Efficacy and Safety of Testosterone Treatment in Men: An Evidence Report for a Clinical Practice Guideline by the American College of Physicians

Susan J. Diem, MD, MPH; Nancy L. Greer, PhD; Roderick MacDonald, MS; Lauren G. McKenzie, MPH; Philipp Dahm, MD, MHSc; Nacide Ercan-Fang, MD; Allison Estrada, MD; Laura S. Hemmy, PhD; Christina E. Rosebush, MPH; Howard A. Fink, MD, MPH; and Timothy J. Wilt, MD, MPH

Background: Testosterone treatment rates in adult men have increased in the United States over the past 2 decades.

Purpose: To assess the benefits and harms of testosterone treatment for men without underlying organic causes of hypogonadism.

Data Sources: English-language searches of multiple electronic databases (January 1980 to May 2019) and reference lists from systematic reviews.

Study Selection: 38 randomized controlled trials (RCTs) of at least 6 months' duration that evaluated transdermal or intramuscular testosterone therapies versus placebo or no treatment and reported prespecified patient-centered outcomes, as well as 20 long-term observational studies, U.S. Food and Drug Administration review data, and product labels that reported harms information.

Data Extraction: Data extraction by a single investigator was confirmed by a second, 2 investigators assessed risk of bias, and evidence certainty was determined by consensus.

Data Synthesis: Studies enrolled mostly older men who varied in age, symptoms, and testosterone eligibility criteria. Testosterone therapy improved sexual functioning and quality of life in men with low testosterone levels, although effect sizes were small (low- to moderate-certainty evidence). Testosterone therapy had little to no effect on physical functioning, depressive symptoms, energy and vitality, or cognition. Harms evidence reported in trials was judged to be insufficient or of low certainty for most harm outcomes. No trials were powered to assess cardiovascular events or prostate cancer, and trials often excluded men at increased risk for these conditions. Observational studies were limited by confounding by indication and contraindication.

Limitation: Few trials exceeded a 1-year duration, minimum important outcome differences were often not established or reported, RCTs were not powered to assess important harms, few data were available in men aged 18 to 50 years, definitions of low testosterone varied, and study entry criteria varied.

Conclusion: In older men with low testosterone levels without well-established medical conditions known to cause hypogonadism, testosterone therapy may provide small improvements in sexual functioning and quality of life but little to no benefit for other common symptoms of aging. Long-term efficacy and safety are unknown.

Primary Funding Source: American College of Physicians. (PROSPERO: CRD42018096585)

Testosterone treatment is approved by the U.S. Food and Drug Administration (FDA) for replacement therapy for men with primary or secondary hypogonadism caused by disorders of the hypothalamus, pituitary gland, or testes, often classified as organic or classical hypogonadism (1). Testosterone treatment in these conditions is considered standard care for the development or maintenance of secondary sexual characteristics.

Testosterone use in the United States, which has tripled in recent years, exceeds that in other countries (2–5). Much of the increase is in men with nonspecific symptoms, such as decreased energy, sexual function, and mobility, who have serum testosterone concentrations below the normal range or in the low-normal range for healthy young men for no apparent reason other than older age or comorbid conditions, such as obesity. Substantial proportions of U.S. men who receive testosterone therapy do not have testosterone levels tested before initiation of therapy (3, 6). The level of baseline testosterone that prompts initiation of such therapy varies widely: In 1 study (3), approximately 20% of men who had their testosterone level measured before initiating therapy had a level above 10.41 nmol/L (300 ng/dL).

We evaluated the efficacy and harms of testosterone treatment in men without established conditions that cause permanent testicular or hypothalamic–pituitary dysfunction (for example, Klinefelter syndrome, orchitis, testicular trauma or radiation, or hypothalamic or pituitary tumors). This systematic review, conducted by the Evidence Synthesis Core at the Minneapolis VA Center for Care Delivery and Outcomes Research, served as the evidence base for a clinical practice guideline from the American College of Physicians. We did not address the appropriate diagnosis

---

See also:

Web-Only
Supplement
CME/MOC activity

and evaluation of hypogonadism. Guidelines by the Endocrine Society (7) and American Urological Association (8) recommend that clinicians measure fasting morning concentrations of total testosterone on 2 occasions to diagnose hypogonadism but vary in their recommendations regarding the threshold at which to categorize a total testosterone level as low.

METHODS

Our protocol was registered in PROSPERO (CRD42018096585).

Data Sources and Searches

We searched MEDLINE, Embase, and the Cochrane Library for peer-reviewed randomized controlled trials (RCTs), cohort studies, and case–control studies published in English and indexed from January 1980 to May 2019. We also searched reference lists from relevant systematic reviews. Search terms included MeSH (Medical Subject Headings) terms and keywords pertaining to testosterone replacement, deficiency, and treatment (Appendix Table 1, available at Annals.org).

Study Selection

Two investigators independently reviewed each study's abstract and full text to determine eligibility. Conflicts were resolved through discussion, with consultation of a third member if necessary.

We included RCTs that assessed transdermal or intramuscular formulations of testosterone in adult men with at least 6 months of active treatment. Studies were included if they had a control group with placebo or no testosterone treatment, reported our efficacy outcomes of interest (sexual function, physical function, quality of life, mood [depression], fractures, energy or vitality, and cognition), and provided measurements of baseline total testosterone. We included 1 trial of 24 weeks' duration, TOM (Testosterone in Older Men with Mobility Limitations) (9, 10), because of its significance with respect to cardiovascular outcomes. This trial was stopped early because of concern regarding excess cardiovascular adverse events in the testosterone group. To assess serious but infrequent harms, we included observational studies with at least 1 year of follow-up and 500 participants. We excluded studies limited to men with inherited or acquired conditions known to cause permanent hypothalamic, pituitary, or testicular dysfunction and studies in which endogenous testosterone was artificially suppressed. We excluded studies evaluating oral testosterone because its use is contraindicated in men without structural or genetic causes of hypogonadism.

Outcome Classification

We classified the reported scales, domains, and questions into patient-centered constructs to aid in clinical interpretation. For sexual function, we focused on self-reported overall sexual function and erectile function. For physical function, we focused on subjective self-reported physical function as assessed by established instruments and objective physical function as measured by gait speed because of its well-established association with important health outcomes (11, 12). For quality of life, energy or vitality, and depression, we limited outcomes to those assessed using established instruments. Appendix Table 2 (available at Annals.org) summarizes the instruments most commonly used to measure outcomes in the included RCTs. For cognition, we focused on measures of overall cognitive function and 6 established domains (attention, language, verbal memory, visuospatial memory, visuospatial function, and executive function) (Appendix Table 3, available at Annals.org).

Our primary harm outcomes were serious adverse events; a composite of adverse cardiovascular events, defined as cardiovascular death, nonfatal myocardial infarction, acute coronary syndrome, nonfatal stroke, revascularization, or heart failure exacerbation; prostate cancer; and mortality. We also assessed deep venous thrombosis or pulmonary embolism, worsening lower urinary tract symptoms, worsening sleep apnea, and withdrawals due to adverse events. We did not include intermediate measures (such as body composition, metabolic variables, hemoglobin levels, blood pressure, and prostate-specific antigen levels).

Data Extraction and Quality Assessment

Data extraction was completed by 1 reviewer and verified by a second. We extracted baseline characteristics of each study and population, including study design, age, race, comorbid conditions, and testosterone levels. Two investigators assessed risk of bias of RCTs using a modified Cochrane approach based on the following elements: sequence generation, allocation concealment, blinding, incomplete outcome data (attrition), and selective reporting (13). A high-risk study would have critical flaws in these elements or very high attrition (≥30%). Two investigators assessed risk of bias of observational studies; using a modification of the Agency for Healthcare Research and Quality approach, they reviewed studies for elements of selection, detection, attrition, and reporting bias (14).

Data Synthesis and Analysis

We organized tables by study design and testosterone formulation and sorted trials by risk-of-bias assessment. We pooled results from trials deemed to have low or medium risk of bias if populations and outcome measures were clinically comparable. Data were analyzed in R (R Foundation) (15, 16). Data for continuous efficacy outcomes were pooled using the Hartung–Knapp–Sidik–Jonkman method for random-effects models to calculate standardized mean differences (SMDs) with corresponding 95% CIs (17). However, this method for meta-analysis has been shown to underestimate uncertainty compared with the fixed-effects model when the number of trials is small, particularly if fewer than 5, and when no between-study variance exists ($\tau^2 = 0$) (17). In such cases, we meta-analyzed results with the fixed-effects model. We interpreted SMDs on the basis of Cohen's definition of small (0.2), medium (0.5), and large (0.8) effects (18). Scale scores with different directionality were adjusted to ensure that all scales pointed

in the same direction when pooled. Categorical harm outcomes data were pooled using the Peto odds ratio (OR) method. Harms data were stratified by duration of follow-up in intervals of less than 12 months, 12 to less than 24 months, and 24 months or more. Absolute event rates and 95% CIs for the primary harm outcomes were pooled for each study group using the Freeman–Tukey double arcsine transformation (19). For all pooled analyses, we assessed the magnitude of statistical heterogeneity with the I^2 statistic ($I^2 > 75\%$ may indicate substantial heterogeneity) (20). Results were also stratified and assessed by baseline testosterone level and formulation. In sensitivity analyses, we included trials that were rated as having high risk of bias or inadequate blinding.

Assessment of evidence certainty for our primary efficacy and harm outcomes was based on methods developed by the GRADE (Grading of Recommendations Assessment, Development and Evaluation) Working Group (21, 22). Two trained research associates graded certainty of evidence for each outcome as high, moderate, low, or very low by evaluating 4 critical domains (risk of bias, consistency, directness, and precision). Discrepancies in ratings of risk of bias and certainty of evidence were resolved by discussion, and a final determination was made through a consensus that included the principal investigator.

Role of the Funding Source

This review was funded by a contract with the American College of Physicians. The American College of Physicians Clinical Guidelines Committee assisted in the development of key questions, study inclusion criteria, and outcome measures of interest but was not involved in data collection, analysis, or manuscript preparation.

RESULTS

We identified 38 RCTs (in 62 articles [9, 10, 23–54, 55–82]) and 20 observational studies (in 21 articles [83–103]) that met inclusion criteria (**Figure 1**). Baseline characteristics are reported in **Supplement Table 1** (RCTs) and **Supplement Table 2** (observational studies)

Figure 1. Evidence search and selection.
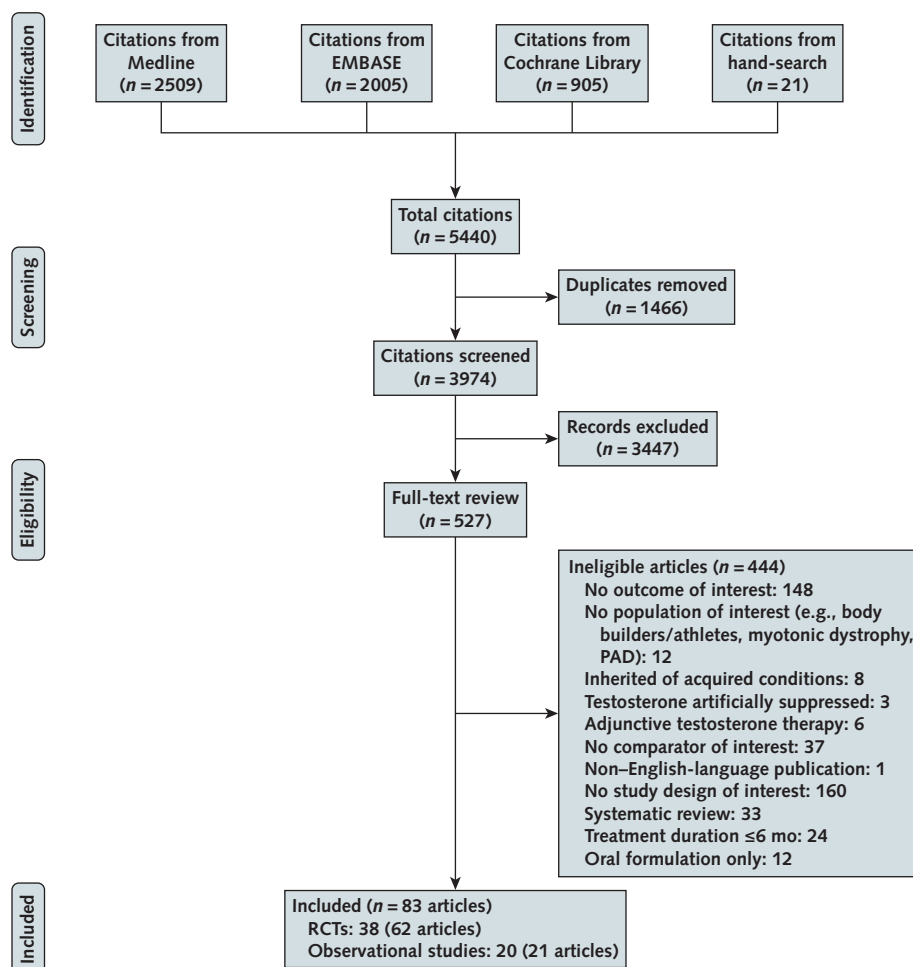
PAD = peripheral artery disease; RCT = randomized controlled trial.

Table. Certainty of Evidence: Testosterone Therapy in Men

Outcome	Participants, *n*	RCTs, *n*	Absolute Effect (95% CI)	
			Testosterone	**Placebo**
Global sexual function (IIEF or sexual function subscale of AMS)	1140	7	–	–
Erectile function (IIEF-5 or IIEF-ED)	1299	7	–	–
Physical function: self-reported measures (SF-36-PF or PASE)	1029	5	–	–
Physical function: objective measures (e.g., 6-min walk)	1063	7	–	–
Quality of life (AMS)	1043	7	7.0 points	3.6 points
ACE	2415	14	2.3% (0.9% to 4.1%)§	1.5% (0.8% to 2.5%)§
Serious adverse events	2268	8	13.2% (3.6% to 27.4%)§	12.8% (4.7% to 23.9%)§
Prostate cancer (follow-up range, 6 mo–3 y)	2143	10	0.3% (0.0% to 1.2%)§	0.4% (0.0% to 1.9%)§
Mortality (follow-up range, 6 mo–3 y)	2727	12	0.4% (0.07% to 1.0%)§	1.5% (0.5% to 2.9%)§

ACE = adverse cardiovascular events; AMS = Aging Males' Symptoms; IIEF = International Index of Erectile Function; IIEF-5 = 5-item version of IIEF; IIEF-ED = IIEF erectile dysfunction domain; OR = odds ratio; PASE = Physical Activity Scale for the Elderly; RCT = randomized controlled trial; SF-36-PF = Short Form-36 Health Survey physical function subscale; SMD = standardized mean difference.
* Where certainty is high, we are very confident that the true effect lies close to the estimate of the effect. Where certainty is moderate, we are moderately confident in the effect estimate: The true effect is likely to be close to the estimate of the effect, but there is a possibility that it is substantially different. Where certainty is low, our confidence in the effect estimate is limited: The true effect may be substantially different from the estimate of the effect. Where certainty is very low, we have very little confidence in the effect estimate: The true effect is likely to be substantially different from the estimate of the effect.
† Downgraded for study limitations, particularly very high attrition.
‡ Downgraded for imprecision (wide CIs).
§ Absolute event rates with random-effects models and Freeman–Tukey double arcsine variance-stabilizing transformation.
‖ CIs were very wide, and there were very few events; trials were not powered to detect group differences for this outcome. Ascertainment of prostate cancer also varied across the trials.

(available at Annals.org). We evaluated efficacy using results from the RCTs and safety using results from the RCTs and observational studies.

RCTs
Overview

Sample sizes ranged from 10 (78) to 790 (32, 66, 71, 74, 82), and 15 trials enrolled at least 100 men. Most trials had less than 12 months of follow-up (*k* = 24); 3 studies reported follow-up to 36 months (25, 27, 49, 70, 72, 73, 81). Sixteen trials were from the United States, 14 were from Europe, and 8 were from Australia or Asia. Twenty-four studies reported at least partial industry sponsorship.

Trials varied in their inclusion criteria. Most required a testosterone level below a stated threshold (*k* = 34); 29 based enrollment on total testosterone level, 2 on free testosterone level, and 3 on bioavailable testosterone, with variability in the method used to measure testosterone. Seven trials enrolled participants on the basis of a total testosterone level of at most 10.41 nmol/L (300 ng/dL), the threshold recommended by the American Urological Association to categorize a man as hypogonadal. Trials also varied in required symptoms for inclusion: Some required the presence of specific symptoms attributed to hypogonadism, whereas others did not. Two specifically required sexual symptoms (32, 60, 66, 71, 74, 82), 4 required physical or mobility limitations (9, 10, 32, 53, 65, 66, 71, 74, 75, 82), and others either required no specific symp-

toms or required the presence of at least 1 symptom attributed to hypogonadism.

Mean baseline total testosterone levels were 10.41 nmol/L (300 ng/dL) or less in 20 studies and less than 9.54 nmol/L (275 ng/dL) in 11 trials (26, 30, 32, 41–46, 48, 50, 51, 58, 60, 63, 64, 66, 71, 74, 77, 79, 80, 82). Two reported only free testosterone levels (55, 67). Baseline testosterone level was greater than 13.88 nmol/L (400 ng/dL) in 5 studies (23, 24, 34, 40, 57, 76), 3 of which were restricted to men with specific underlying medical conditions and 1 of which was rated as having high risk of bias. Only 13 trials required 2 fasting morning testosterone levels, and of these only 2 (29, 30, 32, 66, 71, 74, 82) required 2 morning specimens with a testosterone level of 10.41 nmol/L (300 ng/dL) or less.

The mean age across trials was 66 years. Eight studies restricted age to 65 years or older, and none were limited to (or predominantly enrolled) men younger than 50 years. Only 5 trials had an age threshold as low as 18 years. Race/ethnicity, comorbid conditions, and functional measures were infrequently reported (**Supplement Table 1**).

Doses and formulations of testosterone varied: 19 trials used a transdermal formulation, whereas 19 used intramuscular injections. Nine adjusted dosing to achieve a targeted testosterone level (9, 10, 27, 32, 33, 35, 36, 42–47, 49, 66, 70–75, 82), and 29 used a fixed dose.

Table–Continued

SMD or Relative Effect (95% CI)	Certainty of Evidence*	What Happens
SMD, 0.35 SD higher (0.23 SD higher to 0.46 SD higher)	Moderate†	Testosterone probably improves global sexual function by a small amount on the basis of IIEF or the sexual function subscale of the AMS over a mean duration of 53 wk.
SMD, 0.27 SD higher (0.09 SD higher to 0.44 SD higher)	Low†‡	Testosterone may improve erectile function by a small amount on the basis of IIEF-5 or IIEF-ED over a mean duration of 58 wk.
SMD, 0.15 SD higher (0.19 SD lower to 0.50 SD higher)	Low†‡	Testosterone may make little to no difference in self-reported measures of physical function (on the basis of SF-36-PF or PASE) over a mean duration of 78 wk.
SMD, 0.14 SD higher (0.02 SD higher to 0.27 SD higher)	Low†	Testosterone may make little to no difference in objective measures of physical function (e.g., 6-min walk) over a mean duration of 60 wk.
SMD, 0.33 SD lower (0.50 SD lower to 0.16 SD lower)	Low†‡	Testosterone may improve quality of life by a small amount on the basis of the AMS scale over a mean duration of 35 wk.
Peto OR, 1.22 (0.66 to 2.23)	Low†‡	Testosterone may increase ACE by a small amount or make no difference in ACE over a mean duration of 68 wk.
Peto OR, 0.94 (0.73 to 1.21)	Moderate†	Testosterone probably makes little to no difference in serious adverse events over a mean duration of 56 wk.
Peto OR, 0.97 (0.35 to 2.69)	Very low†‖	–
Peto OR, 0.47 (0.25 to 0.89)	Low†‡	Testosterone may slightly reduce mortality on the basis of very few events, but the trials were not powered to detect group differences for this outcome.

Risk of bias was high in 11 trials, medium in 19, and low in 8 (**Supplement Table 1**). Many trials categorized as having high risk of bias had very high attrition rates (range, 30% to 56%). Two were open-label trials with no control group (55, 67), and another used a placebo gel while the active intervention was administered intramuscularly (26).

Ten RCTs were limited to men with specific underlying conditions: heart failure (56, 62); chronic stable angina (59); chronic kidney disease requiring dialysis (31); cirrhosis (69); chronic obstructive pulmonary disease (76); respiratory, immune, or inflammatory disease requiring long-term glucocorticoid therapy (34); Alzheimer disease (78); mild cognitive impairment (33); and low circulating levels of insulin-like growth factor (40).

Efficacy Outcomes

We report overall findings for each outcome, excluding the 10 studies in special populations as well as those trials judged to have high risk of bias or inadequate blinding. We describe the results for analyses that pool data from trials of both intramuscular and transdermal preparations of testosterone. **Supplement Table 3** (available at Annals.org) shows results by formulation and for trials with a mean baseline testosterone level less than 10.41 nmol/L (300 ng/dL).

The **Table** summarizes certainty of evidence for the primary efficacy outcomes; **Supplement Table 4** (available at Annals.org) shows these assessments by testosterone formulation.
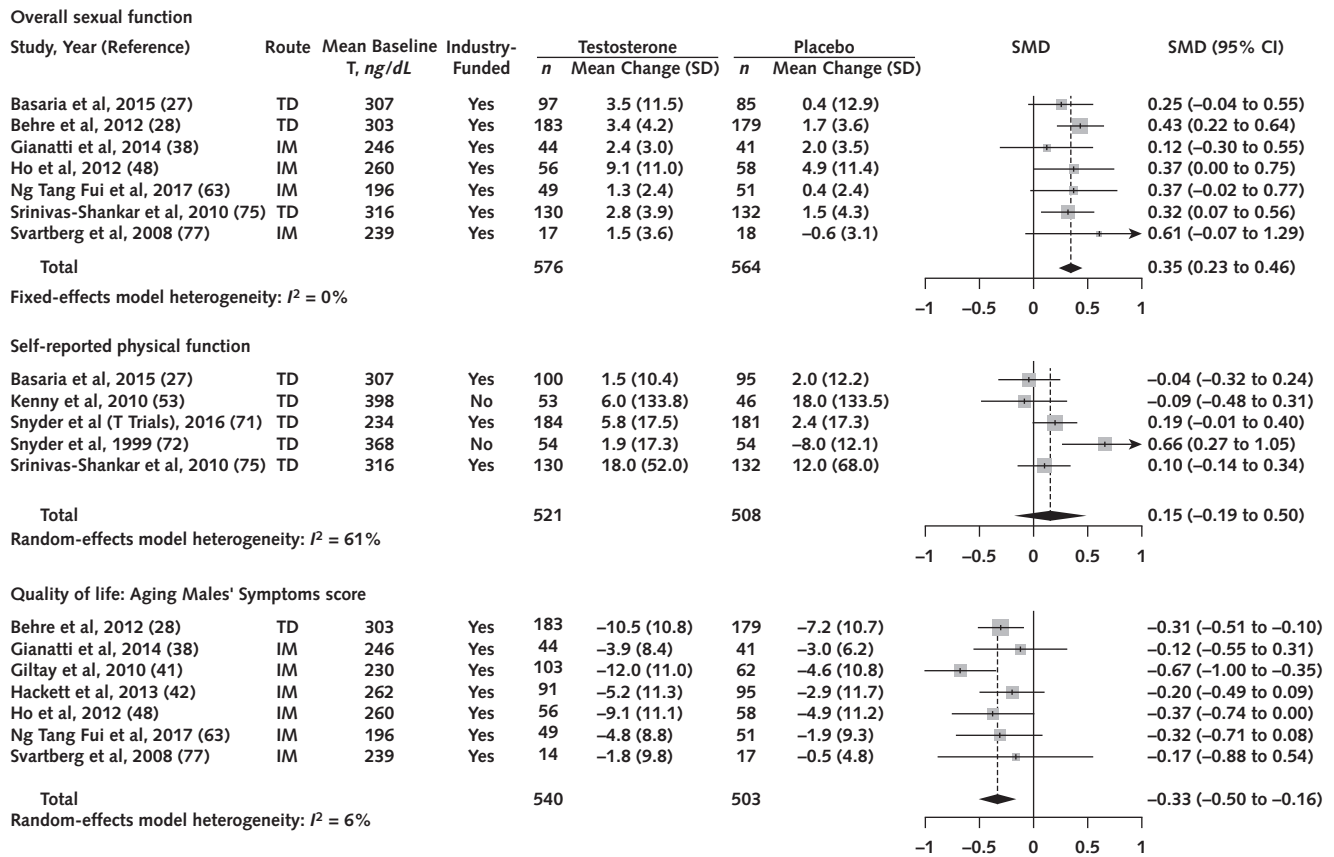
The effect of testosterone did not differ significantly by testosterone formulation (intramuscular vs. transdermal), although some outcomes had few or no studies available for these indirect comparisons. When analyses were restricted to trials with a mean baseline testosterone level less than 10.41 nmol/L (300 ng/dL), results were similar to those of the primary analyses. Data for efficacy were limited in men with higher baseline testosterone levels, and only 5 trials in the primary

analyses had mean baseline levels of at least 10.41 nmol/L (300 ng/dL); of these, 3 had mean baseline levels less than 11.10 nmol/L (320 ng/dL). Results did not differ in analyses that included trials judged to have high risk of bias or inadequate blinding (**Supplement Table 3**).

Sexual Function. **Supplement Table 5** (available at Annals.org) summarizes sexual function outcomes. Studies varied in required level of total or free testosterone and baseline symptoms: Some required low libido or other sexual symptoms, some required a range of other symptoms attributable to hypogonadism, and some did not require any symptoms. In studies that reported baseline sexual functioning, participants on average did report clinically significant sexual dysfunction at baseline, as assessed by various measures (**Supplement Table 1**). The most commonly reported sexual function outcome was score on the sexual function subscale of the Aging Males' Symptoms (AMS) scale. For erectile function, the 5-item International Index of Erectile Function questionnaire and its erectile function domain were most commonly used.

Overall or Global Sexual Function. We pooled 7 RCTs (*n* = 1140) that evaluated the effect of testosterone treatment on global measures of sexual function (**Figure 2** [*top*] and **Supplement Table 5**). Compared with placebo, testosterone treatment was associated with an overall small improvement in global sexual function (SMD, 0.35 [95% CI, 0.23 to 0.46]; I^2 = 0%; moderate-certainty evidence). In the Testosterone Trials, the primary sexual function outcome–sexual activity, as assessed by the Psychosexual Daily Questionnaire question 4–increased more with testosterone treatment than with placebo (effect size, 0.45 [CI, 0.30 to 0.60]) (71). Sexual desire, as measured by the sexual desire score of the Derogatis Interview for Sexual Functioning in Men–II, also increased more in men treated with testosterone (effect size, 0.44 [CI, 0.32 to 0.56]). The effect sizes for these outcomes in the Testosterone

Figure 2. Primary efficacy outcomes for testosterone treatment vs. placebo (SMDs in mean change from baseline).

Overall sexual function

Study, Year (Reference)	Route	Mean Baseline T, ng/dL	Industry-Funded	Testosterone n	Mean Change (SD)	Placebo n	Mean Change (SD)	SMD	SMD (95% CI)
Basaria et al, 2015 (27)	TD	307	Yes	97	3.5 (11.5)	85	0.4 (12.9)		0.25 (−0.04 to 0.55)
Behre et al, 2012 (28)	TD	303	Yes	183	3.4 (4.2)	179	1.7 (3.6)		0.43 (0.22 to 0.64)
Gianatti et al, 2014 (38)	IM	246	Yes	44	2.4 (3.0)	41	2.0 (3.5)		0.12 (−0.30 to 0.55)
Ho et al, 2012 (48)	IM	260	Yes	56	9.1 (11.0)	58	4.9 (11.4)		0.37 (0.00 to 0.75)
Ng Tang Fui et al, 2017 (63)	IM	196	Yes	49	1.3 (2.4)	51	0.4 (2.4)		0.37 (−0.02 to 0.77)
Srinivas-Shankar et al, 2010 (75)	TD	316	Yes	130	2.8 (3.9)	132	1.5 (4.3)		0.32 (0.07 to 0.56)
Svartberg et al, 2008 (77)	IM	239	Yes	17	1.5 (3.6)	18	−0.6 (3.1)		0.61 (−0.07 to 1.29)
Total				576		564			0.35 (0.23 to 0.46)

Fixed-effects model heterogeneity: $I^2 = 0\%$

Self-reported physical function

Study, Year (Reference)	Route	Mean Baseline T	Industry-Funded	Testosterone n	Mean Change (SD)	Placebo n	Mean Change (SD)	SMD (95% CI)
Basaria et al, 2015 (27)	TD	307	Yes	100	1.5 (10.4)	95	2.0 (12.2)	−0.04 (−0.32 to 0.24)
Kenny et al, 2010 (53)	TD	398	No	53	6.0 (133.8)	46	18.0 (133.5)	−0.09 (−0.48 to 0.31)
Snyder et al (T Trials), 2016 (71)	TD	234	Yes	184	5.8 (17.5)	181	2.4 (17.3)	0.19 (−0.01 to 0.40)
Snyder et al, 1999 (72)	TD	368	No	54	1.9 (17.3)	54	−8.0 (12.1)	0.66 (0.27 to 1.05)
Srinivas-Shankar et al, 2010 (75)	TD	316	Yes	130	18.0 (52.0)	132	12.0 (68.0)	0.10 (−0.14 to 0.34)
Total				521		508		0.15 (−0.19 to 0.50)

Random-effects model heterogeneity: $I^2 = 61\%$

Quality of life: Aging Males' Symptoms score

Study, Year (Reference)	Route	Mean Baseline T	Industry-Funded	Testosterone n	Mean Change (SD)	Placebo n	Mean Change (SD)	SMD (95% CI)
Behre et al, 2012 (28)	TD	303	Yes	183	−10.5 (10.8)	179	−7.2 (10.7)	−0.31 (−0.51 to −0.10)
Gianatti et al, 2014 (38)	IM	246	Yes	44	−3.9 (8.4)	41	−3.0 (6.2)	−0.12 (−0.55 to 0.31)
Giltay et al, 2010 (41)	IM	230	Yes	103	−12.0 (11.0)	62	−4.6 (10.8)	−0.67 (−1.00 to −0.35)
Hackett et al, 2013 (42)	IM	262	Yes	91	−5.2 (11.3)	95	−2.9 (11.7)	−0.20 (−0.49 to 0.09)
Ho et al, 2012 (48)	IM	260	Yes	56	−9.1 (11.1)	58	−4.9 (11.2)	−0.37 (−0.74 to 0.00)
Ng Tang Fui et al, 2017 (63)	IM	196	Yes	49	−4.8 (8.8)	51	−1.9 (9.3)	−0.32 (−0.71 to 0.08)
Svartberg et al, 2008 (77)	IM	239	Yes	14	−1.8 (9.8)	17	−0.5 (4.8)	−0.17 (−0.88 to 0.54)
Total				540		503		−0.33 (−0.50 to −0.16)

Random-effects model heterogeneity: $I^2 = 6\%$

IM = intramuscular; SMD = standardized mean difference; T = testosterone level; TD = transdermal.

Trials were consistent with a clinically meaningful improvement in sexual desire and activity for men treated with testosterone (82).

Erectile Function. Testosterone treatment improved erectile function compared with placebo, although the pooled effect was small (7 trials; n = 1299; SMD, 0.27 [CI, 0.09 to 0.44]; $I^2 = 13\%$; low-certainty evidence) (**Table** and **Supplement Figure 1** [available at Annals.org]).

Physical Function. Study populations included men with mobility limitations (9, 10, 71) or frailty (75), but most trials did not specify physical function limitations as entry criteria. Self-reported measures of physical function included the Short Form-36 Health Survey physical function subscale and the Physical Activity Scale for the Elderly. Gait speed was most commonly measured by the 6-minute walk test.

Subjective or Self-Reported Physical Function. Five RCTs (n = 1029) provided self-reported measures of physical function (**Figure 2** [*middle*] and **Supplement Table 6** [available at Annals.org]) that could be pooled. Testosterone treatment did not improve subjective physical function compared with placebo (SMD, 0.15 [CI, −0.19 to 0.50]; $I^2 = 61\%$; low-certainty evidence) (**Table**).

Objective Physical Function. Seven trials (n = 1063) evaluated gait speed (**Supplement Table 7**, available at Annals.org). Testosterone treatment was associated with a less-than-small improvement in objective physi-

cal function as measured by gait (SMD, 0.14 [CI, 0.02 to 0.27]; $I^2 = 0\%$; low-certainty evidence) (**Table** and **Supplement Figure 2** [available at Annals.org]).

The Physical Function Trial of the Testosterone Trials was limited to men who reported difficulty walking or climbing stairs and had a baseline gait speed less than 1.2 m/s on the 6-minute walk test (71). The primary outcome was the proportion of men who increased the distance walked in 6 minutes by at least 50 m. At 12 months, 20% of men assigned to testosterone achieved this threshold, versus 12% of men assigned to placebo (adjusted OR, 1.42 [CI, 0.83 to 2.45]). In the overall Testosterone Trials, including all men regardless of baseline gait speed, 21% and 13% of men in the testosterone and placebo groups, respectively, had increased the distance they walked in 6 minutes by at least 50 m at 12 months (adjusted OR, 1.76 [CI, 1.21 to 2.57]).

Quality of Life. We pooled 7 RCTs (n = 1043) that reported quality of life as an outcome using the AMS scale (**Supplement Tables 8** and **9**, available at Annals.org). The weighted mean total score at baseline on the AMS scale for the 6 trials that reported mean baseline AMS score was 43 points (scale, 17 to 85 points), indicating moderate severity of symptoms. Testosterone treatment was associated with a small improvement in quality of life, as measured by the AMS scale (SMD, −0.33 [CI, −0.50 to −0.16]; $I^2 = 6\%$, low-certainty evi-

dence) (**Table** and **Figure 2** [*bottom*]).

The weighted mean change in AMS score from baseline for the testosterone group was 7.0 points, compared with 3.6 points in the placebo group (weighted mean difference, −3.3 [CI, −5.2 to −1.3]; I^2 = 32%). Thus, men allocated to testosterone on average moved from moderate to mild symptom severity, while the symptoms of men allocated to placebo on average remained moderately severe.

Vitality or Fatigue. Three RCTs (*n* = 665) reported vitality and fatigue outcomes that could be pooled (**Supplement Table 10**, available at Annals.org). Testosterone was associated with a less-than-small improvement in self-reported fatigue or vitality (SMD, 0.17 [CI, 0.01 to 0.32]; I^2 = 0%) (**Supplement Figure 3**, available at Annals.org).

Mood. We pooled 5 RCTs (*n* = 872) that evaluated the effect of testosterone on measures of depressive symptoms (**Supplement Table 11**, available at Annals.org). None required depression or depressive symptoms for enrollment, and in general, most men enrolled did not have significant depressive symptoms at baseline (29, 30, 41–44, 51, 71, 77, 81). Trials measured depressive symptoms with various instruments (Beck Depression Inventory, Geriatric Depression Scale, Hospital Anxiety and Depression Scale, and Patient Health Questionnaire-9). Testosterone treatment was associated with a less-than-small improvement in depressive symptoms (SMD, −0.19 [CI, −0.32 to −0.05]; I^2 = 0%) (**Supplement Figure 4**, available at Annals.org).

Fractures. The RCTs rarely reported fractures (**Supplement Table 7**). Overall, 6 fractures were reported in testosterone groups and 8 in placebo groups during treatment periods ranging from 6 to 24 months (31, 48, 61, 65, 71, 74, 75, 79, 80). During follow-up periods of 6 to 12 months, an additional 3 fractures were reported in the testosterone groups and 5 in the placebo groups.

Cognitive Function. Nine studies reported cognitive outcomes (**Supplement Tables 12** and **13**, available at Annals.org). Because of variability in patients, scale scores, and domains assessed, we did not pool results. Follow-up periods ranged from 6 to 36 months.

One study enrolled persons with Alzheimer disease (78), and another enrolled those with mild cognitive impairment (33). The former was a pilot study that randomly assigned 10 participants to testosterone or placebo; the authors described the findings as preliminary but encouraging, noting the small sample size (75). The study of mild cognitive impairment reported 10 measures of verbal memory, visuospatial memory, visuospatial function, and language; many of the measures had multiple outcomes (a total of 18 outcomes). The pattern of change in these outcomes over the course of the study differed by treatment group for only 1 outcome, and the authors noted that the study may have been underpowered (33).

Of the 7 remaining studies, 5 included fewer than 45 completers and were underpowered for cognitive outcomes (30, 52, 68, 77, 81). A study enrolling 280 cognitively normal men reported that long-term treatment with testosterone did not improve cognitive function (verbal memory, visuospatial memory, language,

and executive function domains) (49). The Cognitive Function Trial (66), part of the Testosterone Trials (71), enrolled 493 men with age-associated memory impairment and reported no association between testosterone treatment and improved cognitive outcomes.

Harm Outcomes

The **Table** summarizes certainty of evidence for the primary harm outcomes. **Supplement Table 14** (available at Annals.org) provides results by formulation limited to trials with a mean baseline testosterone level less than 10.41 nmol/L (300 ng/dL). **Supplement Table 4** shows certainty of evidence for the primary harm outcomes by testosterone formulation. Results did not vary significantly by follow-up duration (<12 months, 12 to <24 months, and ≥24 months) (**Supplement Table 15**, available at Annals.org).

Adverse Cardiovascular Events. Fourteen trials (*n* = 2415) reported cardiovascular events as adverse events (**Supplement Table 16**, available at Annals.org). Most of the trials excluded men with advanced heart failure or a recent history of myocardial infarction or stroke, and none were designed to adequately assess cardiovascular risk of testosterone therapy. Cardiovascular event definitions varied between studies and were often not prespecified; event adjudication was rarely described or done. Pooled risk for adverse cardiovascular outcomes did not differ between groups (Peto OR, 1.22 [CI, 0.66 to 2.23]; I^2 = 18%; low-certainty evidence) (**Table** and **Figure 3** [*top*]). Incidence of cardiovascular events was 2.3% (CI, 0.9% to 4.1%) in the testosterone group, compared with 1.5% (CI, 0.8% to 2.5%) in the placebo group (**Supplement Figure 5**, available at Annals.org).
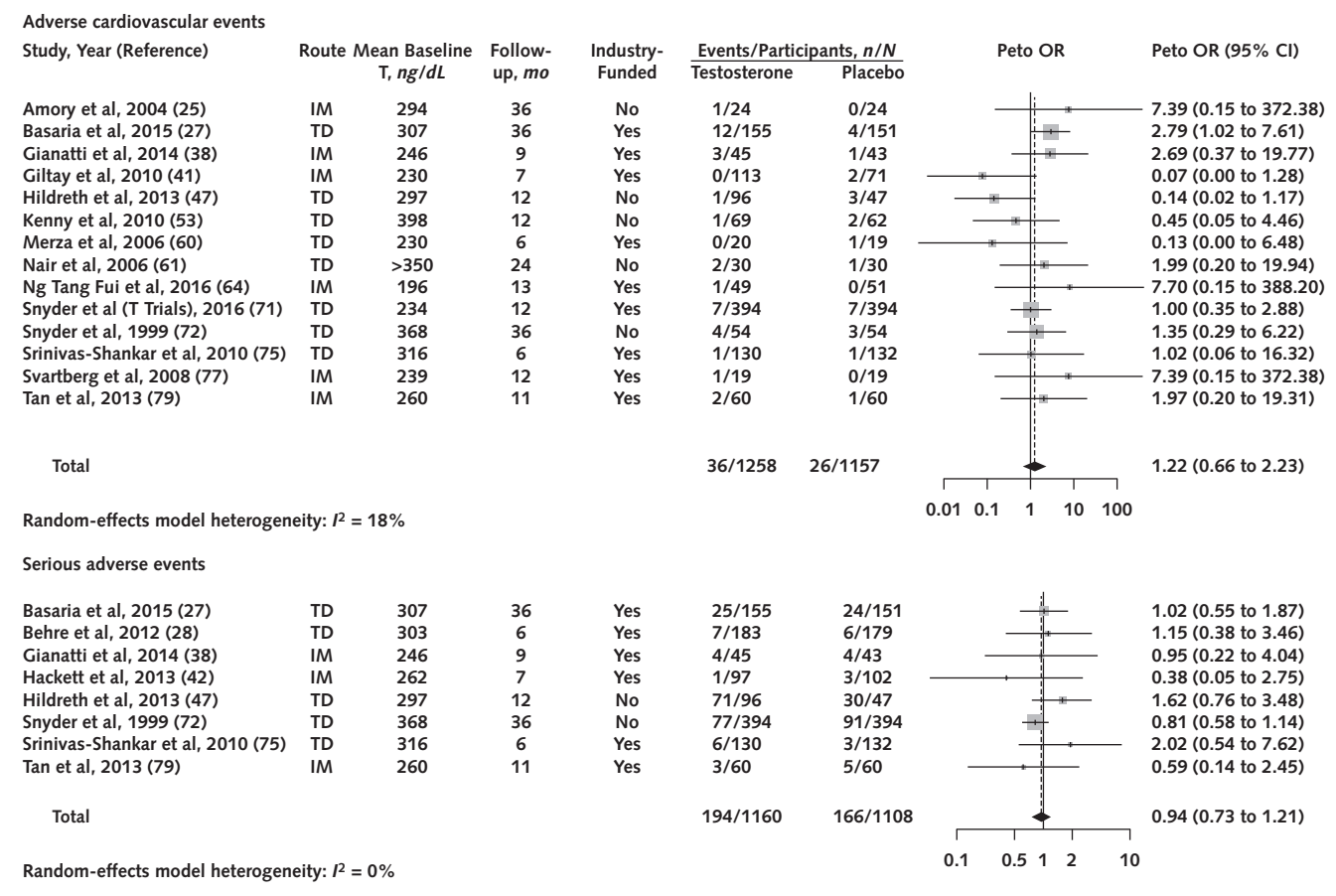
The TOM trial (9, 10) was stopped early because of excess cardiovascular adverse events in the testosterone group. When events were limited to our cardiovascular events of interest, 7 (7%) occurred in the testosterone group versus 1 (1%) in the placebo group. This trial was not included in our pooled analyses because of high risk of bias based on attrition rate.

Serious Adverse Events and Withdrawals Due to Adverse Events. Eight trials (*n* = 2268) reported serious adverse events (**Supplement Table 17**, available at Annals.org). Ascertainment, definition, and adjudication of these events were highly variable. Incidence was similar between groups: 13.2% (CI, 3.6% to 27.4%) of men assigned to testosterone and 12.8% (CI, 4.7% to 23.9%) of those assigned to placebo had a serious adverse event (**Supplement Figure 6**, available at Annals.org). The Peto OR was 0.94 (CI, 0.73 to 1.21) (I^2 = 0%; moderate-certainty evidence) (**Figure 3** [*bottom*] and **Table**).

Similarly, withdrawals due to adverse events did not differ between men assigned to testosterone treatment and those assigned to placebo (5.1% vs. 5.3%; Peto OR, 0.92 [CI, 0.65 to 1.28]; I^2 = 0%) (**Supplement Figure 7**, available at Annals.org).

Venous Thromboembolism. Few venous thromboembolism events were reported in any trial (0.6% in testosterone groups and 0.5% in placebo groups) (**Supplement Table 16**).

Figure 3. Primary harm outcomes from RCTs for testosterone treatment vs. placebo.

Adverse cardiovascular events

Study, Year (Reference)	Route	Mean Baseline T, *ng/dL*	Follow-up, *mo*	Industry-Funded	Events/Participants, *n/N* Testosterone	Placebo	Peto OR	Peto OR (95% CI)
Amory et al, 2004 (25)	IM	294	36	No	1/24	0/24		7.39 (0.15 to 372.38)
Basaria et al, 2015 (27)	TD	307	36	Yes	12/155	4/151		2.79 (1.02 to 7.61)
Gianatti et al, 2014 (38)	IM	246	9	Yes	3/45	1/43		2.69 (0.37 to 19.77)
Giltay et al, 2010 (41)	IM	230	7	Yes	0/113	2/71		0.07 (0.00 to 1.28)
Hildreth et al, 2013 (47)	TD	297	12	No	1/96	3/47		0.14 (0.02 to 1.17)
Kenny et al, 2010 (53)	TD	398	12	No	1/69	2/62		0.45 (0.05 to 4.46)
Merza et al, 2006 (60)	TD	230	6	Yes	0/20	1/19		0.13 (0.00 to 6.48)
Nair et al, 2006 (61)	TD	>350	24	No	2/30	1/30		1.99 (0.20 to 19.94)
Ng Tang Fui et al, 2016 (64)	IM	196	13	Yes	1/49	0/51		7.70 (0.15 to 388.20)
Snyder et al (T Trials), 2016 (71)	TD	234	12	Yes	7/394	7/394		1.00 (0.35 to 2.88)
Snyder et al, 1999 (72)	TD	368	36	No	4/54	3/54		1.35 (0.29 to 6.22)
Srinivas-Shankar et al, 2010 (75)	TD	316	6	Yes	1/130	1/132		1.02 (0.06 to 16.32)
Svartberg et al, 2008 (77)	IM	239	12	Yes	1/19	0/19		7.39 (0.15 to 372.38)
Tan et al, 2013 (79)	IM	260	11	Yes	2/60	1/60		1.97 (0.20 to 19.31)
Total					36/1258	26/1157		1.22 (0.66 to 2.23)

Random-effects model heterogeneity: $I^2 = 18\%$

0.01 0.1 1 10 100

Serious adverse events

Study, Year (Reference)	Route	Mean Baseline T	Follow-up	Industry-Funded	Testosterone	Placebo	Peto OR	Peto OR (95% CI)
Basaria et al, 2015 (27)	TD	307	36	Yes	25/155	24/151		1.02 (0.55 to 1.87)
Behre et al, 2012 (28)	TD	303	6	Yes	7/183	6/179		1.15 (0.38 to 3.46)
Gianatti et al, 2014 (38)	IM	246	9	Yes	4/45	4/43		0.95 (0.22 to 4.04)
Hackett et al, 2013 (42)	IM	262	7	Yes	1/97	3/102		0.38 (0.05 to 2.75)
Hildreth et al, 2013 (47)	TD	297	12	No	71/96	30/47		1.62 (0.76 to 3.48)
Snyder et al, 1999 (72)	TD	368	36	No	77/394	91/394		0.81 (0.58 to 1.14)
Srinivas-Shankar et al, 2010 (75)	TD	316	6	Yes	6/130	3/132		2.02 (0.54 to 7.62)
Tan et al, 2013 (79)	IM	260	11	Yes	3/60	5/60		0.59 (0.14 to 2.45)
Total					194/1160	166/1108		0.94 (0.73 to 1.21)

Random-effects model heterogeneity: $I^2 = 0\%$

0.1 0.5 1 2 10

IM = intramuscular; OR = odds ratio; RCT = randomized controlled trial; T = testosterone level; TD = transdermal.

Prostate Cancer and Lower Urinary Tract Symptoms. Trials typically excluded men with a history of prostate cancer or a prostate-specific antigen value above a predetermined level (most commonly >4.0 μg/L). Ten studies (*n* = 2143) reported cases of prostate cancer (**Supplement Table 16**). Prostate cancer ascertainment varied, and no studies were powered to detect a difference in rate between treatment groups. Prostate cancer incidence was less than 1% during the trials in all men regardless of treatment group (Peto OR, 0.97 [CI, 0.35 to 2.69]; $I^2 = 1\%$; very-low-certainty evidence) (**Table** and **Supplement Figures 8** and **9** [available at Annals.org]).

Three trials reported the number of men who developed clinically significant lower urinary tract symptoms, defined by study authors as a score above 19 (71) or above 21 (27, 49) on the International Prostate Symptom Score or as a score of at least 20 on the American Urological Association symptom scale (47) (**Supplement Table 16**). Clinically significant lower urinary tract symptoms were reported for 42 of 645 men (6.5%) assigned to testosterone and 30 of 592 men (5.1%) assigned to placebo (Peto OR, 1.36 [CI, 0.35 to 5.30]; $I^2 = 78\%$) (**Supplement Figure 10**, available at Annals.org).

Mortality. Twelve studies (*n* = 2727) reported mortality (**Supplement Table 16**). Duration of follow-up varied from 24 weeks to 3 years. None were powered to detect a difference in mortality between treatment groups, and studies typically excluded persons at the highest risk for death. Incidence of death was 0.4% (CI, 0.07% to 0.99%) in men treated with testosterone, compared with 1.5% (CI, 0.48% to 2.89%) in the placebo group (Peto OR, 0.47 [CI, 0.25 to 0.89]; $I^2 = 0$; low-certainty evidence) (**Table** and **Supplement Figures 11** and **12** [available at Annals.org]).

Special Populations

Two trials were limited to men with obesity; these trials were included in the overall analyses. One (63, 64) reported similar effects on quality of life and sexual function to those in our pooled analyses. An older, smaller trial (57) reported improvement in well-being and energy using nonstandard questionnaires.

Five trials included only men with diabetes mellitus or metabolic syndrome (26, 37–39, 41–46, 50, 51); 2 were rated as having high risk of bias and were not included in the overall pooled analyses. For outcomes reported in these studies of diabetes or metabolic

syndrome, results were similar to the overall pooled effects.

No trials of testosterone replacement in opioid users met our inclusion criteria.

The 10 trials done in populations with specific disease conditions were not included in our pooled analyses (31, 33, 34, 40, 56, 59, 62, 69, 76, 78). Three trials were judged to be at high risk of bias (31, 56, 69). Few reported any significant improvements in patient-centered outcomes (Supplement Tables 5 to 13, 16, and 17).

Observational Studies

Twenty observational studies (15 retrospective cohorts, 1 prospective cohort, 2 case–control studies, and 2 prospective registry studies) reported harms as outcomes of interest. Two had low, 14 medium, and 4 high risk of bias. Most reported cardiovascular end points or mortality; few reported thromboembolic disease or prostate cancer. Studies were heterogeneous in their inclusion criteria and patient populations, including baseline levels of testosterone and comorbid conditions. Duration and type of testosterone treatment, duration of follow-up, and adjustment for potential confounders also varied (Supplement Table 2). Mean or median follow-up ranged from 0.73 years (93) to 10.3 years (90), although 1 study assessed outcomes only during postoperative hospitalization (83). Five studies reported at least partial industry sponsorship.

Most studies did not identify an increased risk for death or cardiovascular events associated with testosterone, and some reported decreased risk (92, 97). One retrospective cohort study of veterans who had coronary angiography, had a testosterone level less than 10.41 nmol/L (300 ng/dL), and were treated with testosterone found that participants had an increased risk for a combined end point of all-cause mortality, myocardial infarction, and ischemic stroke compared with those not treated with testosterone, after adjustment for potential confounders (hazard ratio, 1.29 [CI, 1.05 to 1.58]) (101).

No study reported an increased risk for prostate cancer. A single retrospective cohort study (87) reported an increased risk for obstructive sleep apnea in men treated with testosterone. Testosterone was not associated with increased risk for pulmonary embolism or deep venous thrombosis (83, 87, 88, 95, 98) in the few observational studies that reported that outcome.

Among 8 observational studies reporting results by formulation, results were mixed, with no consistent finding of increased risk for harm for one formulation over the other.

DISCUSSION

Our analysis of data from 38 RCTs found that intramuscular or transdermal testosterone therapy resulted in small improvements in sexual functioning and quality of life but had little to no effect on physical functioning, depressive symptoms, energy and vitality, and cognition. Quality of life, when reported as an outcome, was typically measured using the AMS scale; the observed effect may have been driven primarily by effects on the sexual subscale. These findings are limited to men without underlying medical conditions recognized to cause hypogonadism, such as testicular failure or hypothalamic–pituitary injury. Evidence for most outcomes was of low or moderate certainty. None of the trials had adequate power to assess risk for adverse cardiovascular events, prostate cancer, thromboembolic disease, or death—the harms of most clinical interest. The observational studies were limited by likely unmeasured confounding due to indication or contraindication and must be interpreted with caution. Few trials were longer than 1 year, limiting conclusions about the benefits or harms of longer-term treatment. Participants were typically aged 60 years or older; white; and without recent cardiovascular events, history of prostate cancer, or elevated levels of prostate-specific antigen. This limits the generalizability of our findings.

We included RCTs of at least 6 months' duration and prioritized patient-centered outcomes over intermediate measures, such as body composition or metabolic variables. Important heterogeneity existed in the entry criteria for trials, including testosterone levels, presence of symptoms attributable to hypogonadism, comorbid conditions, and extent of evaluation for underlying causes of low testosterone levels. However, we believe that the resulting mix of patients likely reflects clinical practice, especially among primary care physicians who provide most evaluation and treatment of these patients. The largest randomized placebo-controlled trial of testosterone therapy for older men, the Testosterone Trials, enrolled men on the basis of 2 fasting morning levels of total testosterone that averaged less than 9.54 nmol/L (275 ng/dL), no underlying conditions known to cause hypogonadism, and the presence of symptoms associated with hypogonadism; the findings from the several coordinated trials that make up the Testosterone Trials are generally consistent with our findings.

We did not find differences in reported outcomes by testosterone formulation (intramuscular vs. transdermal). However, patient characteristics, testosterone levels, and outcome reporting varied in studies of intramuscular versus transdermal formulations. No clinical trial that met our other entry criteria directly compared an intramuscular versus a transdermal formulation of testosterone. Furthermore, we did not systematically search for or include studies that directly compared different formulations without a control group.

The FDA recently approved an oral formulation of testosterone undecanoate for use in the United States (104). Our review did not evaluate the evidence for efficacy or harms of oral testosterone. The FDA specifically stated that the oral formulation is contraindicated in "[m]en with hypogonadal conditions, such as 'age-related hypogonadism,' that are not associated with structural or genetic etiologies," citing demonstrated increases in blood pressure and lack of established efficacy (105).

We limited our review of harm outcomes to serious adverse events, major adverse cardiovascular events, withdrawals due to adverse events, prostate cancer, worsening lower urinary tract symptoms, venous thromboembolism, worsening or development of sleep apnea, and mortality. Other potential risks of testosterone therapy are recognized but were outside the scope of this review. These include but are not limited to polycythemia, elevated prostate-specific antigen levels, increased blood pressure, gynecomastia, skin reaction to transdermal products, testicular atrophy, infertility or azoospermia, and fluid retention, as well as risk for transfer to others of the transdermal preparations, a concern owing to risk for virilization in women or children (106). Because of concern about inadequate data regarding harms of testosterone treatment in older men with age-related hypogonadism, the FDA has required companies that manufacture these products to conduct a controlled clinical trial to evaluate the effects of testosterone therapy on cardiovascular outcomes (1). This trial, TRAVERSE (Testosterone Replacement Therapy for Assessment of Long-term Vascular Events and Efficacy ResponSE in Hypogonadal Men), began enrollment in May 2018 and will follow participants for up to 5 years for cardiovascular safety and prostate safety, as well as efficacy outcomes (107).

Because deaths were few and entry criteria for RCTs excluded persons at highest risk for death, we cannot make definitive conclusions about testosterone's effect on mortality. However, our findings do not suggest an increased risk for death with testosterone treatment.

Our findings are generally consistent with those of other systematic reviews (108–125), with occasional exceptions (for example, the trial by Huo and colleagues [126]), despite variable inclusion and exclusion criteria for study selection, various outcomes assessed, and methodological differences. Most have found low- to moderate-certainty evidence of small beneficial effects on sexual function, little to no evidence of benefit for other clinical efficacy outcomes, and inadequate evidence to make definitive conclusions about cardiovascular and other long-term harms.

In conclusion, intramuscular or transdermal testosterone treatment in men with low testosterone levels not associated with specific, well-established medical conditions known to cause primary or secondary hypogonadism may result in small improvements in sexual function and self-reported quality of life but little to no benefit for other common symptoms of aging, including fatigue or decreased energy, reduced physical function, and reduced cognition. Evidence is inadequate about the long-term benefits or serious harms of testosterone treatment.

From Minneapolis VA Health Care System and University of Minnesota School of Medicine, Minneapolis, Minnesota (S.J.D., P.D., N.E., A.E., L.S.H., H.A.F., T.J.W.); Minneapolis VA Health Care System, Minneapolis, Minnesota (N.L.G., R.M., L.G.M.); and Minneapolis VA Health Care System and University of Minnesota School of Public Health, Minneapolis, Minnesota (C.E.R.).

Corresponding Author: Susan J. Diem, MD, MPH, One Veterans Drive (111-0), Minneapolis, MN 55417; e-mail, sdiem @umn.edu.

Current author addresses and author contributions are available at Annals.org.

References

1. **U.S. Food and Drug Administration.** FDA Drug Safety Communication: FDA cautions about using testosterone products for low testosterone due to aging; requires labeling change to inform of possible increased risk of heart attack and stroke with use. 2015. Accessed at www.fda.gov/media/91048/download on 18 November 2019.
2. **Jasuja GK, Bhasin S, Rose AJ.** Patterns of testosterone prescription overuse. Curr Opin Endocrinol Diabetes Obes. 2017;24:240-5. [PMID: 28248754] doi:10.1097/MED.0000000000000336
3. **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;99:835-42. [PMID: 24423353] doi:10 .1210/jc.2013-3570
4. **Nguyen CP, Hirsch MS, Moeny D, et al.** Testosterone and "age-related hypogonadism" — FDA concerns. N Engl J Med. 2015;373: 689-91. [PMID: 26287846] doi:10.1056/NEJMp1506632
5. **Baillargeon J, Urban RJ, Ottenbacher KJ, et al.** Trends in androgen prescribing in the United States, 2001 to 2011. JAMA Intern Med. 2013;173:1465-6. [PMID: 23939517] doi:10 .1001/jamainternmed.2013.6895
6. **Baillargeon J, Urban RJ, Kuo YF, et al.** Screening and monitoring in men prescribed testosterone therapy in the U.S., 2001–2010. Public Health Rep. 2015;130:143-52. [PMID: 25729103]
7. **Bhasin S, Brito JP, Cunningham GR, et al.** Testosterone therapy in men with hypogonadism: an Endocrine Society clinical practice guideline. J Clin Endocrinol Metab. 2018;103:1715-44. [PMID: 29562364] doi:10.1210/jc.2018-00229
8. **Mulhall JP, Trost LW, Brannigan RE, et al.** Evaluation and management of testosterone deficiency. Linthicum, MD: American Urological Association; 2018. Accessed at www.auanet.org/guidelines /testosterone-deficiency-guideline on 23 July 2019.
9. **Basaria S, Coviello AD, Travison TG, et al.** Adverse events associated with testosterone administration. N Engl J Med. 2010;363:109-22. [PMID: 20592293] doi:10.1056/NEJMoa1000485
10. **Travison TG, Basaria S, Storer TW, et al.** Clinical meaningfulness of the changes in muscle performance and physical function associated with testosterone administration in older men with mobility limitation. J Gerontol A Biol Sci Med Sci. 2011;66:1090-9. [PMID: 21697501] doi:10.1093/gerona/glr100
11. **Abellan van Kan G, Rolland Y, Andrieu S, et al.** Gait speed at usual pace as a predictor of adverse outcomes in community-dwelling older people an International Academy on Nutrition and

Aging (IANA) Task Force. J Nutr Health Aging. 2009;13:881-9. [PMID: 19924348]

12. Studenski S, Perera S, Patel K, et al. Gait speed and survival in older adults. JAMA. 2011;305:50-8. [PMID: 21205966] doi:10.1001/jama.2010.1923

13. Higgins JPT, Green S, eds. Cochrane Handbook for Systematic Reviews of Interventions. Version 5.1.0. The Cochrane Collaboration; 2011. Accessed at http://handbook.cochrane.org on 23 July 2019.

14. Viswanathan M, Ansari MT, Berkman ND, et al. Assessing the risk of bias of individual studies in systematic reviews of health care interventions. In: Methods Guide for Effectiveness and Comparative Effectiveness Reviews. Rockville: Agency for Healthcare Research and Quality; 2008. [PMID: 22479713]

15. Schwarzer G. Meta: an R package for meta-analysis. R News. 2007;7:40-5.

16. Viechtbauer W. Conducting meta-analyses in R with the metafor package. J Stat Softw. 2010;36:1-48.

17. Veroniki AA, Jackson D, Bender R, et al. Methods to calculate uncertainty in the estimated overall effect size from a random-effects meta-analysis. Res Synth Methods. 2019;10:23-43. [PMID: 30129707] doi:10.1002/jrsm.1319

18. Cohen J. Statistical Power Analysis for the Behavioral Sciences. 2nd ed. Hillsdale, NJ: Lawrence Erlbaum Associates; 1988.

19. Freeman MF, Tukey JW. Transformations related to the angular and the square root. Ann Math Stat. 1950;21:607-11.

20. Higgins JP, Thompson SG, Deeks JJ, et al. Measuring inconsistency in meta-analyses. BMJ. 2003;327:557-60. [PMID: 12958120]

21. GRADEpro GDT: GRADEpro Guideline Development Tool [computer program]. Hamilton, Ontario, Canada: McMaster University and Evidence Prime; 2015.

22. Schünemann H, Brożek J, Guyatt G, et al, eds. GRADE handbook for grading quality of evidence and strength of recommendations. GRADE Working Group; 2013. Accessed at https://gdt.gradepro.org/app/handbook/handbook.html on 23 July 2019.

23. Allan CA, Forbes EA, Strauss BJ, et al. Testosterone therapy increases sexual desire in ageing men with low-normal testosterone levels and symptoms of androgen deficiency. Int J Impot Res. 2008; 20:396-401. [PMID: 18528400] doi:10.1038/ijir.2008.22

24. Allan CA, Strauss BJ, Burger HG, et al. Testosterone therapy prevents gain in visceral adipose tissue and loss of skeletal muscle in nonobese aging men. J Clin Endocrinol Metab. 2008;93:139-46. [PMID: 17940111]

25. Amory JK, Watts NB, Easley KA, et al. Exogenous testosterone or testosterone with finasteride increases bone mineral density in older men with low serum testosterone. J Clin Endocrinol Metab. 2004;89: 503-10. [PMID: 14764753]

26. Aversa A, Bruzziches R, Francomano D, et al. Efficacy and safety of two different testosterone undecanoate formulations in hypogonadal men with metabolic syndrome. J Endocrinol Invest. 2010;33: 776-83. [PMID: 20220293] doi:10.3275/6903

27. Basaria S, Harman SM, Travison TG, et al. Effects of testosterone administration for 3 years on subclinical atherosclerosis progression in older men with low or low-normal testosterone levels: a randomized clinical trial. JAMA. 2015;314:570-81. [PMID: 26262795] doi:10.1001/jama.2015.8881

28. Behre HM, Tammela TL, Arver S, et al; European Testogel Study Team. A randomized, double-blind, placebo-controlled trial of testosterone gel on body composition and health-related quality-of-life in men with hypogonadal to low-normal levels of serum testosterone and symptoms of androgen deficiency over 6 months with 12 months open-label follow-up. Aging Male. 2012;15:198-207. [PMID: 22834649] doi:10.3109/13685538.2012.699562

29. Borst SE, Yarrow JF, Conover CF, et al. Musculoskeletal and prostate effects of combined testosterone and finasteride administration in older hypogonadal men: a randomized, controlled trial. Am J Physiol Endocrinol Metab. 2014;306:E433-42. [PMID: 24326421] doi:10.1152/ajpendo.00592.2013

30. Borst SE, Yarrow JF, Fernandez C, et al. Cognitive effects of testosterone and finasteride administration in older hypogonadal men.

Clin Interv Aging. 2014;9:1327-33. [PMID: 25143719] doi:10.2147/CIA.S61760

31. Brockenbrough AT, Dittrich MO, Page ST, et al. Transdermal androgen therapy to augment EPO in the treatment of anemia of chronic renal disease. Am J Kidney Dis. 2006;47:251-62. [PMID: 16431254]

32. Budoff MJ, Ellenberg SS, Lewis CE, et al. Testosterone treatment and coronary artery plaque volume in older men with low testosterone. JAMA. 2017;317:708-16. [PMID: 28241355] doi:10.1001/jama.2016.21043

33. Cherrier MM, Anderson K, Shofer J, et al. Testosterone treatment of men with mild cognitive impairment and low testosterone levels. Am J Alzheimers Dis Other Demen. 2015;30:421-30. [PMID: 25392187] doi:10.1177/1533317514556874

34. Crawford BA, Liu PY, Kean MT, et al. Randomized placebo-controlled trial of androgen effects on muscle and bone in men requiring long-term systemic glucocorticoid treatment. J Clin Endocrinol Metab. 2003;88:3167-76. [PMID: 12843161]

35. Dias JP, Melvin D, Shardell M, et al. Effects of transdermal testosterone gel or an aromatase inhibitor on prostate volume in older men. J Clin Endocrinol Metab. 2016;101:1865-71. [PMID: 26950683] doi:10.1210/jc.2016-1111

36. Dias JP, Melvin D, Simonsick EM, et al. Effects of aromatase inhibition vs. testosterone in older men with low testosterone: randomized-controlled trial. Andrology. 2016;4:33-40. [PMID: 26588809] doi:10.1111/andr.12126

37. Gianatti EJ, Dupuis P, Hoermann R, et al. Effect of testosterone treatment on glucose metabolism in men with type 2 diabetes: a randomized controlled trial. Diabetes Care. 2014;37:2098-107. [PMID: 24804695] doi:10.2337/dc13-2845

38. Gianatti EJ, Dupuis P, Hoermann R, et al. Effect of testosterone treatment on constitutional and sexual symptoms in men with type 2 diabetes in a randomized, placebo-controlled clinical trial. J Clin Endocrinol Metab. 2014;99:3821-8. [PMID: 24978674] doi:10.1210/jc.2014-1872

39. Gianatti EJ, Hoermann R, Lam Q, et al. Effect of testosterone treatment on cardiac biomarkers in a randomized controlled trial of men with type 2 diabetes. Clin Endocrinol (Oxf). 2016;84:55-62. [PMID: 26120052] doi:10.1111/cen.12842

40. Giannoulis MG, Sonksen PH, Umpleby M, et al. The effects of growth hormone and/or testosterone in healthy elderly men: a randomized controlled trial. J Clin Endocrinol Metab. 2006;91:477-84. [PMID: 16332938]

41. Giltay EJ, Tishova YA, Mskhalaya GJ, et al. Effects of testosterone supplementation on depressive symptoms and sexual dysfunction in hypogonadal men with the metabolic syndrome. J Sex Med. 2010;7:2572-82. [PMID: 20524974] doi:10.1111/j.1743-6109.2010.01859.x

42. Hackett G, Cole N, Bhartia M, et al. Testosterone replacement therapy with long-acting testosterone undecanoate improves sexual function and quality-of-life parameters vs. placebo in a population of men with type 2 diabetes. J Sex Med. 2013;10:1612-27. [PMID: 23551886] doi:10.1111/jsm.12146

43. Hackett G, Cole N, Bhartia M, et al; BLAST Study Group. Testosterone replacement therapy improves metabolic parameters in hypogonadal men with type 2 diabetes but not in men with coexisting depression: the BLAST study. J Sex Med. 2014;11:840-56. [PMID: 24308723] doi:10.1111/jsm.12404

44. Hackett G, Cole N, Bhartia M, et al; BLAST Study Group. The response to testosterone undecanoate in men with type 2 diabetes is dependent on achieving threshold serum levels (the BLAST study). Int J Clin Pract. 2014;68:203-15. [PMID: 24355040] doi:10.1111/ijcp.12235

45. Hackett G, Cole N, Saghir A, et al. Testosterone undecanoate improves sexual function in men with type 2 diabetes and severe hypogonadism: results from a 30-week randomized placebo-controlled study. BJU Int. 2016;118:804-13. [PMID: 27124889] doi:10.1111/bju.13516

46. Hackett G, Cole N, Saghir A, et al. Testosterone replacement therapy: improved sexual desire and erectile function in men with

type 2 diabetes following a 30-week randomized placebo-controlled study. Andrology. 2017;5:905-13. [PMID: 28771964] doi:10.1111/andr.12399

47. Hildreth KL, Barry DW, Moreau KL, et al. Effects of testosterone and progressive resistance exercise in healthy, highly functioning older men with low-normal testosterone levels. J Clin Endocrinol Metab. 2013;98:1891-900. [PMID: 23533227] doi:10.1210/jc.2012-3695

48. Ho CC, Tong SF, Low WY, et al. A randomized, double-blind, placebo-controlled trial on the effect of long-acting testosterone treatment as assessed by the Aging Male Symptoms scale. BJU Int. 2012;110:260-5. [PMID: 22093057] doi:10.1111/j.1464-410X.2011.10755.x

49. Huang G, Wharton W, Bhasin S, et al. Effects of long-term testosterone administration on cognition in older men with low or low-to-normal testosterone concentrations: a prespecified secondary analysis of data from the randomised, double-blind, placebo-controlled TEAAM trial. Lancet Diabetes Endocrinol. 2016;4:657-65. [PMID: 27377542] doi:10.1016/S2213-8587(16)30102-4

50. Jones TH, Arver S, Behre HM, et al; TIMES2 Investigators. Testosterone replacement in hypogonadal men with type 2 diabetes and/or metabolic syndrome (the TIMES2 study). Diabetes Care. 2011;34:828-37. [PMID: 21386088] doi:10.2337/dc10-1233

51. Kalinchenko SY, Tishova YA, Mskhalaya GJ, et al. Effects of testosterone supplementation on markers of the metabolic syndrome and inflammation in hypogonadal men with the metabolic syndrome: the double-blinded placebo-controlled Moscow study. Clin Endocrinol (Oxf). 2010;73:602-12. [PMID: 20718771] doi:10.1111/j.1365-2265.2010.03845.x

52. Kenny AM, Bellantonio S, Gruman CA, et al. Effects of transdermal testosterone on cognitive function and health perception in older men with low bioavailable testosterone levels. J Gerontol A Biol Sci Med Sci. 2002;57:M321-5. [PMID: 11983727]

53. Kenny AM, Kleppinger A, Annis K, et al. Effects of transdermal testosterone on bone and muscle in older men with low bioavailable testosterone levels, low bone mass, and physical frailty. J Am Geriatr Soc. 2010;58:1134-43. [PMID: 20722847] doi:10.1111/j.1532-5415.2010.02865.x

54. Kenny AM, Prestwood KM, Gruman CA, et al. Effects of transdermal testosterone on bone and muscle in older men with low bioavailable testosterone levels. J Gerontol A Biol Sci Med Sci. 2001;56:M266-72. [PMID: 11320105]

55. Konaka H, Sugimoto K, Orikasa H, et al; EARTH study group. Effects of long-term androgen replacement therapy on the physical and mental statuses of aging males with late-onset hypogonadism: a multicenter randomized controlled trial in Japan (EARTH Study). Asian J Androl. 2016;18:25-34. [PMID: 25761833] doi:10.4103/1008-682X.148720

56. Malkin CJ, Pugh PJ, West JN, et al. Testosterone therapy in men with moderate severity heart failure: a double-blind randomized placebo controlled trial. Eur Heart J. 2006;27:57-64. [PMID: 16093267]

57. Mårin P, Holmäng S, Gustafsson C, et al. Androgen treatment of abdominally obese men. Obes Res. 1993;1:245-51. [PMID: 16350577]

58. Marks LS, Mazer NA, Mostaghel E, et al. Effect of testosterone replacement therapy on prostate tissue in men with late-onset hypogonadism: a randomized controlled trial. JAMA. 2006;296:2351-61. [PMID: 17105798]

59. Mathur A, Malkin C, Saeed B, et al. Long-term benefits of testosterone replacement therapy on angina threshold and atheroma in men. Eur J Endocrinol. 2009;161:443-9. [PMID: 19542238] doi:10.1530/EJE-09-0092

60. Merza Z, Blumsohn A, Mah PM, et al. Double-blind placebo-controlled study of testosterone patch therapy on bone turnover in men with borderline hypogonadism. Int J Androl. 2006;29:381-91. [PMID: 16390499]

61. Nair KS, Rizza RA, O'Brien P, et al. DHEA in elderly women and DHEA or testosterone in elderly men. N Engl J Med. 2006;355:1647-59. [PMID: 17050889]

62. Navarro-Peñalver M, Perez-Martinez MT, Gómez-Bueno M, et al. Testosterone replacement therapy in deficient patients with chronic heart failure: a randomized double-blind controlled pilot study. J Cardiovasc Pharmacol Ther. 2018;23:543-50. [PMID: 29929385] doi:10.1177/1074248418784020

63. Ng Tang Fui M, Hoermann R, Prendergast LA, et al. Symptomatic response to testosterone treatment in dieting obese men with low testosterone levels in a randomized, placebo-controlled clinical trial. Int J Obes (Lond). 2017;41:420-6. [PMID: 28028318] doi:10.1038/ijo.2016.242

64. Ng Tang Fui M, Prendergast LA, Dupuis P, et al. Effects of testosterone treatment on body fat and lean mass in obese men on a hypocaloric diet: a randomised controlled trial. BMC Med. 2016;14:153. [PMID: 27716209]

65. O'Connell MD, Roberts SA, Srinivas-Shankar U, et al. Do the effects of testosterone on muscle strength, physical function, body composition, and quality of life persist six months after treatment in intermediate-frail and frail elderly men? J Clin Endocrinol Metab. 2011;96:454-8. [PMID: 21084399] doi:10.1210/jc.2010-1167

66. Resnick SM, Matsumoto AM, Stephens-Shields AJ, et al. Testosterone treatment and cognitive function in older men with low testosterone and age-associated memory impairment. JAMA. 2017;317:717-27. [PMID: 28241356] doi:10.1001/jama.2016.21044

67. Shigehara K, Sugimoto K, Konaka H, et al. Androgen replacement therapy contributes to improving lower urinary tract symptoms in patients with hypogonadism and benign prostate hypertrophy: a randomised controlled study. Aging Male. 2011;14:53-8. [PMID: 21171937] doi:10.3109/13685538.2010.518178

68. Sih R, Morley JE, Kaiser FE, et al. Testosterone replacement in older hypogonadal men: a 12-month randomized controlled trial. J Clin Endocrinol Metab. 1997;82:1661-7. [PMID: 9177359]

69. Sinclair M, Grossmann M, Hoermann R, et al. Testosterone therapy increases muscle mass in men with cirrhosis and low testosterone: a randomised controlled trial. J Hepatol. 2016;65:906-13. [PMID: 27312945] doi:10.1016/j.jhep.2016.06.007

70. Snyder PJ, Peachey H, Berlin JA, et al. Effect of transdermal testosterone treatment on serum lipid and apolipoprotein levels in men more than 65 years of age. Am J Med. 2001;111:255-60. [PMID: 11566454]

71. Snyder PJ, Bhasin S, Cunningham GR, et al; Testosterone Trials Investigators. Effects of testosterone treatment in older men. N Engl J Med. 2016;374:611-24. [PMID: 26886521] doi:10.1056/NEJMoa1506119

72. Snyder PJ, Peachey H, Hannoush P, et al. Effect of testosterone treatment on bone mineral density in men over 65 years of age. J Clin Endocrinol Metab. 1999;84:1966-72. [PMID: 10372695]

73. Snyder PJ, Peachey H, Hannoush P, et al. Effect of testosterone treatment on body composition and muscle strength in men over 65 years of age. J Clin Endocrinol Metab. 1999;84:2647-53. [PMID: 10443654]

74. Snyder PJ, Kopperdahl DL, Stephens-Shields AJ, et al. Effect of testosterone treatment on volumetric bone density and strength in older men with low testosterone: a controlled clinical trial. JAMA Intern Med. 2017;177:471-9. [PMID: 28241231] doi:10.1001/jamainternmed.2016.9539

75. Srinivas-Shankar U, Roberts SA, Connolly MJ, et al. Effects of testosterone on muscle strength, physical function, body composition, and quality of life in intermediate-frail and frail elderly men: a randomized, double-blind, placebo-controlled study. J Clin Endocrinol Metab. 2010;95:639-50. [PMID: 20061435] doi:10.1210/jc.2009-1251

76. Svartberg J, Aasebø U, Hjalmarsen A, et al. Testosterone treatment improves body composition and sexual function in men with COPD, in a 6-month randomized controlled trial. Respir Med. 2004;98:906-13. [PMID: 15338805]

77. Svartberg J, Agledahl I, Figenschau Y, et al. Testosterone treatment in elderly men with subnormal testosterone levels improves body composition and BMD in the hip. Int J Impot Res. 2008;20:378-87. [PMID: 18480825] doi:10.1038/ijir.2008.19

78. **Tan RS, Pu SJ.** A pilot study on the effects of testosterone in hypogonadal aging male patients with Alzheimer's disease. Aging Male. 2003;6:13-7. [PMID: 12809076]

79. **Tan WS, Low WY, Ng CJ, et al.** Efficacy and safety of long-acting intramuscular testosterone undecanoate in aging men: a randomised controlled study. BJU Int. 2013;111:1130-40. [PMID: 23651425] doi:10.1111/bju.12037

80. **Tong SF, Ng CJ, Lee BC, et al.** Effect of long-acting testosterone undecanoate treatment on quality of life in men with testosterone deficiency syndrome: a double blind randomized controlled trial. Asian J Androl. 2012;14:604-11. [PMID: 22635164] doi:10.1038/aja.2011.178

81. **Vaughan C, Goldstein FC, Tenover JL.** Exogenous testosterone alone or with finasteride does not improve measurements of cognition in healthy older men with low serum testosterone. J Androl. 2007;28:875-82. [PMID: 17609296]

82. **Wang C, Stephens-Shields AJ, DeRogatis LR, et al.** Validity and clinically meaningful changes in the Psychosexual Daily Questionnaire and Derogatis Interview for Sexual Function Assessment: results from the Testosterone Trials. J Sex Med. 2018;15:997-1009. [PMID: 29960633] doi:10.1016/j.jsxm.2018.05.008

83. **Argalious MY, You J, Mao G, et al.** Association of testosterone replacement therapy and the incidence of a composite of postoperative in-hospital mortality and cardiovascular events in men undergoing noncardiac surgery. Anesthesiology. 2017;127:457-65. [PMID: 28816783] doi:10.1097/ALN.0000000000001757

84. **Baillargeon J, Urban RJ, Zhang W, et al.** Testosterone replacement therapy and hospitalization rates in men with COPD. Chron Respir Dis. 2019;16:1479972318793004. [PMID: 30205698] doi:10.1177/1479972318793004

85. **Baillargeon J, Urban RJ, Kuo YF, et al.** Risk of myocardial infarction in older men receiving testosterone therapy. Ann Pharmacother. 2014;48:1138-44. [PMID: 24989174]

86. **Cheetham TC, An J, Jacobsen SJ, et al.** Association of testosterone replacement with cardiovascular outcomes among men with androgen deficiency. JAMA Intern Med. 2017;177:491-9. [PMID: 28241244] doi:10.1001/jamainternmed.2016.9546

87. **Cole AP, Hanske J, Jiang W, et al.** Impact of testosterone replacement therapy on thromboembolism, heart disease and obstructive sleep apnoea in men. BJU Int. 2018;121:811-8. [PMID: 29383868] doi:10.1111/bju.14149

88. **Debruyne FM, Behre HM, Roehrborn CG, et al; RHYME Investigators.** Testosterone treatment is not associated with increased risk of prostate cancer or worsening of lower urinary tract symptoms: prostate health outcomes in the Registry of Hypogonadism in Men. BJU Int. 2017;119:216-24. [PMID: 27409523] doi:10.1111/bju.13578

89. **Dedania VS, Zacks DN, Pan W, et al.** Testosterone supplementation and retinal vascular disease. Retina. 2018;38:2247-52. [PMID: 29040227] doi:10.1097/IAE.0000000000001869

90. **Eisenberg ML, Li S, Herder D, et al.** Testosterone therapy and mortality risk. Int J Impot Res. 2015;27:46-8. [PMID: 25078049] doi:10.1038/ijir.2014.29

91. **Etminan M, Skeldon SC, Goldenberg SL, et al.** Testosterone therapy and risk of myocardial infarction: a pharmacoepidemiologic study. Pharmacotherapy. 2015;35:72-8. [PMID: 25582846] doi:10.1002/phar.1534

92. **Hackett G, Cole N, Mulay A, et al.** Long-term testosterone therapy in type 2 diabetes is associated with reduced mortality without improvement in conventional cardiovascular risk factors. BJU Int. 2019;123:519-29. [PMID: 30216622] doi:10.1111/bju.14536

93. **Li H, Mitchell L, Zhang X, et al.** Testosterone therapy and risk of acute myocardial infarction in hypogonadal men: an administrative health care claims study. J Sex Med. 2017;14:1307-17. [PMID: 29110802] doi:10.1016/j.jsxm.2017.09.010

94. **Loeb S, Folkvaljon Y, Damber JE, et al.** Testosterone replacement therapy and risk of favorable and aggressive prostate cancer. J Clin Oncol. 2017;35:1430-6. [PMID: 28447913] doi:10.1200/JCO.2016.69.5304

95. **Maggi M, Wu FC, Jones TH, et al; RHYME Investigators.** Testosterone treatment is not associated with increased risk of adverse cardiovascular events: results from the Registry of Hypogonadism in Men (RHYME). Int J Clin Pract. 2016;70:843-52. [PMID: 27774779] doi:10.1111/ijcp.12876

96. **McClintock TR, Valovska MI, Kwon NK, et al.** Testosterone replacement therapy is associated with an increased risk of urolithiasis. World J Urol. 2019. [PMID: 30903351] doi:10.1007/s00345-019-02726-6

97. **Saad F, Doros G, Haider KS, et al.** Hypogonadal men with moderate-to-severe lower urinary tract symptoms have a more severe cardiometabolic risk profile and benefit more from testosterone therapy than men with mild lower urinary tract symptoms. Investig Clin Urol. 2018;59:399-409. [PMID: 30402573] doi:10.4111/icu.2018.59.6.399

98. **Sharma R, Oni OA, Chen G, et al.** Association between testosterone replacement therapy and the incidence of DVT and pulmonary embolism: a retrospective cohort study of the Veterans Administration database. Chest. 2016;150:563-71. [PMID: 27179907] doi:10.1016/j.chest.2016.05.007

99. **Sharma R, Oni OA, Gupta K, et al.** Normalization of testosterone level is associated with reduced incidence of myocardial infarction and mortality in men. Eur Heart J. 2015;36:2706-15. [PMID: 26248567] doi:10.1093/eurheartj/ehv346

100. **Shores MM, Smith NL, Forsberg CW, et al.** Testosterone treatment and mortality in men with low testosterone levels. J Clin Endocrinol Metab. 2012;97:2050-8. [PMID: 22496507] doi:10.1210/jc.2011-2591

101. **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;310:1829-36. [PMID: 24193080] doi:10.1001/jama.2013.280386

102. **Wallis CJ, Lo K, Lee Y, et al.** Survival and cardiovascular events in men treated with testosterone replacement therapy: an intention-to-treat observational cohort study. Lancet Diabetes Endocrinol. 2016;4:498-506. [PMID: 27165609] doi:10.1016/S2213-8587(16)00112-1

103. **Walsh TJ, Shores MM, Krakauer CA, et al.** Testosterone treatment and the risk of aggressive prostate cancer in men with low testosterone levels. PLoS One. 2018;13:e0199194. [PMID: 29933385] doi:10.1371/journal.pone.0199194

104. **U.S. Food and Drug Administration.** FDA approves new oral testosterone capsule for treatment of men with certain forms of hypogonadism. 27 March 2019. Accessed at www.fda.gov/news-events/press-announcements/fda-approves-new-oral-testosterone-capsule-treatment-men-certain-forms-hypogonadism on 31 July 2019.

105. **U.S. Food and Drug Administration.** Highlights of prescribing information: Jatenzo. 2019. Accessed at www.accessdata.fda.gov/drugsatfda_docs/label/2019/206089s000lbl.pdf on 29 July 2019.

106. **U.S. Food and Drug Administration.** Highlights of prescribing information: AndroGel. 2013. Accessed at www.accessdata.fda.gov/drugsatfda_docs/label/2013/021015s036lbl.pdf on 29 July 2019.

107. A study to evaluate the effect of testosterone replacement therapy (TRT) on the incidence of major adverse cardiovascular events (MACE) and efficacy measures in hypogonadal men (TRAVERSE). ClinicalTrials.gov; 8 May 2018. Accessed at https://clinicaltrials.gov/ct2/show/record/NCT03518034 on 29 July 2019.

108. **Albert SG, Morley JE.** Testosterone therapy, association with age, initiation and mode of therapy with cardiovascular events: a systematic review. Clin Endocrinol (Oxf). 2016;85:436-43. [PMID: 27124404] doi:10.1111/cen.13084

109. **Algeffari M, Jayasena CN, MacKeith P, et al.** Testosterone therapy for sexual dysfunction in men with type 2 diabetes: a systematic review and meta-analysis of randomized controlled trials. Diabet Med. 2018;35:195-202. [PMID: 29171080] doi:10.1111/dme.13553

110. **Borst SE, Shuster JJ, Zou B, et al.** Cardiovascular risks and elevation of serum DHT vary by route of testosterone administration: a systematic review and meta-analysis. BMC Med. 2014;12:211. [PMID: 25428524] doi:10.1186/s12916-014-0211-5

111. **Boyle P, Koechlin A, Bota M, et al.** Endogenous and exogenous testosterone and the risk of prostate cancer and increased prostate-

specific antigen (PSA) level: a meta-analysis. BJU Int. 2016;118:731-41. [PMID: 26779889] doi:10.1111/bju.13417

112. Corona G, Isidori AM, Buvat J, et al. Testosterone supplementation and sexual function: a meta-analysis study. J Sex Med. 2014; 11:1577-92. [PMID: 24697970] doi:10.1111/jsm.12536

113. Corona G, Rastrelli G, Morgentaler A, et al. Meta-analysis of results of testosterone therapy on sexual function based on International Index of Erectile Function scores. Eur Urol. 2017;72:1000-11. [PMID: 28434676] doi:10.1016/j.eururo.2017.03.032

114. Cui Y, Zong H, Yan H, et al. The effect of testosterone replacement therapy on prostate cancer: a systematic review and meta-analysis. Prostate Cancer Prostatic Dis. 2014;17:132-43. [PMID: 24445948] doi:10.1038/pcan.2013.60

115. Elliott J, Kelly SE, Millar AC, et al. Testosterone therapy in hypogonadal men: a systematic review and network meta-analysis. BMJ Open. 2017;7:e015284. [PMID: 29150464] doi: 10.1136/bmjopen-2016-015284

116. 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: 2560-75. [PMID: 20525906] doi:10.1210/jc.2009-2575

117. 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;82:29-39. [PMID: 17285783]

118. Hua JT, Hildreth KL, Pelak VS. Effects of testosterone therapy on cognitive function in aging: a systematic review. Cogn Behav Neurol. 2016;29:122-38. [PMID: 27662450] doi:10.1097/WNN.0000000000000104

119. Kohn TP, Mata DA, Ramasamy R, et al. Effects of testosterone replacement therapy on lower urinary tract symptoms: a systematic review and meta-analysis. Eur Urol. 2016;69:1083-90. [PMID: 26874809] doi:10.1016/j.eururo.2016.01.043

120. Loo SY, Chen BY, Yu OHY, et al. Testosterone replacement therapy and the risk of stroke in men: a systematic review. Maturitas. 2017;106:31-7. [PMID: 29150164] doi:10.1016/j.maturitas.2017.08.013

121. Nian Y, Ding M, Hu S, et al. Testosterone replacement therapy improves health-related quality of life for patients with late-onset hypogonadism: a meta-analysis of randomized controlled trials. Andrologia. 2017;49. [PMID: 27389320] doi:10.1111/and.12630

122. Ponce OJ, Spencer-Bonilla G, Alvarez-Villalobos N, et al. The efficacy and adverse events of testosterone replacement therapy in hypogonadal men: a systematic review and meta-analysis of randomized, placebo-controlled trials. J Clin Endocrinol Metab. 2018. [PMID: 29562341] doi:10.1210/jc.2018-00404

123. Xu L, Freeman G, Cowling BJ, et al. Testosterone therapy and cardiovascular events among men: a systematic review and meta-analysis of placebo-controlled randomized trials. BMC Med. 2013; 11:108. [PMID: 23597181] doi:10.1186/1741-7015-11-108

124. Dimitropoulos K, Verze P, Van den Broeck T, et al. What are the benefits and harms of testosterone therapy for male sexual dysfunction?—a systematic review. Int J Impot Res. 2019. [PMID: 30918363] doi:10.1038/s41443-019-0131-1

125. Houghton DE, Alsawas M, Barrioneuvo P, et al. Testosterone therapy and venous thromboembolism: a systematic review and meta-analysis. Thromb Res. 2018;172:94-103. [PMID: 30396049] doi:10.1016/j.thromres.2018.10.023

126. Huo S, Scialli AR, McGarvey S, et al. Treatment of men for "low testosterone": a systematic review. PLoS One. 2016;11:e0162480. [PMID: 27655114] doi:10.1371/journal.pone.0162480

Current Author Addresses: Drs. Diem, Greer, Dahm, Ercan-Fang, Estrada, Hemmy, Fink, and Wilt; Mr. MacDonald; Ms. McKenzie; and Ms. Rosebush: One Veterans Drive (111-0), Minneapolis, MN 55417.

Appendix Table 1. Search Strategies

Ovid MEDLINE

1.	exp TESTOSTERONE/ad, ae, df, st, tu, th [Administration & Dosage, Adverse Effects, Deficiency, Standards, Therapeutic Use, Therapy]
2.	"testosterone replacement ".mp.
3.	(testosterone adj2 therapy).mp.
4.	testosterone defic$.mp.
5.	(low adj3 testosterone).mp.
6.	(testosterone adj2 supplement$).mp.
7.	testosterone treatment.mp.
8.	1 or 2 or 3 or 4 or 5 or 6 or 7
9.	limit 8 to (english language and male and humans and yr="1980 -Current
10.	((singl$ or doubl$ or treb$ or tripl$) adj (blind$3 or mask$3)).tw.
11.	(random$ adj (enroll$ or assign$ or allocat$)).mp.
12.	(randomi#ed or placebo$).mp.
13.	exp Randomized Controlled Trial/
14.	trial$.ab.
15.	((randomi#ed or clinical) adj2 trial).mp.
16.	10 or 11 or 12 or 13 or 14 or 15
17.	9 and 16
18.	exp Meta-Analysis/
19.	((meta analy$ or meta-analy$) adj2 review).mp.
20.	(meta analy$ or meta-analy$).mp.
21.	(systematic$ adj2 (review or overview)).mp.
22.	selection criteria.ab.
23.	(data abstraction or data extraction).ab.
24.	exp "REVIEW"/
25.	22 or 23
26.	24 and 25
27.	18 or 19 or 20 or 21 or 26
28.	9 and 27
29.	exp Observational Study/
30.	exp Cohort Studies/
31.	exp Cross-Sectional Studies/
32.	observational stud$.mp.
33.	((follow up or follow-up) adj stud$).mp.
34.	((cross sectional or cross-sectional) adj stud$).mp.
35.	retrospective stud$.mp.
36.	(cohort adj2 stud$).mp. or (cohort adj2 analy$).tw.
37.	29 or 30 or 31 or 32 or 33 or 34 or 35 or 36
38.	9 and 37
39.	(contracepti$ or pregnan$).mp.
40.	exp INFERTILITY/
41.	exp CONTRACEPTION/
42.	39 or 40 or 41
43.	17 not 42
44.	28 not 42
45.	38 not 42

Embase

1.	(testosterone adj2 (deficiency or therapy or treatment or supplement$)).mp
2.	exp androgen deficiency/co, dm, dt, si [Complication, Disease Management, Drug Therapy, Side Effect]
3.	exp androgen therapy/
4.	(low adj3 testosterone).mp.
5.	exp testosterone/ae, ad, do, dt, th [Adverse Drug Reaction, Drug Administration, Drug Dose, Drug Therapy, Therapy]
6.	1 or 2 or 3 or 4 or 5
7.	limit 6 to (human and male and english language and yr="1980 -Current")
8.	((single blind or double blind) adj procedure).mp.
9.	exp placebo/ct, th [Clinical Trial, Therapy]

Appendix Table 1–Continued

10.	(random$ adj (enroll$ or assign$ or allocat$)).mp.
11.	randomi?ed.mp.
12.	((randomi?ed or clinical) adj2 trial$).mp.
13.	trial$.ab.
14.	8 or 9 or 10 or 11 or 12 or 13
15.	7 and 14
16.	exp "systematic review"/ or "systematic review (topic)"/
17.	((meta analy$ or meta-analy$) adj2 review).mp.
18.	(meta analy$ or meta-analy$).mp.
19.	(systematic$ adj2 (review or overview)).mp.
20.	16 or 17 or 18 or 19
21.	7 and 20
22.	((cohort or observational) adj stud$).mp.
23.	((cross sectional or cross-sectional) adj stud$).mp.
24.	retrospective stud$.mp.
25.	(cohort adj analy$).mp.
26.	22 or 23 or 24 or 25
27.	7 and 26
28.	(contracept$ or pregnan$).mp.
29.	exp infertility/ or exp male infertility/
30.	28 or 29
31.	15 not 30
32.	21 not 30
33.	27 not 30
34.	(31 not 32) and (31 not 33)
35.	(32 not 31) and (32 not 33)
36.	(33 not 31) and (33 not 32)

Cochrane Library

1.	MeSH descriptor: [Testosterone]: [Administration & dosage - AD, Adverse effects - AE, Deficiency - DF, Standards - ST, Therapeutic use - TU]
2.	testosterone replacement:ti,ab,kw
3.	testosterone near/2 therapy:ti,ab,kw
4.	testosterone defici*:ti,ab,kw
5.	low near/3 testosterone:ti,ab,kw
6.	testosterone near/2 supplement*:ti,ab,kw
7.	testosterone treatment:ti,ab,kw
8.	{or #1-#7} Publication Year from 1980 to 2018
9.	(men or male*):ti,ab,kw
10.	MeSH descriptor: [Male]
11.	(#9 or #10)
12.	(#8 and #11)
13.	(randomi?ed) or (placebo?):ab
14.	randomized controlled trial:pt
15.	(#13 or #14)
16.	(#12 and #15)
17.	(contracept* or pregnan*):ti,ab,kw
18.	MeSH descriptor: [infertility] 1 tree(s) exploded
19.	(#17 or #18)
20.	(#12 not #19)
21.	(#16 not #19)
22.	(#20 and #21)
23.	not embase:an
24.	not pubmed:an
25.	(#22 or #23)
26.	(#20 and #24)

Measure	Domains	Range of Scores and Scoring System	Minimal Clinical Important Difference	Number of Items	Item Description
Age-Related Hormone Deficiency–Dependent Quality of Life (A-RHDQoL) (127)	1. Family life 2. Social life 3. Work 4. Health 5. Physical appearance 6. Physical capabilities 7. Physical stamina 8. Energy 9. Sex life 10. Sleep 11. Bodily pain 12. Stress tolerance 13. Memory 14. Concentration 15. Travel 16. Holidays and leisure 17. Household tasks 18. Confidence 19. Motivation 20. Society's reaction 21. Worry about future	Domain impact rating: −3 (much better) to 3 (much worse) Domain importance rating: 0 (not important) to 3 (very important Weighted impact: −9 (maximum negative impact) to 9 (maximum positive impact) per domain	NR	42 total (2 items per domain)	Higher indicates better. Each domain is introduced with a hypothetical statement followed by rating of importance. Weighted domain impact score is obtained by multiplying impact rating by corresponding importance rating. Abridged versions include the 21-domain Age-Related Hormone Deficiency–Dependent Quality of Life. Domains determined to be "not applicable" (i.e., family life, work, sex life, and bodily pain) to the patient are not included in weighted score.

Aging Males' Symptoms (AMS) (128)	1. Psychological
	2. Somatovegetative (or somatic)
	3. Sexual
		Scoring system: 1 (no symptoms) to 5 (very severe)
		Overall range, 17–85
		Impairment:
		  ≤26 = no impairment
		  27–36 = mild
		  37–49 = moderate
		  ≥50 = severe
		Psychological range, 5–25
		Impairment:
		  9–12 = moderate
		  ≥12 = severe
		Somatovegetative (or somatic) range, 7–35
		Impairment:
		  13–18 = moderate
		  ≥19 = severe
		Sexual range, 5–25
		Impairment:
		  8–10 = moderate
		  ≥11 = severe
			NR	17 total
				Psychological: 5
				Somatovegetative (or somatic): 7
				Sexual: 5
					Higher indicates worse. Psychological subscale items: discouragement, depression, irritability, anxiety, and nervousness. Somatovegetative (or somatic) subscale items: pain in muscles or joints, sweating (hot flushes), increased need for sleep and sleep disturbances, impaired general well-being, decreased muscular strength, and decreased energy. Sexual subscale items: disturbances in potency, decrease in morning erections, decrease in libido and sexual activity, decrease in beard growth, and the "impression of having passed the zenith of life."

Measure	Domains	Range of Scores and Scoring System	Minimal Clinical Important Difference	Number of Items	Item Description
Arizona Sexual Experience (ASE) (129)	1. Sex drive 2. Arousal 3. Vaginal lubrication/penile secretion 4. Ability to reach orgasm 5. Satisfaction from orgasm	Scoring system: 1 (hyperfunction) to 6 (hypofunction) Overall range, 5–30 (1–6 per domain) Impairment: Total score ≥19, any 1 item with score ≥5, or 3 items with a score ≥4 = dysfunction	NR	5 total (1 item per domain)	Higher indicates worse. Sex-specific versions of scales. Designed to be self- or clinician-administered.
Beck Depression Inventory (BDI) (41, 130)	1. Mood 2. Pessimism 3. Sense of failure 4. Lack of satisfaction 5. Guilty feeling 6. Sense of punishment 7. Self-hate 8. Self-accusations 9. Self-punitive wishes 10. Crying spells 11. Irritability 12. Social withdrawal 13. Indecisiveness 14. Body image 15. Work inhibition 16. Sleep disturbance 17. Fatigability 18. Loss of appetite 19. Weight loss 20. Somatic preoccupation 21. Loss of libido	Scoring system: 0 (better) to 3 (worse) Overall range, 0–63 Impairment: 10–16 = mild depression 17–29 = moderate depression 30–63 = severe depression	NR	21 total (1 item per domain)	Higher indicates worse. The Beck Depression Inventory-IA is a revised version that assesses depressive symptoms with scores ≥10 indicative of mild depressive states.
Derogatis Interview for Sexual Functioning in Men–II sexual desire (DISF-M-II) (71, 82, 131)	Sexual desire	Scoring system: 4 questions scored 0 (low) to 7 (high) and 1 question scored 0 (low) to 5 (high) Range, 0–33 Impairment: ≥20 = decreased libido and desire	A change of ≥5 based on regression, ROC, and empirical CDF anchor-based analyses from the Testosterone Trials results	5 total	Higher indicates better. The Derogatis Interview for Sexual Functioning in Men–II is an updated version of the 25-item interview consisting of 5 domains (sexual cognition/fantasy, sexual arousal, sexual behavior/experience, orgasm, and sexual drive/relationship). Scale can be administered or based on self-report. Distinct gender-keyed versions are also available.
Functional Assessment of Chronic Illness Therapy (FACIT) Fatigue Scale (132)	Fatigue/exhaustion	Scoring system: 0 (very fatigued) to 5 (not fatigued) Range, 0–52 Impairment: <30 = severe	NR	13 total	Higher indicates better. The Functional Assessment of Chronic Illness Therapy Fatigue Scale is one of many different scales that are part of a collection of health-related quality of life questionnaires targeted to the management of chronic illness. Items 7 and 8 on the scale are reverse-scored.

Measure	Domains	Range of Scores and Scoring System	Minimal Clinical Important Difference	Number of Items	Item Description
Geriatric Depression Scale (GDS) (33, 133)	Depression	Scoring system: 1 point per bolded question on scale Range, 0–30	NR	Geriatric Depression Scale: 30 total (1 item per domain) Geriatric Depression Scale–Short Form: 15 total (1 item per domain)	Higher indicates worse. Adapted short form used to assess depressive scores (range, 0–15; ≥5 indicates depression). Questions refer to feelings in past week.

Hospital Anxiety and Depression Scale (HADS) (43, 134)	1. Anxiety
	2. Depression
		Scoring system, anxiety: 0 (not at all/only occasionally) to 3 (often)
		Scoring system, depression: 0 (often/very much) to 3 (hardly/not at all)
		Overall range, 0–42
		Anxiety range, 0–21
		Impairment:
		  8–10 = borderline abnormal
		  ≥11 = abnormal
		Depression range, 0–21
		Impairment:
		  8–10 = borderline abnormal
		  ≥11 = abnormal
			NR	14 total (7 items per domain)	Higher indicates worse. One question each on the anxiety ("I can sit at ease and feel relaxed") and depression ("I feel as if I am slowed down") scales are reverse-scored.

International Index of Erectile Function (IIEF) (26, 82, 135)	1. Erectile function 2. Orgasmic function 3. Sexual desire 4. Intercourse satisfaction 5. Overall satisfaction	Scoring system: 0/1 (severe symptoms) to 5 (mild/no symptoms) Overall range, 4 (or 5)–75 Erectile function range, 0 (or 1)–30 Orgasmic function range, 0–10 Sexual desire range, 2–10 Intercourse satisfaction range, 0–15 Overall satisfaction range, 2–10	The MCID for the erectile function domain is 4 points using an anchor response to question 7: "over the past 4 weeks, when you attempted sexual intercourse how often was it satisfactory for you?"	IIEF: 15 total Erectile function: 6 Orgasmic function: 2 Sexual desire: 2 Intercourse satisfaction: 3 Overall satisfaction: 2 IIEF-5: 5 total IIEF-ED: 6 total	Higher indicates better. Abridged versions include IIEF-5 (5 items focused on erectile function and intercourse satisfaction; range, 5–25) and IIEF-ED (6 items from erectile function domain only; range, 0 [or 1]–30).
Medical Outcomes Study Short Form-36 Health Survey (SF-36) (27, 136)	1. Physical function 2. Bodily pain 3. Vitality 4. Role limitations due to physical problems 5. General health perceptions 6. Emotional well-being/general mental health 7. Social function 8. Role limitations	Scoring system: 0 (severe problems/limited mobility) to 100 (no problems/normal functioning) Range, 0–100 per domain	The MCID for various domains has been determined to be between 3 and 5 for various conditions; the Testosterone Trials measured the percentage of men whose physical function domain score increased by 8 points	36 total Physical function: 10 Bodily pain: 2 Vitality: 4 Role limitations due to physical problems: 3 General health perceptions: 5 Emotional well-being/general mental health: 5 Social function: 2 Role limitations due to emotional problems: 3	Higher indicates better. Each domain scored separately. Abridged versions include the Short Form-12 and Short Form-20.

Continued on following page

Measure	Domains	Range of Scores and Scoring System	Minimal Clinical Important Difference	Number of Items	Item Description
Patient Health Questionnaire-9 (PHQ-9) (137)	Depression	Scoring system: 0 (not at all) to 3 (nearly every day) Range, 0–27 total Depression: ≥ 10–14 = moderate ≥20 = severe	NR	9 total	Higher indicates worse. Scale consists of 9 questions that incorporate criteria from the *Diagnostic and Statistical Manual of Mental Disorders, Fourth Edition*, with other leading major depressive symptoms. Questions ask about problems that include trouble sleeping, energy level, appetite or overeating, feelings of self, concentration, restlessness, and suicidal ideation.
Physical Activity Scale for the Elderly (PASE) (138, 139)	1. Muscle strength/endurance 2. Strenuous sports 3. Moderate sports 4. Light sports 5. Job involving standing or walking 6. Lawn work or yard care 7. Caring for another person 8. Home repairs 9. Heavy housework 10. Light housework 11. Outdoor gardening 12. Walking	Scoring system: 6 items are scored on the basis of hours per day; an additional 6 items are scored 0 (did not engage in activity) to 1 (engaged in activity) over 7-day period Overall range, 0–793	NR	12 total (1 item per domain)	Higher indicates better. Designed for individuals aged ≥65 y. Activity weights range from 20 to 36. Scores may vary seasonally.
Positive and Negative Affect Schedule (PANAS) (140)	1. Positive affect 2. Negative affect	Scoring system: 1 (very slightly/not at all) to 5 (extremely) Overall range, 20–100 (10–50 per domain)	NR	20 total terms (10 per affect)	Higher indicates greater intensity of affect. Scales are rated by time frames (moment, today, past few days, past few weeks, year, and general). Positive affect reflects the extent to which a person feels enthusiastic, active, and alert. Positive affect descriptive terms include attentive, interested, alert, excited, enthusiastic, inspired, proud, determined, strong, and active. Negative affect is a general dimension of subjective distress and unpleasurable engagement that subsumes a variety of aversive mood states, including anger, contempt, disgust, guilt, fear, and nervousness.

Measure	Domains	Range of Scores and Scoring System	Minimal Clinical Important Difference	Number of Items	Item Description
					Negative affect terms include distressed, upset, hostile, irritable, scared, afraid, ashamed, guilty, nervous, and jittery.
Profile of Mood States-Depression (POMS) (33, 141)	Depression	Scoring system: 0 (not at all) to 4 (extremely) Overall range, 0–60 Depression range, 0–60	NR	15 total	Higher indicates worse. Scale measures current and past week experiences. Additional domains include tension/ anxiety, anger/ hostility, vigor/ activity, fatigue, and confusion/ bewilderment.
Psychosocial Daily Questionnaire (PDQ) question 4 (82, 142)	Sexual activity	Scoring system: 0 (none) to 1 (any); weekly value averaged Range, 0–12	A change of ≥0.6 based on regression, ROC, and empirical CDF anchor-based analyses from the Testosterone Trials results	12 total	Higher indicates greater number of activities. Question 4: For all of the items listed, check yes if you have experienced (or are experiencing) today (otherwise check no): sexual daydreams, anticipation of sex, sexual interactions with partner, flirting (by you), orgasm, flirting (by others toward you), ejaculation, intercourse, masturbation, night spontaneous erection, day spontaneous erection, or erection in response to sexual activity. Additional domains include sexual desire and performance.
Well-Being Questionnaire (W-BQ12) (143, 144)	1. Negative well-being 2. Energy 3. Positive well-being	Scoring system: 0 (less) to 12 (more) Overall range, 0–36 (0–12 per domain)	NR	12 total (4 items per domain)	Higher indicates worse (negative well-being); higher indicates better (energy, positive well-being, and general well-being) Adapted from the 22-item Well-Being Questionnaire (W-BQ22). Domains scored separately and summed to produce a general well-being score (0–36).

CDF = cumulative distribution function; IIEF = International Index of Erectile Function; IIEF-5 = 5-item version of IIEF; IIEF-ED = IIEF erectile dysfunction domain; MCID = minimal clinically important differences; NR = not reported; ROC = receiver-operating characteristics.

Appendix Table 3. Cognitive Outcomes

Global: screening tests
- MMSE
- Modified Mini-Mental State (3MS)

Global: multidomain tests
- ADAS-Cog
- Composite cognitive score from battery

Attention
- Reaction Time, Simple
- Reaction Time, Choice
- Digit Span Forward or Total
- Trail Making Test A
- Stroop Color +/or Word
- Knox Cube Imitation Test
- Letter Cancellation

Language/verbal abilities
- Verbal Fluency, Semantic/Category
- Verbal Fluency, Phonemic

Memory: verbal
- Buschke Selective Reminding Test: delayed recall (30 min)
- Rey Auditory Verbal Learning Test (RAVLT): delayed
- Prose/Story Recall: delayed recall (20/30 min)
- WMS-R Verbal Paired Associate: immediate and delayed

Memory: visuospatial
- Benton Visual Retention Test (BVRT)
- Puget Sound Route Learning Test: delayed recall
- Rey O Complex Figure Test (RCFT): delayed recall (30 min)
- WMS-R Visual Paired Associate: immediate and delayed
- Visual Spatial Learning Test (VSLT): delayed recall (30 min)

Visuospatial
- Rey O Complex Figure Test (RCFT): figure copy
- Benton Judgement of Line Orientation (JLO)
- Block Design: construction task
- Card Rotation Test
- Mental Rotation
- Clock Drawing Test (CDT)
- Pentagon Figure Drawing from MMSE
- Picture Completion

Executive
- Digit Span Backward
- Trail Making Test B or B-A

ADAS-Cog = Alzheimer's Disease Assessment Scale cognitive subscale; MMSE = Mini-Mental State Examination; WMS-R = Wechsler Memory Scale Revised.

Web-Only References

127. McMillan CV, Bradley C, Giannoulis M, et al. Preliminary development of a new individualised questionnaire measuring quality of life in older men with age-related hormonal decline: the A-RHDQoL. Health Qual Life Outcomes. 2003;1:51. [PMID: 14613571]

128. Heinemann LAJ, Zimmermann T, Vermeulen A, et al. A new 'aging males' symptoms' rating scale. Aging Male. 1999;2:105-14.

129. McGahuey CA, Gelenberg AJ, Laukes CA, et al. The Arizona Sexual Experience Scale (ASEX): reliability and validity. J Sex Marital Ther. 2000;26:25-40. [PMID: 10693114]

130. Beck AT, Ward CH, Mendelson M, et al. An inventory for measuring depression. Arch Gen Psychiatry. 1961;4:561-71. [PMID: 13688369]

131. Derogatis LR. The Derogatis Interview for Sexual Functioning (DISF/DISF-SR): an introductory report. J Sex Marital Ther. 1997;23:291-304. [PMID: 9427208]

132. Webster K, Cella D, Yost K. The Functional Assessment of Chronic Illness Therapy (FACIT) measurement system: properties, applications, and interpretation. Health Qual Life Outcomes. 2003;1:79. [PMID: 14678568]

133. Yesavage JA, Brink TL, Rose TL, et al. Development and validation of a geriatric depression screening scale: a preliminary report. J Psychiatr Res. 1982-1983;17:37-49. [PMID: 7183759]

134. Zigmond AS, Snaith RP. The Hospital Anxiety and Depression Scale. Acta Psychiatr Scand. 1983;67:361-70. [PMID: 6880820]

135. Rosen RC, Riley A, Wagner G, et al. The International Index of Erectile Function (IIEF): a multidimensional scale for assessment of erectile dysfunction. Urology. 1997;49:822-30. [PMID: 9187685]

136. Ware JE Jr, Sherbourne CD. The MOS 36-item short-form health survey (SF-36). I. Conceptual framework and item selection. Med Care. 1992;30:473-83. [PMID: 1593914]

137. Kroenke K, Spitzer RL, Williams JB. The PHQ-9: validity of a brief depression severity measure. J Gen Intern Med. 2001;16:606-13. [PMID: 11556941]

138. Logan SL, Gottlieb BH, Maitland SB, et al. The Physical Activity Scale for the Elderly (PASE) questionnaire; does it predict physical health? Int J Environ Res Public Health. 2013;10:3967-86. [PMID: 23999546] doi:10.3390/ijerph10093967

139. Washburn RA, Smith KW, Jette AM, et al. The Physical Activity Scale for the Elderly (PASE): development and evaluation. J Clin Epidemiol. 1993;46:153-62. [PMID: 8437031]

140. Watson D, Clark LA, Tellegen A. Development and validation of brief measures of positive and negative affect: the PANAS scales. J Pers Soc Psychol. 1988;54:1063-70. [PMID: 3397865]

141. Haythornthwaite JA, Edwards RR. Profile of Mood States (POMS). 2004. Accessed at www.immpact.org/static/meetings/Immpact4/background/Haythornthwaite.pdf on 23 July 2019.

142. Lee KK, Berman N, Alexander GM, et al. A simple self-report diary for assessing psychosexual function in hypogonadal men. J Androl. 2003;24:688-98. [PMID: 12954659]

143. Bradley C. The W-BQ12. Health Psychology Research; 2009. Accessed at www.healthpsychologyresearch.com/sites/default/files/summaries/w-bq12%20summary%20rev15.9.09.pdf on 23 July 2019.

144. Pouwer F, van der Ploeg HM, Adèr HJ, et al. The 12-item wellbeing questionnaire. An evaluation of its validity and reliability in Dutch people with diabetes. Diabetes Care. 1999;22:2004-10. [PMID: 10587834]