidelines were developed. Systematic reviews have demonstrated the efficacy of testosterone therapy in men with low serum levels of testosterone for only 2 specific indications. Findings from a fair-quality meta-analysis suggest that testosterone therapy is more effective for improving sexual function in hypogonadal men with sexual dysfunction than in the general population of men with sexual dysfunction. Another fair-quality meta-analysis suggested that testosterone therapy can improve some metabolic outcomes in hypogonadal men with type 2 diabetes. No other health benefits of testosterone therapy in hypogonadal men have been clearly demonstrated by systematic reviews. None of the systematic reviews discussed the clinical relevance of their pooled estimates.

Safety of Testosterone Therapy

According to published reports, adverse events in young, hypogonadal men occur at low frequency. An increase in hematocrit, acne and oily skin, fluid retention, testicular atrophy, worsening of lower urinary tract symptoms, subclinical prostate cancer, growth of metastatic prostate cancer, and reduced sperm production and fertility have been reported as adverse effects. Evidence from systematic reviews cited in the Endocrine Society guidelines suggested that testosterone therapy in older men may increase prostate events (nonsignificant estimate), substantially increases the risk of erythrocytosis, and has no clear effect on obstructive sleep apnea, cardiovascular events, or mortality. More recent evidence is summarized in **Appendix VII** and is generally applicable to men ≥ 60 years of age. The newer evidence is consistent with the Society's conclusions regarding the effect on prostate events, erythrocytosis, and sleep apnea, but the most recent pooled estimates for prostate events continue to be nonsignificant and

very imprecise. Newer evidence does not resolve the uncertainty about risk of cardiovascular events but does suggest that there may be an elevated risk in older men and men with a history of cardiovascular disease. Newer evidence regarding the effect of testosterone therapy on mortality suggests a *protective* effect, but this evidence comes from a single observational study (Shores et al., 2012) and cannot be considered conclusive. In response to 2 recently published, very large population studies showing an elevated cardiovascular risk in men receiving testosterone therapy (see **Appendix VII**), the Food and Drug Administration (FDA) announced a decision in 2014 to reassess the safety of FDA-approved testosterone products. Additionally, the FDA issued a new requirement in 2014 that manufacturers include a general warning in the drug labeling of all approved testosterone products about the risk of blood clots in the veins due to postmarket reports of venous blood clots unrelated to polycythemia (another term for erythrocytosis). Because these clots occur in the veins, this new warning is not related to the FDA's ongoing evaluation of the possible risk of stroke, heart attack, and death in patients taking testosterone products.

Contraindications: Given the observations of prostate and breast cancer growth and the fact that testosterone therapy is a hormonal therapy, testosterone therapy is contraindicated in men with either form of cancer. Although evidence from randomized trials is lacking, some experts believe that testosterone therapy is safe after treatment for prostate cancer when certain signs of remission or cure are evident. Guidelines advise that testosterone therapy should also be avoided in other conditions because of its known or potential ability to worsen the disease: elevated hematocrit (> 50%), untreated severe obstructive sleep apnea, severe lower urinary tract symptoms, and uncontrolled or poorly controlled heart failure.

Summary of Description of Testosterone Testing

Serum testosterone measurements can target total testosterone, free testosterone, or bioavailable testosterone. Most of the circulating testosterone is bound to sex hormone binding globulin (SHBG), while approximately 0.5% to 3% is *free*. *Bioavailable testosterone* refers to unbound testosterone plus testosterone that is bound loosely to albumin. Albumin-bound testosterone is assumed to be bioavailable. Serum testosterone measurements should begin with total testosterone and should take into account the considerable intraindividual variation in testosterone measurements. Levels vary within the day, and from day to day. Levels are typically approximately 30% higher between the hours of 8:00 a.m. and 10:00 a.m. compared with later in the day. Studies have also shown that approximately 30% of men with an initial testosterone level in the mildly hypogonadal range have a normal result on repeat measurement. Thus, serum testosterone measurements should be made early in the morning and testosterone therapy should not be considered without a diagnosis of androgen deficiency based on 2 measurements of total testosterone on different days. Lastly, testing during acute illness or during a time of decompensation of chronic illness is not advised since testosterone levels may be temporarily depressed during such times.

Guidelines recommend that free or bioavailable testosterone be measured when total testosterone levels are close to the lower limit of the normal range and when altered SHBG levels are suspected. The potential for an underestimation of low serum testosterone increases with increasing age if measurement is restricted to total testosterone. However, it is also possible for total testosterone to be low but free or bioavailable testosterone to be normal.

Analytic Validity

Analytic validity refers to the ability of a test to accurately detect the target analyte. It might be thought of as the technical accuracy of the test. The analytic validity of testosterone testing methods is generally hampered by the following: possible interference from similar analytes; differences in the reference populations used by different laboratories and lack of control for medical comorbidities; diurnal, seasonal, and age-related variations; and variation from day to day in the same individual. As already noted, some of the intraindividual variability in test results can be addressed by standardizing the time of day when blood samples are drawn (early morning) and requiring confirmation of low levels by drawing and testing blood samples on 2 different days.

An additional problem is the impracticality of using the most accurate methods in typical laboratories. Automated enzyme-linked immunoassays for measurement of *total* testosterone are available in most hospital laboratories. However, there is variation in results between different commercial assays and therefore variation in normal cutoff values even for the same reference population. Quality control programs sponsored by the Centers for Disease Control and Prevention (CDC) and the College of American Pathologists (CAP) allow laboratories to assess the comparability of their assay results with results obtained by gold standard method for a reference set of serum samples. Participating laboratories can then recalibrate their test results. Participation in these 2 quality control programs is voluntary and not required for accreditation. Experts consider automated assays to have sufficient accuracy to distinguish between men who have serum levels consistent with hypogonadism and men who have adequate total levels of serum testosterone even though precision is poor for measurements in men whose true levels are close to the lower limit of a normal range. Free and bioavailable testosterone can be calculated by various formulas on the basis of total testosterone and SHBG assays.

Clinical Validity

Clinical validity refers to the ability of a test to discriminate between individuals who do and do not have a clinical condition, or to make accurate clinical predictions. In other words, a test has clinical validity if it enables valid clinical decision making. No generally recognized age- and gender-specific cutoff points have been defined for making clinical predictions based on serum testosterone levels. In symptomatic men, regardless of age, testosterone levels are assessed by comparing them with the normal range for *young men*, based on the known gradual decline of testosterone levels with aging, starting at approximately age 30. The lower limit of *normal for healthy young men* is 280 to 300 ng/dL (9.7 to 10.4 nmol/L). There is no consensus regarding cutoff values for *free* or *bioavailable* testosterone, but a level of > 225 pmol/L (6 ng/dL) is generally considered normal. The validity of testosterone testing as a predictor of health depends on the target condition and which components of serum testosterone are measured.

Monitoring

There is no consensus on the optimal frequency for monitoring testosterone levels in men receiving testosterone therapy. Experts advise that clinical response as well as testosterone levels be monitored to determine whether adjustments are needed in the therapy.

Policy Context

Considerable controversy and uncertainty exist concerning the diagnostic criteria for hypogonadism, the cutoff value for normal testosterone, and the benefits and harms of treatment. In recent years, new topical formulations of testosterone have been developed and direct-to-consumer marketing may have contributed to the use of testosterone replacement therapy for milder and potentially nonpathological forms of hypogonadism. A recently published analysis of U.S. private and Medicare claims data demonstrated an increase in testosterone testing for the years 2000 to 2010, even though the prevalence of low testosterone levels remained constant. There was also an increase in the percentage of men with low testosterone levels who initiated therapy for the years 2007 to 2011 (Layton et al., 2014). A small minority of men with normal or high testosterone levels initiated therapy during the study period. A substantial proportion of those undergoing testing had conditions that have been shown in systematic reviews to be associated with low testosterone levels: erectile dysfunction (10.4%) and diabetes (15.1%). However, an even greater proportion of tested men were identified in claims records as having fatigue (19.8%), which has *not* been shown to be clearly indicative of low testosterone levels in aging men. Furthermore, the study also analyzed healthcare records from the United Kingdom and found significant differences between the 2 countries in patterns of testing and testosterone replacement therapy. The study's findings are summarized in **Table 5**.

Table 5. Summary of Findings, Testosterone Laboratory Testing and Initiation in the United States and the United Kingdom, 2000 to 2010 (Layton et al., 2014)

Selected Cohort Characteristics	Men with New Laboratory Testosterone Test*		Men Who Initiated Testosterone Therapy	
	United States	United Kingdom	United States	United Kingdom
# men	1,114,329	66,140	410,119	6858
Mean age (yrs)	50.2	52.7	52.1	54.1
Hypogonadism/low testosterone diagnosis (% of cohort)	9.7%	0.2%	39.9%	11.8%
Sexual dysfunction†	0.1%	47.4%	0.2%	23.6%
Erectile dysfunction drug use	10.4%	15.3%	14.6%	15.4%
Fatigue	19.8%	0.7%	20.4%	1.0%
Hypertension	28.7%	2.5%	32.7%	2.5%
Diabetes	15.1%	3.2%	19.6%	5.9%
Change in incidence of new testosterone test, 2000-2010	From 39.6 to	From 13.0 to 46.4		

Selected Cohort Characteristics	Men with New Laboratory Testosterone Test*		Men Who Initiated Testosterone Therapy	
	United States	United Kingdom	United States	United Kingdom
(# men per 10,000 person-yrs)	170.0			
Change in low testosterone results, 2000-2011 (% of cohort)	Remained constant (slightly >20%)‡	From 18.9% to 26.7% ‡		
Change in % of men with low testosterone levels§ who initiated testosterone therapy	From 36% to 43%, 2007-2011	Remained constant (~10%), 2000-2011		
% of men w/ normal or high testosterone levels who initiated testosterone therapy	Varied (4%-9%), (2007-2011)	Remained constant (~1%), 2000-2010		
% testosterone therapy recipients who did not have a testosterone measurement within 180 days prior			32.7%	53.8%

*The therapy initiation and test claim cohorts were overlapping.

†The authors believed the discrepancies between the 2 countries were likely due to undercoding since the prevalence of erectile dysfunction medication use was similar.

‡Missing data for 95.9% of U.S. men and 3.1% of UK men, but analysis showed no differences between men with and without data.

§Laboratory results were classified as *low*, *normal*, or *high* according to results flags or assay-specific reference ranges, as reported in the laboratory records. If results flags or reference ranges were not available, the study authors classified results as follows: low, < 300 ng/dL (10.4 nmol/L); normal, 300 to 849 ng/dL (10.4-25.4 nmol/L); and high, ≥ 850 ng/dL (29.5 nmol/L).

The Centers for Medicare & Medicaid Services (CMS) has no National Coverage Determination (NCD) applicable to testosterone testing or treatment of hypogonadism. The Medicare Benefit Policy Manual includes no statements on testosterone testing or hypogonadism. U.S. Preventive Services Task Force (USPSTF) has issued no recommendations on testosterone testing. A number of commercial assays for testosterone testing have been cleared for marketing by the FDA and are described in recent decision summaries as being appropriate for the diagnosis and treatment of disorders involving androgens, including primary and secondary hypogonadism. FDA approvals of testosterone products are restricted to use in men with low testosterone levels who also have an associated medical condition characterized by one of the following (CDER, 2014a):

1) Primary hypogonadism (congenital or acquired): testicular failure due to conditions such as cryptorchidism, bilateral torsion, orchitis, vanishing testes syndrome, orchiectomy, Klinefelter's syndrome, chemotherapy, or toxic damage from alcohol or heavy metals. These

men usually have low serum testosterone concentrations and gonadotropins (FSH, LH) above the normal range.

Or

2) Hypogonadotropic hypogonadism (congenital or acquired): idiopathic gonadotropin or luteinizing hormone-releasing (LHRH) deficiency or pituitary-hypothalamic injury from tumors, trauma, or radiation. These men have low testosterone serum concentrations but have gonadotropins in the normal or low range.

Minutes from the September 17, 2014, meeting of the FDA Bone, Reproductive and Urologic Drugs Advisory Committee & Drug Safety and Risk Management Advisory Committee include a recognition that even though product labeling does not include "age-related hypogonadism" or "aging," product use data show that the real-world population of testosterone therapy users consists of middle-aged men with "low T" (CDER, 2014a). Furthermore, the FDA recently announced a decision to reassess the safety of FDA-approved testosterone products (CDER, 2014b).

A review of the evidence on the risks and benefits of testing, taking into account what is known about the benefits of treatment, is warranted to guide coverage policy on testosterone testing.

Rationale for This Report

The foregoing sections show that most of the factors that would strengthen a case for testosterone testing have not been clearly established:

- A recognized case definition of hypogonadism that includes a definition of low serum testosterone: There is no consensus on a case definition of symptomatic androgen deficiency because of the lack of a standard symptom profile and the lack of a standard cutoff value for normal serum testosterone. However, certain organic causes of abnormal testosterone levels are not disputed, nor are the consequences of low testosterone levels in these situations.
- Screening tools, symptoms, or health conditions that can identify individuals at high risk of low testosterone levels:
    - General symptom scales have been developed for diagnosing hypogonadism in men without testosterone testing, but these tools have low specificity (and thus low positive predictive value) for low serum testosterone.
    - Factors such as health status, physical performance, and psychological symptoms have not been shown to be good predictors of low testosterone levels.
    - The simultaneous presence of multiple symptoms of sexual dysfunction have been shown to have a strong association with low testosterone in a large population study.
    - Certain pathological conditions raise suspicion of low testosterone because of the biological rationale or limited evidence suggesting elevated prevalence of low testosterone in patients who have those conditions, but clear associations in terms of risk ratios or sensitivity and specificity have not been reported for most such conditions, according to the literature reviewed for the present report. The clearest relationships are with diabetes and metabolic syndrome. However, the nature of the causal relationship between low testosterone levels and chronic disease remains uncertain.

- Analytic validity (technical accuracy) of testosterone testing methods: Lack of uniform reference ranges across laboratories presents some limitations on confidence in analytic validity. Although gold standard methods are impractical for most laboratories, the literature reflects a consensus that usual methods are adequate for distinguishing between hypogonadal and eugonadal men when considered in the context of signs and symptoms. The CDC and CAP administer national quality assurance programs.
- Clinical validity (clinical performance) of testosterone testing: Associations of low levels of serum testosterone with certain chronic diseases may suggest that testosterone testing could identify individuals at risk of disease or disease-related events. However, definitive cutoff values for identifying men with low enough testosterone levels to pose a clinically relevant overall health risk have not been established.
- The effectiveness of testosterone replacement therapy: There was evidence from a recent meta-analysis of RCTs that in men with symptoms of sexual dysfunction, testosterone therapy improves sexual function and is more effective in men with low testosterone levels at baseline. This finding implies that testosterone testing may have clinical utility for identifying men with sexual dysfunction who are most likely to benefit from testosterone therapy. There was also evidence from a meta-analysis of randomized controlled trials (RCTs) that testosterone therapy improves some but not all metabolic factors in men who have both hypogonadism and type 2 diabetes. However, another larger meta-analysis, published after the draft version of the present report, cast doubt on the effectiveness of testosterone therapy for diabetes-related outcomes. For other health conditions, the effectiveness of testosterone replacement therapy has either not been proven or a strong association between low endogenous testosterone and the indication for therapy has not been established.
- The safety of testosterone replacement therapy: Evidence regarding the intermediate- and long-term safety of testosterone therapy is conflicting. The potential for adverse events must be weighed against the potential benefits, especially in older men. Given the uncertainties associated with the meaning of low serum testosterone, the effectiveness of testosterone therapy, and the safety of testosterone therapy, there is a need for empirical evidence that clinical decision making informed by knowledge of testosterone test results confers health benefits without unacceptable adverse effects. The Key Questions for the present report were designed to guide a search for such empirical evidence.

See also the **Washington State Utilization and Cost Data.**

Summary of Review Objectives and Methods

Review Objectives

Population: Adult men

Interventions: Measurement of circulating total, free, or bioavailable testosterone as an initial assessment of possible hypogonadism

Comparisons: Investigation and clinical management of symptoms or health problems without the use of testosterone testing

Outcomes: Outcomes such as symptom improvement; general health outcomes (e.g., osteoporosis, chronic disease, mortality); clinical management decisions; potential harms resulting from testosterone treatment decisions; cost and cost-effectiveness.

Key Questions

1. Is there evidence that testosterone testing improves outcomes?

   1a. Does the impact on outcomes vary according to age, race/ethnicity, baseline testosterone levels, treatment status, or clinical history?

   1b. What is the minimum interval required to assess a change in testosterone status in untreated and treated individuals?

2. What are the potential harms of testosterone testing, including potential subsequent harms resulting from treatment decisions?
3. What are the costs and cost-effectiveness of testosterone testing?

Analytic Framework

See **TECHNICAL REPORT**, **Methods**.

Methods

In addition to the formal methods described in this section, guidance on search dates and potentially missed studies was sought through personal communication with a clinical expert. See the **Methods** section of the **TECHNICAL REPORT**, **Appendix I**, and **Appendix II** for additional details.

Preplanned Search Strategy and Selection Criteria

A topic-scoping search of core databases, PubMed, and the websites of relevant specialty societies was conducted on June 29, 2014, for systematic reviews, meta-analyses, economic evaluations, and practice guidelines published in the last 10 years. Systematic reviews were selected if they reviewed studies considered eligible for answering the Key Questions or if they provided useful background information. Initial searches for clinical studies were conducted 0n September 7, 2014 (PubMed) and September 10, 2014 (Embase). A PubMed search for cost studies and economic evaluations was conducted on September 25, 2014 (in addition to the June 29 search of the National Health Service Economic Evaluation Database [NHS-EED]). The bibliographies of review articles and practice guidelines were also searched.

Update searches for direct evidence pertaining to the Key Questions were conducted on December 3, 2014.

Inclusion/Exclusion Criteria for Direct Evidence Pertaining to the Key Questions

Inclusion Criteria

- Consistency with the Population-Interventions-Comparator-Outcomes (PICO) statement.

- Observational studies with a control or comparator group, nonrandomized trials, and randomized controlled or comparator trials (RCTs) designed to measure the effectiveness of testosterone testing (Key Questions #1, #1a), the effectiveness of repeat testosterone testing (Key Question #1b), or adverse events as consequences of testosterone testing (Key Question #2). Systematic reviews of such studies were also eligible.
- Longitudinal studies designed to assess the time required for a clinically relevant change in serum testosterone levels to occur (relevant to Key Question #1b).
- Any study reporting the cost outcomes of or an economic evaluation of testosterone testing (Key Question #3).

Exclusion Criteria

- Studies that were inconsistent with the PICO statement.
- Publications that did not meet the study designs specified in inclusion criteria.
- Studies published prior to 1990, identified as the date after which testing technologies have not changed substantially.
- Economic evaluations and cost studies published prior to September 2003 (studies published earlier than 10 years prior to the search date were considered to have minimal applicability to the current economy).

Post Hoc Search for Selected Recent Indirect Evidence

No studies with direct evidence relevant to the Key Questions were identified. For in-depth assessment of the most promising indirect evidence, subpopulations were identified where (a) systematic reviews or observational studies provided evidence of an association with low testosterone and reported findings in terms that allowed an assessment of clinical validity or strength of association *and* (b) systematic reviews provided positive evidence of treatment efficacy. Two subpopulations met these criteria:

- Men with type 2 diabetes or metabolic syndrome.
- Men with symptoms of sexual dysfunction.

A post hoc search (see **Appendix I** for details) was conducted to identify studies in the selected subpopulations (diabetes/metabolic syndrome, sexual dysfunction) that were published more recently than the observational research and systematic reviews cited in the **CLINICAL BACKGROUND** section. Studies were selected in accordance with the selection criteria used in the systematic reviews. The purpose of the post hoc search and analysis was to determine whether recent indirect evidence would add to or conflict with the indirect evidence identified initially for the background section of this technical report. See the METHODS section of the TECHNICAL REPORT for details regarding selection criteria for the more recent evidence.

Quality Assessment

The process used by Hayes for assessing the quality of primary studies and bodies of evidence is in alignment with the methods recommended by the Grading of Recommendations, Assessment,

Development, and Evaluation (GRADE) Working Group. Like the GRADE Working Group, Hayes uses the phrase *quality of evidence* to describe bodies of evidence in the same manner that other groups, such as the Agency for Healthcare Research and Quality (AHRQ), use the phrase *strength of evidence*. A tool created for internal use at Hayes was used to guide interpretation and critical appraisal of economic evaluations. This tool is based on best practices as identified in the literature and addresses issues such as the reliability of effectiveness estimates, transparency of the report, quality of analysis (e.g., the inclusion of all relevant costs, benefits, and harms), generalizability/applicability, and conflicts of interest. The Rigor of Development domain of the Appraisal of Guidelines Research and Evaluation (AGREE) tool, along with a consideration of commercial funding and conflicts of interest among the guideline authors, was used to assess the quality of practice guidelines. Use of the AGREE tool was limited to these areas because they relate most directly to the link between guideline recommendations and evidence. See the **Methods** section of the **TECHNICAL REPORT** and **Appendix II** for details on quality assessment methods.

Summary of Search Results

Systematic Reviews Identified in Topic-Scoping Searches

No systematic reviews designed to directly answer the Key Questions were identified.

Primary Studies with Direct Evidence Pertaining to the Key Questions

No primary studies designed to directly answer the Key Questions were identified.

Evidence of Clinical Associations with Testosterone Levels or the Benefits and Harms of Testosterone Therapy

The largest observational studies and most recent systematic reviews were selected and are summarized in the **CLINICAL BACKGROUND** section. The following selections were made:

- Seven (7) observational studies measuring the association between testosterone levels and signs or symptoms (see **Appendix III**).
- Nine (9) systematic reviews of the association between testosterone levels and health outcomes or conditions (see **Appendix V**).
- Seven (7) systematic reviews evaluating the effectiveness of testosterone therapy (see **Appendix VI**). An additional systematic review was published after the draft version of the present report was published and has been added to this final report.
- Six (6) systematic reviews evaluating the adverse events associated with testosterone therapy (see **Appendix VII**).

Recent Indirect Evidence Identified Through a Targeted Post Hoc Search

- 15 studies that would potentially meet the inclusion criteria of 3 recent systematic reviews evaluating the association between testosterone levels and metabolic syndrome or diabetes.
- 5 studies that would potentially meet the inclusion criteria of a recent systematic review evaluating the effectiveness of testosterone therapy in men with type 2 diabetes.

- 4 population studies for comparison of results with a large 2010 population study evaluating the association between low testosterone levels and symptoms of sexual dysfunction.
- No RCTs that would potentially meet the inclusion criteria of a recent systematic review evaluating the effectiveness of testosterone therapy for improving sexual function.

Practice Guidelines

Four relevant practice guidelines published by U.S. groups in the last 10 years were identified. An additional guideline from the American Diabetes Association was reviewed after the identification of subpopulations for the post hoc analysis.

Findings

Key Question #1: Is there evidence that testosterone testing improves outcomes? Key Question #1a. Does the impact on outcomes vary according to age, race/ethnicity, baseline testosterone levels, treatment status, or clinical history?

The literature search identified no studies that were designed to compare outcomes between individuals who were tested for low testosterone and similar individuals whose symptoms and health problems were managed without knowledge of serum testosterone levels. Furthermore, no studies with direct evidence were cited in the practice guidelines that were reviewed.

Indirect Evidence

Key Question #1

See the section on **POST HOC SEARCH AND ANALYSIS** for a more detailed and critical discussion of the indirect evidence regarding the utility of testosterone testing in men with sexual dysfunction or diabetes/metabolic syndrome. Indirect evidence is strongest for these 2 subpopulations.

The post hoc analysis has the following implications for Key Question #1: A positive but very weak inference regarding the utility of testosterone testing in men with sexual dysfunction can be drawn. This inference is based on evidence of a strong association of low testosterone levels with the simultaneous presence of multiple sexual symptoms. The inference also reflects evidence that testosterone therapy improves sexual symptoms in men who have sexual dysfunction and low testosterone levels. Although there is a strong association of low testosterone levels with metabolic syndrome and significant differences in serum testosterone between men with and without type 2 diabetes, the effectiveness of testosterone therapy for improving outcomes related to type 2 diabetes remains uncertain.

Key Question #1a

Two key findings from the post hoc analysis are particularly pertinent to Key Question #1a: Testosterone therapy *to improve sexual symptoms* is more effective in men with low testosterone levels at baseline than in men with eugonadal levels, and more effective in men with diabetes, regardless of baseline

testosterone levels. Most participants in the studies included in the systematic reviews of testosterone *to improve diabetes-related outcomes* had low levels of testosterone, with means close to or somewhat below the lower bound of typical reference ranges. Some, but not all, studies specifically enrolled men with classic symptoms of hypogonadism. However, neither of the 2 reviews analyzed whether the effect on metabolic outcomes actually varied according to baseline testosterone level. Other systematic reviews have suggested that the effect of testosterone therapy does *not* depend on baseline testosterone levels for improvement in intermediate cardiovascular measures or depression.

Additional Potentially Policy-Relevant Information, Testosterone Levels and Metabolic Factors

Additional support for a link between low testosterone levels and metabolic factors was provided by 3 studies demonstrating an association between body mass index (BMI), which could be considered a modifiable factor, and low testosterone levels. Furthermore, a recent meta-analysis demonstrated that weight loss after bariatric surgery and a low-calorie diet were both associated with an increase in testosterone levels. Such evidence lends further support to a link between hypogonadism and type 2 diabetes, but is also consistent with expert recommendations to consider evaluating for and addressing obesity and chronic disease prior to recommending testosterone therapy.

Key Question #1b. What is the minimum interval required to assess a change in testosterone status in untreated and treated individuals?

The literature search identified no studies designed to estimate the optimal interval for serial testosterone testing.

Other Potentially Policy-Relevant Information

A very small study of 282 men selected on the basis of age (60 to 82 years at follow-up) compared changes in serum testosterone levels and changes in responses to the Androgen Deficiency in the Aging Male (ADAM) questionnaire over a 5-year period. Total testosterone and bioavailable testosterone levels declined, but only 1 of 10 ADAM symptoms (less strong erection) changed over that time period. These findings suggest that clinically meaningful changes in testosterone levels may not occur due to age alone within 5 or fewer years.

The Endocrine Society Guidelines on testosterone therapy include a *weak* recommendation that serum levels of testosterone be checked at 3 to 6 months *following initiation of therapy* to assess whether serum levels have reached the normal range. However, the guidelines characterize the supporting evidence as being of very low quality. A review article acknowledged that there is insufficient evidence to allow a conclusion about the usefulness of monitoring serum levels in patients undergoing testosterone replacement therapy. However, other experts point out that there is no evidence that testosterone levels beyond the normal range provide health benefits, which implies a rationale for monitoring serum levels to assure that they remain in the normal range and do not unnecessarily expose the patient to adverse effects.

Key Question #2: What are the potential harms of testosterone testing, including potential subsequent harms resulting from treatment decisions?

The literature search identified no studies that were designed to assess harms resulting from a clinical strategy of testosterone testing. Furthermore, no studies with direct evidence were cited in the practice guidelines that were reviewed.

Other Potentially Policy-Relevant Information

Testosterone testing is a blood test and poses no significant procedural harms beyond those associated with phlebotomy.

Indirect Evidence: Safety of Testosterone Treatment

As discussed in the **CLINICAL BACKGROUND** section, adverse events were described in guidelines and a review article as occurring at a low frequency in *young* men on testosterone therapy, but quantitative data were not reported in these sources.

The Endocrine Society's guidelines state that testosterone therapy in *older* men may increase prostate events (nonsignificant estimate), substantially increases the risk of erythrocytosis, and has no clear effect on obstructive sleep apnea, cardiovascular events, or mortality. The guidelines' review of safety evidence for older men is consistent with systematic reviews and key studies published since the guidelines were developed. Large observational studies (incidentally identified) and the most comprehensive systematic reviews (systematically identified as part of topic scoping) that have been published since the guidelines suggested that testosterone therapy:

- May increase prostate events (nonsignificant estimate). The pooled RR of a composite prostate outcome was 1.41 (95% CI, 0.93 to 2.14; low heterogeneity; 15 studies).
- Substantially increases the risk of erythrocytosis. The pooled RR of erythrocytosis was 3.15 (95% CI, 1.56 to 6.35; low heterogeneity; 11 studies).
- Has no clear effect on obstructive sleep apnea. Small case series reported inconsistent findings.
- Has no clear effect on cardiovascular events (no overall effect; conflicting evidence for subpopulations defined by older age or existing cardiovascular disease [CVD]). The adjusted hazard ratio (HR) for the presence of coronary artery disease was 1.29 (95% CI, 1.04 to 1.58) in a population cohort study of 8709 men who had undergone coronary angiography. Another retrospective population cohort study (55,593 men) estimated that the risk of myocardial infarction within 90 days of treatment initiation was doubled, but only in men > 65 years of age or in younger men with a history of heart disease. In a meta-analysis of 75 RCTs, the overall risk of any cardiovascular event was 1.07 (95% CI, 0.69 to 1.65; 31 studies). Study participants were generally selected on the basis of age. In contrast to the implications of 2 population studies, the meta-analysis found no variation in adverse effects according to subpopulations defined by age or CVD at baseline, and no variation according to baseline testosterone levels was detected.

- Has an uncertain effect on all-cause mortality. A recent study suggested a *protective* effect (adjusted HR, 0.61; 95% CI, 0.42 to 0.88), but this evidence comes from a single observational study and cannot be considered conclusive.

Follow-up was typically 2 or 3 years. Thus, the long-term safety of testosterone therapy is unknown. See **Appendix VI** for additional details. As noted in the **POLICY CONTEXT** statement, the FDA is reassessing the safety of FDA-approved testosterone products because of conflicting findings regarding long-term safety.

Key Question #3: What are the costs and cost-effectiveness of testosterone testing?

No studies evaluating the cost or cost-effectiveness of testosterone testing were identified in the literature.

Post Hoc Analysis of Indirect Evidence for the Effectiveness of Testosterone Testing (Key Questions #1, #1a)

Since no studies were identified in the search for direct evidence to answer the Key Questions, an in-depth review of some of the indirect evidence covered in the **CLINICAL BACKGROUND** section was conducted. For subpopulations where the evidence met certain criteria, the quality of key indirect evidence was appraised and an additional targeted search was conducted for recent indirect evidence . Since this was a post hoc analysis and follow-up search for a subgroup of conditions, it does not constitute a full systematic search for all potentially relevant indirect data on this topic. The following 2 conditions were both required for in-depth consideration of testing in a subpopulation:

- A systematic review of evidence supported an association between symptoms, a health condition, or the consequences of a health condition and endogenous testosterone levels or in the absence of a systematic review, at least 1 large observational study demonstrated a strong association.
- The most recent systematic review evidence suggested that testosterone therapy may be effective.

The value of the indirect evidence is based on 2 assumptions:

- Testosterone testing could not have clinical utility unless low testosterone levels were shown to be associated with clinical characteristics that can be used to select patients for testing and possible testosterone therapy. Such associations would ideally be quantified in terms that allow assessment of the discriminatory power of serum testosterone measurements (i.e., risk ratios, estimates of risk ratios [odds ratios], sensitivity and specificity, or calculations related to sensitivity and specificity).
- Testosterone testing would have greatest clinical utility, i.e., have the greatest potential to improve health outcomes, in those populations where testosterone therapy was effective.

Negative evidence for either set of evidence could rule out the clinical utility of testosterone testing. The evidence collected as part of the background review did not identify any subpopulations for which an association between endogenous testosterone levels and symptoms or health was disproven or for which effectiveness of testosterone therapy was disproven. *Positive* evidence for clinical associations

and treatment effectiveness does not definitively confirm the clinical utility of a test but strengthens the possibility that testing could be useful.

The background research for the current report suggested that a relationship between symptoms or health conditions and testosterone levels, as well as the efficacy of testosterone treatment, has been suggested by population-based studies and/or systematic reviews in these 2 subpopulations:

- Men with type 2 diabetes or metabolic syndrome.
- Men with symptoms of sexual dysfunction.

A post hoc search and analysis were designed to further explore these 2 indications as having the strongest possibility for the usefulness of testosterone testing. For testosterone testing in men with diabetes or metabolic syndrome, the post hoc analysis was based on the following:

- 2 systematic reviews of the link between testosterone levels and metabolic outcomes.
- 1 systematic review of the link between testosterone levels and type 2 diabetes.
- 2 systematic reviews of the effectiveness of testosterone therapy in men with type 2 diabetes. In 1 review, study inclusion criteria required that participants also have hypogonadism (defined by the presence of ≥ 3 sexual symptoms as well as low testosterone levels). The other systematic review was published following release of the draft version of the present report.

The systematic review evidence was supplemented by a search for studies published subsequent to the systematic reviews.

For testosterone testing in men with sexual dysfunction, the post hoc analysis was based on the following:

- A single large population study (the European Male Aging Study [EMAS]) , which was identified in the full literature search designed to produce direct evidence for the Key Questions. The EMAS was also cited in review articles as one of the major population studies. An additional and more targeted effort was made to identify relevant studies that were published since the EMAS study.
- 1 systematic review (Corona et al., 2014a), supplemented by a search for studies published subsequent to this systematic review.

See **Appendix I** for details of the post hoc search.

Findings of Post Hoc Analysis: Men with Type 2 Diabetes or Metabolic Syndrome

Association of Type 2 Diabetes/Metabolic Syndrome with Testosterone Levels

A good-quality meta-analysis of 32 studies computed a relative risk (RR) for metabolic syndrome of 0.38 (95% CI, 0.25 to 0.50), comparing higher with lower levels of total testosterone. Free testosterone and SHBG were also found to be lower in men with metabolic syndrome. Another meta-analysis of fair quality reported very similar findings, expressed in terms of mean difference rather than RR. A third, fair-quality, meta-analysis of 33 studies found that mean levels of total testosterone were significantly lower in men with type 2 diabetes than in men without diabetes. Findings are summarized in **Table 6**.

A systematic search for observational studies published since these 3 systematic reviews did not identify any evidence that was inconsistent with the conclusions of the systematic reviews. However, the new studies were not reviewed in detail or critically appraised; thus, the strength with which new studies confirm the conclusions of the systematic reviews is unknown.

Effect Modifiers (Pertinent to Key Question #1a)

Sensitivity analyses suggested that differences in serum testosterone were greater in:

- Younger men (age < 55 years), metabolic syndrome vs no metabolic syndrome (trend, *P*=0.08)
- Co-diagnosis of type 2 diabetes, metabolic syndrome vs no metabolic syndrome
- Younger men, type 2 diabetes vs no diabetes
- Greater BMI and greater body mass, type 2 diabetes vs no diabetes

Other analyses showed *no change* in serum testosterone differences due to:

- Erectile dysfunction, metabolic syndrome vs no metabolic syndrome
- Erectile dysfunction, type 2 diabetes vs no diabetes

These analyses suggest that the effect of type 2 diabetes on testosterone levels (or the effect of testosterone levels on type 2 diabetes; the direction of causality is unknown) is greater in younger men with greater body mass index (BMI), and the effect of metabolic syndrome on testosterone levels (or the other way around) is greater when there is concomitant type 2 diabetes. However, the presence erectile dysfunction does not influence the association between low testosterone and type 2 diabetes or metabolic syndrome although it may be an independent risk factor.

Table 6. Summary of Indirect Evidence: Association Between Low Testosterone and Type 2 Diabetes or Metabolic Syndrome(see Appendix V for additional detail)

Key: ED, erectile dysfunction; FT, free testosterone; MA, meta-analysis; MetS, metabolic syndrome; nmol, nanomole; NR, not reported; RR, relative risk; SHBG, sex hormone binding globulin; SR, systematic review; TT, total testosterone

Study and Quality	Findings	Comments
MetS		
Brand et al. (2011) MA of 32 studies (26 cross-sectional, 5 prospective cohort, 1 case-control) Mean age, 14 to >70 yrs by study (2 studies involved adolescents)	***Mean differences:*** TT: –2.64 nmol/L (95% CI, –2.95 to –2.32) FT also lower in men with MetS ***Unadjusted RR of MetS:*** Higher vs lower TT: 0.38 (CI, 0.25 to 0.50) (13 studies) Higher vs lower FT: 0.64 (CI, 0.41-1.01); slight inconsistency in results across studies (7 studies)	High heterogeneity in main analyses but generally good consistency of results across studies. Study quality poor-fair by authors' standards but key aspects of study quality did not affect differences. No clear evidence of publication bias.

Study and Quality	Findings	Comments
Good-quality SR	Higher vs lower SHBG: 0.29 (CI, 0.21-0.41) (10 studies)	
Corona et al. (2011a) MA of 20 studies (13 cross-sectional, 3 longitudinal) Mean age, 52-58 yrs by study. Fair-quality SR	***Mean difference in serum TT in cross-sectional studies, cases minus controls (nmol/L):*** Overall: –2.85 (95% CI, –3.34 to –2.36) (13 studies) No ED: –2.60 (CI, –3.15 to –2.06) (19 studies) ED: –3.51 (CI, –4.48. to –2.53) (5 studies) ***Mean difference in serum TT, prospective cohort studies, cases minus controls (nmol/L):*** –2.17 (CI, –2.41 to –1.94; I2=96.2%) (number of studies NR)	High heterogeneity in main analysis but no serious inconsistency in direction of study results. Study quality not reported. No evidence of publication bias.
Type 2 Diabetes		
Corona et al. (2011b) MA of 33 studies (28 cross-sectional studies): 11,831 men; mean age 38-72 yrs by study 5 prospective cohort studies: 2700 men; mean age, 45-67 yrs by study Fair-quality SR	***Mean difference in serum TT, cross-sectional studies, cases minus controls (nmol/L):*** Overall: –2.99 (95% CI, –3.59 to –2.40) (24 studies) No ED: –3.00 (CI, –3.61 to –2.38) (19 studies) ED: –2.99 (CI, –3.59 to –2.40) (5 studies) ***Mean difference in serum TT, prospective cohort studies, cases minus controls (nmol/L):*** –2.08 (CI, –3.57 to –0.59) (5 studies) FT and SHBG were also significantly lower in men with MetS.	Moderate-high heterogeneity in main analysis but no serious inconsistency in direction of study results. Study quality not reported. No evidence of publication bias.

Effectiveness of Testosterone Therapy in Men with Type 2 Diabetes

A fair-quality meta-analysis of 5 randomized controlled trials (RCTs) found that the mean difference in levels of fasting plasma glucose, fasting serum insulin, glycated hemoglobin (HbA1c), and triglycerides favored men who underwent testosterone therapy. No effect on body fat or blood pressure was detected. The authors did not discuss whether differences were clinically important. Individual study differences suggested negligible to substantial improvement in glucose control but individual study data suggested negligible to substantial between-group differences at study endpoint. The mean values for control groups indicate that patients were under treatment for their diabetes. There was no subgroup analysis of whether the effects of testosterone therapy depended on the baseline level of glucose control. All study participants had both late-onset hypogonadism (defined by sexual symptoms) and type 2 diabetes. The authors note that given the short follow-up periods (3 to 12 months), definitive

conclusions about the metabolic effects of testosterone therapy could not be made. Findings are summarized in **Table 7**.

An abstract review of findings from systematically identified RCTs published since the systematic review neither confirmed nor refuted the findings of the meta-analysis. However, the new studies were not reviewed in detail or critically appraised; thus, the strength with which new studies confirm the conclusions of the systematic reviews is unknown. After release of the draft version of the present report, another systematic review and meta-analysis of testosterone therapy in men with diabetes was published. This review, which allowed inclusion of studies conducted in men with metabolic syndrome who did not yet have a diagnosis of diabetes, was based on 7 RCTs, 1 involving men with metabolic syndrome and 6 involving men with diabetes. The review authors stated that all diabetic study participants were currently being treated with antidiabetic medication and had adequate glucose control. The results of this meta-analysis were less positive than those of the slightly earlier review. All pooled estimates were statistically nonsignificant, except for a subgroup analysis of improvement in insulin resistance that included only studies using a more conventional but less rigorous measurement technique. However, the direction of individual study results with respect to insulin was consistently positive. The authors described the observed improvement in insulin resistance as modest. They concluded that their results do not support the routine use of testosterone therapy to improve glucose metabolism in men with relatively well-controlled type 2 diabetes and/or metabolic syndrome. However, they pointed out that since metabolic syndrome and type 2 diabetes are slowly progressive conditions, it is possible that long-term testosterone therapy might have a more pronounced effect on metabolic factors.

In both reviews, the direction of individual study results varied for glucose control (HbA1c). The lack of statistical significance in the pooled estimate of the Grossman review is noteworthy since this analysis was based on 6 trials (603 patients) and should perhaps be given greater weight; the Cai analysis was based on only 3 trials (102 patients), 2 of which were excluded from the Grossman review because of lack of placebo control. The Cai review observed significant effects on measures that were not analyzed in the Grossman review: fasting plasma glucose and fasting serum insulin. The conclusions of both meta-analyses should be considered preliminary. The included studies varied with regard to whether any changes in antidiabetic medication were allowed during the study. Even where medication changes were discouraged, analysis in the individual studies generally did not control for group differences in compliance.

Table 7. Summary of Indirect Evidence: Effectiveness of Testosterone Therapy in Men with Type 2 Diabetes and Hypogonadism (see Appendix VI for additional detail)

Key: DM, diabetes mellitus; FPG, fasting plasma glucose; FSI, fasting serum insulin; f/u, follow-up; HbA1c, glycated hemoglobin; HOMA-IR, Homeostasis Model Assessment for Insulin Resistance; LOH, late-onset hypogonadism; MA, meta-analysis; MetS, metabolic syndrome; mmol, millimole; mIU, milli-international unit; SR, systematic review; RCT, randomized controlled trial; SMD, standardized mean difference; (T)T, (total) testosterone; tx, treatment (or therapy)

Study and Quality	Findings	Comments
Cai et al. (2014) MA of 5 RCTs (351 patients) to evaluate	***Mean difference, control minus T therapy:*** FPG (nmol/L): –1.10 (CI, –1.88 to –0.31) (5 RCTs). Individual studies showed differences of 0.17-2.0 (1.8%-25% relative improvement); control means, 6.0-9.18.	Study quality was 5 to 7 on a 0-8 scale. Some inconsistency (no effect vs benefit) in study

Study and Quality	Findings	Comments
effects of T therapy in men w/ LOH* and type 2 DM (mean age 44-64 yrs; f/u, 3-12 mos). 2 RCTs compared w/ no tx rather than placebo Fair-quality SR	FSI (mIU): –2.73 (CI, –3.62 to –1.84) (4 RCTs). Individual studies showed differences of 0.15-2.88 (0.8-33.2% relative improvement); control means, 5.79-18.85. HbA1c (%): –0.87 (CI, –1.32 to –0.42) (3 RCTs). Individual studies showed differences of 0.7%-1.3% (0.3%-13.1% relative improvement); control means, 6.27%-9.9%. Triglycerides (mmol/L): –0.35 (CI, –0.62 to –0.07) (4 RCTs). Individual studies showed differences of 0.2 favoring controls (1 study) and 0.29-0.82 favoring T therapy (3 studies); control means 1.7-2.56. Low heterogeneity in each analysis, except for FPG (I^2=61%), which was explained by differences in treatment regimen. No effect on body fat or blood pressure (3 RCTs each).	results for FSI and HbA1c.
Grossman et al. (2014) MA of 7 double-blind, placebo-controlled RCTs (833 patients; mean age 44-64 yrs; mean BL TT, 8.6-10.1 nmol/L at BL; f/u 3-12 mos) in men w/ MetS (1 RCT) or type 2 DM (6 RCTs)	***Mean difference, control minus T tx:*** HOMA-IR: –0.26 (CI, –1.09 to 0.57) (7 RCTs [3 included in MA by Cai et al.]; high heterogeneity, I^2=76%). Significant results in trials using more conventional, less rigorous technique. SMD was –0.34 (CI, –0.51 to –0.16) HbA1c (%): –0.15 (CI, –0.39 to 0.10) (6 RCTs [3 included in MA by Cai et al.]; high heterogeneity, I^2=77%). No difference in larger trials. Modest effect on lipids (MA not possible) and no effect on triglycerides or blood pressure.	Study quality (according to 25-item CONSORT reporting checklist for RCTs) was 14-24 across studies (1 point per item). Study results were consistent in direction for HOMA-IR but not for HbA1c. No evidence of publication bias in funnel plot, but authors were aware of 1 unpublished trial reporting no effect on HbA1c.

*LOH defined as TT < 3.2 ng/mL (320 ng/dL; 11 nmol/L), or TT < 3.2 ng/mL and FT < 220 pmol/L, *and* ≥ 3 sexual symptoms.

Findings of Post Hoc Analysis: Men with Symptoms of Sexual Dysfunction

Association of Sexual Dysfunction with Testosterone Levels

The EMAS employed cluster analysis to show that 3 symptoms of sexual dysfunction collectively had a strong association with testosterone levels. The 3 symptoms were awakening with full erection less frequently, erectile dysfunction, and thinking about sex less frequently. Depending on the cutoff value and whether levels of both total and free testosterone were considered, strong associations (RR estimate ≥ 2.0) were demonstrated. Another smaller observational study found an association of low total testosterone with erectile dysfunction but not with sex drive. Findings are summarized in **Table 8**.

Studies identified in a systematic search for observational studies published since the EMAS did not conduct a comparable analysis. The recently published studies were not reviewed in detail or critically appraised; thus, the implications of their results with respect to the findings of the EMAS are not fully known.

Effect Modifiers: Neither study evaluated patient factors that might alter the relationship between low testosterone and sexual symptoms.

Table 8. Summary of Indirect Evidence: Association Between Low Testosterone and Sexual Dysfunction (see Appendix III for additional detail)

Key: BMI, body mass index; ED, erectile dysfunction; EMAS, European Male Aging Study; FT, free testosterone; MA, meta-analysis; MetS, metabolic syndrome; ng, nanogram; nmol, nanomole; OR, odds ratio; pmol, picomole; SR, systematic review; T, testosterone; TT, total testosterone

Study and Quality	Findings	Comments
Gades et al. (2008) Random sample, 414 men, age ≥50 yrs, Olmstead County, MN Fair quality	***Lower TT:*** Associated w/ ED but not w/ sex drive *BT*: No association	No analysis similar to the EMAS (assessment of association w/ *simultaneous* presence of >1 symptom). Estimates adjusted for age; men w/ comorbidities excluded.
Wu et al. (2010) (EMAS) Random general population sample, 3369 men, ages 40-79 yrs, Europe Good quality	***Optimal cutoff levels:*** *TT:* 8.0-11 nmol/L (231-317 ng/dL), depending on symptom *FT:* 160-280 pmol/L (0.16-0.28 nmol/L; 5-8 ng/dL), depending on symptom *OR (odds of presence of 3 symptoms, low T level vs high T level):* 1.64-2.24, depending on cutoff chosen and whether free as well as total T was low; strong associations. Nonsignificant ORs if <3 symptoms present.	Estimates adjusted for age, BMI, # coexisting illnesses.

Effectiveness of Testosterone Therapy in Men with Sexual Dysfunction

The most comprehensive systematic review of testosterone therapy for sexual dysfunction reported large pooled effect sizes (0.75 to 0.82) suggesting that testosterone therapy as a stand-alone therapy reduced overall erectile dysfunction, sexual-related erectile dysfunction, and poor libido, and improved orgasm score. (NOTE: Effect sizes > 0.80 are by convention considered to be *large*.) The systematic review did not clarify how men were selected for the included studies or describe the typical baseline degree of sexual dysfunction. In a subgroup analysis, the effect sizes for men with low testosterone were even larger (1.00 to 1.26) than the overall effect. In studies where testosterone therapy was an add-on to phosphodiesterase type 5 inhibitor (PDE5i), which is a medication for erectile dysfunction, testosterone therapy was effective in uncontrolled studies but not in placebo-controlled trials. Findings are summarized in **Table 9**.

Effect Modifiers:

Subgroup and regression analyses showed that treatment effects:

- Were greater in men with lower baseline levels of testosterone.
- Were greater in men who had a diagnosis of type 2 diabetes.
- Did not depend on age or duration of therapy.

A systematic search for RCTs published since the systematic review yielded no studies that would potentially meet the review's inclusion criteria.

Table 9. Summary of Indirect Evidence: Effectiveness of Testosterone Therapy for Treating Sexual Dysfunction (see Appendix VI for additional detail)

Key: CI, confidence interval; ED, erectile dysfunction; FT, free testosterone; MA, meta-analysis; MetS, metabolic syndrome; NR, not reported; PDE5i, phosphodiesterase type 5 inhibitor; SMD, standardized mean difference; SR, systematic review; T, testosterone

Study and Quality	Findings	Comments
Corona et al. (2014a) MA of 41 (1930 men), T supplementation vs placebo or T supplementation as add-on to PDE5i vs placebo, in men w/ ED (ages 19-83 yrs; mean f/u, 27 wks for T vs placebo and 12 wks for T as add-on) Fair-quality SR	T VS PLACEBO TRIALS *SMD in (>0 favors T therapy):* Overall ED Overall: 0.82 (CI, 0.47-1.17; I^2=87.73%) (24 studies) Hypogonadal men: 1.23 (CI, 0.72-1.72) (5 studies) Sexual-related ED Overall: 0.75 (CI, 0.37-1.12) (19 studies) Hypogonadal men: 1.26 (0.47-2.06) (# studies NR) Poor libido Overall: 0.81 (CI, 0.47-1.17) (17 studies) Hypogonadal: 1.00 (CI, 0.47-1.53) (# studies NR) Low orgasm score Overall: 0.68 (CI, 0.34-1.02) Meta-regression and linear regression analysis based on low vs high baseline T levels rather than diagnosis of hypogonadism showed modification of effect on ED and orgasm; trend for libido. T AS ADD-ON Positive effect in uncontrolled studies but not in placebo-controlled. (12 studies; I^2=96.41%)	High heterogeneity in linear regression analyses. Evidence of publication bias and a greater effect in industry-sponsored trials. Study quality NR, and no information on consistency of results across studies.

Summary of Post Hoc Analysis

A positive but very weak inference regarding the utility (Key Question #1) of testosterone testing in men with sexual dysfunction can be drawn. This inference is based on evidence of a strong association of low testosterone levels with the simultaneous presence of multiple sexual symptoms. The inference also reflects evidence that testosterone therapy improves sexual symptoms in men who have sexual dysfunction and low testosterone levels. Although there is a strong association of low testosterone levels with metabolic syndrome, and significant differences in serum testosterone between men with and without type 2 diabetes, the effectiveness of testosterone therapy for improving outcomes related to type 2 diabetes remains uncertain.

Key positive findings from analyses of effect modifiers suggested that:

- Diabetes is a stronger predictor of low testosterone levels in younger compared with older men.
- The effectiveness of testosterone therapy for improving sexual dysfunction was greater in men who had a diagnosis of type 2 diabetes before treatment, regardless of their baseline testosterone levels.

The quality of the indirect evidence was low to moderate with respect to the objectives of the studies, taking into account risk of bias in individual studies where reported by review authors, the quantity of data, consistency of study findings, and the precision of pooled estimates. As evidence for Key Questions #1 and #1a of the current report, the evidence is of low quality at best because of its lack of direct applicability to the intervention of interest (testing) and the Key Question (impact of testing on outcomes).

The clinical implications of the links between testosterone levels and sexual dysfunction or diabetes/metabolic syndrome are unclear. Authors generally did not discuss the relevance of the magnitude of testosterone level differences in observational studies, present risk differences, or discuss the clinical relevance of benefits demonstrated in testosterone therapy trials. However, the following statements can be made regarding the use of testosterone therapy in clinical practice:

- For men with both hypogonadism (defined by low serum testosterone as well as sexual symptoms) and type 2 diabetes, improvement in glucose control ranged from negligible to substantial, and average improvement in insulin resistance was modest in populations where baseline glucose control was, on average, adequate or close to adequate.
- For men with sexual dysfunction:
    - Testosterone therapy as a *stand-alone* therapy was found to be effective in improving symptoms of sexual dysfunction.
    - Testosterone therapy as an *add-on* therapy was not clearly shown to be effective for men receiving established therapy for erectile dysfunction.
    - The *comparative effectiveness* of testosterone therapy as an alternative to erectile dysfunction medications has not been systematically addressed. The authors of the systematic review of testosterone therapy in men with sexual dysfunction stated in their concluding remarks that testosterone therapy administered to men with hypogonadism

is mainly considered to be a hormonal therapy targeting the condition of hypogonadism rather than a symptomatic treatment of sexual dysfunction.

In addition to the uncertainty regarding clinical application of reviewed evidence, another concern is that follow-up in the testosterone therapy trials was too short (3 to 12 months) to allow a strong conclusion that testing for and treating low testosterone levels in men with metabolic syndrome, type 2 diabetes, or sexual dysfunction would be beneficial and safe in the long term.

Practice Guidelines

The search of the core sources and relevant specialty groups identified 4 guidelines published by U.S. groups within the past 10 years. An additional guideline from the American Diabetes Association was reviewed after the choice of subpopulations for the post hoc analysis. Details, by guideline, are presented in **Appendix VIII**. See also **Practice Guidelines** in the **TECHNICAL REPORT** for additional background information on some guidelines and a description of guidelines that were reviewed but found not to contain relevant recommendations.

The Endocrine Society

The most comprehensive guidelines on testosterone testing and testosterone therapy have been published by the Endocrine Society. They were considered to be of fair quality (5 on a scale of 0 to 7). The Endocrine Society differentiates between subpopulations in which testosterone should be tested if hypogonadal symptoms are present (e.g., men with type 2 diabetes or end-stage renal disease) and subpopulations in whom screening might be reasonable regardless of symptoms (e.g., pituitary mass, use of glucocorticoids or opioids, or human immunodeficiency virus (HIV)-associated weight loss). However, the Society recommends addressing reversible illness and drug usage before testing. The guidelines recommend *against* screening for low testosterone in a general population, even an older population, or routinely offering testosterone therapy to older men. For older men who have *symptoms* suggestive of age-related hypogonadism, the Society recommends discussing the risks and benefits of testosterone therapy with the patient. The Society characterizes all of the evidence supporting testing and treatment recommendations to be of *low or very low quality* and considers all of its positive testing and screening recommendations to be *weak.*

As noted in the previous paragraph, the Society recommends taking signs and symptoms into account for subpopulations that have not been shown to have a high prevalence of low testosterone levels. A consensus process was used to compile a list of signs and symptoms considered to be relatively specific to androgen deficiency and another list of signs and symptoms that are less specific. The 2 lists are presented in **Table 3** and in **Appendix VIII.** Greater consideration should be given to testing if the more specific signs and symptoms are present.

American College of Physicians (ACP) and American Urological Association (AUA)

Guidelines from the ACP on erectile dysfunction contain no recommendation for or against routine testosterone testing because of insufficient evidence. Two sets of guidelines produced by the AUA include expert opinion recommendations in support of testosterone testing in men with abnormal semen analysis, impaired sexual findings, or clinical findings (not specified) suggestive of a specific endocrinopathy. All 3 of these guidelines were judged to be of good quality.

See the **PRACTICE GUIDELINES** section of the **TECHNICAL REPORT** for information on guidelines that were reviewed but found not to have relevant recommendations.

Selected Payer Policies

At the direction of Washington State HCA, the coverage policies for the following organizations were reviewed: Aetna, Centers for Medicare & Medicaid Services (CMS), Oregon Health Evidence Review Commission (HERC), GroupHealth, and Regence Blue Cross/Blue Shield. The following highlights information on coverage policies:

- Aetna: No coverage policy for testosterone testing per se was identified in the Clinical Policy Bulletin. However, Aetna considers androgenic anabolic steroids (e.g., testosterone) to be medically necessary for a number of covered indications. Symptomatic androgen deficiency in men is one of the covered indications. The policy specifically requires assessment of 2 total testosterone levels to establish the medical necessity of testosterone replacement for symptomatic androgen deficiency. Two morning samples drawn between 8:00 a.m. and 10:00 a.m. obtained on different days are required. The policy defines androgen deficiency as a total testosterone level < 200 nanograms per deciliter (ng/dL) or a low normal testosterone level (≥ 200 ng/dL but < 500 ng/dL) plus elevated sex hormone binding globulin (SHBG). Testing details are not specified for any of the indications other than symptomatic androgen deficiency, and symptomatic is not defined.
- CMS: No CMS National Coverage Determination (NCD) was identified for testosterone testing.
- Other Payers: No coverage policy for serum testosterone testing was identified on the websites for GroupHealth, Oregon HERC, or Regence. Regence has concluded that that salivary hormone testing is investigational; this includes salivary testing for testosterone.

See **Selected Payer Policies** in the **TECHNICAL REPORT** for additional details and links to policy documents.

Overall Summary and Discussion

Evidence-Based Summary Statement

Testing for and treating abnormally low levels of testosterone is controversial in the absence of known pathology of or injury to the hypothalamus-pituitary-testicular axis. Evidence demonstrates that age-related declines are, in fact, *normal*, and the symptoms used to characterize age-related hypogonadism are nonspecific to low testosterone. At this time, there is no direct empirical evidence that testing for low testosterone in any subpopulation leads to better health outcomes. The most recent Endocrine Society guidelines on testosterone therapy characterize the quality of evidence for all testing and treatment recommendations as low or very low.

The searches conducted for the present report identified 2 subpopulations for which there was both evidence of strong associations between endogenous testosterone and symptoms or health conditions and, at the time of the draft version of the present report, evidence of the efficacy of treatment with exogenous testosterone:

- Men who have sexual symptoms and low testosterone.
- Men who have type 2 diabetes and symptomatic hypogonadism.

However, a systematic review published after the draft report raised considerable doubt about the effectiveness of testosterone therapy for improvement of metabolic factors. Furthermore, several caveats must be stated:

- Systematic review authors did not comment on the clinical relevance of typical gains associated with testosterone therapy for improvement of sexual dysfunction.
- The long-term safety of testosterone testing has not been demonstrated and is a cause for concern regardless of the indication for testing and treatment.
- The American Urological Association (AUA) guidelines stress that testosterone therapy should not be substituted for conventional treatment for erectile dysfunction dysfunction in men with normal levels of testosterone. There is no evidence that testosterone therapy is superior to or enhances the effectiveness of conventional medication for erectile dysfunction.
- Endocrine Society guidelines state that testosterone testing in men with diabetes is for the purpose of addressing symptomatic hypogonadism, not of treating diabetes.

The literature reflects concerns regarding the analytic validity of techniques for testosterone testing due to variability within individuals, between laboratories, and between different assay kits. Certain practices help improve the reliability and accuracy of testing in clinical laboratories: establishment of a reference range specific to the population served by a particular laboratory, laboratory participation in quality control programs, measurement early in the day, and observance of recommendations that abnormal results from measurements on 2 different days are necessary for a diagnosis of low testosterone.

An overall summary of information and indirect evidence is summarized in **Table 10**. Gaps in the evidence and report limitations are presented following the table.

Table 10. Summary of Information and Evidence Reviewed in This Report

Key: ACP, American College of Physicians; AUA, American Urological Association; BL, baseline; CI, confidence interval; COPD, chronic pulmonary obstructive disease; CVD, cardiovascular disease; ED, erectile dysfunction; FDA, Food and Drug Administration; FSH, follicle stimulating hormone; HIV, human immunodeficiency virus; LH, luteinizing hormone; obs, observational; OR, odds ratio; RR, relative risk; SR, systematic review; sx, symptoms; T, testosterone; TRT, testosterone replacement therapy; TT, total testosterone

Guideline Recommendations	Relevant Policies	Evidence of an Association with Low Testosterone Levels (see Appendices III & V)	Efficacy of TRT (see Appendix VI)
PRIMARY HYPOGONADISM (Examples include chromosomal or genetic disorders, toxin exposure or chemotherapy, orchitis due to mumps or an autoimmune disorder, trauma, hemochromatosis, medications that inhibit androgen biosynthesis, damage due to varicocele)			
---	FDA-approved indications for TRT. Aetna covers anabolic steroids for constitutional delay in growth, delayed male puberty, Klinefelter's syndrome w/ hypogonadism, microphallus.	---	---
SECONDARY CONGENITAL HYPOGONADISM			
Testing warranted, even in the absence of symptoms (Endocrine Society)[1]:			
A mass, radiation, or disease of sellar region (a depression in the middle line of the upper surface of the sphenoid bone in which the pituitary gland is lodged).	Secondary acquired hypogonadism is an FDA-approved indication for TRT. In addition to the indications described in column 1, FDA approval extends to hypogonadism caused by radiation and trauma.	---	---
Use of medications that affect T production or metabolism (e.g., glucocorticoids, anabolic steroids, or opioids).		Medication use: 3 of 4 studies showed inverse association, opioid dose vs TT (1 SR).	---
HIV-associated weight loss.		HIV: Higher than expected prevalence	HIV: Positive effect *only when T levels*

[1] The Endocrine Society recommends addressing reversible illness and drug usage before considering testosterone therapy.

Guideline Recommendations	Relevant Policies	Evidence of an Association with Low Testosterone Levels (see Appendices III & V)	Efficacy of TRT (see Appendix VI)
	Aetna covers anabolic steroids for AIDS wasting syndrome.	in men undergoing HIV therapy; association w/ poorer HIV outcomes (cited in Endocrine Society guidelines).	*are normal*. Authors described benefits as small, and evidence as insufficient (1 good SR).
Osteoporosis or low-trauma fracture (especially in a young man). (TRT recommended for improving bone health.) (Endocrine Society).		Osteoporosis: Potentially strong association (OR, 1.76-2.77) (1 large obs study).	Osteoporosis/fractures: Very limited evidence of effect on osteoporosis; no evidence regarding effect on fracture incidence (2 good SRs).
Testing may be warranted in the presence of characteristic symptoms (see Table 3) (Endocrine Society):			
Type 2 diabetes	Secondary acquired hypogonadism is an FDA-approved indication for TRT. Aetna covers anabolic steroids for AIDS wasting syndrome.	Type 2 diabetes: Inverse strong association (RR for TT, 0.38; CI, 0.25-0.50) w/ metabolic syndrome (1 good SR, confirmed by significant difference in another fair SR). Significant differences, type 2 diabetes vs no diabetes (1 fair SR).	Type 2 diabetes: Some positive effects in smaller (fair-quality) SR; no significant effects in larger SR except in subgroup analysis. The 2 reviews were heterogeneous w/ respect to inclusion criteria and outcome measures. Most study participants had low testosterone levels at BL and sx of hypogonadism (1 fair SR, 1 good SR).
End-stage renal disease and maintenance hemodialysis		Chronic kidney disease: Some evidence of association (cited in Endocrine Society guidelines, no specifics).	---
Moderate to severe COPD		COPD: Higher than expected prevalence of low T (studies cited in Endocrine Society guidelines).	---
T testing may be considered for infertility (expert opinion) (AUA)	---	---	---
OTHER CHRONIC DISEASE			
---	---	Inverse weak association (RR, 0.8-1.00)	Positive effect on intermediate CVD

Guideline Recommendations	Relevant Policies	Evidence of an Association with Low Testosterone Levels (see Appendices III & V)	Efficacy of TRT (see Appendix VI)
		w/ CVD (2 large obs studies).	outcomes; no discernible pattern of outcomes according to BL TT (1 fair SR).
AGE			
No routine testing or treatment based on age alone. Consider severity of symptoms (Endocrine Society, ACP).	---	---	---
SYMPTOMS OF ANDROGEN DEFICIENCY			
Test if more specific symptoms, consider testing if less specific symptoms (see **Table 3**). First exclude reversible illness, drugs that can deplete T levels, nutritional deficiency. TRT *recommended* for maintenance of secondary sex characteristics and improving sexual function and sense of well-being (Endocrine Society).	Aetna covers anabolic steroids for symptomatic androgen deficiency.	---	---
PHYSICAL SYMPTOMS			
Considered a less specific symptom (Endocrine Society).	---	Inverse association but no evidence of strong association, i.e., no calculation of OR or RR (2 large obs studies).	---
PSYCHOLOGICAL SYMPTOMS			
Considered a less specific symptom (Endocrine Society).	---	Mixed findings (2 large obs studies).	Positive effect on depression, regardless of hypogonadal status (1 good SR).
SEXUAL SYMPTOMS			
Considered a relatively specific symptom. TRT is an *option* for low libido or ED w/	---	Strong association (up to OR 2.24 with simultaneous presence of 3 specific	Positive effect in men with ED, especially in men who were

Guideline Recommendations	Relevant Policies	Evidence of an Association with Low Testosterone Levels (see Appendices III & V)	Efficacy of TRT (see Appendix VI)
proviso that established therapies be tried first for ED (Endocrine Society). No recommendation for or against testing or TRT in men with ED (ACP). Test in men with impaired sexual function (AUA).		symptoms if both TT and FT are considered, confirmed by cluster analysis) (1 large obs study).	hypogonadal or had type 2 DM at BL (1 fair SR).
GENERAL HEALTH ASSESSMENT			
---	---	Health status, NS association (1 obs study). All-cause mortality, inverse but weak and NS association (1 large obs study).	---
CANCER			
Do not diagnose androgen deficiency during acute or subacute illness (Endocrine Society).	Aetna covers anabolic steroids for weight loss from cancer.	Mixed findings, cancer-related symptoms (2 obs studies). No clear relationship w/ cancer outcomes (1 obs study).	---

Gaps in the Evidence

- Additional prospective epidemiological studies to establish the temporal relationship between low testosterone levels and health outcomes or conditions and to quantify the strength of associations in terms of relative risks or odds ratios.
- Clinical trials designed to compare health outcomes between individuals whose symptoms and health conditions are managed with the knowledge of serum testosterone levels and individuals who are managed without that information.
- Large, long-term studies designed to assess the durability of benefits and the risks associated with testosterone replacement therapy.
- Analyses of the cost and cost-effectiveness of testosterone testing in specific subpopulations.

Limitations of This Report

The post hoc search and analysis was limited in several ways. The selection of subpopulations for which evidence of both clinical associations with endogenous testosterone and testosterone treatment effectiveness was available was not an entirely systematic process. It assumed that the most important observational evidence establishing a link between testosterone levels and symptoms or clinical conditions would have been identified in the literature search designed to identify direct evidence for the Key Questions, in review articles and practice guidelines, or in the topic-scoping search for systematic reviews. A more thorough search for evidence of clinical associations with endogenous testosterone would have used a slightly different original search strategy. Secondly, the process assumed that subpopulations in which treatment effectiveness had been most clearly demonstrated would be those subpopulations for which there were systematic reviews of treatment trials. It is possible that the evidence supports a conclusion that other subpopulations also respond to testosterone therapy even though no systematic reviews have been published for these subpopulations.

Nevertheless, the identification of sexual dysfunction and metabolic disorder or diabetes as the potentially strongest indications for testosterone therapy was supported by the more common use of these clinical conditions, rather than others, to define subgroup analyses in population studies and systematic reviews. Furthermore, an expert consensus process conducted by the Endocrine Society identified sexual symptoms as relatively specific to symptomatic androgen deficiency and the Society identified type 2 diabetes as one of the conditions in which testing may be warranted if characteristic symptoms of hypogonadism are present. Lastly, studies published more recently than systematic reviews were not reviewed in detail or critically appraised, but the process for identifying these studies was systematic and judgments regarding their implications for the conclusions of the systematic reviews made no assumptions regarding study quality.

TECHNICAL REPORT

Clinical Background

Low Testosterone and Hypogonadism

Testosterone is the principal androgen secreted by the testis. Additional types of androgen are produced by the adrenal gland (Ho, 2011). A recent analysis of the *general* population of men in the United States reported the following estimates of the prevalence of low testosterone regardless of symptoms that might signal a clinical syndrome: 9.0% in men aged 45 to 54 years, 16.5% in men aged 55 to 64 years, and 18.3% in men aged 65 to 74 years. These estimates were derived from the National Health and Nutrition Examination Survey III (NHANESIII), which defined low testosterone levels as < 300 nanograms per deciliter (ng/dL) (10.4 nanomoles per liter [nmol/L]) (Moskovic et al., 2013).

Hypogonadism is defined as a clinical syndrome resulting from a failure of the testes to produce physiological levels of testosterone and/or a normal number of spermatozoa. The causes of hypogonadism represent disruption of ≥ 1 levels of the hypothalamic-pituitary-testicular axis. Primary hypogonadism is caused by abnormalities at the testicular level, whereas secondary hypogonadism is caused by defects in both the testis and the pituitary (Bhasin et al., 2010; Pantalone and Faiman, 2012; Shea et al., 2014). **Table 1** lists possible causes of primary and secondary hypogonadism.

Table 1. Known or Possible Causes of Hypogonadism

Key: DM, diabetes mellitus; HIV, human immunodeficiency virus; OSA obstructive sleep apnea

Primary Hypogonadism (Due To Testicular Abnormalities)	• Karyotype (chromosomal) disorders • Toxin exposure or chemotherapy • Congenital defects (anorchia, cryptorchidism) • Orchitis due to mumps or an autoimmune disorder • Trauma • Hemochromatosis* • Medications that inhibit androgen biosynthesis • An increase in the temperature of the testicular region due to nearby tissue abnormalities (e.g., varicocele)
Secondary Hypogonadism **(Due To Defects In Both Testis And Pituitary)**	• Several genetic abnormalities/mutations, which are usually diagnosed in childhood or adolescence • Use of opioids, corticosteroids, progestins, or androgenic steroids • Acute illness • HIV infection

	• Obesity and related conditions (OSA, insulin resistance, type 2 DM)
	• Chronic illness (e.g., cirrhosis, renal failure, rheumatoid arthritis)
	• Alcohol abuse
	• Anorexia nervosa
	• Severe primary hypothyroidism
	• Hemochromatosis*
	• Hyperprolactinemia†
	• Estrogen excess due to exposure or testicular or adrenal tumors
	• Pubertal delay
	• Disorders of the pituitary affecting release of gonotrophic hormones:
	  ○ Sellar mass or cyst
	  ○ Infiltrative lesions
	  ○ Metastatic lesion
	  ○ Head trauma
	  ○ Radiation exposure
	  ○ Surgery
	  ○ Stalk severance
	  ○ Pituitary apoplexy
	• Aging

Source: Pantalone and Faiman (2012).

*A hereditary metabolic disorder that can result in iron deposits in the hypothalamus, pituitary, and or testes.

†Excess serum levels of the hormone prolactin, which can cause infertility of sexual dysfunction in men.

The correct classification of hypogonadism as primary or secondary is relevant since impairment of spermatogenesis (infertility) can be corrected in patients with secondary hypogonadism but not in most patients with primary hypogonadism. Furthermore, secondary hypogonadism can indicate a pituitary tumor, other disorders related to the pituitary gland, genetic disorders, or systemic illness (Bhasin et al., 2010; Pantalone and Faiman, 2012).

Age-Related Hypogonadism (Symptomatic Androgen Deficiency)

Low serum testosterone alone does not constitute a diagnosis of androgen deficiency or clinical hypogonadism. Diagnosis of a clinical condition requires the presence of certain characteristic symptoms, as well as abnormally low serum testosterone. The literature generally distinguishes between hypogonadism diagnosed on the basis of signs and symptoms associated with aging (plus low serum testosterone levels), and hypogonadism due to a disorder (congenital or acquired) of the

hypothalamus, pituitary, or testis. The latter category has been referred to by some writers as "organic" hypogonadism. *Symptomatic androgen deficiency*, or simply *androgen deficiency*, are other terms often used in place of age-related hypogonadism, which is a form of secondary hypogonadism. Whether age-related hypogonadism represents a pathological condition per se is a matter of controversy. The prevalence of symptomatic androgen deficiency depends not only on age, but also on which symptoms are considered for the diagnosis and the cutoff value assumed for defining normal serum testosterone levels. There is no definitive cutoff value for normal testosterone. Serum levels for men of all ages are typically compared with reference ranges for young men, which suggest cutoff values of 280 to 300 ng/dL (9.7 to 10.4 nmol/L). There is also no consensus on a standard symptom profile (Ho, 2011; Pantalone and Faiman, 2012; Watts et al., 2012). **Table 2** lists some terms that are sometimes used to describe age-related decline in serum testosterone, along with the inadequacies of those terms.

Table 2. Common Terminology for Androgen Deficiency in Older Men

Term	Limitations
Male menopause, andropause	No correlation with female menopause
Late-onset hypogonadism	Decline begins in middle age and is gradual
Testosterone deficiency syndrome	Symptoms are nonspecific
Androgen decline in the aging male (ADAM)	No standard case definition

Characteristic symptoms of age-related hypogonadism are not highly specific and may reflect any of a variety of clinical factors. The latest guideline on testosterone therapy from the Endocrine Society includes a consensus- and experience-based list of the symptoms and signs suggestive of androgen deficiency (see **Table 3**). These include specific symptoms and signs such as reduced libido, very small or shrinking testes, and signs of osteoporosis (height loss, low-trauma fracture). The list also includes less specific symptoms and signs such as decreased energy, depressed mood, and increased body fat or body mass index (BMI). Other sources present similar lists without differentiating according to specificity (ASRM, 2006; McGill et al., 2012).

Table 3. Symptoms Considered to Be Associated with Low Testosterone Levels

More Specific Signs and Symptoms	Less Specific Signs and Symptoms
• Incomplete or delayed sexual development; eunuchoidism • Reduced sexual desire (libido) and activity • Decreased spontaneous erections • Breast discomfort, gynecomastia • Loss of body (axillary and pubic) hair, reduced shaving • Very small (especially <5 mL) or shrinking	• Decreased energy, motivation, initiative, and self-confidence • Feeling sad or blue, depressed mood, dysthymia • Poor concentration and memory • Sleep disturbance, increased sleepiness • Mild anemia (normochromic, normocytic, in the female range)

More Specific Signs and Symptoms	Less Specific Signs and Symptoms
testes • Inability to father children • Low or zero sperm count • Height loss, low-trauma fracture, low bone mineral density • Hot flushes, sweats	• Reduced muscle bulk and strength • Increased body fat or body mass index • Diminished physical or work performance

Source: The Endocrine Society Guidelines on *Testosterone Therapy in Adult Men with Androgen Deficiency Syndromes* (Bhasin et al., 2010). Symptoms were identified on the basis of guideline panelists' experience.

Prevalence of Age-Related Hypogonadism

Epidemiological studies have demonstrated that testosterone levels decline with age at an estimated rate of 1% to 2% per year and a substantial proportion of older men have levels below the lower limit of the normal range for young, healthy men. Serum levels begin to decline when men are in their mid-40s. By 1 estimate, 30% of men in their 70s have total testosterone values in the abnormally low range and 50% of men in their 70s have free testosterone values in the abnormally low range, irrespective of symptoms of androgen deficiency, when cutoff points are based on normal reference ranges for young men (McGill et al., 2012). However, estimates of *symptomatic* androgen deficiency are considerably smaller.

Two large population studies are often cited for estimates of symptomatic androgen deficiency in the general population of middle-aged to elderly men. These studies are called the European Male Aging Study (EMAS) (Wu et al., 2010) and the Massachusetts Male Aging Study (MMAS) (Araujo et al., 2004). A more recent population study (Araujo et al., 2007), based on the Boston Area Community Health Survey, used a definition of hypogonadism that was more consistent with the current guidelines of the Endocrine Society. The 3 studies' definitions of symptomatic androgen deficiency and prevalence estimates are compared in **Table 4**. The 2 studies that used consensus-based definitions (MMAS and Boston Area Community Health Survey) reported higher prevalence than the study that used an empirically derived definition based on an analysis of the strength of association between symptoms and testosterone levels (EMAS). It is noteworthy that the 2007 study by Araujo et al. reported a prevalence of 24% for abnormally low total testosterone but only a 5.6% prevalence of symptomatic androgen deficiency.

Table 4. Prevalence of Hypogonadism (Symptomatic Androgen Deficiency) in General Population Studies

Key: ng, nanogram; nmol nanomole; pmol, picomole; T, testosterone

European Male Aging Study (EMAS) (Wu et al., 2010)	Massachusetts Male Aging Study (MMAS) (Araujo et al., 2004)	Boston Area Community Health Survey (Araujo et al., 2007)
3369 men (ages 40 to 70 yrs) receiving primary care from 8 European centers	1691 men (ages 40-70 yrs at baseline), selected on the basis of stratified randomization from communities in the vicinity of Boston	1475 men (ages 30-79 yrs, mean 47.3) randomly selected
Empirically derived definition based on analysis of strongest associations: 3 symptoms of sexual dysfunction: less frequent awakening w/ full erection diminished ability to get and keep an erection (erectile dysfunction) less frequent thoughts about sex *And* total T level <11 nmol/L (317 ng/dL) *And* free T <220 pmol/L (6 ng/dL)	*Definition based on 2001 consensus meeting and data available from the study:** ≥3 of the following signs/symptoms: loss of libido erectile dysfunction depression lethargy inability to concentrate sleep disturbance irritability depressed mood *And* total T <200 ng/dL (6.9 nmol/L) *Or* total T 200-400 ng/dL (6.9-19.9 nmol/L) *and* free T <8.91 ng/dL (300 pmol/L)	*Definition based on 2010 consensus lists and data available from the study:†* ≥1 specific symptoms: depressed libido erectile dysfunction osteoporosis/osteoporotic fracture *Or* ≥2 nonspecific signs or symptoms: lethargy sleep disturbance depressed mood low physical performance *And* total T <300 ng/dL (10.4 nmol/L) *And* free T <5 ng/dL (200 pmol/L)
Prevalence: Overall: 2% Ages 40-49 yrs: 0.1% Ages 50-59 yrs: 0.6% Ages 60-69 yrs: 3% Ages 70-79 yrs: 5%	*Prevalence at baseline:* Overall: 6.0% Ages 40-49 yrs: 4.1% Ages 50-59 yrs: 4.5% Ages 60-70 yrs: 9.4% (The study also found that after a mean follow-up of 8.8 yrs, greater prevalence was associated w/ greater age.)	*Prevalence:* Overall: 5.6% Age <70 yrs: 3.1% to 7.0% by decade Age ≥70 yrs: ≥ 18.4%

*The 2001 Second Annual Andropause Consensus Meeting defined 12 characteristic signs and symptoms, which were approved by the Endocrine Society. The 4 signs/symptoms missing from the MMAS were osteoporosis, loss of muscle strength, regression of secondary sex characteristics, and decreased interest in activities. See **Table 3** or **Appendix III** for the most current lists of signs and symptoms recognized by the Endocrine Society as being associated with hypogonadism.

†The study had data for 3 of 15 items in the Endocrine Society's list of more specific signs/symptoms and 4 of 9 items in the Society's list of less specific signs/symptoms. See **Appendix III** for the most current lists of signs and symptoms recognized by the Endocrine Society as being associated with hypogonadism.

Screening Tools for Diagnosing Androgen Deficiency Without Testosterone Testing

Self-report case detection instruments and structured interview formats may be used to identify individuals who could be classified as having androgen deficiency on the basis of symptoms alone. Self-report questionnaires include the Androgen Deficiency in Aging Males (ADAM) questionnaire and the Aging Males' Symptoms (AMS) Rating Scale for use in general populations. The ANDROTEST and PsychoANDROTEST for men with sexual dysfunction are examples of structured interviews. These tools do not have high specificity, and neither the clinical utility nor the cost-effectiveness of using these instruments prior to or instead of serum testosterone testing is known (Bhasin et al., 2010; Ho, 2011; McGill et al., 2012; Corona et al., 2013b). The details of these tools are presented in **Appendix IV**.

Selecting Patients for Testosterone Testing

Guidelines recommend against screening in general populations (Bhasin et al., 2010). The purpose of testing for and treating low serum testosterone is generally to improve symptoms associated with low testosterone per se (e.g., erectile dysfunction or fatigue). Studies have also investigated the ability of testosterone therapy to reduce the adverse consequences of the medical condition associated with low testosterone (e.g., to reduce cardiovascular events in men with heart failure or to improve glucose control in men with diabetes). For both purposes, the utility of testing depends on the ability to identify clinical populations in which testing would have high diagnostic yield. Thus, investigators have attempted to identify symptoms, signs, and medical conditions that characterize men most likely to have low levels of testosterone.

Evidence for Associations Between Low Testosterone Levels and Signs or Symptoms of Hypogonadism

The strength of association between the signs and symptoms described in **Table 3** and low serum testosterone, and the causal nature of those relationships, are uncertain. **Appendix III** presents the results of observational studies designed to explore specific associations. The clearest association between low serum testosterone and symptoms is with symptoms of sexual dysfunction. At least 2 observational studies have demonstrated such associations (Gades et al., 2008; Wu et al., 2010). The study by Wu et al. (n=3369), referred to in the section on *Prevalence of Age-Related Hypogonadism*, found a strong association between low levels of total and free testosterone and the simultaneous presence of 3 particular sexual symptoms (Wu et al., 2010). American College of Physicians (ACP) guidelines on management of erectile dysfunction cited estimates of the prevalence of low serum testosterone in men with erectile dysfunction in the range of 12.5% to 35% (Qaseem et al., 2009). The citation that involved a study population of all ages (mean age, 52 years) reported a low testosterone prevalence of 15% for men with erectile dysfunction (El-Sakka et al., 2005). This estimate is considerably higher than estimates of 2% to 6% for the general middle-aged to older male population (see **Table 4**).

Four observational studies reported inconclusive evidence regarding poor health status, poor physical performance, or increased psychological symptoms as characteristic of populations likely to have low levels of testosterone (Wu et al., 2010; Tajar et al., 2012; Dev et al., 2014; Hsu et al., 2014). Evidence of an association with health status was conflicting. In general, there was an inverse relationship between testosterone levels and physical performance and psychological symptoms, but the strength of associations was either not reported or was reported to be weak[2]. The relationship between low

[2] Strength of association is best assessed in terms of a relative risk (RR) based on an assumed cutoff value or an estimate of relative risk, i.e., an odds ratio (OR). Wu and colleagues used a technique called clustering analysis,

testosterone levels and psychological outcomes such as depression is controversial and without a clear biological rationale (Zarrouf et al., 2009; Corona et al., 2013b).

Given the uncertainty regarding the contribution of low serum testosterone to the symptoms associated with age-related hypogonadism, the Endocrine Society guidelines make only weak recommendations to test in presence of symptoms that are thought to be associated with hypogonadism. In many cases, testosterone testing should be preceded by other forms of testing such as a comprehensive metabolic panel, complete blood count, and thyroid-stimulating hormone (TSH) level. In severely obese patients complaining of fatigue, evaluation of obstructive sleep apnea may be warranted before testing for low testosterone (Bhasin et al., 2010; Pantalone and Faiman, 2012).

Medical Conditions Raising Suspicion of Low Serum Testosterone

According to the Endocrine Society's guideline on testosterone therapy and other expert sources, testosterone measurement may be warranted in the following situations (Bhasin et al., 2010; McGill et al., 2012; Pantalone and Faiman, 2012):

- A mass in, radiation of, or disease of sellar region (a depression in the upper surface of the sphenoid bone in which the pituitary gland sits).*
- Use of medications that affect testosterone production or metabolism (e.g., glucocorticoids, anabolic steroids, or opioids).*
- Human immunodeficiency virus (HIV)-associated weight loss.*
- Osteoporosis or low-trauma fracture (especially in a young man).*
- Type 2 diabetes.
- End-stage renal disease and maintenance hemodialysis.
- Moderate to severe chronic obstructive pulmonary disease.

*According to the Endocrine Society, these conditions warrant testing for low testosterone even in the absence of characteristics symptoms. Consideration of testing in men with the other conditions should be based on the presence of concomitant characteristic symptoms (see **Table 3**).

The rationale for assuming that men with these conditions are likely to have low testosterone varies by condition. Pathology associated with the pituitary, which produces the hormones (luteinizing hormone [LH] and follicle-stimulating hormone [FSH]) that regulate testosterone production, can be expected to affect testosterone levels (Pantalone and Faiman, 2012). Certain medications are known to suppress the hypothalamic-pituitary gonadal axis and small studies have suggested an association between androgen deficiency and opioid or glucocorticoid use (Bhasin et al., 2010; Corona et al., 2013b). A systematic review identified 4 studies measuring the association between opioid use and testosterone levels; 3 studies showed an inverse relationship (higher opioid dose was associated with lower testosterone levels) and 1 study showed no relationship (McWilliams et al., 2014). No systematic reviews

---

which looked at the stability of relationships between the training set and the validation set. This analysis showed sexual symptoms, but not physical or psychological symptoms, to show a consistently close relationship with low testosterone levels. The authors then computed ORs for sexual symptoms. Tajar and colleagues expressed results in terms of differences in prevalence, and Hsu et al. reported a statistically nonsignificant OR.

investigating the effects of glucocorticoids or other medications were identified. A meta-analysis evaluating the impact of low testosterone levels on bone health found that the odds ratio (OR) for fracture, comparing men with a history of drug-induced hypogonadism with men with no history of hypogonadism, was statistically significant but smaller (1.53) than the OR comparing natural hypogonadism versus no hypogonadism (2.77), but confidence intervals for the 2 estimates overlapped considerably, suggesting no true difference.

The literature reviewed for the present report did not provide a biological rationale for the link between the chronic diseases in the Endocrine Society list and low testosterone levels. However, there is some evidence in support of certain links. Primary studies cited in review articles and guidelines have suggested an association between androgen deficiency and chronic kidney disease (Bhasin et al., 2010; Corona et al., 2013b). The prevalence of hypogonadism has been reported at 22% to 69% in men with chronic obstructive pulmonary disease (COPD) (Corona et al., 2013b). The Endocrine Society Guidelines cited sources for estimates of 20% to 25% for the prevalence of low testosterone levels in HIV-infected men on highly active antiretroviral therapy. Furthermore, low testosterone levels have been shown to be associated with weight loss, progression to acquired immune deficiency syndrome (AIDS), wasting, depression, and loss of muscle mass and exercise capacity in men infected with HIV (Bhasin et al., 2010).

Systematic reviews have shown low testosterone to be associated with metabolic syndrome (Brand et al., 2011; Corona et al., 2011a) and with type 2 diabetes (Corona et al., 2011b). Other reviews have shown an association between low testosterone levels and osteoporosis (Drake et al., 2012) and cardiovascular disease (Corona et al., 2011c; Ruige et al., 2011) as well as cardiovascular mortality (Araujo et al., 2011). A systematic review of studies evaluating the relationship between hypogonadism and prespecified cancer-related outcomes found no evidence of a relationship (Vigano et al., 2010). The most recent systematic reviews evaluating a link between endogenous testosterone and health conditions are summarized in **Appendix V**.

The Clinical Significance of Low Testosterone: Association of Low Serum Levels of Testosterone with Health Outcomes

The long-term *consequences* of low testosterone levels are largely unknown for the largest subpopulations in which low testosterone levels are common: older men and men with chronic illness (Bhasin et al., 2010). Some associations reflected in **Appendix V** suggest testosterone deficiency could be a predisposing factor to poorer health, but there is no consensus on the extent to which low serum testosterone is a true health problem. A meta-analysis of 11 studies computed a relative risk (RR) of 1.35 (95% CI, 1.13 to 1.62) for the association between lower versus higher testosterone levels and all-cause mortality (Araujo et al., 2011). The size of the RR does not suggest a strong association (RR ≥ 2 or RR ≤ 0.5) and the pooled estimate had high statistical heterogeneity. Although low serum testosterone has been shown to be associated with some chronic diseases and symptom complexes, strong associations have not been demonstrated except for the relationship with metabolic syndrome (Brand et al., 2011) and sexual dysfunction (Wu et al., 2010). Some health conditions, such as obesity and type 2 diabetes, may be the *cause* rather than the result of low testosterone levels (see also the following section on **Environmental and Modifiable Lifestyle Influences on Testosterone Levels, Independent of Aging**). Some systematic review authors (Araujo et al., 2011; Brand et al., 2011; Corona et al., 2011c; Ruige et al., 2011; Corona et al., 2013b) surmised that low testosterone may simply be a marker for concurrent, even occult, disease, rather than a causative agent even though a temporal relationship is suggested by some of the analyses. Other review authors have speculated that in hypogonadal men with a chronic comorbidity such as metabolic syndrome, either testosterone therapy could improve the comorbidity or

improvement in the comorbidity could raise testosterone levels (Corona et al., 2011a). Other narrative and systematic review authors have acknowledged that current knowledge does not allow a conclusion as to whether improving testosterone levels could improve a comorbid chronic condition or improving the comorbidity could raise testosterone levels (Corona et al., 2011a; Oskui et al., 2013). For other indications such as sexual dysfunction, this ambiguity does not exist.

Lastly, systematic reviews investigating the link between low testosterone and medical conditions have not included an adjustment for health behavior, an issue discussed in the next section.

Environmental and Modifiable Lifestyle Influences on Testosterone Levels

Epidemiological studies have demonstrated a decline in average circulating testosterone levels over recent decades, independent of age. Other evidence suggests several possible contributors to this secular trend: increases in body mass index (BMI) (adiposity is linked with diminished testosterone levels), smoking cessation, health behaviors, traumatic life events, environmental toxins, and changes in assay technology. Increasing polypharmacy is also a conceivable contributor (Travison et al., 2009). The following 5 studies typify the evidence of a link (or lack thereof) between environmental or lifestyle factors and low testosterone levels:

- An analysis of 3453 men aged 65 to 83 years residing in Australia compared a set of 8 health-related behaviors with blood test results. Blood samples were collected a mean of 5.7 years after the lifestyle assessments had been made (Yeap et al., 2009). Lifestyle score was based on smoking status, exercise, alcohol use, consumption of fish, consumption of meat, adding salt to food, BMI, and use of skim milk. Higher (better) adjusted lifestyle scores were associated with higher levels of both total testosterone and sex hormone binding globulin (SHBG). The odds of having a total testosterone measurement in the lowest quartile was reduced by 63% for men with a lifestyle score of ≥ 7 out of a possible 8: OR, 0.37 (95% CI, 0.18 to 0.77).
- In a study of 375 men aged 45 to 85 years, smoking was correlated with *higher* levels of total testosterone and free testosterone, while higher BMI and higher age were correlated with *lower* levels of both total testosterone and free testosterone (Ponholzer et al., 2005). However, no correlation was detected between physical activity, social activity, nicotine and alcohol consumption, stress level, or sleep quality and serum levels of testosterone.
- A prospective cohort study of men residing in Massachusetts found that some lifestyle- and life event-related factors were associated with declines in testosterone levels when controlling for aging (Travison et al., 2007). A multiple regression analysis simultaneously controlled for a number of factors. A statistically significant or near-significant adjusted percentage decline in total testosterone levels over time was observed for the following factors: 10-year increase in age (10.1%); incident diabetes (4.9%; *P*=0.05), incident hypertension (4.7%), incident use of ≥ 6 medications (6.0%), BMI increase of 1 kilogram per square millimeter (kg/m^2) (1.9%), incident widowhood (11.1%), smoking cessation (8.6%), loss of employment (3.2%). Similar trends were observed for free testosterone and SHBG.
- An analysis of the Boston Area Community Health Survey data yielded no statistically significant ORs for smoking or other lifestyle factors (Hall et al., 2008). The other factors investigated were not identified.
- A meta-analysis of the association between metabolic syndrome and testosterone levels found that the association was stronger if the criteria for diagnosing metabolic syndrome weighted

abdominal obesity more heavily than other factors rather than treating all factors the same (Brand et al., 2011).

It is noteworthy that 3 of the studies found a link between low testosterone and BMI, but the studies did not establish the direction of effect. Androgen deprivation is known to increase fat mass and insulin, while weight loss increases both insulin sensitivity and testosterone levels (Grossman et al., 2014). The authors of the listed studies, as well as review articles, suggested promotion of healthy behaviors as a potential treatment for age-related declines in serum testosterone prior to consideration of medical treatment of low testosterone. Nutritional deficiencies, eating disorders, and excessive exercise can also contribute to low levels of testosterone (McGill et al., 2012; Pantalone and Faiman, 2012).

Medical Treatment of Low Testosterone Levels

Testosterone replacement therapy is considered the medical treatment of choice when fertility is not required. Since serum testosterone downregulates the hypothalamus-pituitary axis, exogenous testosterone actually inhibits sperm production. Reversible conditions should be addressed as possible causes of low serum testosterone and corrected where possible before initiating testosterone therapy. Testosterone therapy comes in various formulations: intramuscular, subcutaneous, transdermal, buccal, and oral. Different forms have different advantages and disadvantages (McGill et al., 2012; Pantalone and Faiman, 2012).

Guidelines suggest that testosterone therapy can be offered for maintenance of secondary sex characteristics, improvement of sexual function, promotion of a sense of well-being, or correction of the effects of HIV treatment or high doses of glucocorticoids. The evidence of efficacy varies considerably by indication. Clinicians are advised to also consider erectile dysfunction as the reason for low libido and to consider established treatments (treatment other than testosterone therapy) for erectile dysfunction before offering testosterone therapy, especially in older men because of the risk of adverse treatment effects (Bhasin et al., 2010; McGill et al., 2012; Pantalone and Faiman, 2012; Corona et al., 2013a; Corona et al., 2014a).

Efficacy of Testosterone Therapy

The Endocrine Society guideline considered the evidence of testosterone treatment efficacy for all indications to be of low quality (Bhasin et al., 2010). There is no evidence that the effectiveness of testosterone therapy varies by formulation (ASRM, 2006). The benefits and safety of testosterone therapy in asymptomatic men have not been demonstrated (McGill et al., 2012).

See **Appendix VI** for a summary of the most recent systematic reviews of the efficacy of exogenous testosterone. Some of the systematic reviews had not been published at the time the Endocrine Society guidelines were developed. Systematic reviews have demonstrated the efficacy of testosterone therapy for only 2 specific indications. Findings from a fair-quality meta-analysis suggest that testosterone therapy is more effective for improving sexual function in hypogonadal men with sexual dysfunction than in the general population of men with sexual dysfunction (Corona et al., 2014a). Another fair-quality meta-analysis suggested that testosterone therapy can improve some metabolic outcomes in hypogonadal men with type 2 diabetes (Cai et al., 2014). No other health benefits of testosterone therapy in hypogonadal men have been clearly demonstrated by systematic reviews. None of the systematic reviews discussed the clinical relevance of their pooled estimates. The following is a synopsis of the evidence identified in systematic reviews, review articles, and practice guidelines.

- Secondary Sex Characteristics: No systematic reviews addressed this indication. Endocrine Society guidelines report that in young men, non–placebo-controlled trials have shown testosterone therapy to be associated with improvements in various measures of sexual activity and interest, an increase in androgen-related hair growth, and improvements in physical measures related to body composition (fat-free mass, muscle strength, and bone mineral density [BMD]). Other positive effects have been observed with respect to mood, energy, and well-being. No other evidence was cited in the review articles or other practice guidelines reviewed for the present report.
- Sexual Function: According to the most recent systematic review of the topic, testosterone therapy is generally effective in improving sexual function with low testosterone levels (Corona et al., 2014a). Study groups represented men of all ages. Effects were consistently larger in subgroups of hypogonadal men, compared with overall populations. (The review was judged to be of fair quality in post hoc analysis.)
- Metabolic Outcomes: A meta-analysis of randomized controlled trials (RCTs) involving middle-aged hypogonadal men with type 2 diabetes suggested an effect on measures of glucose control and lipids (triglyceride) but no effect on body fat or blood pressure (Cai et al., 2014). (The review was judged to be of fair quality in post hoc analysis.)
- Depression: According to 1 systematic review, testosterone therapy has an antidepressive effect in middle-aged to elderly men who have a diagnosis of depression. However, no important difference between the hypogonadal subgroup and the overall population was observed (Zarrouf et al., 2009).
- Bone Health: Two systematic reviews found no RCTs evaluating the effectiveness of testosterone therapy for prevention of fractures (Tracz et al., 2006a; MacLean et al., 2008). Participants represented a wide range of ages. The review by Tracz and colleagues identified 8 RCTs that evaluated the effect on BMD. BMD outcomes were mixed, and the authors acknowledged that the evidence permits only a weak inference regarding the effect of testosterone therapy on fracture prevention.
- Cancer-Related Outcomes: No systematic reviews addressed this indication, and no evidence was cited in the review articles or practice guidelines reviewed for the present report. Prostate cancer and breast cancer are considered contraindications for testosterone therapy (Bhasin et al., 2010).
- Cardiovascular Disease-Related Outcomes: A systematic review identified 6 RCTs demonstrating a positive effect of testosterone therapy on intermediate outcomes (treadmill test duration, time to 1 millimeter [mm] ST segment depression) in men with coronary heart disease but included no analysis of whether outcomes differed between hypogonadal men and other study participants. The studies varied as to whether baseline levels of testosterone were normal, abnormally low, or not reported; no pattern of results according to baseline testosterone level could be discerned (Corona et al., 2011c).
- Outcomes Related to Chronic Kidney Disease: No systematic reviews addressed the effectiveness of testosterone therapy in patients with chronic kidney disease.
- COPD-Related Outcomes: No systematic reviews evaluated the effectiveness of testosterone therapy in men with COPD through meta-analysis. However, 1 review cited 2 small studies suggesting that testosterone therapy can improve lean body mass in patients with COPD, but

neither the symptom profile or baseline testosterone levels of the patients in those studies were described (Corona et al., 2013b).

- HIV-Related Outcomes: A meta-analysis of RCTs of testosterone therapy in HIV-infected men found that testosterone treatment, compared with placebo, was associated with greater gains in lean body mass and total body weight when baseline testosterone levels were normal but not when baseline levels were borderline or hypogonadal (Johns et al., 2005). Authors described the benefits as small and stated that the evidence was insufficient for recommending the treatment.
- Global Measures of Well-Being: No systematic reviews assessed the impact of testosterone therapy on quality of life (QOL), functional status, physical performance, or overall mental health.
- Use of Medications that Can Suppress Testosterone Production: The Endocrine Society guidelines cited a small, open-label trial in which testosterone therapy was found to improve sexual function, well-being, and mood in men who had opioid-induced androgen deficiency. The same guidelines also cited 2 placebo-controlled trials in which testosterone therapy was associated with greater gains in lean body mass as well as greater decreases in fat mass in men receiving glucocorticoid treatment for bronchial asthma or COPD. The symptom profile of the study populations was not described, and the baseline testosterone levels in the studies of testosterone therapy for men taking glucocorticoids were not described (Bhasin et al., 2010).

Safety of Testosterone Therapy

According to published reports, adverse events in young, hypogonadal men occur at low frequency. An increase in hematocrit, acne and oily skin, fluid retention, testicular atrophy, worsening of lower urinary tract symptoms, subclinical prostate cancer, growth of metastatic prostate cancer, and reduced sperm production and fertility have been reported as adverse effects. Other less common adverse events are suspected of being associated with testosterone therapy: gynecomastia (breast enlargement), male pattern balding, growth of breast cancer, new or worsening heart failure, and induction or worsening of obstructive sleep apnea (Bhasin et al., 2010; McGill et al., 2012).

Evidence from systematic reviews cited in the Endocrine Society guidelines suggested that testosterone therapy in older men may increase prostate events (nonsignificant estimate), substantially increases the risk of erythrocytosis, and has no clear effect on obstructive sleep apnea, cardiovascular events, or mortality (Bhasin et al., 2010). More recent evidence is summarized in **Appendix VII** and is generally applicable to men ≥ 60 years of age. The newer evidence is consistent with the Society's conclusions regarding the effect on prostate events, erythrocytosis, and sleep apnea, but the most recent pooled estimate for prostate events continues to be nonsignificant and very imprecise. Newer evidence does not resolve the uncertainty about risk of cardiovascular events but does suggest that there may be an elevated risk in older men and men with a history of cardiovascular disease. Newer evidence regarding the effect of testosterone therapy on mortality suggests a *protective* effect, but this evidence comes from a single observational study (Shores et al., 2012) and cannot be considered conclusive.

In response to 2 of the studies described in **Appendix VII** (Vigen et al., 2013; Finkle et al., 2014), the Food and Drug Association (FDA) recently announced a decision to reassess the safety of FDA-approved testosterone products (CDER, 2014b). Additionally, the FDA issued a new requirement in 2014 that manufacturers include a general warning in the drug labeling of all approved testosterone products about the risk of blood clots in the veins due to postmarket reports of venous blood clots unrelated to

polycythemia (another term for erythrocytosis). Because these clots occur in the veins, this new warning is not related to the FDA's ongoing evaluation of the possible risk of stroke, heart attack, and death in patients taking testosterone products.

Contraindications: Given the observations of prostate and breast cancer growth and the fact that testosterone therapy is a hormonal therapy, testosterone therapy is contraindicated in men with either form of cancer. Although evidence from randomized trials is lacking, some experts believe that testosterone therapy is safe after treatment for prostate cancer when certain signs of remission or cure are evident. Guidelines advise that testosterone therapy should also be avoided in other conditions because of its known or potential ability to worsen the disease: elevated hematocrit (> 50%), untreated severe obstructive sleep apnea, severe lower urinary tract symptoms, and uncontrolled or poorly controlled heart failure (Bhasin et al., 2010).

Description of Testosterone Testing

Serum testosterone measurements can target total testosterone, free testosterone, or bioavailable testosterone. Most of the circulating testosterone is bound to SHBG, while approximately 0.5% to 3% is *free*. *Bioavailable testosterone* refers to unbound testosterone plus testosterone that is bound loosely to albumin. Albumin-bound testosterone is assumed to be bioavailable. Serum testosterone measurements should begin with total testosterone and should take into account the considerable intraindividual variation in testosterone measurements. Levels are typically approximately 30% higher between the hours of 8:00 a.m. and 10:00 a.m. compared with later in the day. Studies have also shown that approximately 30% of men with an initial testosterone level in the mildly hypogonadal range have a normal result on repeat measurement. Therefore, serum testosterone measurements should be made early in the morning and testosterone therapy should not be considered without a diagnosis of androgen deficiency based on 2 measurements of total testosterone on different days. Lastly, testing during acute illness or during a time of decompensation of chronic illness is not advised since testosterone levels may be temporarily depressed during such times. Relevant chronic illnesses include coronary artery disease, heart failure, and diabetes. The recommended target of testosterone therapy is the middle of a normal range for healthy young men (Bhasin et al., 2010; McGill et al., 2012; Pantalone and Faiman, 2012).

Guidelines recommend that free or bioavailable testosterone be measured when total testosterone levels are close to the lower limit of the normal range and when altered SHBG levels are suspected, as may be the case in older men and men with obesity, diabetes mellitus, cachexia (disease-related wasting), malnutrition, advanced cirrhosis, acromegaly, hypothyroidism, or nephrotic syndrome (a disorder characterized by deficiency of albumin in the blood and excess excretion of protein). SHBG levels increase with age, which diminishes the levels of bioavailable testosterone at a rate greater than the rate at which total testosterone declines. Thus, the potential for an underestimation of low serum testosterone increases with increasing age if measurement is restricted to total testosterone. However, it is also possible for total testosterone to be low but free or bioavailable testosterone to be normal (ASRM, 2006; Bhasin et al., 2010; Ho, 2011; Pantalone and Faiman, 2012).

Following a determination of abnormally low serum testosterone, best practice requires measurement of serum LH and FSH levels to distinguish between primary (testicular) and secondary (pituitary-hypothalamic) hypogonadism, as well as other tests for possible causes of primary or secondary hypogonadism. LH controls the production of testosterone, whereas FSH controls the production of

sperm. Additional laboratory tests and/or imaging are recommended for cases of secondary hypogonadism in order to evaluate etiology and exclude diagnoses such as pituitary neoplasia, hyperprolactinemia (over-production of prolactin, a hormone produced by the pituitary gland that can cause infertility in excess), hemochromatosis (a hereditary disorder of metabolism in which iron-containing pigments collect in tissue), obstructive sleep apnea, and genetic disorders. A karyotype is recommended for cases of primary hypogonadism of unknown etiology to rule out Klinefelter's syndrome (Bhasin et al., 2010; Pantalone and Faiman, 2012).

Analytic Validity

Analytic validity refers to the ability of a test to accurately detect the target analyte. It might be thought of as the technical accuracy of the test. The analytic validity of testosterone testing methods is generally hampered by the following (Paduch et al., 2014; Shea et al., 2014): possible interference from similar analytes; differences in the reference populations used by different laboratories and lack of control for medical comorbidities; diurnal, seasonal, and age-related variations; and variation from day to day in the same individual. As already noted, some of the intraindividual variability in test results can be addressed by standardizing the time of day when blood samples are drawn (early morning) and requiring confirmation of low levels by drawing and testing blood samples on 2 different days.

An additional problem is the impracticality of using the most accurate methods in typical laboratories. The gold standard technique for *total* testosterone measurement is liquid chromatography-tandem mass spectrometry (LC-MSMS) assay. However, the cost and complexity of mass spectrometry are considered prohibitive for most laboratories (McGill et al., 2012; Paduch et al., 2014). A study conducted in 2004 found that of 6 immunoassay kits currently in wide use in clinical laboratories, 2 automated assays and 1 manual assay had acceptable accuracy for measuring total testosterone according to a standard definition: results within ± 20% of results from LC-MSMS in over 60% of samples. The authors concluded that these assays were capable of distinguishing eugonadal from hypogonadal males if adult male reference ranges were established in each individual laboratory (Wang et al., 2004). More recent review articles and guidelines also state that automated immunoassays usually have sufficient accuracy to distinguish between men who have serum levels consistent with hypogonadism and men who have adequate total levels of serum testosterone, and that automated enzyme-linked immunoassays are currently available in most hospital laboratories. However, these authors point out that such assays have poor performance for precisely measuring testosterone levels in men whose true levels are close to the lower limit of a normal range. Although the normative range for healthy young men varies across laboratories, lower limits are commonly in the range of 280 to 300 ng/dL (9.8 to 10.4 nmol/L) (Bhasin et al., 2010; Ho, 2011; McGill et al., 2012; Paduch et al., 2014).

The reference standard for measurement of *free* testosterone and *bioavailable* testosterone is centrifugal ultrafiltration or equilibrium dialysis, which is also not likely to be available in local laboratories. Free testosterone can be directly measured through the use of immunoassay methods (Brand et al., 2011; Morales et al., 2012). Alternatively, free and bioavailable testosterone can be calculated by various formulas on the basis of total testosterone and SHBG assays. One formula yields the free testosterone index (FTI), also known as the free androgen index (FAI), but this formula has given way to a more complex methods of converting the measured concentrations of total testosterone and SHBG into an index of free or bioavailable testosterone. Like total testosterone, SHBG is routinely measured by commercially available immunoassay kits (Shea et al., 2007; Bhasin et al., 2010; Ho, 2011; Ho and Beckett, 2011; McGill et al., 2012; Corona et al., 2013b; Paduch et al., 2014).

The Centers for Disease Control and Prevention (CDC) administers a voluntary Hormone Standardization Program to help laboratories maintain and enhance the quality and comparability of measurement results obtained from steroid hormone and vitamin D tests. The Program includes publication of laboratories that have demonstrated accuracy in their performance of testosterone testing (CDC, 2014). The College of American Pathologists also provides quality control programs (CAP, 2015). Data from the CAP program for year 2004 showed that across 11 immunoassay kits used to test the same serum samples, median values ranged from 215 to 345 ng/dL (Wang et al., 2004). Data from 72 labs participating in the 2011 CAP program showed good consistency within manufacturer but wide variation across 6 manufacturers: mean levels from just under 4 nmol/L to just over 16 nmol/L (Morales et al., 2012). However, both the CDC and CAP programs offer laboratories the opportunity to compare their own assay results for a reference set of serum samples, supplied by the quality control program, with results obtained by a gold standard method for the same samples. Participating laboratories can then recalibrate their test results. Participation in these 2 quality control programs is voluntary and not required for accreditation.

Measurement of salivary testosterone has been proposed as an alternative to serum testosterone, but measurement of salivary testosterone is not a standard practice. Other methods that have not yet proven satisfactory for routine practice are direct analog assays for free testosterone and ammonium sulfate precipitation for bioavailable testosterone (Shea et al., 2007; Ho and Becket, 2011).

Clinical Validity

Clinical validity refers to the ability of a test to discriminate between individuals who do and do not have a clinical condition, or to make accurate clinical predictions. In other words, a test has clinical validity if it enables valid clinical decision making. No generally recognized age- and gender-specific cutoff points have been defined for making clinical predictions based on serum testosterone levels (Shea et al., 2007). In symptomatic men, regardless of age, testosterone levels are assessed by comparing them with the normal range for *young men,* based on the known gradual decline of testosterone levels with aging, starting at approximately age 30 years. The lower limit of *normal for healthy young men* is 280 to 300 ng/dL (9.7 to 10.4 nmol/L) (ASRM, 2006; Bhasin et al., 2010). The panel assembled for the current Endocrine Society guidelines was divided between 2 options for symptomatic older men (Bhasin et al., 2010):

- Treat only when levels of *total* testosterone are < 300 ng/dL (10.4 nmol/L) because of the association between testosterone at those levels and typical symptoms of androgen deficiency.
- Treatment only when levels of *total* testosterone are < 200 ng/dL (6.9 nmol/L) because randomized trials have suggested that testosterone therapy is ineffective in men with pretreatment values of 300 ng/dL.

There is no consensus regarding cutoff values for *free* or *bioavailable* testosterone, but a level of > 225 picomoles per liter (pmol/L) (6 ng/dL) is generally considered normal (McGill et al., 2012). Compared with total testosterone levels, bioavailable testosterone levels have the stronger relationship with some health effects (ASRM, 2006); see **Appendix V** for examples. Thus, the validity of testosterone testing as a predictor of health depends on the target condition and which components of serum testosterone are measured.

Monitoring

There is no consensus on the optimal frequency for monitoring testosterone levels in men receiving testosterone therapy. Experts advise that clinical response as well as testosterone levels be monitored to determine whether adjustments are needed in the therapy (ASRM, 2006; Bhasin et al, 2010; McGill et al., 2012).

Washington State Agency Utilization and Costs

Figure 1a. PEBB/UMP Testosterone Tests and Treatments, Costs and Counts, 2010-2013

Public Employees Benefits (PEBB) /Uniform Medical Plan (UMP)	2010	2011	2012	2013	4 Yr Overall **	Avg Annual % Chg	
Populations							
Avg Annual Members	213,487	212,596	212,684	222,339		1.4%	
Men	96,435	96,283	95,581	101,144		0.2%	*
Testosterone Tests							
Paid $ 84402 [1]	$92,270	$66,432	$70,811	$75,560	$305,073	-10.0%	*
Paid $ 84403 [2]	$217,732	$190,031	$198,658	$195,132	$801,553	-5.4%	*
Testing Paid $ (PEBB Primary)	**$310,002**	**$256,463**	**$269,469**	**$270,692**	**$1,106,626**	**-6.6%**	*
Test Counts - 84402	2627	2678	2849	3343	11497	6.4%	*
Test Counts - 84403	6226	7133	7470	8353	29182	8.0%	*
Total Test Counts	**8,853**	**9,811**	**10,319**	**11,696**	**40,679**	**7.6%**	*
Count of Patients (Women)	1600	1520	1465	1611	4929	-1.2%	*
Count of Patients (Men)	2861	3231	3552	4075	9470	9.9%	*
Total Patient Counts	**4461**	**4751**	**5017**	**5686**	**14399**	**6.5%**	*
Avg Tests per patient	2.0	2.1	2.1	2.1	2.8		
Avg Tests per pt (Men)	2.0	2.0	2.0	2.0	2.9		
Avg Paid $ 84402(PEBB Primary -93%)	$36	$27	$26	$24	$28	-11.6%	
Avg Paid $ 84403 (PEBB Primary)	$37	$29	$29	$26	$30	-10.9%	
Testosterone Supplementation							
Paid $ Injections	$12,862	$14,497	$21,807	$24,623	$73,789	17.5%	*
Paid $ Testosterone Pharmaceuticals	$782,110	$972,702	$836,191	$1,153,242	$3,744,245	9.3%	*
Total Paid for Treatments (PEBB Prim.)	**$794,972**	**$987,199**	**$857,998**	**$1,177,865**	**$3,818,034**	**9.5%**	*
Avg Paid per pt (PEBB Primary)	$744	$841	$636	$749	$1,460	2.1%	
Avg Paid per pt (PEBB Primary - Men)	$754	$852	$642	$754	$1,438	1.9%	

*Average Annual Percent Change adjusted for population.

**Four-year patient counts are unique patients over the 4 years reported, and are not necessarily the total of annual patient counts.

[1] 84402 Assay of Testosterone

[2] 84403 Assay of Total Testosterone. In most cases, Assay of Total Testosterone was the only test given (62.9% of encounters). In 5.7% of encounters, 84402, Assay of Testosterone test was given alone. In 31.4% of encounters, both tests occurred on the same day. Test 84402 is slightly more likely to be used for women, accounting for 35% of testing in women, compared with 26% of testing in men.

Figure 1b. L&I Testosterone Tests and Treatments, Costs and Counts, 2010-2013

Labor and Industries (L&I)		2010	2011	2012	2013	4 Yr Overall **	Avg Annual % Chg	
Populations	Avg Annual Members	122,712	121,043	121,660	123,159		0.1%	
Testosterone Tests								
	Paid $ 84402[1] Assay of Testosterone	$856	$555	$904	$901	$3215	-5.1%	*
	Paid $ 84403[1] Assay of Total Testosterone	$2787	$2817	$2827	$2684	$11,116	-1.5%	*
	Total Paid for Testing	**$3643**	**$3372**	**$3731**	**$3585**	**$14,331**	**-0.9%**	*
	Test Counts - 84402 Assay of Testosterone	18	11	18	19	66	-6.2%	*
	Test Counts - 84403 Assay of Total Testosterone	55	56	56	54	221	-0.8%	*
	Total Test Counts	**73**	**67**	**74**	**73**	**287**	**-0.4%**	*
	Total Patient Counts (Men only)	**51**	**46**	**48**	**48**	**169**	**-2.3%**	*
	Avg Tests per patient	1.4	1.5	1.5	1.5	1.7		
	Avg Paid per 84402 Assay of Testosterone	$48	$50	$50	$47	$49	0.0%	
	Avg Paid per 84403 Assay of Total Testosterone	$51	$50	$50	$50	$50	-0.6%	
Testosterone Treatments								
	Paid $ Injection HCPCS	$749	$599	$919	$793	$3,060	-2.0%	*
	Paid $ Testosterone Pharmaceuticals	$14,691	$14,195	$20,844	$24,464	$74,195	14.4%	*
	Total Paid for Treatments	**$15,440**	**$14,794**	**$21,763**	**$25,257**	**$77,254**	**13.8%**	*
	Avg Paid per patient	$303	$322	$453	$526	$457	21.1%	

*Average Annual Percent Change adjusted for population.

**Four-year patient counts are unique patients over the 4 years reported, and are not necessarily the total of annual patient counts.

[1] 84402 Assay of Testosterone

[2] 84403 Assay of Total Testosterone.

NOTE: Agency payment schedules do not include testosterone test CPTs 84402 and 84403.

Figure 2. Agency Fee Schedules

Current pricing for Testosterone Testing as available on agency web sites:

CPT Codes	CPT Code Descriptions	Current Agency Fees (Allowed)		
		PEBB/UMP*	L&I†	Medicaid‡
84402	**Assay of testosterone**	$35.00	$48.64	$27.86
84403	**Assay of total testosterone**	$28.00	$49.31	$28.25

*Regence Blue Shield Provider Fee Schedule – effective January 1 2013, MD/DO/DPM Provider rates, Maximum Allowable fee, http://www.hca.wa.gov/ump/documents/Regence_Professional_Fee_Schedule_Jan_2013.pdf. Accessed 12/15/2014. Payment based on the Regence Fee Schedule is subject to all of the terms and conditions of the applicable Regence BlueShield provider agreement, member benefits, Regence BlueShield policies, and all published Regence BlueShield administrative guidelines. Therefore, the appearance of fees for particular procedure codes does not guarantee coverage. Some providers may have contracted fees at different rates.

†Washington State Labor and Industries Fee Schedules and Payment Policies (MARFS), Fee Schedules and Payment Policies for: 2014, http://www.lni.wa.gov/apps/FeeSchedules/. Accessed 12/15/2014.

‡Washington State Medicaid Rates Development Fee Schedule, July 1, 2014 Physician and Related Services Fee Schedule (Updated October 1, 2014), http://www.hca.wa.gov/medicaid/rbrvs/pages/index.aspx#P. Accessed 12/15/2014.

Figure 3. PEBB Testosterone Test Patients by Age and Gender, 2010-2013

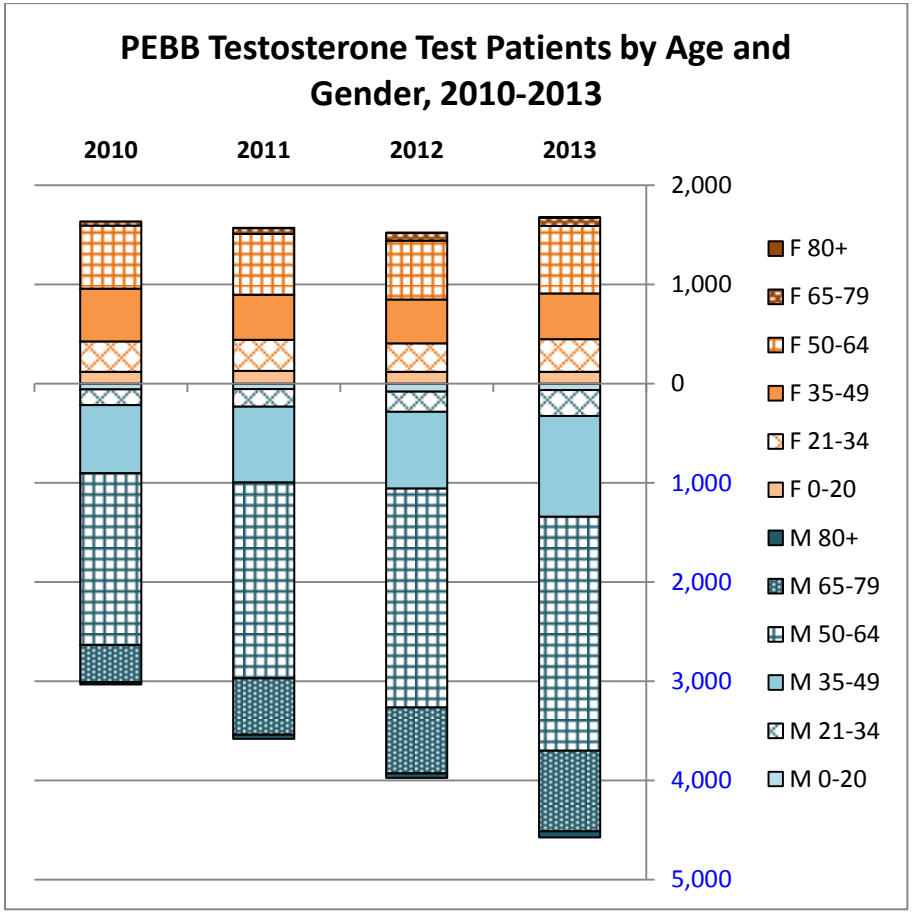
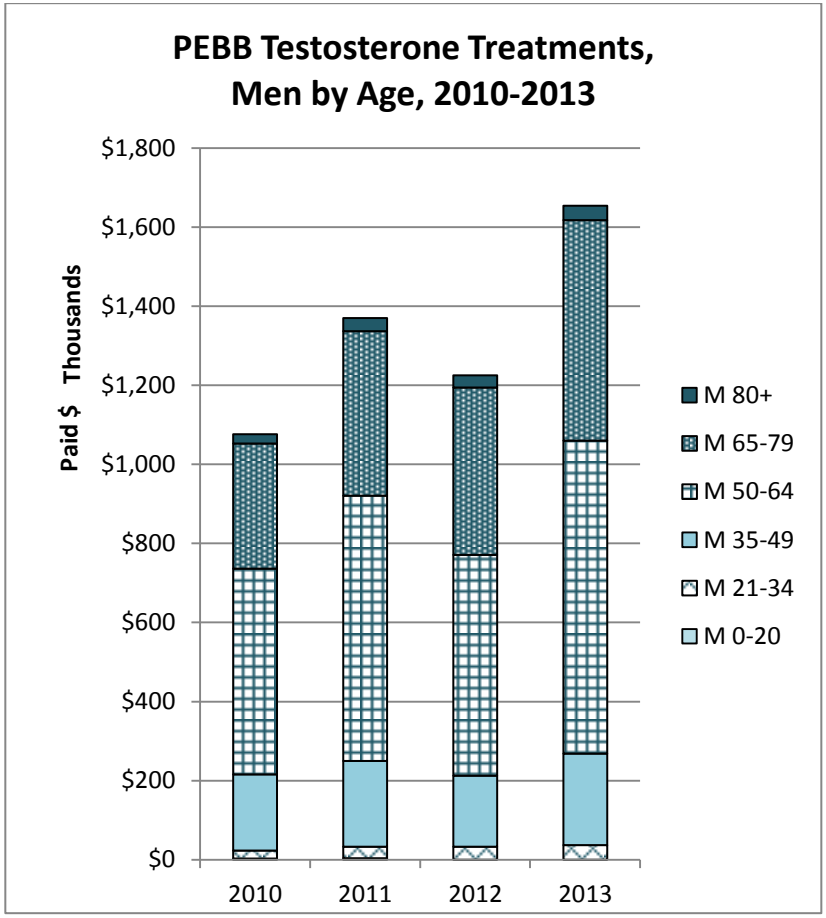


Figure 3a. PEBB Testosterone Treatment Patients by Age and Gender, Men only, 2010-2013



NOTE: All L&I Testosterone Test patients were men, with 40% between 35-49, and 43% between 50-64.

Figure 4a. PEBB Test Counts per Patient, 2010-2013

Number Of Test Dates	Patient Count
20+	25
10 - 19	142
5 - 9	917
4	597
3	1,132
2	2,568
1	10,140

Figure 4b. L&I Test Counts per Patient, 2010-2013

Number Of Test Dates	Patient Count
5+	3
4	2
3	7
2	22
1	135

Figure 5a. PEBB Top Diagnoses by Paid $ for Testosterone Tests, 2010-2013

Diagnosis Code	Top 20 Diagnoses of 1272	Paid $	Percent of Total $
Total		$1,118,096	
257.2	TESTICULAR HYPOFUNC NEC	$176,175	15.8%
780.79	MALAISE AND FATIGUE NEC	$104,311	9.3%
V70.0	ROUTINE MEDICAL EXAM	$74,659	6.7%
244.9	HYPOTHYROIDISM NOS	$41,551	3.7%
627.2	SYMPT FEM CLIMACT STATE	$39,892	3.6%
272.4	HYPERLIPIDEMIA NEC/NOS	$37,869	3.4%
607.84	IMPOTENCE, ORGANIC ORIGN	$30,218	2.7%
250	DMII WO CMP NT ST UNCNTR	$28,627	2.6%
185	MALIGN NEOPL PROSTATE	$28,355	2.5%
401.9	HYPERTENSION NOS	$20,227	1.8%
401.1	BENIGN HYPERTENSION	$16,964	1.5%
799.81	DECREASED LIBIDO	$16,031	1.4%
272.2	MIXED HYPERLIPIDEMIA	$15,626	1.4%
272	PURE HYPERCHOLESTEROLEM	$14,149	1.3%
268.9	VITAMIN D DEFICIENCY NOS	$14,100	1.3%
626.4	IRREGULAR MENSTRUATION	$14,093	1.3%
253.4	ANTER PITUITARY DIS NEC	$11,890	1.1%
259.9	ENDOCRINE DISORDER NOS	$11,500	1.0%
V58.69	LONG-TERM USE MEDS NEC	$11,477	1.0%
250.02	DMII WO CMP UNCNTRLD	$10,293	0.9%

Figure 5b. PEBB Top Diagnoses, Paid $ for Testosterone Tests (Men), 2010-2013

Diagnosis Code	Top 20 Diagnoses of 1,039 (Men)	Paid $	% Total $
Total		$796,612	
257.2	TESTICULAR HYPOFUNC NEC	$175,886	22.1%
780.79	MALAISE AND FATIGUE NEC	$73,737	9.3%
V70.0	ROUTINE MEDICAL EXAM	$63,501	8.0%
272.4	HYPERLIPIDEMIA NEC/NOS	$32,404	4.1%
607.84	IMPOTENCE, ORGANIC ORIGN	$30,218	3.8%
185	MALIGN NEOPL PROSTATE	$28,355	3.6%
250	DMII WO CMP NT ST UNCNTR	$26,655	3.3%
401.9	HYPERTENSION NOS	$18,697	2.3%
244.9	HYPOTHYROIDISM NOS	$16,706	2.1%
401.1	BENIGN HYPERTENSION	$15,128	1.9%
272.2	MIXED HYPERLIPIDEMIA	$13,545	1.7%
272	PURE HYPERCHOLESTEROLEM	$12,517	1.6%
799.81	DECREASED LIBIDO	$11,959	1.5%
253.4	ANTER PITUITARY DIS NEC	$11,838	1.5%
250.02	DMII WO CMP UNCNTRLD	$8,815	1.1%
V58.69	LONG-TERM USE MEDS NEC	$8,458	1.1%
268.9	VITAMIN D DEFICIENCY NOS	$7,483	0.9%
V76.44	SCRN MALIG NEOP-PROSTATE	$6,199	0.8%
302.72	INHIBITED SEX EXCITEMENT	$5,633	0.7%
259.9	ENDOCRINE DISORDER NOS	$5,237	0.7%

Figure 5c. PEBB Top Diagnoses, Paid $ for Testosterone Tests (Women), 2010-2013

Diagnosis Code	Top 20 Diagnoses of 573 (Women)	Paid $	% Total $
Total		$321,289	
627.2	SYMPT FEM CLIMACT STATE	$39,845	12.4%
780.79	MALAISE AND FATIGUE NEC	$30,574	9.5%
244.9	HYPOTHYROIDISM NOS	$24,845	7.7%
626.4	IRREGULAR MENSTRUATION	$14,093	4.4%
V70.0	ROUTINE MEDICAL EXAM	$11,158	3.5%
626	ABSENCE OF MENSTRUATION	$9,403	2.9%
256.4	POLYCYSTIC OVARIES	$8,909	2.8%
704	ALOPECIA NOS	$8,209	2.6%
704.1	HIRSUTISM	$7,846	2.4%
268.9	VITAMIN D DEFICIENCY NOS	$6,617	2.1%
259.9	ENDOCRINE DISORDER NOS	$6,263	1.9%
627.9	MENOPAUSAL DISORDER NOS	$5,891	1.8%
V72.31	ROUTINE GYN EXAMINATION	$5,557	1.7%
272.4	HYPERLIPIDEMIA NEC/NOS	$5,439	1.7%
626.1	SCANTY MENSTRUATION	$4,767	1.5%
799.81	DECREASED LIBIDO	$4,072	1.3%
706.1	ACNE NEC	$3,962	1.2%
625.4	PREMENSTRUAL TENSION	$3,774	1.2%
626.2	EXCESSIVE MENSTRUATION	$3,431	1.1%
V58.69	LONG-TERM USE MEDS NEC	$3,019	0.9%

Figure 5d. L&I Top Diagnoses by Paid $ for Testosterone Tests, 2010-2013

Diagnosis Code	Top 20 Diagnoses of 573 (Women)	Paid $	% Total $
Total		$14,331	
257.2	OTHER TESTICULAR HYPOFUNCTION	$3,248	22.7%
V72.60	LABORATORY EXAMINATION UNSPECIFIED	$1,084	7.6%
780.79	OTHER MALAISE AND FATIGUE	$1,055	7.4%
607.84	IMPOTENCE OF ORGANIC ORIGIN	$713	5.0%
V58.69	LONG-TERM (CURRENT) USE OF OTHER MEDICAT	$520	3.6%
722.10	DISPLCMT LUMBAR INTERVERT DISC W/O MYELO	$358	2.5%
338.4	CHRONIC PAIN SYNDROME	$256	1.8%
311	DEPRESSIVE DISORDER NOT ELSEWHERE CLASSI	$255	1.8%
244.9	UNSPECIFIED HYPOTHYROIDISM	$253	1.8%
296.33	MAJ DPRSV D/O RECUR EPIS SEV W/O PSYCHOT	$206	1.4%
722.4	DEGENERATION OF CERVICAL INTERVERTEBRAL	$206	1.4%
846.0	SPRAIN AND STRAIN OF LUMBOSACRAL	$156	1.1%
799.81	DECREASED LIBIDO	$155	1.1%
401.1	ESSENTIAL HYPERTENSION, BENIGN	$155	1.1%
344.1	PARAPLEGIA	$154	1.1%
V70.0	ROUTINE GENERAL MEDICAL EXAM@HEALTH CARE	$151	1.1%
847.2	LUMBAR SPRAIN AND STRAIN	$150	1.0%
724.2	LUMBAGO	$148	1.0%
453.40	AC VENUS EMBO & THROMB UNSPEC DEEP VES L	$139	1.0%
596.54	NEUROGENIC BLADDER, NOS	$110	0.8%

Related Medical Codes

Code Type	Codes	Short Description
CPTs	84402	Assay of testosterone
	84403	Assay of total testosterone
HCPCS	J1060-J1090	Testosterone cypionate injections
	J3120-J3150	Testosterone injections, enanthate, propionate, suspension
ICD9	302.50-53	Transgender diagnoses

Testosterone Pharmaceuticals by NDC

NDC	Brand	Label	Route
9034702	DEPO-TESTOSTERONE	DEPO-TESTOSTERONE 100 MG/ML VL	INTRAMUSC
9041701	DEPO-TESTOSTERONE	DEPO-TESTOSTERONE 200 MG/ML	INTRAMUSC
9041702	DEPO-TESTOSTERONE	DEPO-TESTOSTERONE 200 MG/ML	INTRAMUSC
51842501	ANDROGEL	ANDROGEL 1%(2.5G) GEL PACKET	TRANSDERM
51842530	ANDROGEL	ANDROGEL 1%(2.5G) GEL PACKET	TRANSDERM
51845030	ANDROGEL	ANDROGEL 1%(5G) GEL PACKET	TRANSDERM
51848888	ANDROGEL	ANDROGEL 1% GEL PUMP	TRANSDERM
115703701	METHITEST	METHITEST 10 MG TABLET	ORAL
187090101	TESTRED	TESTRED 10 MG CAPSULE	ORAL
187090201	ANDROID	ANDROID 10 MG CAPSULE	ORAL
574082001	TESTOSTERONE CYPIONATE	TESTOSTERONE CYP 200 MG/ML	INTRAMUSC
574082010	TESTOSTERONE CYPIONATE	TESTOSTERONE CYP 200 MG/ML	INTRAMUSC
574082105	TESTOSTERONE ENANTHATE	TESTOSTERONE ENAN 200 MG/ML	INTRAMUSC
591322126	TESTOSTERONE ENANTHATE	TESTOSTERONE ENAN 200 MG/ML	INTRAMUSC
591322379	TESTOSTERONE CYPIONATE	TESTOSTERON CYP 2,000 MG/10 ML	INTRAMUSC
781307370	TESTOSTERONE CYPIONATE	TESTOSTERONE CYP 100 MG/ML	INTRAMUSC
781307470	TESTOSTERONE CYPIONATE	TESTOSTERONE CYP 200 MG/ML	INTRAMUSC
781307471	TESTOSTERONE CYPIONATE	TESTOSTERONE CYP 200 MG/ML	INTRAMUSC
21695011230	ANDROGEL	ANDROGEL 1%(5G) GEL PACKET	TRANSDERM
35356005810	DEPO-TESTOSTERONE	DEPO-TESTOSTERONE 100 MG/ML VL	INTRAMUSC
35356037605	ANDROGEL	ANDROGEL 1%(5G) GEL PACKET	TRANSDERM
43773100102	TESTOPEL	TESTOPEL 75 MG PELLETS	IMPLANT
43773100103	TESTOPEL	TESTOPEL 75 MG PELLETS	IMPLANT
43773100104	TESTOPEL	TESTOPEL 75 MG PELLETS	IMPLANT
52544046954	ANDRODERM	ANDRODERM 2.5 MG/24HR PATCH	TRANSDERM
52544046960	ANDRODERM	ANDRODERM 2.5 MG/24HR PATCH	TRANSDERM
52544047030	ANDRODERM	ANDRODERM 5 MG/24HR PATCH	TRANSDERM

NDC	Brand	Label	Route
52544047054	ANDRODERM	ANDRODERM 5 MG/24HR PATCH	TRANSDERM
54569213100	TESTOSTERONE CYPIONATE	TESTOSTERONE CYP 200 MG/ML	INTRAMUSC
54569530100	DEPO-TESTOSTERONE	DEPO-TESTOSTERONE 100 MG/ML VL	INTRAMUSC
54569533900	ANDROGEL	ANDROGEL 1%(5G) GEL PACKET	TRANSDERM
54569559500	TESTIM	TESTIM 1% (50MG) GEL	TRANSDERM
54868021600	DEPO-TESTOSTERONE	DEPO-TESTOSTERONE 200 MG/ML	INTRAMUSC
54868021601	DEPO-TESTOSTERONE	DEPO-TESTOSTERONE 200 MG/ML	INTRAMUSC
54868079600	DEPO-TESTOSTERONE	DEPO-TESTOSTERONE 100 MG/ML VL	INTRAMUSC
54868361800	TESTOSTERONE CYPIONATE	TESTOSTERONE CYP 200 MG/ML	INTRAMUSC
54868361801	TESTOSTERONE CYPIONATE	TESTOSTERONE CYP 200 MG/ML	INTRAMUSC
54868370400	ANDRODERM	ANDRODERM 2.5 MG/24HR PATCH	TRANSDERM
54868479200	ANDROGEL	ANDROGEL 1%(2.5G) GEL PACKET	TRANSDERM
54868481000	ANDROGEL	ANDROGEL 1%(5G) GEL PACKET	TRANSDERM
54868498900	TESTIM	TESTIM 1% (50MG) GEL	TRANSDERM
54868501600	DELATESTRYL	DELATESTRYL 200 MG/ML VIAL	INTRAMUSC
54868581400	ANDROGEL	ANDROGEL 1% GEL PUMP	TRANSDERM
54868603200	ANDRODERM	ANDRODERM 5 MG/24HR PATCH	TRANSDERM
55056306001	STRIANT	STRIANT 30 MG MUCOADHESIVE	BUCCAL
58016096700	TESTRED	TESTRED 10 MG CAPSULE	ORAL
58016096730	TESTRED	TESTRED 10 MG CAPSULE	ORAL
58016096760	TESTRED	TESTRED 10 MG CAPSULE	ORAL
58016096790	TESTRED	TESTRED 10 MG CAPSULE	ORAL
65628002001	FIRST-TESTOSTERONE	FIRST 2% TESTOSTERONE OINT	TRANSDERM
65628002101	FIRST-TESTOSTERONE MC	FIRST-TESTOSTERONE MC 2% CR	TRANSDERM
66887000105	TESTIM	TESTIM 1% (50MG) GEL	TRANSDERM
67979050140	DELATESTRYL	DELATESTRYL 200 MG/ML VIAL	INTRAMUSC
16590071930	ANDROGEL	ANDROGEL 1%(5G) GEL PACKET	TRANSDERM
16590085330	TESTIM	TESTIM 1% (50MG) GEL	TRANSDERM
63481018316	FORTESTA	FORTESTA 10 MG GEL PUMP	TRANSDERM
2197590	AXIRON	AXIRON 30 MG/ACTUATION SOLN	TRANSDERM
51846233	ANDROGEL	ANDROGEL 1.62% GEL PUMP	TRANSDERM
54569533901	ANDROGEL	ANDROGEL 1%(5G) GEL PACKET	TRANSDERM
52544007654	ANDRODERM	ANDRODERM 2 MG/24HR PATCH	TRANSDERM
52544007660	ANDRODERM	ANDRODERM 2 MG/24HR PATCH	TRANSDERM
52544007730	ANDRODERM	ANDRODERM 4 MG/24HR PATCH	TRANSDERM
52544007754	ANDRODERM	ANDRODERM 4 MG/24HR PATCH	TRANSDERM
35356075830	TESTIM	TESTIM 1% (50MG) GEL	TRANSDERM

NDC	Brand	Label	Route
54569633700	ANDROGEL	ANDROGEL 1.62% GEL PUMP	TRANSDERM
51846230	ANDROGEL	ANDROGEL 1.62%(2.5G) GEL PCKT	TRANSDERM
51846231	ANDROGEL	ANDROGEL 1.62%(1.25G) GEL PCKT	TRANSDERM
52244003060	STRIANT	STRIANT 30 MG MUCOADHESIVE	BUCCAL
143972601	TESTOSTERONE CYPIONATE	TESTOSTERON CYP 2,000 MG/10 ML	INTRAMUSC
143975001	TESTOSTERONE ENANTHATE	TESTOSTERON ENAN 1,000 MG/5 ML	INTRAMUSC
62756001640	TESTOSTERONE CYPIONATE	TESTOSTERON CYP 2,000 MG/10 ML	INTRAMUSC
54868640600	AXIRON	AXIRON 30 MG/ACTUATION SOLN	TRANSDERM
62756001540	TESTOSTERONE CYPIONATE	TESTOSTERONE CYP 200 MG/ML	INTRAMUSC
62756001740	TESTOSTERONE CYPIONATE	TESTOSTERON CYP 1,000 MG/10 ML	INTRAMUSC
9052001	DEPO-TESTOSTERONE	DEPO-TESTOSTERONE 200 MG/ML	INTRAMUSC
9052010	DEPO-TESTOSTERONE	DEPO-TESTOSTERONE 200 MG/ML	INTRAMUSC
67979051143	AVEED	AVEED 750 MG/3 ML VIAL	INTRAMUSC

Review Objectives and Analytic Framework

PICO

The scope of this report is defined as:

Population: Adult men.

Interventions: Measurement of circulating total, free, or bioavailable testosterone as an initial assessment of possible hypogonadism.

Comparisons: Investigation and clinical management of symptoms or health problems without the use of testosterone testing.

Outcomes: Outcomes such as: symptom improvement; general health outcomes (e.g., osteoporosis, chronic disease, mortality); clinical management decisions; potential harms resulting from testosterone treatment decisions; cost and cost-effectiveness.

Key Questions

The following key questions will be addressed:

1. Is there evidence that testosterone testing improves outcomes?
1a. Does the impact on outcomes vary according to age, race/ethnicity, baseline testosterone levels, treatment status, or clinical history?
1b. What is the minimum interval required to assess a change in testosterone status in untreated and treated individuals?
2. What are the potential harms of testosterone testing, including potential subsequent harms resulting from treatment decisions?
3. What are the costs and cost-effectiveness of testosterone testing?

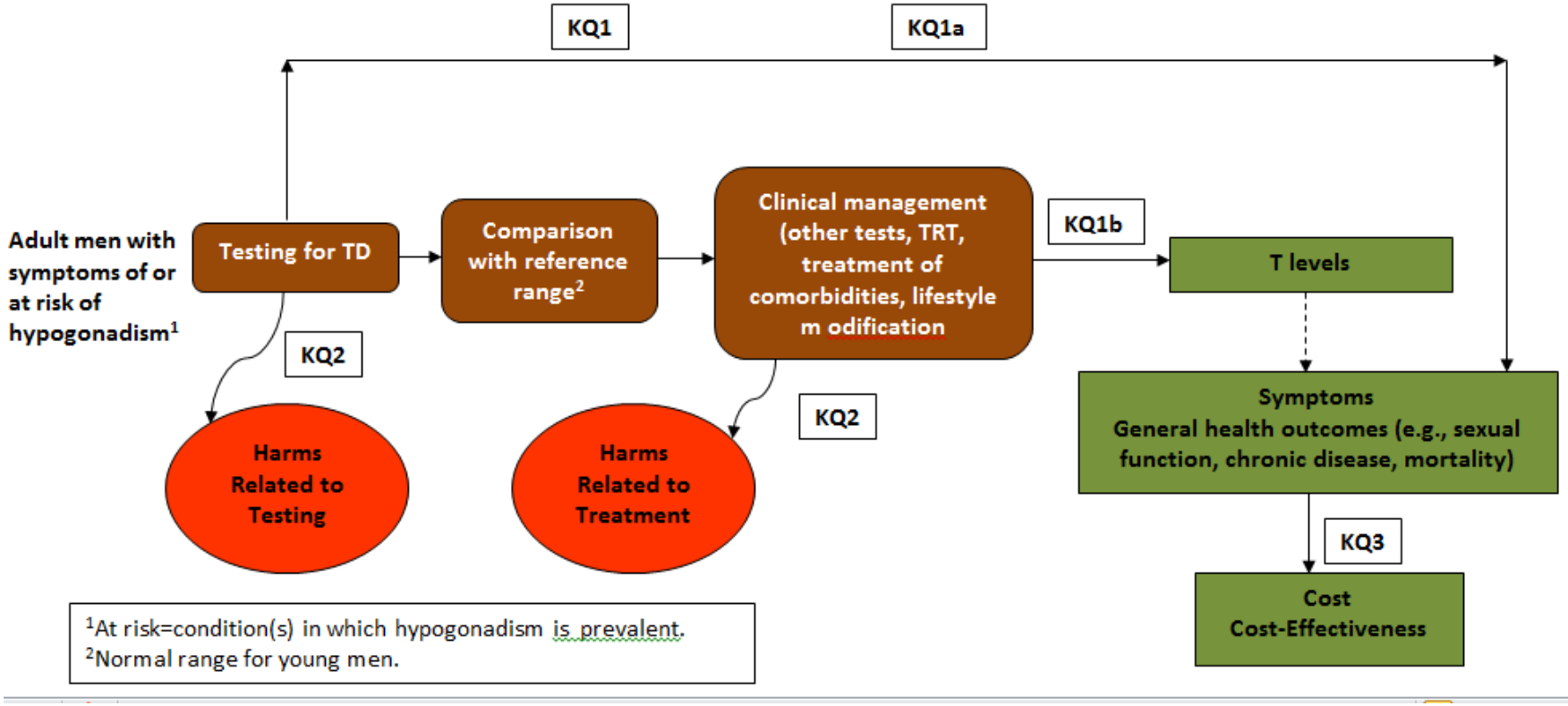
Analytic Framework

The relationships between the key questions are depicted in Figure 1.

Figure 1. Analytic Framework: Testosterone Testing
(Key Questions referenced by number in the graphic)

1. Is there evidence that testosterone testing improves outcomes?
   1a. Does the impact on outcomes vary according to age, race/ethnicity, baseline testosterone levels, treatment status, or clinical history?
   1b. What is the minimum interval required to assess a change in testosterone status in untreated and treated individuals?
2. What are the potential harms of testosterone testing, including potential subsequent harms resulting from treatment decisions?
3. What are the costs and cost-effectiveness of testosterone testing?

Key: T(D), testosterone (deficiency); TRT, testosterone replacement therapy



Methods

In addition to the formal methods described in this section, guidance on search dates and potentially missed studies was sought through personal communication with a clinical expert.

Preplanned Search Strategy and Selection Criteria

See **Appendix I** for additional search details.

Search for Systematic Reviews and Guidelines (Topic-Scoping Search)

These sources were searched on June 29, 2014, for systematic reviews, meta-analyses, economic evaluations, and practice guidelines published in the last 10 years:

- Core online databases such as the Agency for Healthcare Research and Quality (AHRQ), Centre for Reviews and Dissemination (CRD) (York University), and National Guidelines Clearinghouse (NGC). NOTE: The CRD databases include the National Health Service Economic Evaluation Database (NNS-EED) and searches include a search for Cochrane Reviews.
- Websites of relevant professional societies.
- PubMed, using filters for Practice Guidelines, Guidelines, Meta-analyses, and Systematic Reviews.

Systematic reviews were selected if they reviewed studies considered eligible for answering the Key Questions or if they provided useful background information on: (a) the association of testosterone levels with symptoms considered characteristic of hypogonadism; (b) the association of testosterone levels with health outcomes and conditions such as mortality or metabolic syndrome; (c) the effectiveness of testosterone therapy in men with hypogonadism; or (d) adverse events associated with testosterone therapy. For systematic reviews that could provide background information, the most recent 1 or 2 systematic reviews on a particular topic were selected.
Search for Primary Studies

Initial searches were conducted in on September 7, 2014 (PubMed) and on September 10, 2014 (Embase). An update search for direct evidence pertaining to the Key Questions was conducted on December 3, 2014. A PubMed search for cost studies and economic evaluations was conducted on September 25, 2014 (in addition to the June 29 search of the NHS-EED. Specific search strings are documented in **Appendix I**. The bibliographies of review articles and practice guidelines were also searched.

Inclusion/Exclusion Criteria for Direct Evidence Pertaining to the Key Questions

Inclusion Criteria

- Consistency with the Population-Interventions-Comparator-Outcomes (PICO) statement.
- Observational studies with a control or comparator group, nonrandomized trials, and randomized controlled trials (RCTs) designed to measure the effectiveness of testosterone testing (Key Questions #1, #1a), the effectiveness of repeat testosterone testing (Key Question #1b), or adverse events as consequences of testosterone testing (Key Question #2). Systematic reviews of such studies were also eligible.

- Longitudinal studies designed to assess the time required for a clinically relevant change in serum testosterone levels to occur to (relevant to Key Question #1b).
- Any study reporting the cost outcomes of or an economic evaluation of testosterone testing (Key Question #3).

Exclusion Criteria

- Studies that were inconsistent with the PICO statement.
- Publications that did not meet the study designs specified in inclusion criteria.
- Studies published prior to 1990, identified as the date after which testing technologies have not changed substantially.
- Economic evaluations and cost studies published prior to September 2003 (studies published earlier than 10 years prior to the search date were considered to have minimal applicability to the current economy).

Post Hoc Search for Selected Recent Indirect Evidence

No studies with direct evidence relevant to the Key Questions were identified. In its guidelines, the Endocrine Society characterizes all recommendations concerning testing to be supported by low- or very-low-quality evidence, and the guidelines did not cite any direct evidence of the effectiveness, safety, or cost implications of testosterone testing. For more in-depth assessment of the most promising indirect evidence, subpopulations were identified where (a) systematic reviews or observational studies provided evidence of an association with low testosterone and reported findings in terms that allowed an assessment of clinical validity or strength of association *and* (b) systematic reviews provided positive evidence of treatment efficacy. Two subpopulations met these criteria:

- Men with type 2 diabetes or metabolic syndrome
- Men with symptoms of sexual dysfunction

A post hoc search (see **Appendix I** for details) was conducted to identify studies in the selected subpopulations (diabetes/metabolic syndrome, sexual dysfunction) that were published more recently than the observational research and systematic reviews cited in the **CLINICAL BACKGROUND** section. Studies were selected in accordance with the selection criteria used in the systematic reviews. The purpose of the post hoc search and analysis was to determine whether recent indirect evidence would add to or conflict with the indirect evidence identified initially for the background section of this technical report.

Inclusion Criteria

- Observational studies assessing the discriminatory power of serum testosterone measurements. Studies were selected if they reported sensitivity and specificity, diagnostic odds ratios (ORs), relative risks (RR), hazard ratio (HR), or prevalence ratio for (a) low testosterone versus sexual dysfunction or (b) low testosterone versus metabolic syndrome, diabetes, or measures of glucose control (published 2010 or later). These criteria were designed to identify evidence of clinical associations that were published later than the relevant systematic reviews (systematically identified during topic scoping) or observational studies (incidentally identified).

- RCTs assessing the effectiveness of testosterone therapy in men with type 2 diabetes and a diagnosis of hypogonadism* (published 2013 or later) or metabolic syndrome† (published 2010 or later) or symptoms of sexual dysfunction (published 2013 or later). These criteria reflect the end of search dates and the inclusion criteria of the most comprehensive systematic reviews of testosterone therapy in these subpopulations at the time of the topic-scoping search.

*Defined as total testosterone < 3.2 ng/mL (320 ng/dL) or 11 nmol/L and free testosterone < 64 picograms per milliliter (pg/mL) or 220 pmol/L and ≥ 3 sexual symptoms.

†Defined as the presence of ≥ 3 of the following 5 components: obesity (based on waist circumference, waist-to-hip ratio or body mass index [BMI]), elevated triglyceride levels, low high-density lipoprotein cholesterol levels, impaired glucose metabolism (based on fasting glucose or insulin levels, presence of insulin resistance, or diagnosis of diabetes) and hypertension (based on systolic and diastolic blood pressure measurements).

Exclusion Criteria

Observational studies that did not include men both with and without hypogonadism or symptomatic androgen deficiency (typically defined on basis of low testosterone plus symptoms) plus both men with and without the symptoms or health condition of interest, to assure that the discriminatory power of testing could be ascertained.

Quality Assessment

Clinical Studies

Appendix II outlines the process used by Hayes for assessing the quality of individual primary studies and the quality of bodies of evidence. This process is in alignment with the methods recommended by the Grading of Recommendations Assessment, Development and Evaluation (GRADE) Working Group. Quality checklists for individual studies address study design, integrity of execution, completeness of reporting, and the appropriateness of the data analysis approach. Individual studies are labeled as *good*, *fair*, *poor*, or *very poor*. For individual studies included in systematic reviews, this report relies on the quality assessment by review authors. To aid in interpreting the assessment by review authors, a systematic review quality checklist, the Assessment of Multiple Systematic Reviews (AMSTAR) tool (Shea et al., 2007), was used.

Like the GRADE Working Group, Hayes uses the phrase *quality of evidence* to describe bodies of evidence in the same manner that other groups, such as Agency for Healthcare Research and Quality (AHRQ), use the phrase *strength of evidence*. The Hayes Evidence-Grading Guides assure that assessment of the quality of bodies of evidence takes into account the following considerations:

- Methodological quality of individual studies, with an emphasis on the risk of bias within studies.
- Applicability to the population(s), intervention(s), comparator(s), and outcome(s) of interest (PICO) statement.
- Consistency of the results across studies.
- Quantity of data (number of studies and sample sizes).
- Publication bias, if relevant information or analysis is available.

NOTE: Two terms related to applicability are *directness* and *generalizability*. *Directness* refers to how applicable the evidence is to the outcomes of interest (i.e., health outcomes versus surrogate or intermediate outcomes) or to the comparator of interest (indirect comparison of 2 treatments versus head-to-head trials). *Generalizability* usually refers to whether study results are applicable to real-world practice. If the setting is not specified in a PICO statement, the issue of generalizability to real-world settings is not typically treated as an evidence quality issue. Another term used by some organizations is *imprecision*, which refers to findings based on such a small quantity of data that the confidence interval surrounding a pooled estimate includes both clinically important benefits and clinically important harms, or such a small quantity of data that any results other than large statistically significant effects should be considered unreliable.

Bodies of evidence for particular outcomes are labeled as being of *high*, *moderate*, or *low quality*, or they are deemed to be *insufficient* to permit conclusions. These labels can be interpreted in the following manner:

High: Suggests that we can have high confidence that the evidence found is reliable, reflecting the true effect, and is very unlikely to change with the publication of future studies.
Moderate: Suggests that we can have reasonable confidence that the results represent the true direction of effect but that the effect estimate might well change with the publication of new studies.
Low: We have very little confidence in the results obtained, which often occurs when the quality of the studies is poor, the results are mixed, and/or there are few available studies. Future studies are likely to change the estimates and possibly the direction of the results.
Insufficient: Suggests no confidence in any result found, which often occurs when there is a paucity of data or the data are such that we cannot make a statement on the findings.

Economic Evaluations

A tool created for internal use at Hayes was used to guide interpretation and critical appraisal of economic evaluations. This tool was based on best practices as identified in the literature and addresses issues such as the reliability of effectiveness estimates, transparency of the report, quality of analysis (e.g., the inclusion of all relevant costs, benefits, and harms), generalizability/applicability, and conflicts of interest. Sources are listed in **Appendix II**.

Guidelines

The Rigor of Development domain of the Appraisal of Guidelines Research and Evaluation (AGREE) tool (AGREE Enterprise, 2013), along with a consideration of the items related to commercial funding and conflicts of interest among the guideline authors, was used to assess the quality of practice guidelines. Use of the AGREE tool was limited to these domains because they relate most directly to the link between guideline recommendations and evidence (AGREE Enterprise, 2013).

Search Results

Systematic Reviews Identified in Topic-Scoping Searches

No systematic reviews designed to directly answer the Key Questions were identified.

Primary Studies with Direct Evidence Pertaining to the Key Questions

No primary studies designed to directly answer the Key Questions were identified.

Evidence of Clinical Associations with Testosterone Levels or the Benefits and Harms of Testosterone Therapy

The largest observational studies and most recent systematic reviews were selected for summarization in the **CLINICAL BACKGROUND** section. The following selections were made:

- Seven (7) observational studies measuring the association between testosterone levels and signs or symptoms (see **Appendix III**).
- Nine (9) systematic reviews of the association between testosterone levels and health outcomes or conditions (see **Appendix V**).
- Seven (7) systematic reviews evaluating the effectiveness of testosterone therapy (see **Appendix VI**). An additional systematic review was published after the draft version of the present report was published and has been added.
- Six (6) systematic reviews evaluating the adverse events associated with testosterone therapy (see **Appendix VII**).

Some systematic reviews are represented in more than 1 Appendix.

Recent Indirect Evidence Identified Through a Targeted Post Hoc Search

- Fifteen (15) studies that would potentially meet the inclusion criteria of 3 recent systematic reviews evaluating the association between testosterone levels and metabolic syndrome or diabetes.
- Five (5) studies that would potentially meet the inclusion criteria of a recent systematic review evaluating the effectiveness of testosterone therapy in men with type 2 diabetes.
- Four (4) population studies for comparison of results with a large 2010 population study evaluating the association between low testosterone levels and symptoms of sexual dysfunction.
- No RCTs that would potentially meet the inclusion criteria of a recent systematic review evaluating the effectiveness of testosterone therapy for improving sexual function.

Practice Guidelines

Four relevant practice guidelines, published by U.S. groups in the last 10 years, were identified.

Literature Review

Key Question #1: Is there evidence that testosterone testing improves outcomes? Key Question #1a. Does the impact on outcomes vary according to age, race/ethnicity, baseline testosterone levels, treatment status, or clinical history?

The literature search identified no studies that were designed to compare outcomes between individuals who were tested for low testosterone and similar individuals whose symptoms and health problems were managed without knowledge of serum testosterone levels. Furthermore, no studies with direct evidence were cited in the practice guidelines that were reviewed.

Indirect Evidence

Key Question #1

See the section on **POST HOC SEARCH AND ANALYSIS** for a detailed discussion of the indirect evidence regarding the utility of testosterone testing in men with sexual dysfunction or diabetes/metabolic syndrome. Indirect evidence was strongest for these 2 subpopulations.

The post hoc analysis has the following implications for Key Question #1: A positive but very weak inference regarding the utility (Key Question #1) of testosterone testing in men with sexual dysfunction can be drawn. This inference is based on evidence of a strong association of low testosterone levels with the simultaneous presence of multiple sexual symptoms. The inference also reflects evidence that testosterone therapy improves sexual symptoms in men who have sexual dysfunction and low testosterone levels. Although there is a strong association of low testosterone levels with metabolic syndrome and significant differences in serum testosterone between men with and without type 2 diabetes, the effectiveness of testosterone therapy for improving outcomes related to type 2 diabetes remains uncertain.

Key Question #1a

Two key findings from the Post Hoc Analysis are particularly pertinent to Key Question #1a: Testosterone therapy *to improve sexual symptoms* is more effective in men with low testosterone levels at baseline than in men with eugonadal levels, and more effective in men with diabetes, regardless of baseline testosterone levels (Corona et al., 2014a). Most participants in the studies included in the systematic reviews of testosterone *to improve diabetes-related outcomes* had low levels of testosterone, with means close to or somewhat below the lower bound of typical reference ranges. Some, but not all, studies specifically enrolled men with classic symptoms of hypogonadism (Cai et al., 2014; Grossman et al., 2014). However, neither of the 2 reviews analyzed whether the effect on metabolic outcomes actually varied according to baseline testosterone level. Other systematic reviews have suggested that the effect of testosterone therapy does *not* depend on baseline testosterone levels for improvement in intermediate cardiovascular measures (Corona et al., 2011c) or depression (Zarrouf et al., 2009).

Additional Potentially Policy-Relevant Information, Testosterone Levels and Metabolic Factors

Additional support for a link between low testosterone levels and metabolic factors was provided by 3 studies demonstrating an association between body mass index (BMI), which could be considered a

modifiable factor, and low testosterone levels (Ponholzer et al., 2005; Travison et al., 2007; Yeap et al., 2009). Furthermore, a recent meta-analysis demonstrated that weight loss after bariatric surgery and a low-calorie diet were both associated with an increase in testosterone levels (Corona et al., 2013a). Such evidence lends further support to a link between hypogonadism and type 2 diabetes, but is also consistent with expert recommendations to consider evaluating for and addressing obesity and chronic disease prior to recommending testosterone therapy (Bhasin et al., 2010; McGill et al., 2012; Pantalone and Faiman, 2012; Corona et al., 2013a).

Key Question #1b. What is the minimum interval required to assess a change in testosterone status in untreated and treated individuals?

The literature search identified no studies designed to estimate the optimal interval for serial testosterone testing.

Other Potentially Policy-Relevant Information

A very small study of 282 men selected on the basis of age (60 to 82 years at follow-up) compared changes in serum testosterone levels and changes in responses to the Androgen Deficiency in the Aging Male (ADAM) questionnaire over a 5-year period. Total testosterone and bioavailable testosterone levels declined, but only 1 of 10 ADAM symptoms (less strong erection) changed over that time period (Holm et al., 2011). These findings suggest that clinically meaningful changes in testosterone levels may not occur due to age alone within ≤ 5 years.

The Endocrine Society Guidelines on testosterone therapy include a *weak* recommendation that serum levels of testosterone be checked at 3 to 6 months *following initiation of therapy* to assess whether serum levels have reached the normal range. However, the guidelines characterize the supporting evidence as being of very low quality. A review article acknowledged that there is insufficient evidence to allow a conclusion about the usefulness of monitoring serum levels in patients undergoing testosterone replacement therapy (Ho, 2011). However, other experts point out that there is no evidence that testosterone levels beyond the normal range provide health benefits (ASRM, 2006), which implies a rationale for monitoring serum levels to assure that they remain in the normal range and do not unnecessarily expose the patient to adverse effects.

Key Question #2: What are the potential harms of testosterone testing, including potential subsequent harms resulting from treatment decisions?

The literature search identified no studies that were designed to assess harms resulting from a clinical strategy of testosterone testing. Furthermore, no studies with direct evidence were cited in the practice guidelines that were reviewed.

Other Potentially Policy-Relevant Information

Testosterone testing is a blood test and poses no significant procedural harms beyond those associated with phlebotomy.

Indirect Evidence

As discussed in the **CLINICAL BACKGROUND** section, adverse events were described in guidelines and a review article as occurring at a low frequency in young men on testosterone therapy, but quantitative data were not reported in these sources.

The Endocrine Society's guidelines stated that testosterone therapy may increase prostate events (nonsignificant estimate); substantially increases the risk of erythrocytosis; and has no clear effect on obstructive sleep apnea, cardiovascular events, or mortality (Bhasin et al., 2010). The guidelines' review of safety evidence for older men is consistent with systematic reviews and key studies published since the guidelines were developed. Large observational studies (incidentally identified) and the most comprehensive systematic reviews (systematically identified as part of topic scoping) that have been published since the guidelines suggested that testosterone therapy:

- May increase prostate events (nonsignificant estimate). The pooled relative risk (RR) of a composite prostate outcome was 1.41 (95% CI, 0.93 to 2.14; low heterogeneity; 15 studies) (Fernández-Balsells et al., 2010).
- Substantially increases the risk of erythrocytosis. The pooled RR of erythrocytosis was 3.15 (95% CI, 1.56 to 6.35; low heterogeneity; 11 studies) (Fernández-Balsells et al., 2010).
- Has no clear effect on obstructive sleep apnea. Small case series reported inconsistent findings (Hanafy, 2007).
- Has no clear effect on cardiovascular events (no overall effect; conflicting evidence for subpopulations defined by older age or existing cardiovascular disease [CVD]). The adjusted hazard ratio (HR) for the presence of coronary artery disease was 1.29 (95% CI, 1.04 to 1.58) in a population cohort study of 8709 men who had undergone coronary angiography (Vigen et al., 2013). Another retrospective population cohort study (55,593 men) estimated that the risk of myocardial infarction within 90 days of treatment initiation was doubled, but only in men > 65 years of age or in younger men with a history of heart disease (Finkle et al., 2014). In a meta-analysis of 75 RCTs, the overall risk of any cardiovascular event was 1.07 (95% CI, 0.69 to 1.65; 31 studies) (Corona et al., 2014b). In contrast to the implications of the 2 population studies, the meta-analysis found no variation in adverse effects according to subpopulations defined by age or CVD at baseline, and no variation according to baseline testosterone levels was detected.
- Has an uncertain effect on all-cause mortality. A recent study suggested a *protective* effect (adjusted HR, 0.61; 95% CI, 0.42 to 0.88), but this evidence comes from a single observational study (Shores et al., 2012) and cannot be considered conclusive.

Follow-up was typically 2 or 3 years. Thus, the long-term safety of testosterone therapy is unknown. See **Appendix VI** for more detail. As noted in the **POLICY CONTEXT** statement, the Food and Drug Administration (FDA) is reassessing the safety of FDA-approved testosterone products because of conflicting findings regarding long-term safety (CDER, 2014b).

Key Question #3: What are the costs and cost-effectiveness of testosterone testing?

No studies evaluating the cost or cost-effectiveness of testosterone testing were identified in the literature.

Post Hoc Analysis of Indirect Evidence for the Effectiveness of Testosterone Testing (Key Questions #1, #1a)

Since no studies were identified in the search for direct evidence to answer the Key Questions, a more in-depth review of some of the indirect evidence covered in the **CLINICAL BACKGROUND** section was conducted. For subpopulations where the evidence met certain criteria (see following description), the quality of key indirect evidence was appraised and an additional search was conducted for primary data published after the included systematic reviews. The following 2 conditions were required for more in-depth consideration of testing in a subpopulation:

- Systematic review evidence supported an association between symptoms, a health condition, or the consequences of a health condition and endogenous testosterone levels; or in the absence of a systematic review, ≥ 1 large observational study demonstrated a strong association.
- Systematic review evidence suggested that testosterone therapy is effective.

The rationale for these conditions is based on 2 assumptions:

- Testosterone testing could not have clinical utility unless low testosterone levels were shown to be associated with clinical characteristics that can be used to select patients for testing and possible testosterone therapy. Such associations would ideally be quantified in terms that allow assessment of the discriminatory power of serum testosterone measurements (i.e., risk ratios, estimates of risk ratios [odds ratios], sensitivity and specificity, or calculations related to sensitivity and specificity).
- Testosterone testing would have greatest clinical utility, i.e., have the greatest potential to improve health outcomes, in those populations where testosterone therapy was effective.

The rationale for the value of this approach includes the fact that a finding of "negative evidence" (i.e., a finding of no association or no evidence of effective treatment) for either condition could rule out the clinical utility of testosterone testing. The evidence collected as part of the background review did not identify any subpopulations for which an association between endogenous testosterone levels and symptoms or health was disproven or for which effectiveness of testosterone therapy was disproven. Conversely, *positive* evidence for clinical associations and treatment effectiveness does not definitively confirm the clinical utility of a test but strengthens the possibility that testing could be useful.

The background research for the current report suggested that a relationship between symptoms or health conditions and endogenous testosterone levels, as well as the efficacy of testosterone treatment, has been suggested by population-based studies and/or systematic reviews in these 2 subpopulations:

- Men with type 2 diabetes or metabolic syndrome.
- Men with symptoms of sexual dysfunction.

A post hoc search and analysis were designed to further explore these 2 indications as having the strongest possibility for the usefulness of testosterone testing. For testosterone testing in men with diabetes or metabolic syndrome, the post hoc analysis was based on 2 systematic reviews of the link between testosterone levels and metabolic outcomes (Brand et al., 2011; Corona et al., 2011b), 1 systematic review of the link between testosterone levels and type 2 diabetes (Corona et al., 2011a), and 1 systematic review of the effectiveness of testosterone therapy in men with type 2 diabetes (Cai et al., 2014), supplemented by a search for studies published subsequent to the systematic reviews. After

release of the draft version of the present report, an additional systematic review of testosterone therapy in men with type 2 diabetes was published (Grossman et al., 2014) and is included in this final report.

For testosterone testing in men with sexual dysfunction, the key evidence regarding clinical associations was a single large population study (the European Male Aging Study [EMAS]) (Wu et al., 2010), which was identified incidentally in the full literature search designed to produce direct evidence for the Key Questions. The EMAS was also cited in review articles as one of the major population studies. An additional and more targeted effort was made to identify relevant studies that were published since the EMAS study. Treatment effectiveness in men with sexual dysfunction was assessed by a recent systematic review (Corona et al., 2014a), supplemented by a search for studies published subsequent to this systematic review.

See **Appendix I** for details of the post hoc search.

Findings of Post Hoc Analysis: Men with Type 2 Diabetes or Metabolic Syndrome

Association of Type 2 Diabetes/Metabolic Syndrome with Testosterone Levels

Three meta-analyses established a link between low levels of testosterone and metabolic syndrome (often a precursor to diabetes) or between low levels of testosterone and diabetes.

Brand et al. (2011): Association with Metabolic Syndrome

A good-quality meta-analysis of 32 studies computed an RR for metabolic syndrome of 0.38 (95% CI, 0.25 to 0.50), comparing higher with lower levels of total testosterone. Free testosterone and sex hormone binding globulin (SHBG) were also found to be lower in men with metabolic syndrome. There was high heterogeneity in these analyses (Brand et al., 2011). A single quality rating was not assigned to each study, but several factors were prespecified as quality indicators. Two of 6 indicators (use of fasting blood samples and morning collection of blood samples) were met by most studies. Most studies did not adjust for confounders; 4 studies adjusted for age only. The authors considered longitudinal (prospective cohort) studies to be of higher quality than cross-sectional studies; only 5 of the 32 studies were longitudinal. Population-based patient selection was considered to be a quality indicator, and about half the studies were not population based. Most studies did not clarify whether participants taking hormonal therapy were included. By the authors' standards, typical study quality appeared to be poor to fair. However, sensitivity (subgroup) analyses by metaregression showed that adjustment for whether or not men with type 2 diabetes were clearly excluded and longitudinal versus cross-sectional design did not affect differences in mean serum levels. The direction of findings across studies was consistent, except for slight inconsistency in estimates of relationships involving free testosterone.

Effect Modifiers: Sensitivity analysis also suggested that differences in total testosterone were somewhat greater in younger men (< 55 years), but only a trend toward statistical significance (P=0.08) was detected; this evidence is pertinent to Key Question #1a.

Corona et al. (2011a): Association with Metabolic Syndrome

Another meta-analysis (fair quality) of a somewhat smaller number of studies (20 studies) also found that men with metabolic syndrome had significantly lower total testosterone and free testosterone

levels (Corona et al., 2011a). The difference in mean total testosterone between men with and without metabolic syndrome was –2.85 nanomoles per liter (nmol/L) (95% CI, –3.34 to –2.36) according to cross-sectional studies and –2.17 nmol/L (CI, –2.41 to –1.94) according to longitudinal (prospective cohort) studies. These estimates compare with a pooled estimate of –2.64 nmol/L (95% CI, –2.95 to –2.32) for all observational studies in the review by Brand et al. (2011). Although the findings in the review by Corona et al. were not reported in the terms specified as inclusion criteria for the post hoc search and analysis (diagnostic accuracy or an estimate of RR), they are useful as corroboration of the findings of the meta-analysis by Brand and colleagues. Study quality was not reported.

Effect Modifiers: Corona and colleagues also provided subgroup analyses that suggested (a) greater total testosterone differences between men with and without metabolic syndrome in the presence of type 2 diabetes and (b) no effect of the presence of erectile dysfunction on the relationship between total testosterone levels and the presence of metabolic syndrome.

Corona et al. (2011b): Association with Type 2 Diabetes

A fair-quality meta-analysis (Corona et al., 2011b) of observational studies reported results for the relationship between testosterone levels and diabetes that closely paralleled those reported in the 2 meta-analyses of metabolic syndrome and testosterone levels (Brand et al., 2011; Corona et al., 2011a). Although results were not expressed in terms of a RR, the systematic review was considered because of the close relationship between metabolic syndrome and diabetes. The difference in mean total testosterone between men with and without diabetes was –2.99 nmol/L (95% CI, –3.59 to –2.40) according to cross-sectional studies and –2.08 nmol/L (CI, –3.57 to –0.59) according to longitudinal (prospective cohort) studies. There was no serious inconsistency in individual study results. Individual study quality was not reported.

Effect Modifiers: The presence of erectile dysfunction had no effect on the relationship between testosterone levels and diabetes, but younger age and greater body mass index (BMI) were associated with greater differences in testosterone levels between men with and without diabetes.

Additional Primary Studies:

A systematic search for observational studies (the post hoc search) published since these 3 systematic reviews yielded 15 studies that would potentially meet the reviews' inclusion criteria. All 15 studies reported an inverse relationship between total testosterone and metabolic syndrome or type 2 diabetes and/or an inverse relationship between SHBG and metabolic syndrome or type 2 diabetes. Thus, more newly published studies do not suggest any inconsistency with the conclusions of the most recent systematic reviews. However, the studies were not reviewed in detail or critically appraised; thus, the strength with which new studies confirm the conclusions of the systematic reviews is unknown.

Effectiveness of Testosterone Therapy in Men with Type 2 Diabetes

Two meta-analyses suggested that testosterone therapy may improve some metabolic factors in men with diabetes as well as hypogonadism, but the evidence was not conclusive.

Cai et al. (2014)

A fair-quality meta-analysis of 5 randomized controlled trials (RCTs) (351 participants) found that the mean difference in levels of fasting plasma glucose, fasting serum insulin, glycated hemoglobin (HbA1c), and triglycerides favored men who underwent testosterone therapy (Cai et al., 2014).. However, no effect on body fat or blood pressure was detected. The authors did not discuss whether differences were clinically important but individual study data suggested negligible to substantial between-group differences at study endpoint. The mean values for control groups indicate that patients were under treatment for their diabetes. There was no subgroup analysis of whether the effects of testosterone therapy depended on the baseline level of glucose control. All study participants had both late-onset hypogonadism and type 2 diabetes. A diagnosis of hypogonadism was defined in terms of total testosterone levels and free testosterone levels and also required the presence of at least 3 sexual symptoms. Study quality was judged by the review authors to be fair on average. For fasting serum insulin and glycated hemoglobin (HbA1c), there was some inconsistency in individual study results, with some studies showing a benefit and others showing no effect. The authors note that given the short follow-up periods (3 to 12 months), definitive conclusions about the metabolic effects of testosterone therapy could not be made.

Effect Modifiers: The review included no analysis of any patient characteristics, including baseline testosterone level, that might alter treatment effect.

Additional Primary Studies:

A systematic search for RCTs published since the systematic review by Cai et al. (2014) yielded 5 studies that would potentially meet the review's inclusion criteria. All of the trials enrolled men with type 2 diabetes but did not necessarily enroll men who also had a diagnosis of hypogonadism according to the criteria specified by the review authors . Two studies reported positive results for measures of glucose control or insulin resistance (fasting plasma glucose, fasting serum insulin, and/or HbA1c), which would be consistent with the review findings. Two studies did not report glucose control outcomes as primary outcomes, and the fifth study observed no effect on glucose control. Thus, an abstract review of newer RCTs neither confirms nor refutes the findings of Cai et al. (2014).

Grossman et al. (2014) (published after publication of Draft Report)

After publication of the draft version of the present report, another systematic review and meta-analysis of testosterone therapy in men with diabetes was published (Grossman et al., 2014). This review, which allowed inclusion of studies conducted in men with metabolic syndrome who did not yet have a diagnosis of diabetes, was based on 7 RCTs, 1 involving men with metabolic syndrome and 6 involving men with diabetes. The review authors stated that all diabetic study participants were currently being treated with antidiabetic medication and had adequate glucose control. The results of this meta-analysis were less positive than those included in the review by Cai et al. (2014):

- A pooled estimate for glucose control, measured in terms of HbA1c, slightly favored testosterone therapy but was statistically nonsignificant (the estimate was statistically significant in the Cai review), and results were inconsistent across individual trials (as in the Cai review). The magnitude of the pooled estimate for improvement in HbA1c was very small compared with mean baseline values across studies.
- As in the Cai review, improvement in insulin resistance was suggested in individual studies and in the pooled estimate, but the pooled estimate was nonsignificant except when analysis was

restricted to use of a more conventional but less rigorous measurement technique. The authors described the observed improvement in insulin resistance as modest.

- In contrast to the Cai review, no effect on triglycerides was demonstrated.
- In agreement with the Cai review, no effect on blood pressure was detected.

The authors of the most recent review concluded that their results do not support the routine use of testosterone therapy to improve glucose metabolism in men with relatively well-controlled type 2 diabetes and/or metabolic syndrome. However, they pointed out that since metabolic syndrome and type 2 diabetes are slowly progressive conditions, it is possible that long-term testosterone therapy might have a more pronounced effect on metabolic factors.

The Grossman analysis differed from the Cai analysis in the following ways:

- Only placebo-controlled RCTs were included in the Grossman review; 2 of 5 RCTs in the Cai review compared testosterone therapy with no treatment. The implications of this difference are unclear, considering the objective nature of the outcome measures.
- The Cai review required that study participants have low testosterone and classic symptoms of hypogonadism, whereas the Grossman review did not limit study selection in that way. However, the range of mean baseline levels of total testosterone across studies was nearly identical in the 2 reviews, and most of the studies included in the Grossman review reported that participants had symptomatic hypogonadism.

Effect Modifiers: The review included no analysis of any patient characteristics, including baseline testosterone level, that might alter treatment effect.

Cai Versus Grossman Review of Testosterone Therapy in Men with Diabetes

In both reviews, the direction of individual study results varied for glucose control (HbA1c). The lack of statistical significance in the pooled estimate of the Grossman review is noteworthy since this analysis was based on 6 trials (603 patients) and should perhaps be given greater weight; the Cai analysis was based on only 3 trials (102 patients), 2 of which were excluded from the Grossman review because of lack of placebo control. The Cai review observed significant effects on measures that were not analyzed in the Grossman review: fasting plasma glucose and fasting serum insulin. The conclusions of both meta-analyses should be considered preliminary. The included studies varied with regard to whether any changes in antidiabetic medication were allowed during the study. Even where medication changes were discouraged, analysis in the individual studies generally did not control for group differences in compliance.

Findings of Post Hoc Analysis: Men with Symptoms of Sexual Dysfunction

Association of Sexual Dysfunction with Testosterone Levels

The EMAS was a large cross-sectional study of 3369 men, ages 40 to 70 years, residing in Europe (Wu et al., 2010). It was considered a good-quality study for evaluating the strength of clinical associations. The sample was divided into a training set and a validation set. Cluster analysis was used to assess the stability of associations between the training and validation set. Cluster analysis showed 3 symptoms of physical performance and 3 psychological symptoms to collectively have weak associations with

testosterone levels. However, 3 symptoms of sexual dysfunction collectively had a strong association with testosterone levels. The 3 symptoms were awakening with full erection less frequently, erectile dysfunction, and thinking about sex less often. Odds ratios (ORs) were calculated for the sexual dysfunction symptoms. After control for confounding, ORs for the presence of all 3 sexual dysfunction symptoms were statistically significant and ranged from 1.64 to 2.24, depending on the cutoff value assumed and whether free as well as total testosterone was low. Another observational study (Gades et al., 2008) found an association of low total testosterone with erectile dysfunction but not with sex drive. The study by Gades et al. was smaller (n=414) than the study by Wu et al. and did not attempt to analyze the association of low testosterone with the simultaneous presence of multiple symptoms.

A systematic search for observational studies published since the analysis by Wu et al. (2010) identified 4 studies. Three studies detected no association between low testosterone levels and any 1 symptom or form of sexual dysfunction, while a fourth study found an inverse relationship between testosterone levels and 2 aspects of erectile dysfunction. The studies with negative results do not directly contradict the findings of Wu et al. since the more recent studies did not evaluate an association with the simultaneous presence of erectile dysfunction plus other forms of sexual dysfunction, as was the case in the Wu study. However, the search revealed that the results reported by Wu et al. have not been replicated in newer studies. The recently published studies were not reviewed in detail or critically appraised; thus, the implications of their results with respect to the findings of the Wu study are not fully known.

Effect Modifiers: Neither study evaluated patient factors that might alter the relationship between low testosterone and sexual symptoms.

Effectiveness of Testosterone Therapy in Men with Sexual Dysfunction

The most comprehensive systematic review of testosterone therapy for sexual dysfunction identified 41 studies evaluating the effect of testosterone therapy on sexual dysfunction (Corona et al., 2014a). The review was considered to be of fair quality. The systematic review did not clarify how men were selected for the included studies or describe the typical baseline degree of sexual dysfunction. For overall study populations, large pooled effect sizes (0.75 to 0.82) suggested that testosterone therapy as a stand-alone therapy reduced overall erectile dysfunction, sexual-related erectile dysfunction, and poor libido; and improved orgasm score. (NOTE: Effect sizes > 0.80 are by convention considered to be *large*.) In a subgroup analysis, the effect sizes for men with low testosterone were even larger (1.00 to 1.26), suggesting that baseline testosterone level could be used to identify men most likely to benefit from testosterone therapy. In studies where testosterone therapy was an add-on to phosphodiesterase type 5 inhibitor (PDE5i) therapy, which is a medication for erectile dysfunction, testosterone therapy was effective in uncontrolled studies but not in placebo-controlled trials. The quality of the studies and the consistency of findings across studies were unclear.

Effect Modifiers: Sensitivity analyses (meta-regression) confirmed the relationship between baseline testosterone levels and treatment effect and also showed that treatment effects were greater in men who had a diagnosis of diabetes. Age and duration of therapy did not influence treatment effect. The authors detected evidence of publication bias, greater effects in industry-sponsored trials, and statistical heterogeneity across studies.

Additional Primary Studies: A systematic search for RCTs published since the systematic review yielded no studies that would potentially meet the review's inclusion criteria. No search was conducted for

more recent non-RCTs since the systematic review found heterogeneity of results when comparing uncontrolled studies with placebo-controlled studies, at least for testosterone therapy as an add-on therapy.

Summary of Post Hoc Analysis

A positive but very weak inference regarding the utility (Key Question #1) of testosterone testing in men with sexual dysfunction can be drawn. This inference is based on evidence of a strong association of low testosterone levels with the simultaneous presence of multiple sexual symptoms. The inference also reflects evidence that testosterone therapy improves sexual symptoms in men who have sexual dysfunction and low testosterone levels. Although there is a strong association of low testosterone levels with metabolic syndrome and significant differences in serum testosterone between men with and without type 2 diabetes, the effectiveness of testosterone therapy for improving outcomes related to type 2 diabetes remains uncertain.
Subgroup analyses (Key Question #1a) suggested that diabetes is a stronger predictor of low testosterone levels in younger compared with older men and in men with greater BMI, and that metabolic syndrome is a stronger predictor of low testosterone levels in men with type 2 diabetes than in men without diabetes. The presence of erectile dysfunction does *not* influence the relationship between either diabetes or metabolic syndrome and testosterone levels. The systematic review of the effectiveness of testosterone therapy in men with sexual dysfunction found that treatment effects were greater in men who had a diagnosis of type 2 diabetes before treatment, regardless of their baseline testosterone levels, but age did not influence treatment effects.

The quality of the indirect evidence was low to moderate with respect to the objectives of the studies, taking into account risk of bias in individual studies where reported by review authors, the quantity of data, consistency of study findings, and the precision of pooled estimates. As evidence for Key Questions #1 and #1a of the current report, the evidence is of low quality at best because of its lack of direct applicability to the intervention of interest (testing) and the Key Question (impact of testing on outcomes).

The clinical implications of the links between testosterone levels and sexual dysfunction or diabetes/metabolic syndrome are unclear. Authors generally did not discuss the relevance of the magnitude of testosterone level differences in observational studies, present risk differences, or discuss the clinical relevance of benefits demonstrated in testosterone therapy trials. For men with both hypogonadism (defined by low serum testosterone as well as sexual symptoms) and type 2 diabetes, improvement in glucose control ranged from negligible to substantial, and average improvement in insulin resistance was modest in populations where baseline glucose control was, on average, adequate or close to adequate. For men with sexual dysfunction, testosterone therapy as a stand-alone therapy was found to be effective in improving symptoms of sexual dysfunction, but no systematic assessment of testosterone therapy as an alternative to erectile dysfunction medications was available. The authors of the systematic review of testosterone therapy in men with sexual dysfunction stated in their concluding remarks that testosterone therapy administered to men with hypogonadism is mainly considered to be a hormonal therapy targeting the condition of hypogonadism rather than a symptomatic treatment of sexual dysfunction.

In addition to the uncertainty regarding clinical application of the reviewed evidence, another concern is that follow-up in the testosterone therapy trials was too short (3 to 12 months) to allow a strong

conclusion that testing for and treating low testosterone levels in men with metabolic syndrome, type 2 diabetes, or sexual dysfunction would be beneficial and safe in the long term.

Practice Guidelines

Four relevant practice guidelines published by U.S. groups in the last 10 years were initially identified. An additional guideline from the American Diabetes Association was reviewed after the choice of subpopulations for the post hoc analysis. **Appendix VIII** presents the recommendations of each guideline.

The Endocrine Society

The most comprehensive guidelines on testosterone testing and testosterone therapy have been published by the Endocrine Society (Bhasin et al., 2010). They were considered to be of fair quality (5 on a scale of 0 to 7). The Endocrine Society differentiates between subpopulations in which testosterone should be tested if hypogonadal symptoms are present (e.g., a patient with end-stage renal disease and unexplained weight loss) and subpopulations in whom screening might be reasonable regardless of symptoms (e.g., pituitary mass, use of glucocorticoids or opioids, or human immunodeficiency virus ([HIV]-associated weight loss). However, the Society recommends addressing reversible illness and drug usage before testing. The guidelines recommend *against* screening for low testosterone in a general population, even an older population, or routinely offering testosterone therapy to older men. For older men who have *symptoms* suggestive of age-related hypogonadism, the Society recommends discussing the risks and benefits of testosterone therapy with the patient. The Society characterizes all of the evidence supporting testing and treatment recommendations to be of *low or very low quality*. The Society also considers all of its testing and screening recommendations to be *weak*, except for a strong recommendation to confirm a low testosterone measurement with repeat testing.

As noted in the previous paragraph, the Society recommends taking signs and symptoms into account for subpopulations that have not been shown to have a high prevalence of low testosterone levels. A consensus process was used to compile a list of signs and symptoms considered to be relatively specific to androgen deficiency and another list of signs and symptoms that are less specific. These 2 lists are presented in **Table 3** and in **Appendix VIII**. Greater consideration should be given to testing if the more specific signs and symptoms are present.

The guideline was based on 3 commissioned systematic reviews (Boloña et al., 2007; Haddad et al., 2007; Fernández-Balsells et al., 2010) and other available systematic reviews and primary studies. More recent systematic reviews on the same topic have been published since publication of the Boloña and Haddad reviews, but the Fernández-Balsells review is described in **Appendix VII**. Where necessary, a consensus development process was developed to formulate recommendations in the absence of evidence. The guideline Task Force graded the evidence according to the Grading of Recommendations, Assessment, Development, and Evaluation (GRADE) Working Group guidelines.

American College of Physicians (ACP) and American Urological Association (AUA)

Guidelines from the ACP on erectile dysfunction contain no recommendation for or against routine testosterone testing because of insufficient evidence (Qaseem et al., 2009). Guidelines produced by the AUA include expert opinion recommendations in support of testosterone testing in men with abnormal

semen analysis, impaired sexual findings, or clinical findings (not specified) suggestive of a specific endocrinopathy (AUA, 2010a; AUA, 2010b). All 3 of these guidelines were judged to be of good quality.

Additional Information not Included in Appendix VIII

American Society for Reproductive Medicine

A statement from the ASRM on treatment of androgen deficiency in the aging male (ASRM, 2006) was not included in the review of practice guidelines. No guideline process was described and the statement refers frequently to other guidelines. The guidance provided in this statement is very consistent with the Endocrine Society's guidelines on *Testosterone Therapy in Adult Men with Androgen Deficiency Syndromes*.

American Urological Association (AUA)

The AUA's updated guideline on *Erectile Dysfunction* makes no recommendations concerning testing for low testosterone levels in men with erectile dysfunction (Montague et al., 2005). The update panel considered a review of testosterone therapy in men with erectile dysfunction to be beyond the scope of the guideline but did include a recommendation *against* testosterone therapy for treatment of erectile dysfunction in men with normal serum testosterone levels.

The Endocrine Society

The Society's clinical practice guideline on *Diagnosis and Treatment of Hyperprolactinemia* was found to have no recommendations regarding testosterone testing or regarding testosterone therapy based on measurement of testosterone levels (Melmed et al., 2011).

National Institute for Health and Care Excellence (NICE)

NICE guidance documents on *Assessment and Treatment for People with Fertility Problems* (NICE, 2013) and *Osteoporosis: Assessing the Risk of Fragility Fracture* (NICE, 2012) include no recommendations regarding testosterone testing or therapy.

U.S. Preventive Services Task Force (USPSTF)

A search on November 15, 2014, yielded no recommendation concerning testosterone testing was found on the USPSTF website (search by keywords *testosterone*, *hypogonadism*).

Selected Payer Policies

Aetna

No coverage policy for testosterone testing per se was identified in the Clinical Policy Bulletin on November 9, 2014 (search by keywords *testosterone*, *hypogonadism*).

However, Aetna considers androgenic anabolic steroids (e.g., testosterone) to be medically necessary for symptomatic androgen deficiency. The full list of indications for which anabolic steroids are considered necessary follows:

- AIDS wasting syndrome; or
- Anemia accompanying renal failure; or
- Bone marrow failure anemias; or
- Breast cancer; or
- Conditions associated with decreased fibrinolytic activity due to anti-thrombin III deficiency or fibrinogen excess (including cutaneous vasculitis, scleroderma of Raynaud's disease, vasculitis of Behcet's disease, complications of deep vein thrombosis such as venous lipodermatosclerosis, other vascular disorders associated with these forms of reduced fibrinolytic activity, and prevention of recurrent venous thrombosis associated with anti-thrombin III deficiency); or
- Constitutional delay in growth (androgenic anabolic steroids); or

- Delayed male puberty (androgenic anabolic steroids); or
- Endometriosis (danazol) (see CPB 0327 – Infertility); or
- Female-to-male gender reassignment; or
- Fibrocystic breast disease or mastalgia (danazol) (see CPB 0512 – Premenstrual Syndrome and Premenstrual Dysphoric Disorder); or
- Growth failure in children with growth hormone deficiency (treatment adjunct); or
- Hereditary angioedema; or
- Hypospadias (testosterone injection as presurgical adjuvant hormonal therapy); or
- Klinefelter's syndrome with hypogonadism (androgenic anabolic steroids); or
- Microphallus (androgenic anabolic steroids); or
- Refractory red cell production anemias (including aplastic anemia, myelofibrosis, myelosclerosis, agnogenic myeloid metaplasia, hypoplastic anemias caused by malignancy or myelotoxic drugs); or
- Severe burn injury; or
- Symptomatic androgen deficiency in men; or
- Weight loss from cancer chemotherapy.

The policy specifically requires assessment of 2 total testosterone levels to determine medical necessity of testosterone replacement for symptomatic androgen deficiency. Two morning samples, drawn between 8:00 a.m. and 10:00 a.m. obtained on different days, are required. The policy defines androgen deficiency as a total testosterone level < 200 ng/dL (6.9 nmol/L) or a low normal testosterone level (≥ 200 ng/dL but < 500 ng/dL) plus elevated sex hormone binding globulin (SHBG). Testing details are not specified for any of the indications other than symptomatic androgen deficiency, and *symptomatic* is not defined.
See Androgens and Anabolic Steroids: Clinical Policy Bulletin No. 0528.

Centers for Medicare & Medicaid Services (CMS)

No CMS National Coverage Determination (NCD) was identified for testosterone testing on November 9, 2014 (search National Coverage Documents by keywords *testosterone*, *hypogonadism* in National Coverage Determinations at: CMS Advanced Search Database). In the absence of an NCD, coverage decisions are left to the discretion of local Medicare carriers. No relevant information was identified by a search on November 9, 2014, of the Your Medicare Coverage site (keywords *testosterone*, *hypogonadism*)

GroupHealth

No coverage policy for testosterone testing was identified in the GroupHealth Clinical Review Criteria on November 9, 2014 (search separately for *testosterone* and *hypogonadism* in the alphabetical list).

Oregon Health Evidence Review Commission (HERC)

No coverage policy for testosterone testing was identified in the HERC Coverage Guidances (search Completed Guidances by keywords *testosterone*, *hypogonadism*).

Regence Group

No general coverage policy for testosterone testing was identified in the Regency Medical Policy on November 9, 2014 (search by keywords *testosterone*, *hypogonadism*).

However, Regence has concluded that that salivary hormone testing is investigational; this includes salivary testing for testosterone. Saliva tests can be ordered by consumers over the Internet, according to the statement.

See Salivary Hormone Testing for Aging and Menopause: Regence Medical Policy No. 36.

References

American Diabetes Association (ADA). Standards of medical care in diabetes—2014. *Diabetes Care*. 2014;37(Suppl 1):S14-S80. Available at: http://care.diabetesjournals.org/content/37/Supplement_1/S14. Accessed December 3, 2014.

American Society for Reproductive Medicine (ASRM). Practice Committee. Treatment of androgen deficiency in the aging male. *Fertil Steril*. 2006;86(5 Suppl 1):S236-S240. PMID: 17055832.

American Urological Association (AUA). *The Evaluation of the Azoospermic Male: AUA Best Practice Statement*. 2010a. Available at: http://www.auanet.org/common/pdf/education/clinical-guidance/Male-Infertility-b.pdf. Accessed July 9, 2014.

American Urological Association (AUA). *The Optimal Evaluation of the Infertile Male: AUA Best Practice Statement*. 2010b. Available at: http://www.auanet.org/common/pdf/education/clinical-guidance/Male-Infertility-d.pdf. Accessed July 9, 2014.

Appraisal of Guidelines Research & Evaluation (AGREE) Enterprise. Appraisal of Guidelines for Research and Evaluation II Instrument. Updated September 2013. Available at: http://www.agreetrust.org/wp-content/uploads/2013/10/AGREE-II-Users-Manual-and-23-item-Instrument_2009_UPDATE_2013.pdf. Accessed November 26, 2014.

Araujo AB, Dixon JM, Suarez EA, Murad MH, Guey LT, Wittert GA. Clinical review: endogenous testosterone and mortality in men: a systematic review and meta-analysis. *J Clin Endocrinol Metab*. 2011;96(10):3007-3019. PMCID: 3200249.

Araujo AB, Esche GR, Kupelian V, et al. Prevalence of symptomatic androgen deficiency in men. *J Clin Endocrinol Metab*. 2007;92(11):4241-4247. PMID: 17698901.

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(12):5920-5926. PMID: 15579737.

Bhasin S, Cunningham GR, Hayes FJ, et al.; Endocrine Society Task Force. Testosterone therapy in men with androgen deficiency syndromes: an Endocrine Society clinical practice guideline. *J Clin Endocrinol Metab*. 2010;95(6):2536-2559. PMID: 20525905.

Boloña ER, Uraga MV, Haddad RM, et al. Testosterone use in men with sexual dysfunction: a systematic review and meta-analysis of randomized placebo-controlled trials. *Mayo Clin Proc*. 2007;82(1):20-28. PMID: 17285782.

Brand JS, van der Tweel I, Grobbee DE, Emmelot-Vonk MH, van der Schouw YT. Testosterone, sex hormone-binding globulin and the metabolic syndrome: a systematic review and meta-analysis of observational studies. *Int J Epidemiol*. 2011;40(1):189-207. PMID: 20870782.

Cai X, Tian Y, Wu T, Cao CX, Li H, Wang KJ. Metabolic effects of testosterone replacement therapy on hypogonadal men with type 2 diabetes mellitus: a systematic review and meta-analysis of randomized controlled trials. *Asian J Androl*. 2014;16(1):146-152. PMID: 24369149.

Calof OM, Singh AB, Lee ML, et al. Adverse events associated with testosterone replacement in middle-aged and older men: a meta-analysis of randomized, placebo-controlled trials. *J Gerontol A Biol Sci Med Sci*. 2005;60(11):1451-1457. PMID: 16339333.

Center for Drug Evaluation and Research (CDER). Testosterone Replacement Therapy. Bone, Reproductive and Urologic Drugs Advisory Committee & Safety and Risk Management Advisory Committee. September 17, 2014a. Food and Drug Administration [website]. Available at: http://www.fda.gov/downloads/advisorycommittees/committeesmeetingmaterials/drugs/reproductivehealthdrugsadvisorycommittee/ucm416461.pdf. Accessed November 4, 2014.

Center for Drug Evaluation and Research (CDER). FDA Drug Safety Communication: FDA evaluating risk of stroke, heart attack and death with FDA-approved testosterone products. Updated June 20, 2014b. Food and Drug Administration [website]. Available at: http://www.fda.gov/drugs/drugsafety/ucm383904.htm. Accessed November 4, 2014.

Centers for Disease Control and Prevention (CDC). Laboratory Quality Assurance and Standardization Programs. Hormone Standardization. Updated October 10, 2014. Available at: http://www.cdc.gov/labstandards/hs.html. Accessed November 17, 2014.

Chueh KS, Huang SP, Lee YC, et al. The comparison of the aging male symptoms (AMS) scale and androgen deficiency in the aging male (ADAM) questionnaire to detect androgen deficiency in middle-aged men. *J Androl*. 2012;33(5):817-823. PMID: 22240559.

College of American Pathologists (CAP). Quality Assurance Programs. 2015. Available at: http://www.cap.org/web/home/lab/quality-assurance-programs?_adf.ctrl-state=fgxvngkao_25&_afrLoop=339071935724599. Accessed January 25, 2015.

Corona G, Isidori AM, Buvat J, et al. Testosterone supplementation and sexual function: a meta-analysis study. *J Sex Med*. 2014a;e11(6):1577-1592. PMID: 24697970.

Corona G, Maseroli E, Rastrelli G, et al. Cardiovascular risk associated with testosterone-boosting medications: a systematic review and meta-analysis. *Expert Opin Drug Saf*. 2014b;13(10):1327-1351. PMID: 25139126.

Corona G, Monami M, Rastrelli G, et al. Testosterone and metabolic syndrome: a meta-analysis study. *J Sex Med*. 2011a;8(1):272-283. PMID: 20807333.

Corona G, Monami M, Rastrelli G, et al. Type 2 diabetes mellitus and testosterone: a meta-analysis study. *Int J Androl*. 2011b;34(6 Pt 1):528-540. PMID: 20969599.

Corona G, Rastrelli G, Maggi M. Diagnosis and treatment of late-onset hypogonadism: systematic review and meta-analysis of TRT outcomes. *Best Pract Res Clin Endocrinol Metab*. 2013b;27(4):557-579. PMID: 24054931.

Corona G, Rastrelli G, Monami M, et al. Body weight loss reverts obesity-associated hypogonadotropic hypogonadism: a systematic review and meta-analysis. *Eur J Endocrinol*. 2013a;168(6):829-843. PMID: 23482592.

Corona G, Rastrelli G, Monami M, et al. Hypogonadism as a risk factor for cardiovascular mortality in men: a meta-analytic study. *Eur J Endocrinol*. 2011c;165(5):687-701. PMID: 21852391.

Cui Y, Zhang Y. The effect of androgen-replacement therapy on prostate growth: a systematic review and meta-analysis. *Eur Urol*. 2013;64(5):811-822. PMID: 23567065.

Dev R, Hui D, Del Fabbro E, et al. Association between hypogonadism, symptom burden, and survival in male patients with advanced cancer. *Cancer*. 2014;120(10):1586-1593. PMID: 24577665.

Drake MT, Murad MH, Mauck KF, et al. Clinical review. Risk factors for low bone mass-related fractures in men: a systematic review and meta-analysis. *J Clin Endocrinol Metab*. 2012;97(6):1861-1870. PMID: 22466344.

El-Sakka AI, Hassoba HM, Sayed HM, Tayeb KA. Pattern of endocrinal changes in patients with sexual dysfunction. *J Sex Med*. 2005;2(4):551-558. PMID: 16422853.

Fernández-Balsells MM, Murad MH, Lane M, et al. Clinical review 1: adverse effects of testosterone therapy in adult men: a systematic review and meta-analysis. *J Clin Endocrinol Metab*. 2010;95(6):2560-2575. PMID: 20525906.

Finkle WD, Greenland S, Ridgeway GK, et al. Increased risk of non-fatal myocardial infarction following testosterone therapy prescription in men. *PloS One*. 2014;9(1):e85805. PMCID: 3905977.

Food and Drug Administration (FDA). Drug Safety and Availability >> FDA adding general warning to testosterone products about potential for venous blood clots. June 19, 2014. Available at: http://www.fda.gov/Drugs/DrugSafety/ucm401746.htm?source=govdelivery&utm_medium=email&utm_source=govdelivery. Accessed January 22, 2015.

Gades NM, Jacobson DJ, McGree ME, et al. The associations between serum sex hormones, erectile function, and sex drive: the Olmsted County Study of Urinary Symptoms and Health Status among Men. *J Sex Med*. 2008;5(9):2209-2220. PMCID: 2586969.

Grossmann M, Hoermann R, Wittert G, Yeap BB. Effects of testosterone treatment on glucose metabolism and symptoms in men with type 2 diabetes and the metabolic syndrome: a systematic review and meta-analysis of randomized controlled clinical trials. *Clin Endocrinol (Oxf)*. 2014. Epub ahead of print. December 29, 2014. PMID: 25557752.

Haddad RM, Kennedy CC, Caples SM, et al. Testosterone and cardiovascular risk in men: a systematic review and meta-analysis of randomized placebo-controlled trials. *Mayo Clin Proc*. 2007;82(1):29-39. PMID: 17285783.

Hall SA, Esche GR, Araujo AB, et al. Correlates of low testosterone and symptomatic androgen deficiency in a population-based sample. *J Clin Endocrinol Metab*. 2008;93(10):3870-3877. PMCID: 2579652.

Hanafy HM. Testosterone therapy and obstructive sleep apnea: is there a real connection? *J Sex Med*. 2007;4(5):1241-1246. PMID: 17645445.

Ho CK. Testosterone testing in adult males. *Malays J Pathol*. 2011;33(2):71-81. PMID: 22299206.

Holm AC, Fredrikson MG, Theodorsson E, et al. Change in testosterone concentrations over time is a better predictor than the actual concentrations for symptoms of late onset hypogonadism. *Aging Male*. 2011;14(4):249-256. PMID: 21627562.

Hsu B, Cumming RG, Blyth FM, et al. Longitudinal and cross-sectional relationships of circulating reproductive hormone levels to self-rated health and health-related quality of life in community-dwelling older men. *J Clin Endocrinol Metab*. 2014;99(5):1638-1647. PMID: 24552220.

Johns KK, Beddall MJ, Corrin RC. Anabolic steroids for the treatment of weight loss in HIV-infected individuals. *Cochrane Database Syst Rev*. 2005;(4):CD005483. PMID: 16235407.

Layton JB, Li D, Meier CR, et al. Testosterone lab testing and initiation in the United Kingdom and the United States, 2000 to 2011. *J Clin Endocrinol Metab*. 2014;99(3):835-842. PMID: 24423353.

MacLean C, Newberry S, Maglione M, et al. Systematic review: comparative effectiveness of treatments to prevent fractures in men and women with low bone density or osteoporosis. *Ann Intern Med*. 2008;148(3):197-213. PMID: 18087050.

McGill JJ, Shoskes DA, Sabanegh ES. Androgen deficiency in older men: indications, advantages, and pitfalls of testosterone replacement therapy. *Cleve Clin J Med*. 2012;79(11):797-806. PMID: 23125330.

McWilliams K, Simmons C, Laird BJ, Fallon MT. A systematic review of opioid effects on the hypogonadal axis of cancer patients. *Support Care Cancer*. 2014;22(6):1699-1704. PMID: 24633592.

Melmed S, Casanueva FF, Hoffman AR, et al. Diagnosis and treatment of hyperprolactinemia: an Endocrine Society clinical practice guideline. *J Clin Endocrinol Metab*. 2011;96(2):273-288. PMID: 24633592.

Montague DK, Jarow JP, Broderick GA, et al.; Erectile Dysfunction Guideline Update Panel. Chapter 1: The management of erectile dysfunction: an AUA update. *J Urol*. 2005;174(1):230-239. PMID: 15947645.

Moore C, Huebler D, Zimmermann T, Heinemann LA, Saad F, Thai DM. The Aging Males' Symptoms scale (AMS) as outcome measure for treatment of androgen deficiency. *Eur Urol*. 2004;46(1):80-87. PMID: 15183551.

Mohr BA, Bhasin S, Kupelian V, Araujo AB, O'Donnell AB, McKinlay JB. Testosterone, sex hormone-binding globulin, and frailty in older men. *J Am Geriatr Soc*. 2007;55(4):548-555. PMID: 17397433.

Morales A, Collier CP, Clark AF. A critical appraisal of accuracy and cost of laboratory methodologies for the diagnosis of hypogonadism: the role of free testosterone assays. *Can J Urol*. 2012;19(3):6314-6318. PMID: 22704323.

Morley JE, Charlton E, Patrick P, et al. Validation of a screening questionnaire for androgen deficiency in aging males. *Metabolism*. 2000;49(9):1239-1242. PMID: 11016912.

Moskovic DJ, Araujo AB, Lipshultz LI, Khera M. The 20-year public health impact and direct cost of testosterone deficiency in U.S. men. *J Sex Med*. 2013;10(2):562-569. PMID: 23035926.

National Institute for Health and Care Excellence (NICE). *Assessment and Treatment for People with Fertility Problems.* 2013. Available at: http://www.nice.org.uk/guidance/cg156/resources/guidance-fertility-pdf. Accessed July 9, 2014.

National Institute for Health and Care Excellence (NICE). *Osteoporosis: assessing the risk of fragility fracture*. 2012. Available at: http://www.nice.org.uk/guidance/cg146/resources/guidance-osteoporosis-assessing-the-risk-of-fragility-fracture-pdf. Accessed July 9, 2014.

Oskui PM, French WJ, Herring MJ, Mayeda GS, Burstein S, Kloner RA. Testosterone and the cardiovascular system: a comprehensive review of the clinical literature. *J Am Heart Assoc*. 2013;2(6):e000272. PMID: 24242682.

Paduch DA, Brannigan RE, Fuchs EF, Kim ED, Marmar JL, Sandlow JI. The laboratory diagnosis of testosterone deficiency. *Urology*. 2014;83(5):980-988. PMID: 24548716.

Pantalone KM, Faiman C. Male hypogonadism: more than just a low testosterone. *Cleve Clin J Med*. 2012;79(10):717-725. PMID: 23027731.

Ponholzer A, Plas E, Schatzl G, et al. Relationship between testosterone serum levels and lifestyle in aging men. *Aging Male Off J Int Soc Study Aging Male*. 2005;8(3-4):190-193. PMID: 16390745.

Qaseem A, Snow V, Denberg TD, et al.; Clinical Efficacy Assessment Subcommittee of the American College of Physicians. Hormonal testing and pharmacologic treatment of erectile dysfunction: a clinical practice guideline from the American College of Physicians. *Ann Intern Med*. 2009;151(9):639-649. PMID: 19884625.

Ruige JB, Mahmoud AM, De Bacquer D, Kaufman J-M. Endogenous testosterone and cardiovascular disease in healthy men: a meta-analysis. *Heart Br Card Soc*. 2011;97(11):870-875. PMID: 21177660.

Schaap LA, Pluijm SMF, Deeg DJH, et al. Low testosterone levels and decline in physical performance and muscle strength in older men: findings from two prospective cohort studies. *Clin Endocrinol (Oxf)*. 2008;68(1):42-50. PMID: 17666086.

Shea BJ, Bouter LM, Peterson J, et al. External validation of a measurement tool to assess systematic reviews Shea BJ, Bouter LM, Peterson J, et al. External validation of a measurement tool to assess systematic reviews (AMSTAR). *PloS One*. 2007;2(12):e1350. PMID: 18159233.

Shea JL, Wongt PY, Chen Y. Free testosterone: clinical utility and important analytical aspects of measurement. *Adv Clin Chem*. 2014;63:59-84. PMID: 24783351.

Shores MM, Smith NL, Forsberg CW, Anawalt BD, Matsumoto AM. Testosterone treatment and mortality in men with low testosterone levels. *J Clin Endocrinol Metab*. 2012;97(6):2050-2058. PMID: 22496507.

Tajar A, Huhtaniemi IT, O'Neill TW, et al. Characteristics of androgen deficiency in late-onset hypogonadism: results from the European Male Aging Study (EMAS). *J Clin Endocrinol Metab*. 2012;97(5):1508-1516. PMID: 22419720.

Tracz MJ, Sideras K, Boloña ER, et al. Testosterone use in men and its effects on bone health. A systematic review and meta-analysis of randomized placebo-controlled trials. *J Clin Endocrinol Metab*. 2006;91(6):2011-2016. PMID: 16720668.

Travison TG, Araujo AB, Hall SA, McKinlay JB. Temporal trends in testosterone levels and treatment in older men. *Curr Opin Endocrinol Diabetes Obes*. 2009;16(3):211-217. PMID: 19396984.

Travison TG, Araujo AB, Kupelian V, O'Donnell AB, McKinlay JB. The relative contributions of aging, health, and lifestyle factors to serum testosterone decline in men. *J Clin Endocrinol Metab*. 2007;92(2):549-555. PMID: 17148559.

Vigano A, Piccioni M, Trutschnigg B, Hornby L, Chaudhury P, Kilgour R. Male hypogonadism associated with advanced cancer: a systematic review. *Lancet Oncol*. 2010;11(7):679-684. PMID: 20541464.

Vigen R, O'Donnell CI, Barón AE, et al. Association of testosterone therapy with mortality, myocardial infarction, and stroke in men with low testosterone levels. *JAMA*. 2013;310(17):1829-1836. PMID: 24193080.

Wang C, Catlin DH, Demers LM, Starcevic B, Swerdloff RS. 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(2):534-543. PMID: 14764758.

Watts NB, Adler RA, Bilezikian JP, et al. Osteoporosis in men: an Endocrine Society clinical practice guideline. *J Clin Endocrinol Metab*. 2012;97(6):1802-1822. PMID: 22675062.

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(2):123-135. PMID: 20554979.

Yeap BB, Almeida OP, Hyde Z, et al. Healthier lifestyle predicts higher circulating testosterone in older men: the Health In Men Study. *Clin Endocrinol (Oxf)*. 2009;70(3):455-463. PMID: 18691270.

Zarrouf FA, Artz S, Griffith J, Sirbu C, Kommor M. Testosterone and depression: systematic review and meta-analysis. *J Psychiatr Pract*. 2009;15(4):289-305. PMID: 19625884.

APPENDICES

Appendix I. Search Strategy

INITIAL SEARCH, SYSTEMATIC REVIEWS AND PRACTICE GUIDELINES (conducted June 29, 2014)

Initially, evidence for this report was obtained by searching for systematic reviews and guidelines that had been published in the past 10 years. Searches were conducted in the following databases using the term *testosterone*: Agency for Healthcare Research and Quality (AHRQ), Blue Cross Blue Shield TEC Assessments, Canadian Agency for Drugs and Technology in Health (CADTH), Centre for Reviews and Dissemination (CRD) (York University), Hayes Knowledge Center, Institute for Clinical Systems Improvement (ICSI), National Institute for Health Research Health Technology Assessment (NIHR HTA) Programme (UK), U.S. Preventive Services Task Force (USPSTF), National Guidelines Clearinghouse (NGC), National Institute for Health and Care Excellence (NICE), and Veterans Affairs Technology Assessment Program (VA TAP). (NOTE: The CRD search strategy includes a search for Cochrane Reviews.)

The websites for the American Academy of Family Physicians, American College of Physicians, the American Urological Association, and the American Gerontological Society, and the Endocrine Society, were also searched since guidelines for both older men and older women were not identified for these groups in a search of the NGC.
Additional systematic reviews were selected from a search of the PubMed database using filters for Practice Guidelines, Guidelines, Meta-Analyses, and Systematic Reviews, according to this search:

testosterone or hypogonadism

Filters: Meta-Analysis; Systematic Reviews; Publication date from 2009/01/01 to 2014/12/31; English

PRIMARY CLINICAL STUDIES

The following search was conducted in PubMed on September 7, 2014:

1. "low testosterone" or "Testosterone/deficiency"[Majr] or "Testosterone/blood"[Majr] or "Hypogonadism/blood"[Majr] or “Andropause”[Majr]
2. screen[tiab] or screening[All Fields] or monitor[tiab] or monitoring[All Fields] or test[All Fields] or testing[All Fields] or follow-up[All Fields] or interval[tiab]
3. 1 AND 2
4. ("Longitudinal Studies"[Mesh] or "Cohort Studies"[Mesh] or "Case-Control Studies"[Mesh]) or ((((clinical[Title/Abstract] AND trial[Title/Abstract]) OR clinical trials as topic[MeSH Terms] OR clinical trial[Publication Type] OR random*[Title/Abstract] OR random allocation[MeSH Terms] OR therapeutic use[MeSH Subheading])))
5. 3 AND 4
6. ("addresses"[Publication Type] OR "autobiography"[Publication Type] OR "bibliography"[Publication Type] OR "biography"[Publication Type] OR "book illustrations"[Publication Type] OR "classical article"[Publication Type] OR "clinical

conference"[Publication Type] OR "collected works"[Publication Type] OR "comment"[Publication Type] OR "congresses"[Publication Type] OR "consensus development conference"[Publication Type] OR "consensus development conference, nih"[Publication Type] OR "dictionary"[Publication Type] OR "directory"[Publication Type] OR "duplicate publication"[Publication Type] OR "editorial"[Publication Type] OR "ephemera"[Publication Type] OR "festschrift"[Publication Type] OR "guideline"[Publication Type] OR "historical article"[Publication Type] OR "in vitro"[Publication Type] OR "interactive tutorial"[Publication Type] OR "interview"[Publication Type] OR "lectures"[Publication Type] OR "legal cases"[Publication Type] OR "legislation"[Publication Type] OR "letter"[Publication Type] OR "news"[Publication Type] OR "newspaper article"[Publication Type] OR "overall"[Publication Type] OR "patient education handout"[Publication Type] OR "periodical index"[Publication Type] OR "personal narratives"[Publication Type] OR "pictorial works"[Publication Type] OR "popular works"[Publication Type] OR "portraits"[Publication Type] OR "practice guideline"[Publication Type] OR "review"[Publication Type] OR "scientific integrity review"[Publication Type] OR "video audio media"[Publication Type] OR "webcasts"[Publication Type])

5 NOT 6

Filters: Publication date from 1990/01/01 to 2014/12/31; Humans; Male, Adult: 19+ years

The following search was conducted in OVID-Embase on September 10, 2014:

1. (Testosterone or hypogonadism or andropause).mp. [mp=ti, ab, sh, hw, tn, ot, dm, mf, dv, kw, nm, kf, px, rx, ui]
2. (screen or monitor or test or interval).ab,ti.
3. (screening or monitoring or testing or follow-up).af,ab,ti.
4. 2 or 3
5. 1 and 4
6. limit 5 to "therapy (maximizes sensitivity)"
7. (longitudinal studies or cohort studies or case-control studies).sh.
8. 5 and 7
9. 6 or 8
10. limit 9 to (book or book series or conference abstract or conference paper or conference proceeding or "conference review" or editorial or letter or note or "review" or short survey or trade journal or addresses or autobiography or bibliography or biography or case reports or clinical conference or congresses or consensus development conference or consensus development conference, nih or dataset or dictionary or directory or interactive tutorial or interview or lectures or legal cases or legislation or meta-analysis or multicenter study or news or newspaper article or patient education handout or periodical index or video-audio media or webcasts)
11. 9 not 10

Limits: "all adult (19 plus years)"; human; humans; yr="1990 - 2014"

The publication date limit was selected in consultation with the clinical expert assigned to this report.

ECONOMIC EVALUATIONS AND COST STUDIES

The following search was conducted in PubMed on September 25, 2014:

1. "low testosterone" or "Testosterone/deficiency"[Majr] or "Testosterone/blood"[Majr] or "Hypogonadism/blood"[Majr] or “Andropause”[Majr]
2. screen[tiab] or screening[All Fields] or monitor[tiab] or monitoring[All Fields] or test[All Fields] or testing[All Fields] or follow-up[All Fields] or interval[tiab]
3. ((((economic analysis) OR (economic evaluation)))) OR (((((cost AND (analysis OR benefit OR effective* OR consequence OR minimization)))) OR (("Costs and Cost Analysis"[MeSH] OR "Cost-Benefit Analysis"[MeSH]))))
4. 1 and 2 and 3

Filters: Publication date from 2003/09/01 to 2014/12/31; Humans; Male, Adult: 19+ years

UPDATE SEARCHES

Update searches were conducted on November 19 (for systematic reviews) and December 3, 2014, for primary studies designed to provide direct evidence for the Key Questions.

POST HOC SEARCH

The following search was conducted in OVID-MEDLINE and OVID-Embase on September 10, 2014:

1. (testosterone or hypogonadism or andropause).mp. [mp=ti, ab, sh, hw, tn, ot, dm, mf, dv, kw, nm, kf, px, rx, an, ui]
2. (sexual dysfunction or sexual function or libido or erectile dysfunction or erectile function or metabolic syndrome or diabetes).mp. [mp=ti, ab, sh, hw, tn, ot, dm, mf, dv, kw, nm, kf, px, rx, an, ui]
3. 1 and 2
4. limit 3 to (book or book series or conference abstract or conference paper or conference proceeding or "conference review" or editorial or letter or note or "review" or short survey or trade journal or addresses or autobiography or bibliography or biography or case reports or clinical conference or congresses or consensus development conference or consensus development conference, nih or dataset or dictionary or directory or interactive tutorial or interview or lectures or legal cases or legislation or meta-analysis or news or newspaper article or patient education handout or periodical index or video-audio media or webcasts)
5. 3 not 4
6. limit 5 to randomized controlled trial
7. (longitudinal studies or cohort studies or case-control studies).sh.
8. (prevalence or association).m_titl.
9. 7 or 8
10. 5 and 9
11. 6 or 10

12. limit 11 to yr="2010 -Current"
13. remove duplicates from 12

Appendix II. Overview of Evidence Quality Assessment Methods

Clinical Studies

Tools used include internally developed Quality Checklists for evaluating the quality (internal validity) of different types of studies, a checklist for judging the adequacy of systematic reviews used instead of de novo analysis, and Hayes Evidence-Grading Guides for evaluating bodies of evidence for different types of technologies. Hayes methodology is in alignment with the GRADE (Grading of Recommendations, Assessment, Development, and Evaluation) system, which was developed by the GRADE Working Group, an international collaborative body.

Step 1	Individual study appraisal a. Initial rating according to study design *Good:* Randomized Controlled Trials *Fair:* Nonrandomized Trial (controlled, parallel-group, quasirandomized) *Poor:* Observational Analytic Studies (prospective or retrospective trials involving historical controls, pretest-posttest control trial [patients legitimately serve as their own controls], case-control, registry/chart/database analysis involving a comparison group) *Very Poor:* Descriptive Uncontrolled Studies (case reports, case series, cross-sectional surveys [individual-level data], correlation studies [group-level data]) b. Consider the methodological rigor of study execution according to items in a proprietary Quality Checklist c. Repeat for each study
Step 2	Evaluation of each body of evidence by outcome, key question, or application a. Initial quality designation according to best study design in a body of evidence b. Downgrade/upgrade Downgrade factors: Study weaknesses (Quality Checklists), small quantity of evidence, lack of applicability, inconsistency of results, publication bias Possible upgrade factors: Strong association, dose-response effect, bias favoring no effect c. Assign final rating: High-Moderate-Low-Insufficient d. Repeat for each outcome/question/application
Step 3	Evaluation of overall evidence a. Rank outcomes by clinical importance b. Consider overall quality of evidence for each *critical* outcome c. Assign overall rating based on lowest-quality body: High-Moderate-Low-Insufficient
Step 4	Evidence-based conclusion Overall quality of evidence + Balance of benefits and harms

Practice Guidelines (checklist taken from **AGREE Tool** and approach to scoring used in this report)

Rank each item on a scale of 1-7.

Decide on overall quality (1 = lowest to 7 = highest), giving strongest weight to items 7-14 (Rigor of Development Domain) and items 22-23 (Editorial Independence). For qualitative labels:

Very poor = 1
Poor = 2-3
Fair = 4-5
Good = 6-7

1. The overall objective(s) of the guideline is (are) specifically described.
2. The health question(s) covered by the guideline is (are) specifically described.
3. The population (patients, public, etc.) to whom the guideline is meant to apply is specifically described.
4. The guideline development group includes individuals from all relevant professional groups.
5. The views and preferences of the target population (patients, public, etc.) have been sought.
6. The target users of the guideline are clearly defined.
7. Systematic methods were used to search for evidence.
8. The criteria for selecting the evidence are clearly described.
9. The strengths and limitations of the body of evidence are clearly described.
10. The methods for formulating the recommendations are clearly described.
11. The health benefits, side effects, and risks have been considered in formulating the recommendations.
12. There is an explicit link between the recommendations and the supporting evidence.
13. The guideline has been externally reviewed by experts prior to its publication.
14. A procedure for updating the guideline is provided.
15. The recommendations are specific and unambiguous.
16. The different options for management of the condition or health issue are clearly presented.
17. Key recommendations are easily identifiable.
18. The guideline describes facilitators and barriers to its application.
19. The guideline provides advice and/or tools on how the recommendations can be put into practice.
20. The potential resource implications of applying the recommendations have been considered.
21. The guideline presents monitoring and/or auditing criteria.
22. The views of the funding body have not influenced the content of the guideline.
23. Competing interests of guideline development group members have been recorded and addressed.

Economic Evaluations

A tool developed by Hayes for internal use guides interpretation and critical appraisal of economic evaluations. The tool includes a checklist of items addressing issues such as the reliability of effectiveness assumptions, transparency of reporting, quality of analysis, generalizability/applicability, and conflicts of interest. The following publications served as sources of best practice.

Articles

Brunetti M, Shemilt I, Pregno S, et al. GRADE guidelines: 10. Considering resource use and rating the quality of economic evidence. J Clin Epidemiol. 2013;66(2):140-150.

Drummond MF, Jefferson TO. Guidelines for authors and peer reviewers of economic submissions to the BMJ. The BMJ Economic Evaluation Working Party. BMJ. 1996;313(7052):275-283.

Drummond M, Sculpher M. Common methodological flaws in economic evaluations. Med Care. 2005;43(7 Suppl):5-14.

Evers S, Goossens M, de Vet H, van Tulder M, Ament A. Criteria list for assessment of methodological quality of economic evaluations: Consensus on Health Economic Criteria. Int J Technol Assess Health Care. 2005;21(2):240-245.

Gerkens S, Crott R, Cleemput I, et al. Comparison of three instruments assessing the quality of economic evaluations: a practical exercise on economic evaluations of the surgical treatment of obesity. Int J Technol Assess Health Care. 2008;24(3):318-325.

Hutubessy R, Chisholm D, Edejer TT. Generalized cost-effectiveness analysis for national-level priority-setting in the health sector. Cost Eff Resour Alloc. 2003;1(1):8.

Shemilt I, Thomas J, Morciano M. A web-based tool for adjusting costs to a specific target currency and price year. Evid Policy. 2010;6(1):51-59.

Smith KA, Rudmik L. Cost collection and analysis for health economic evaluation. Otolaryngol Head Neck Surg. 2013;149(2):192-199.

Ubel PA, Hirth RA, Chernew ME, Fendrick AM. What is the price of life and why doesn't it increase at the rate of inflation? Arch Intern Med. 2003;163(14):1637-1641.

Books

Drummond MF, O'Brien BJ, Stoddart GL, Torrance GW. Methods for the Economic Evaluation of Health Care Programmes. 2nd Edition. Oxford, UK: Oxford University Press; 1997.

Gold MR, Siegel JE, Russell LB, Weinstein MC, eds. Cost-Effectiveness in Health and Medicine. 1996. Oxford, UK: Oxford University Press; 1996.

Other

Canadian Agency for Drugs and Technologies in Health (CADTH). Guidelines for the Economic Evaluation of Health Technologies. 3rd Edition. Ottawa, Canada: Canadian Agency for Drugs and Technologies in Health; 2006. Available at: http://www.cadth.ca/media/pdf/186_EconomicGuidelines_e.pdf. Accessed August 20, 2014.

Appendix III. Observational Studies Investigating the Association of Signs and Symptoms with Testosterone Level or Hypogonadism Diagnosis

Key: BMI, body mass index; BT, bioavailable testosterone; ECOG, Eastern Cooperative Oncology Group; ED, erectile dysfunction; (M/P)CS, (mental/physical) component summary; ng, nanogram; nmol, nanomole; OR, odds ratio; NS, statistically nonsignificant; QOL, quality of life; RR, relative risk; SF-36, SF-36 Health Survey (QualityMetric Inc.); sig, statistically significant; (T/F)T, (total/free) testosterone

Study/ Patients/ Adjustments	Findings	Relationship*
Health and Functional Status		
Hsu et al. (2014) Convenience sample, 1637 men, age ≥70 yrs, Australia Adjusted for smoking status and a physician diagnosis of 19 disorders such as osteoporosis, dementia, depression, myocardial infarction, and chronic lung or kidney disease.	*Baseline association self-related health status:* Sig for TT and FT *OR (odds of fair-poor-very poor health vs excellent-good health, lowest vs highest quartile total T):* 1.47 (95% CI, 1.04 to 2.06) *2-yr change in T levels:* No association of TT or FT w/ self-rated health or QOL	Inverse (but nonsignificant)
Physical Symptoms		
Mohr et a. (2007) Cross-sectional study of random population sample, 646 men, Adjusted for age, chronic disease (DM, hypertension, heart disease, cancer, arthritis, and depression), lifestyle (drinking, physical activity), and diet (calories, protein intake).	*TT:* No association with frailty phenotype (≥3 of 5 symptoms: weight loss, exhaustion, low physical activity, slowness, weak grip strength). Inverse association (based on serum differences) with grip strength and physical activity; not with other symptoms. *FT:* No significant associations. *SHBG:* Odds of frailty phenotype, higher vs lower levels: OR, 1.25 (CI, 1.06-1.46) per 10-nM increase in levels.	Inverse association with some but not all symptoms.

Study/ Patients/ Adjustments	Findings	Relationship*
	Positive association (based on serum differences) with weight loss, exhaustion, and physical activity; not with other symptoms.	
Wu et al. (2010) (European Male Aging Study) Random general population sample, 3369 men, ages 40-79 yrs, Europe Adjusted for age, BMI, #coexisting illnesses.	*Lower TT:* Association w/ 2 of 3 physical symptoms (limits to vigorous activity, walking, and bending/kneeling/stooping) *Lower FT:* Sig association w/ each physical symptom Weak associations	Inverse, FT only and weak association
Tajar et al. (2012) Cross-sectional study, 2966 men, age 40-79 yrs, Europe	Hypogonadism defined as 3 sexual symptoms (decreased frequency of morning erection, decreased frequency of sexual thoughts, ED) plus TT <317 ng/dL (11 nmol/L) and FT <220 pmol/L (6 ng/dL). Men w/ hypogonadism had significantly lower SF-36 PCS scores, physical activity, and self-reported general health.	Inverse, no RR estimate
Psychological Symptoms†		
Wu et al. (2010) (European Male Aging Study) Random general population sample, 3369 men, ages 40-79 yrs, Europe Adjusted for age, BMI, # coexisting illnesses.	*Lower TT:* No association with any of 3 psychological symptoms (sadness, loss of energy, fatigue) *Lower FT:* Sig association w/ each psychological symptom Weak associations	Inverse, FT only and weak association
Tajar et al. (2012) Cross-sectional study, 2966 men, age 40-79 yrs, Europe	Hypogonadism defined as 3 sexual symptoms (decreased frequency of morning erection, decreased frequency of sexual thoughts, ED) plus TT <317 ng/dL (11 nmol/L) and FT <220 pmol/L (6 ng/dL).	No association

Study/ Patients/ Adjustments	Findings	Relationship*
	No relationship between hypogonadism and SF-36 MCS.	
Sexual Symptoms		
Gades et al. (2008) Random sample, 414 men, age ≥50 yrs, Olmstead County, MN Adjusted for age; men w/ comorbidities excluded.	*Lower TT:* Associated w/ ED but not w/ sex drive *BT*: No association	Inverse, depending on symptoms Limited adjustment for confounding; unclear whether men being treated w/ ED medication were included
Wu et al. (2010) (European Male Aging Study) Random general population sample, 3369 men, ages 40-79 yrs, Europe Adjusted for age, BMI, # coexisting illnesses; men w/ known pituitary or testicular diseases or current use of medications that could affect pituitary or testicular function or sex-steroid clearance were excluded.	*Lower TT:* Associated w/ 2 of 3 symptoms (less frequent awakening w/ full erection; diminished ability to get and keep erection; less frequent thinking about sex) *Lower FT:* Associated w/ each of all 3 symptoms; cluster analysis suggested a stable association between low T and simultaneous presence of the 3 symptoms. *Optimal cutoff levels:* TT: 8.0-11 nmol/L (231-317 ng/dL), depending on symptom. FT: 160-280 pmol/L (0.16-0.28 nmol/L; 5-8 ng/dL), depending on symptom *OR (odds of presence of 3 symptoms, low T level vs high T level):* 1.64-2.24, depending on cutoff chosen and whether free as well as total T was low; strong associations NS ORs if <3 symptoms present.	Inverse, depending on symptom and consideration of TT alone or TT plus FT. Strong association if all factors considered. Unclear whether men being treated w/ ED medication were included.
Cancer Symptoms		
Dev et al. (2014)	*Signs/symptoms significantly associated w/ lower T values:* Worse	Inverse

Study/ Patients/ Adjustments	Findings	Relationship*
Cross-sectional study of 119 men with advanced cancer (median age 64 yrs)	fatigue, weight loss, ECOG performance status: TT, FT, and BT. Increased anxiety, decreased feeling of well-being, increased dyspnea: TT and FT. Anorexia: BT *P*<0.05 for all comparisons	

*An inverse relationship is one where low T levels are associated with a higher prevalence of incidence of the health outcome in question; in other words, a relationship where low T level was found to be a risk factor or marker for poorer health outcomes.

†Authors of a systematic review cited small studies suggesting an inverse association with depression (Zarrouf et al., 2009).

Appendix IV. Screening Tools

Key: FT, free testosterone; TT, total testosterone

Tool	Elements	Uses and Validation
Androgen Deficiency in the Aging Male (ADAM)s	1. Do you have a decrease in libido (sex drive)? 2. Do you have a lack of energy? 3. Do you have a decrease in strength and/or endurance? 4. Have you lost height? 5. Have you noticed a decreased ‘‘enjoyment of life’’? 6. Are you sad and/or grumpy? 7. Are your erections less strong? 8. Have you noted a recent deterioration in your ability to play sports? 9. Are you falling asleep after dinner? 10. Has there been a recent deterioration in your work performance? **Yes** answer to questions 1 or 7 or to any 3 other questions constitutes a positive result.	Screen for low testosterone. 72% sensitivity/27% specificity for detecting low TT 88% sensitivity/60% specificity for detecting BT ≤70 ng/dL (2.4 nmol/L) 67% sensitivity/26% specificity for detecting both low TT and low FT
Aging Male Symptom (AMS) Scale	17 items in 3 domains: Psychological, Somatic, Sexual Each item is self-scored for severity according to a 1 to-5 scale.	Screen for low testosterone. Measure outcomes. 57% sensitivity/49% specificity for detecting low TT 57% sensitivity/48% specificity for detecting both low TT and low FT
ANDROTEST	Designed for men with sexual dysfunction.	Sensitivity and specificity close to 70% for detecting low TT or low FT

Source: Morley et al. (2000); Moore et al. (2004); Corona et al. (2011a).

Appendix V. Systematic Reviews of the Association Between Endogenous Testosterone and Health Outcomes or Conditions

Key: BL, baseline; CVD, cardiovascular disease; DM, diabetes mellitus; HR, hazard ratio; MA, meta-analysis; MetS, metabolic syndrome; ng, nanogram; nmol, nanomole; NR, not reported; OR, odds ratio; RCTs, randomized controlled trials; RR, relative risk; SHBG, sex hormone binding globulin; (T/F)T, (total/free) testosterone

Review/Included Studies/Patients	Findings	Relationship*
Bone Health		
Drake et al. (2012) MA of cohort studies, case-control studies, or RCTs. Studies had to evaluate risk factors for low bone mass or for fracture due to low bone mass; 8 studies addressed hypogonadism. Mean age 50-80 yrs by study	*Odds of osteoporosis, hypogonadism vs no hypogonadism (unadjusted OR):* Overall: 1.76 (CI, 1.37-2.26; I^2=85%) (8 studies) Natural hypogonadism: 2.77 (CI, 1.30-5.87; I^2=51%) (4 studies) Drug-induced hypogonadism: 1.53 (CI, 1.19 to 1.96 I^2=91%)); high heterogeneity (4 studies)	Inverse, strong association depending on subpopulation.
Cancer-Related Outcomes		
Vigano et al. (2010)† SR of 6 studies evaluating relationship between hypogonadism and prespecified cancer-related outcomes in men of any age w/ advanced cancer	No clear relationship of hypogonadism w/ nutritional status, body composition, musculoskeletal physiology, quality of life, symptoms, or functional status.	No relationship, but authors did not consider the evidence strong enough to rule out relationships.
CVD		
Corona et al. (2011c)	*Cross-sectional studies and prospective cohort studies:* Men w/ CVD had significantly lower T levels, according to differences in mean levels. Considerable inconsistency	Inverse but weak association.

Review/Included Studies/Patients	Findings	Relationship*
MA of 54 cross-sectional, 10 prospective cohort, and case-control studies Mean age and f/u NR	across individual studies. *HR (higher vs lower T) of CVD in logistic regression (cross-sectional studies):* = 0.536 (CI, 0.447-0.606) after adjustment for age, BMI, diabetes, and hypertension. Cutoff values varied by study. *Difference in TT level (incidence minus no incidence) (prospective cohort studies):*. All-cause mortality: –1.53 nmol/L (CI, –2.69 to –0.37) (3 studies) CVD mortality: –0.97 nmol/L (CI, –1.55 to –0.40) (3 studies) CVD incidence: –1.66 nmol/L (CI, –5.80 to 2.47) (7 studies)	No control for confounders in analysis of prospective cohort studies.
Ruige et al. (2011) MA of 18 prospective cohort studies or nested case-control studies Mean age NR; mean follow-up 5-15 yrs by study	*Pooled RR (for increase of TT by 1 SD):* Overall: 0.89 (CI, 0.83-0.96) (18 studies) Age <70 yrs: 1.01 (CI, 0.95-1.08) (7 studies) Age >70yrs : 0.84 (CI, 0.76-0.92) Similar findings for FT and BT.	Inverse weak association, older men only.
MetS		
Brand et al. (2011)‡ MA of 32 studies (26 cross-sectional, 5 prospective cohort, 1 case-control)	*Mean differences:* TT and FT were lower in men w/ MetS (high heterogeneity for both analyses). Sensitivity analyses (metaregression) of differences in serum T: Trend (*P*=0.08) toward greater TT difference in	Inverse strong association. No analysis of RR specific to longitudinal studies. Poor to fair studies according to author

Review/Included Studies/Patients	Findings	Relationship*
Mean age 14 to >70 yrs by study (2 studies involved adolescents)	younger (<55 yrs) men. Otherwise, no significant subgrp differences for TT or FT, including BMI <25 vs ≥25, or source for MetS/type 2 DM criteria. *Unadjusted RR of MetS:* Higher vs lower TT: 0.38 (CI, 0.25 to 0.50); high heterogeneity (I^2=86.5%) (13 studies) Higher vs lower FT: 0.64 (CI, 0.41-1.01); high heterogeneity (I^2=86.4%); slight inconsistency in results across studies (7 studies) Higher vs lower SHBG: 0.29 (CI, 0.21-0.41); high heterogeneity (I^2=80.7%) (10 studies) Study-reported RRs (higher vs lower serum T) based on varying cutoff values were converted to a standard scale of effect.	standards. High heterogeneity, but generally good consistency of results across studies. No clear evidence of publication bias. *Quality assessment performed as part of post hoc analysis:* Good
Corona et al. (2011a) MA of 17 studies (13 cross-sectional, 3 longitudinal) Mean age 52-58 yrs by study	*Mean difference in serum TT, cross-sectional studies, cases minus controls (nmol/L):* Overall: –2.85 (95% CI, –3.34 to –2.36; I^2=73.8%) (13 studies) No ED: –2.60 (CI, –3.15 to –2.06; I^2 NR) (19 studies) ED: –3.51 (CI, –4.48. to –2.53; I^2 NR) (5 studies) Metaregression: Greater difference in presence of type 2 DM. Multiple regression: Both MetS and type 2 DM were	Inverse. Some evidence of a temporal relationship between low T level and incident MetS. High heterogeneity but no serious inconsistency in direction of study results. Study quality NR. No evidence of publication bias. *Quality:* Fair.

Review/Included Studies/Patients	Findings	Relationship*
	independent predictors for lower TT. FT and SHBG were also significantly lower in men w/ MetS. *Mean difference in serum TT, prospective cohort studies, cases minus controls (nmol/L):* –2.17 (CI, –2.41 to –1.94; I^2=96.2%) (# studies NR)	
Mortality		
Araujo et al. (2011)‡ 12 cohort studies included in MAs (16,184 patients included in MA of all-cause, 11,831 patients included in MA of CVD mortality) Mean age 61 yrs; mean serum T, 487 ng/dL (16.9 nmol/L); mean f/u 9.7 yrs	*RR of all-cause mortality, lower vs higher T:* 1.35 (CI, 1.13-1.62); 11 studies; significant heterogeneity *RR of CVD mortality:* 1.25 (CI, 0.97-1.60); 7 studies; no significant heterogeneity Study-reported RRs based on varying cutoff values were converted to a standard scale of effect. *Sensitivity analyses (metaregression) of all-cause RR mortality:* Higher RRs in study subgroups defined by: Patient age >60 yrs vs ≤60 yrs Patient BL total T ≤487 ng/dL (16.9 nmol/L) vs >487 ng/dL Study f/u ≤9.6 yrs vs >9.6 yrs Blood sampling other than morning vs morning Authors concluded that effects were driven by differences in underlying health status.	Inverse but nonsignificant.

Review/Included Studies/Patients	Findings	Relationship*
Type 2 DM		
Corona et al. (2011b) MA of 33 studies 28 cross-sectional studies: 11,831 men; mean age 38-72 yrs by study 5 prospective cohort studies: 2700 men; mean age, 45-67 yrs by study	*Mean difference in serum TT, cross-sectional studies, cases minus controls (nmol/L):* Overall: –2.99 (95% CI, –3.59 to –2.40; I^2=68.7%) (24 studies) No ED: –3.00 (CI, –3.61 to –2.38; I^2 NR) (19 studies) ED: –2.99 (CI, –3.59 to –2.40; I^2 NR) (5 studies) Metaregression: Differences diminished w/ age but increased w/ BMI. Multiple regression: Type 2 DM was still significantly associated w/ lower TT levels after adjustment for age and BMI. FT and SHBG were also significantly lower in men w/ MetS. *Mean difference in serum TT, prospective cohort studies, cases minus controls (nmol/L):* –2.08 (CI, –3.57 to –0.59; I^2=72.9%) (5 studies) Men w/ MetS also had significantly lower and FT levels.	Inverse association but no estimate of RR. *Quality assessment performed as part of post hoc analysis:* Fair-quality SR. Moderate-high heterogeneity but no serious inconsistency in direction of study results. Study quality NR. No evidence of publication bias.
Opioid Use		
McWilliams et al. (2014) SR of 4 studies of patients (210 men) w/ cancer, opioids or no opioids for pain.	3 studies showed an inverse relationship between opioid dose and TT; 1 study found no relationship between dose and T level. 1 study also reported worse survival for hypogonadal patients, compared w/ eugonadal patients.	Inverse, 3 studies; none, 1 study.

*An inverse relationship is one where low testosterone levels are associated with a higher prevalence of incidence of the health problem in question; in other words, a relationship where low testosterone level was found to be a risk factor or marker for poorer health outcomes.

†A recent study suggested an association between lower total testosterone levels and survival (adjusted hazard ratio [HR], 1.55) in men with advanced cancer(Dev et al., 2014).

‡Very similar findings were reported by a slightly less recent meta-analysis (Corona et al., 2011a) that included fewer studies. An analysis of the Boston Area Community Health Survey showed a significant association between waist circumference and low testosterone levels, but no association between testosterone levels and weight, BMI, or hip circumference (Hall et al., 2008).

Appendix VI. Systematic Reviews of the Effectiveness of Exogenous Testosterone (Testosterone Therapy) in Men with Hypogonadism

Key: BL, baseline BMD, bone mineral density; BT, bioavailable testosterone; CI, confidence interval; CHD, coronary heart disease; CVD, cardiovascular disease; DM, diabetes mellitus; ED, erectile dysfunction; FPG, fasting plasma glucose; FSI, fasting serum insulin; FT, free testosterone; f/u, follow-up; HAM-D, Hamilton Depression Scale; HbA1c, glycated hemoglobin; HIV, human immunodeficiency virus; HOMA-IR, Homeostasis Model Assessment for Insulin Resistance; ITT, intention-to-treat; LOH, late-onset hypogonadism; MA, meta-analysis; MetS, metabolic syndrome; mIU, milli-international unit; ng, nanogram; nmol, nanomole; NR, not reported; NS, not statistically significant; OP, osteoporosis; PDE5i, phosphodiesterase type 5 inhibitor; pg, picogram; PSA, prostate specific antigen; RCT, randomized controlled trial; SMD, standardized mean difference (represents a standardization of different outcome measures; generally an SMD with absolute value ≥ 0.30 is considered small, ≥ 0.50 is considered moderate, and ≥ 0.80 is considered large); SR, systematic review; T, testosterone; TRT, testosterone replacement therapy; TT, total testosterone

Review/Included Studies/Patients	Findings	Summary and Comments
Bone Health		
Tracz et al. (2006) MA of 8 RCTs (365 patients; mean age 36-75 yrs; mean BL T level 291-646 nmol/L TT or 92 nmol/L BT)	No RCTs evaluated effect on fracture incidence. T therapy was associated w/ a 0 to significant 8% gain in BMD, depending on anatomic site and T formulation. Only 1 RCT (n=34) had positive and statistically significant findings. Largest study (n=108) had results that slightly favored placebo but were NS. Moderate statistical heterogeneity.	No conclusions regarding the effect of T therapy on bone health are possible. Patients were blinded in most trials; median 14% loss to follow-up. *Quality of SR:* Good
MacLean et al. (2008) SR of effectiveness of treatments to prevent fractures in men and women w/ low BMD or OP.	No MAs or RCTs assessed the impact of T therapy on fracture risk.	No confident conclusions regarding the effect of T therapy on bone health are possible. *Quality of SR:* Good
CVD		
Corona et al. (2011c) MA of 6 RCTs (189 patients w/ CHD; mean f/u 23 wks; BL TT 4.2-12.9nmol/L where reported; 9.7-10.4 nmol/L is considered normal)	TRT was positively associated w/ (1) a significant increase in treadmill test duration (2 studies w/ consistent results; BL TT was normal in 1 study and NR in the other), and (2) time to 1 mm ST segment depression (6 studies; significant increase in 2 studies w/ normal or unknown TT at BL, NS increase in 4 studies w/ low or normal BL TT, no increase in 1 study w/ unknown BL TT). Health outcomes NR.	According to pooled estimates, T therapy in men w/ CHD improves intermediate measures of cardiovascular health. No evidence that effectiveness varies by BL TT level . Studies were blinded, had acceptable dropout rates, and used ITT analysis. *Quality of SR:* Fair (lack of documented study quality or explicit consideration in conclusions)

Review/Included Studies/Patients	Findings	Summary and Comments
Depression		
Zarrouf et al. (2009) MA of 7 RCTs of T therapy vs placebo in men w/ a diagnosis of depression (mean ages 30-52 yrs)	*OR (placebo vs T therapy for 50% reduction in HAM-D score (<1 favors T therapy):* Overall: 0.40 (CI, 0.26-0.63) (7 RCTs). No heterogeneity. All but 1 RCT had results favoring T therapy. Hypogonadal men: 0.35 (CI, 0.20-0.60) (4 RCTs)	According to pooled estimates, T therapy had an antidepressant effect in men w/ depression. There was little difference in effect between the hypogonadal subgroup and the overall populations. Whether patients had concurrent treatment w/ antidepressant medications was not discussed. Authors did not comment on quality of studies but all were placebo-controlled and double-blind. *Quality of SR:* Good
HIV-Related Outcomes		
Johns et al. (2009) MA of 8 RCTs (359 men) of anabolic steroids (T or derivatives) for HIV-related wasting in subgroup of men; age 34-42 yrs for all trials in review; mean f/u described by authors as short)	*Mean difference in % change/absolute change (kg) from BL in lean body mass:* Overall: 100.0%/1.47 (CI, 0.38-2.55; I^2=76%) (8 studies).Good consistency in direction of study results. Hypogonadal:62.4%/0.86 (CI, –0.22 to 1.95; I^2=61%) (5 studies) Borderline gonadal: 13.4%/0.70 (CI, –0.74 to 2.14) (1 study) Eugonadal: 24.3%/3.29 (CI, 2.11-4.48; I^2=0) (2 studies) *Mean difference in % change/absolute change (kg) from BL in body weight:* Overall: 100.0%/0.78 (CI, –0.28 to 1.84; I^2=68%) (8 studies). Some inconsistency in direction of results across studies. Hypogonadal:66.5%/0.18 (–1.07 to 1.43;I^2=66%) (6 studies) Borderline gonadal: 10.7%/1.10 (–0.91 to 3.11) (1 study) Eugonadal: 22.8%/2.32 (CI, 1.04-3.60; I^2=0) (2 studies)	According to pooled estimates, anabolic steroids may be effective in treating HIV-related wasting in men only when T levels are normal. Authors described the benefits as small and stated that the evidence was insufficient for recommending the treatment. Authors reported 7 of 13 studies had adequate allocation concealment; mean Jadad score, 4.3 (scale, 1-5). Funnel plot did not suggest publication bias. *Quality of SR:* Good
Metabolic Factors in Men with MetS or DM		
Cai et al. (2014)* MA of 5 RCTs (3 RCTs, vs placebo; 2 RCTs, vs no treatment) 351 patients) to evaluate effects of T therapy in men w/ LOH and type 2 DM (mean age 44-64 yrs; mean BL TT, 8.8-10.4 nmol/L; f/u 3-	*Mean difference, control minus T therapy:* FPG (mmol/L): –1.10 (CI, –1.88 to –0.31) (5 RCTs). Individual studies showed endpoint differences of 0.17-2.0 (1.8%-25% relative improvement); control means 6.0-9.18 at endpoint. FSI (mIU): –2.73 (CI, –3.62 to –1.84) (4 RCTs). Individual studies showed differences of 0.15-2.88 (0.8%-33.2% relative improvement); control means 5.79-18.85 . Direction of results	According to pooled estimates, T therapy in hypogonadal men w/ type 2 DM is effective in improving glucose control and triglyceride levels but not in reducing body fat or blood pressure. Study quality according to SR authors was 5 to 7 on a 0-8 scale.

Review/Included Studies/Patients	Findings	Summary and Comments
12 mos). LOH defined as TT <3.2 ng/mL (300 ng/dL; 11 nmol/L) *or* TT <3.2 ng/mL *and* FT <64 **pg/**mL (220 pmol/L *and* ≥3 sexual symptoms	inconsistent across studies. HbA1c (%): –0.87% (CI, –1.32% to –0.42%) (3 RCTs). Individual studies showed differences of 0.7%-1.3% (0.3%-13.1% relative improvement); control means 6.27%-9.9%. Direction of results inconsistent across studies. Triglycerides (mmol/L): –0.35 (CI, –0.62 to –0.07) (4 RCTs). Individual studies showed differences of 0.2 favoring controls (1 study) and 0.29-0.82 favoring T therapy (3 studies); control means 1.7-2.56 Low heterogeneity in each analysis, except for FPG (I^2=61%), which was explained by differences in treatment regimen. No effect on body fat or blood pressure (3 RCTs each).	*Quality of SR:* Fair (no assessment of publication bias; no explicit consideration of study quality in conclusions, no explicit use of ITT results).
Grossman et al. (2014) MA of 7 double-blind, placebo-controlled RCTs (833 patients) to evaluate effects of T therapy in men w/ type 2 DM (6 RCTs) or MetS (1 RCT) (mean age 44-64 yrs; mean BL TT, 8.6-10.1 nmol/L at BL; f/u 3-12 mos; mean % HbA1c, 5.7%-7.6% at BL, characterized by authors as well-controlled). Sx of hypogonadism not required but **participants in all but 2 studies had sx of hypogonadism.**	*Mean difference, control minus T therapy:* HOMA-IR: –0.26 (CI, –1.09-0.57) (7 RCTs; [3 included in MA by Cai et al.]; high heterogeneity, I^2 = 76%).). Significant results in trials using more conventional, less rigorous technique. SMD was –0.34 (CI, –0.51 to –0.16). Results consistent across studies. HbA1c (%): –0.15 (CI, –0.39 to 0.10) (6 RCTs [3 included in MA by Cai et al.]; high heterogeneity, I^2 = 77%). No difference in larger trials. BL values, 5.7-7.6%. Direction of results inconsistent across studies. Modest effect on lipids (MA not possible) and no effect on triglycerides or blood pressure.	T therapy in men w/ type 2 DM or MetS and T levels near the lower bound of typical normal reference ranges for T levels had no clear effect on insulin resistance or glucose control. Study quality (according to 25-item CONSORT reporting checklist for RCTs) was 14-24 across studies (1 point per item). No evidence of publication bias in funnel plot, but authors were aware of 1 unpublished trial reporting no effect on HbA1c. *Quality of SR:* Good
Sexual Function		
Corona et al. (2014a) MA of 41 RCTs (1930 men), T supplementation vs placebo or T supplementation as add-on to PDE5i vs placebo, in men w/ ED (ages 19-83 yrs; mean f/u 27 wks for T vs placebo and 12 wks for T as add-on)	T VS PLACEBO TRIALS *SMD in (>0 favors T therapy):* Overall ED Overall: 0.82 (CI, 0.47-1.17; I^2=87.73%) (24 studies) Hypogonadal men: 1.23 (CI, 0.72-1.72) (5 studies) Sexual-related ED Overall: 0.75 (CI, 0.37-1.12) (19 studies) Hypogonadal men: 1.26 (0.47-2.06) (# studies NR) Poor Libido Overall: 0.81 (CI, 0.47-1.17) (17 studies)	According to pooled estimates, T therapy was effective in improving sexual function in men w/ ED, especially those who were hypogonadal. High heterogeneity in linear regression analyses. Evidence of publication bias and a greater effect in industry-sponsored trials. No interaction w/ age (effect did not vary by age). Study quality generally good: placebo-controlled studies blinded as appropriate, no studies w/ serious

Review/Included Studies/Patients	Findings	Summary and Comments
	Hypogonadal: 1.00 (CI, 0.47-1.53) (# studies NR) Low Orgasm Score Overall: 0.68 (CI, 0.34-1.02) Good consistency in direction of results across trials. *Meta-regression:* Lower BL T level, hypogonadal status, and BL type 2 DM were each significantly associated w/ treatment effects of greater magnitude for ED. Trend toward significant relationship between hypogonadal status and greater effect on libido. Lower BL T level and hypogonadal status each associated w/ greater effect on orgasm. BL age and duration of treatment did not influence treatment effect. *Multivariate linear regression analysis* w/ adjustment for age confirmed BL T levels and diabetes as effect modifiers for ED and orgasm. *Safety:* No difference in incidence of CVD or PSA increase. T supplementation was associated w/ increase in hematocrit levels (SMD, 0.899; CI, 0.718-1.061) but not w/ increase in risk of pathological levels. T AS ADD-ON Positive effect in uncontrolled studies but not in placebo-controlled studies. (12 studies; I^2=96.41%)	dropout rates, ITT in most studies where appropriate. *Quality of SR:* Fair (lack of documented study quality or explicit consideration in conclusions)

*An earlier systematic review of 4 RCTs found that testosterone replacement therapy reduced FPG, HbA1c, fat mass, and triglycerides but not low-density lipoprotein (LDL), high-density lipoprotein (HDL), blood pressure, or body mass index (BMI) (Corona et al., 2011b).

Appendix VII. Analyses of Adverse Effects of Testosterone Therapy in Adult Men

Key: (A)MI, (acute) myocardial infarction; CAD, coronary artery disease; CV(D), cardiovascular (disease); f/u, follow-up; HR, hazard ratio; IPSS, International Prostate Symptom Score; MA, meta-analysis; MACE, major adverse cardiac event; MH, Mantel Haenszel; ng, nanogram; nmol, nanomole; OR, odds ratio; PSA, prostate-specific antigen; RCT, randomized controlled trial; RR, relative risk; subgrp, subgroup; T, testosterone; WMD, weighted mean difference

Newest Evidence	Findings from Newest Evidence	Findings Cited in Endocrine Society Guidelines	Summary
Cardiovascular Events			
Vigen et al. (2013) Retrospective cohort study (8709 men who had undergone coronary angiography; mean age 64 [T therapy] and 61 [no T therapy] yrs; f/u 2 yrs)	*Cumulative incidence of death, myocardial infarction or stroke (T therapy vs no T therapy):* 25.7% vs 19.9%; risk difference 5.8% (CI, –1.4% to 13.1%) *HR adjusted for presence of CAD (T therapy vs no T therapy):* 1.29 (CI, 1.04-1.58); no treatment-CAD interaction	---	Results suggested an increased risk of adverse cardiac events or mortality In men w/ previous coronary angiography, but results were statistically nonsignificant.
Corona et al. (2014b) MA (75 RCTs, 5464 men; mean age 60 yrs; mean duration of treatment 34 wks); in most studies, men were selected by age; 2 RCTs only enrolled men w/ heart failure	*MH-OR (T therapy vs placebo):* Any cardiovascular event: 1.07 (CI, 0.69-1.65) (31 studies) MACE, overall: 1.01 (CI, 0.57-1.77) (26 studies); considerable inconsistency across studies MACE, TT <12 nmol/L: 0.84 (CI, 0.32-2.23) (12 studies) MACE, mixed population: 1.26 (CI, 0.58-2.73) (14 studies) AMI: 0.68 (CI, 0.30-1.52) (14 studies) Acute coronary syndrome: 0.92 (CI, 0.43-1.97) (15 studies) Stroke: 0.82 (CI, 0.24-2.83) New heart failure: 1.64 (0.25-10.63) CV mortality: 1.14 (0.49-2.66) Metaregression analysis of the effect of T therapy on MACE by subgrp resulted in elevated but NS ORs in subgrps defined by older age and frailty, reduced risk in men with low T level (TT <12 nmol/L) (NS), and reduced risk in men w/ diabetes (MH-OR, 0.19; CI, 0.04-0.85). Overlapping CIs for low T and mixed populations	No clear effect (Calof et al., 2005)	(Comments apply both to Corona et al and Finkle et al.) Results suggested no effect on cardiovascular risk, but estimates were very imprecise and direction of findings for MACE were inconsistent across studies. There is conflicting evidence regarding differential safety according to age and new evidence suggesting greater risk in men w/ a history of CVD.

Newest Evidence	Findings from Newest Evidence	Findings Cited in Endocrine Society Guidelines	Summary
	suggested NS interaction between baseline T level and risk of MACE.		Corona et al. cautioned that the # of RCTs in the MetS subgrp was too small to allow a conclusion regarding the protective effect in men w/ type 2 DM.
Finkle et al. (2014) Retrospective cohort study of large healthcare database (55,593 men; f/u 90 days)	*RR of myocardial infarction in 90 days after initial T prescription vs 1 yr prior to prescription:* Age ≥65 yrs: 2.19 (CI, 1.27-3.77) Age <65 yrs and prior history of heart disease: 2.90 (CI, 1.49-5.62) Age <65 yrs and no history of heart disease: no effect		
Erythrocytosis			
Fernández-Balsells et al. (2010) MA (all study designs eligible; # men and mean age not reported by outcome; generally, age >60 yrs; f/u 3 mos to 3 yrs)	*RR (T therapy vs placebo or no intervention):* Erythrocytosis: 3.15 (CI, 1.56-6.35); low heterogeneity (11 studies) *WMD in change (favors placebo/control):* Hemoglobin: 0.80 g/dL (CI, 0.45-1.14); high heterogeneity (unknown # studies) Hematocrit: 3.18% (CI, 1.35-5.01); high heterogeneity (unknown # studies)	*OR for hematocrit >50%:* 3.69 (95% CI, 1.82 to 7.51) (Calof et al., 2005)	Results showed an increase in the risk of erythrocytosis. Most studies had short f/u and high loss to f/u.
Mortality			
Shores et al. (2012) Retrospective cohort study of men w/ low T levels (1031 men; age >40 yrs; mean age 63 yrs treated and 61 yrs untreated; TT <250 ng/dL (8.7 nmol/L); no history of prostate cancer; mean f/u 41 mos)	*Mortality (treated, untreated):* 10.3%, 20.7% (*P*<0.0001) *Cumulative mortality (treated, untreated):* 3.4 deaths/100 person-yrs vs 5.7 deaths/100 person-yrs *Adjusted HR:* 0.61 (CI, 0.42-0.88) (adjusted for age, body mass index, testosterone level, medical morbidity, diabetes, coronary heart disease)	No clear effect (Calof et al., 2005)	New evidence from a single study suggests a protective effect but is not conclusive.
Obstructive Sleep Apnea			
Hanafy (2007) Narrative review	Only small case series available. Inconsistent findings.	No clear effect on sleep apnea (Calof et al., 2005)	No conclusion possible.

Newest Evidence	Findings from Newest Evidence	Findings Cited in Endocrine Society Guidelines	Summary
Prostate Event			
Fernández-Balsells et al. (2010)* MA (all study designs eligible; # men and mean age not reported by outcome; generally, age >60 yrs; f/u 3 mos to 3 yrs)	*RR (T therapy vs placebo or no intervention):* Composite prostate outcome: 1.41 (CI, 0.93-2.14) (15 studies) Impaired urinary flow: 0.86 (CI, 0.13-5.53) (2 studies) PSA level >4 ng/mL: 1.22 (CI, 0.67-2.21) (9 studies) Significant increase in PSA: 1.56 (CI, 0.87-2.80) All estimates had low heterogeneity. No significant effect on prostatic cancer, the need for prostate biopsy, increase in PSA, IPSS lower urinary tract symptoms, or composite prostate outcome.	*OR for all prostate events:* 1.78 (CI, 1.07 to 2.95) (Calof et al., 2005)	Pooled estimates suggested an increased risk of prostate events, but estimates were nonsignificant and, in the more recent analysis, were imprecise. Most studies had short f/u and high loss to f/u.

*A more recent meta-analysis that calculated weighted mean differences based on a similar set of studies also reported no association with adverse prostate events (Cui and Zhang, 2013).

Appendix VIII. Summary of Practice Guidelines

Key: AD, androgen deficiency; ADT, androgen-deprivation therapy; BMD, bone mineral density; BMI, body mass index; COPD, chronic obstructive pulmonary disease; ED, erectile dysfunction; FSH, follicle-stimulating hormone; HF, heart failure; LH, luteinizing hormone; ng, nanogram; nmol, nanomole; OP, osteoporosis; OSA, obstructive sleep apnea; PSA, prostate-specific antigen; SHBG, sex hormone binding globulin; T, testosterone

Sponsor, Title	Relevant Recommendations (*Strength of Recommendation; Quality of Evidence According to Authors*)				Quality*/Main Limitations
	Screening/Testing	Diagnosis	Testosterone Therapy	Monitoring T Levels During Treatment	
American College of Physicians (ACP) (Qaseem et al., 2009) *Hormonal Testing and Pharmacologic Treatment of Erectile Dysfunction*	No recommendation for or against routine hormonal blood tests (testosterone or prolactin) for management of ED. *Insufficient evidence.*	---	No recommendation for or against hormonal treatment for management of ED. *Insufficient evidence.*	---	6 – Good
American Diabetes Association (ADA) (2014) *Standards of Medical Care in Diabetes—2014* (VII. Assessment of Common Comorbid Conditions)	No specific recommendation, but document states that "obesity is a major confounder" (p. S49) and cites Endocrine Society guidelines and their recommendations to test only in the presence of symptoms.		No specific recommendation, but document states that "evidence for effects of testosterone replacement on outcomes is mixed" (p. S49) and cites Endocrine Society guidelines regarding treatment.		
American Urological Association (AUA) (2010b) *The Optimal Evaluation of the Infertile Male: AUA Best Practice Statement*	Endocrine evaluation if (1) abnormal semen analysis, (2) impaired sexual function, *or* (3) other clinical findings (not specified) suggestive of a specific endocrinopathy. Minimum initial hormonal	---	---	---	6 – Good NOTE: A best practice statement based on expert opinion was issued because the literature search did not identify sufficient

Sponsor, Title	Relevant Recommendations (*Strength of Recommendation; Quality of Evidence According to Authors*)				Quality*/Main Limitations
	Screening/Testing	Diagnosis	Testosterone Therapy	Monitoring T Levels During Treatment	
	evaluation should include T and FSH. All recommendations based on expert opinion due to *insufficient evidence*.				evidence.
American Urological Association (AUA) (2010a) *The Evaluation of the Azoospermic Male: AUA Best Practice Statement*	Minimum initial hormonal evaluation should include T and FSH. All recommendations based on expert opinion due to *insufficient evidence*.	---	---	---	6 – Good NOTE: A best practice statement based on expert opinion was issued because the literature search did not identify sufficient evidence.
The Endocrine Society (Bhasin et al., 2010) *Testosterone Therapy in Adult Men with Androgen Deficiency Syndromes*	<u>Screening</u> Do not screen in general population. *Strong; very low.* Consider case detection by total T measurement in conditions in which there is a high prevalence of low T levels.[3] *Weak; very low.*	Diagnose AD only w/ consistent symptoms and signs and unequivocally low serum T levels. *Strong; very low.* Do not diagnose AD during acute or subacute illness. *Weak; low.*	<u>Recommended for</u> Symptomatic men w/ classical AD syndromes[6] for purposes of including and maintaining secondary sex characteristics and improving sexual function, sense of well-being, and BMD. *Strong; low.*	3-6 mos after initiation of T therapy to assess whether T levels have reached the normal range. *Weak recommendation; very low.*	5 (fair) (search details and study selection criteria not provided).

[3]According to Table 3 in the guidelines: Conditions in which T measurement should be based on characteristic symptoms include type 2 diabetes, end-stage renal disease, and moderate to severe COPD; symptoms such as sexual dysfunction, unexplained weight loss, weakness, or mobility limitation may indicate the need for testing. These symptoms were identified on the basis of panelists' experience rather than population surveys. Conditions in which measurement of T levels may be indicated regardless of symptoms include mass in, radiation of, or disease of sellar region (a depression in the upper surface of the sphenoid

Sponsor, Title	Relevant Recommendations (*Strength of Recommendation; Quality of Evidence According to Authors*)				Quality*/Main Limitations
	Screening/Testing	Diagnosis	Testosterone Therapy	Monitoring T Levels During Treatment	
	Testing to diagnose Test patients w/ more specific signs and symptoms suggestive of AD.[4] Consider testing patients w/ less specific signs and symptoms.[5] Test morning total T. All testing recommendations: *Weak; very low.* Exclude reversible illness, drugs that can deplete T levels, nutritional deficiency (not rated) Confirmatory testing to		Suggested as an option for Low T levels and low libido. *Weak; very low.* Low T levels and ED after evaluation of underlying causes of ED and consideration of established therapies for ED. *Weak; very low.* Case-by-case for older men w/ low T levels on >1 occasion and clinically significant symptoms of AD after discussion of uncertainty about risks and benefits. *Weak; very low.*		

bone in which the pituitary gland sits); medications that affect T production or metabolism (e.g., glucocorticoids and opioids); HIV-associated weight loss; or osteoporosis or low-trauma fracture (especially in a young man).

[6] See More specific signs and symptoms in Footnote 2.

[4] According to Table 1 in the guidelines: A. More specific signs and symptoms (incomplete or delayed sexual development; eunuchoidism; reduced sexual desire (libido) and activity; decreased spontaneous erections; breast discomfort, gynecomastia; loss of body [axillary and pubic] hair, reduced shaving; very small [especially <5 mL] or shrinking testes; inability to father children, low or zero sperm count; height loss, low-trauma fracture, low BMD; hot flushes, sweats.

[5] According to Table 1 in the guidelines: B. Other less specific signs and symptoms (decreased energy, motivation, initiative, and self-confidence; feeling sad or blue, depressed mood, dysthymia; poor concentration and memory; sleep disturbance, increased sleepiness; mild anemia (normochromic, normocytic, in the female range); reduced muscle bulk and strength; increased body fat, BMI; diminished physical or work performance).

Sponsor, Title	Relevant Recommendations (*Strength of Recommendation; Quality of Evidence According to Authors*)				Quality*/Main Limitations
	Screening/Testing	Diagnosis	Testosterone Therapy	Monitoring T Levels During Treatment	
	<u>establish androgen deficiency</u> Repeat measurement of total T. *Strong; low quality.* Measure free or bioavailable serum T in some men w/ total serum T near lower limit of normal and in whom alterations of SHBG are suspected. *Weak; low.* <u>Additional testing to establish etiology</u> LH+FSH		HIV-infected men w/ low T levels and weight loss to promote weight maintenance and gains in lean body mass and muscle strength. *Weak; low.* Low T levels and high doses of glucocorticoids to promote preservation of lean body mass and BMD. *Weak; very low.* <u>Not recommended</u> As a general policy for all older men w/ low T levels. *Strong; very low.* <u>Contraindicated</u> Men w/ history of breast cancer (*weak; very low) or* prostate cancer *(weak; low*). In presence of palpable prostate nodule or induration or PSA 4 ng/mL, or PSA 3 ng/mL and high risk of prostate cancer. *Strong; very low.* Hematocrit >50%, untreated severe OSA, severe lower urinary tract symptoms, uncontrolled/poorly		

Sponsor, Title	Relevant Recommendations (*Strength of Recommendation; Quality of Evidence According to Authors*)				Quality*/Main Limitations
	Screening/Testing	Diagnosis	Testosterone Therapy	Monitoring T Levels During Treatment	
			controlled HF. *Strong; very low.* To improve fertility. *Strong; very low.*		
The Endocrine Society (Watts et al., 2012) *Osteoporosis in Men*	Test for TT in men being evaluated for OP or considered for pharmacological treatment w/ bone-active agents. *Weak; low.* In men w/ a history or physical examination suggesting a specific cause of OP, conduct further testing, e.g., calculated FT or BT. *Weak; low.*		Offer in lieu of "bone drug" for men at borderline high risk for fracture who have serum T levels <200 ng/dL (6.9 nmol/L) on >1 determination, if accompanied by signs or symptoms of AD or "organic" hypogonadism (e.g., due to hypothalamic, pituitary, or specific testicular disorder). *Weak; low.* Consider for men at high risk for fracture w/ T levels <200 ng/dL (6.9 nmol/L) who lack standard indications for T therapy but who have contraindications to approved pharmacological agents for OP. *Weak; low.*		5 (fair) (search details and study selection criteria not provided).

*According to the Rigor of Development domain of the Appraisal of Guidelines Research and Evaluation (AGREE) tool, along with a consideration of commercial funding and conflicts of interest among the guideline authors. Guidelines were scored on scale of 1 to 7 and judged to be good (6-7), fair (4-5), or poor (1-3).