HHS Public Access

Author manuscript

Cogn Behav Neurol. Author manuscript; available in PMC 2017 September 01.

Published in final edited form as:

Cogn Behav Neurol. 2016 September ; 29(3): 122–138. doi:10.1097/WNN.0000000000000104.

Effects of Testosterone Therapy on Cognitive Function in Aging: A Systematic Review

Jeremy T. Hua, MD[*,†], **Kerry L. Hildreth, MD**[†], and **Victoria S. Pelak, MD**[*,‡]

[*]Department of Neurology, University of Colorado School of Medicine, Aurora, Colorado

[†]Department of Medicine, University of Colorado School of Medicine, Aurora, Colorado

[‡]Department of Ophthalmology, University of Colorado School of Medicine, Aurora, Colorado

Abstract

Endogenous testosterone in the aging man has been scrutinized extensively in regard to its effects on performance in many cognitive domains, especially verbal fluency, visuospatial and visuoperceptual abilities, memory, and executive function. Studies of testosterone supplementation have sought to identify potential cognitive improvements in men with and without baseline cognitive impairment, and have had a wide range of results. The variability in outcomes is likely related in part to the lack of consensus on methods for testosterone measurement and supplementation and in part to the disparate measures of cognitive function used in randomized controlled studies. Despite the limitations imposed by such inconsistent methods, promising associations have been found between cognition and testosterone supplementation in both eugonadal men and men with low testosterone levels, with and without baseline cognitive dysfunction. This systematic review highlights the cognitive measures used in and the outcomes of existing studies of testosterone and cognition in aging men. The review suggests that larger studies and a more standardized approach to assessment will be needed before we can fully understand and realize sustained benefits from testosterone supplementation in the elderly male population, particularly given the substantial increase in testosterone supplementation in clinical practice.

Keywords

testosterone; aging; cognition; visuospatial; dementia

From the first, investigations into testosterone (T) and cognition showed that men with low levels of endogenous T perform below normal on tests of verbal fluency (Alexander et al, 1998), visuospatial abilities (Hier and Crowley, 1982), memory (Barrett-Connor et al, 1999; Moffat et al, 2002), executive function (Muller et al, 2005), and attention (Cappa et al, 1988). Studies also showed that T supplementation in men with low T levels and/or hypogonadism (a condition of low T levels; see description in “Gonadal State” below) may improve these cognitive functions (Cherrier et al, 2003; Kenny et al, 2002; Vaughan et al,

Corresponding Author: Victoria S. Pelak, MD, Mail Stop B 185, Academic Office 1, Room 5217, 12631 E. 17th Avenue, Aurora, CO 80045, victoria.pelak@ucdenver.edu, 303-724-2184.

The authors declare no conflicts of interest.

2007). In some studies, nonlinear associations between T level and cognition have suggested that there may be a T level at which cognitive performance is optimally enhanced (Barrett-Connor et al, 1999; Hogervorst et al, 2010; Martin et al, 2007; Matousek and Sherwin, 2010; Muller et al, 2005).

Since men over the age of 40 years have a 1.6% natural decline per year in their total T, these results led to further studies focusing on the relationship among the aging man, T levels, and cognition (Feldman et al, 2002). Although controversial, some evidence shows that healthy men *without* hypogonadism and *without* cognitive impairment may derive cognitive benefits from T supplementation, as shown by tests of attention and executive function (Kenny et al, 2002; Vaughan et al, 2007), visuospatial and visuoperceptual function (Cherrier et al, 2001; Gray et al, 2005; Janowsky et al, 1994), and memory (Cherrier et al, 2001; Janowsky et al, 2000).

Further research into the relationship between T and age- or disease-related cognitive decline (Pike et al, 2009) has not always yielded consistent or generalizable results. This may result, in part, from the participants' variable baseline cognitive function and baseline T levels (low T versus eugonadal), and whether T supplementation or just T levels were tested.

As of the February 2016 completion date of our search, no human study had elucidated the brain mechanisms underlying the cognitive changes potentially associated with endogenous T production or with exogenous T supplementation. Animal studies, however, implicate the numerous androgen receptors in the hippocampus, thalamus, and deep layers of the cerebral cortex, as well as nerve growth factor proteins in the hippocampus (Bimonte-Nelson et al, 2003).

Two notable recent reviews of human research are Holland et al's (2011) review of potential mechanisms of cognitive preservation by sex hormones in men, and Hogervorst's (2013) report on studies of the effects of gonadal hormones on cognitive behavior in both men and women.

Our systematic review, performed using the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) (Moher et al, 2009), adds to this body of literature by focusing on the current state of knowledge about cognition and T in older men and the potential role of T supplementation in age-related cognitive disorders. We organized the information by grouping together studies whose populations were similar in baseline T states, cognitive status, and whether or not T supplementation was investigated.

Part of the challenge in evaluating the literature on this topic is the studies' use of disparate neuropsychological measures to assess different aspects of cognition and cognitive domains. For instance, generalized statements about T and memory without specific distinctions about the memory function tested make it difficult to summarize how T affects the memory domain.

To overcome the challenges of differences in study participants and aspects of the cognitive domains tested, we report the literature with a focus on participant characteristics. To aid in interpreting study results, we provide tables listing the neuropsychological tests used in each

study. We also present factors to consider when interpreting findings and we make recommendations for the future direction of studies on T and cognition in the aging male patient.

Testosterone and Cognition: The Impact of Study Methodology

Although many studies of T and cognition have shown positive associations, others have shown contradictory and variable results. These differences may result from variations in study design:

- Participant characteristics, eg, age, gonadal state, neurologic disorders such as Alzheimer disease
- Method of assessing T level, eg, free T, bioavailable T, total T
- Interventions, eg, T supplementation versus no supplementation
- Methods used to assess cognition

For studies involving T supplementation, additional variables that make it difficult to compare results across studies include the administration route, the dose, and the T level achieved. The methods used to study T and cognition have a significant impact on cognitive outcomes because of important characteristics of T biology (reviewed below) and because these factors exacerbate the difficulty of comparing one study to another.

Before reviewing individual studies, we will illustrate how differences in gonadal state, supplementation, assessment of T levels, and cognitive assessments limit our ability to make generalizations about T and cognition.

Gonadal State

The inconsistent terminology that different authors use in describing gonadal states complicates the interpretation and comparison of studies. The Endocrine Society defines hypogonadism in men as a syndrome resulting from “failure of the testes to produce physiological levels of testosterone (androgen deficiency) and a normal number of spermatozoa due to disruption of one or more levels of the hypothalamic-pituitary-testicular axis” (Bhasin et al, 2010). A diagnosis of hypogonadism requires both consistently and unequivocally low T levels and symptoms attributable to low T. This can be problematic in older men, whose T levels may range from slightly below normal to within the lower range of normal. Many of the symptoms of low T are common in older men and have multiple causes that may or may not relate to low T levels.

Many studies use the term *low T* to describe the participants’ gonadal state. Because each reference laboratory determines its own “normal” T range, the same nominal T level may not be equivalent across studies. Furthermore, the T threshold below which men develop symptoms varies between individuals and among different symptoms (Bhasin et al, 2010). This may also hold true for different cognitive functions. In general, symptoms are more likely at a total T (TT) level below the lower limit of normal for young men (~300 ng/dL), but interpreting or comparing studies of T and cognition requires careful attention to how

each study defines hypogonadism or low T relative to the reference laboratory's defined range.

Testosterone Supplementation

The different routes by which T can be supplemented—oral, intramuscular injection, oral or buccal application, intranasal spray, implantable pellets, and transdermal patches or gels—can influence serum levels. Oral preparations are rarely used now because first pass metabolism by the liver worsens side effects and does little to raise T levels. While oral T undecanoate avoids these problems through lymphatic absorption, both food and high intra- and inter-individual variability in absorption can severely reduce delivery; this form is not available in the US.

In the US, the most commonly used routes are intramuscular injection and transdermal gel or patch. Injections produce strong fluctuations in T levels, with a spike immediately after an injection and a nadir immediately before the next injection. Transdermal preparations allow more stable physiologic T levels and more closely mimic the circadian fluctuations of endogenous T, although these rhythms are significantly attenuated with age (Bremner et al, 1983). Importantly, transdermal absorption can be highly variable even with proper application.

T levels are also influenced by dosage, formulation, timing, and duration of administration. For instance, T levels may fluctuate to below baseline when a man uses a scrotal patch in a 12-hours-on and 12-hours-off regimen. In addition, exogenous T can interfere with endogenous T production by altering feedback pathways. T is produced endogenously through conversion from dihydroepiandrosterone and can be inactivated in the blood when bound to sex hormone binding globulin (SHBG). T is broken down to form several byproducts, notably, estradiol by aromatase, and dihydrotestosterone (DHT) by 5-alpha-reductase. In this environment of dynamic production and degradation, understanding the true impact of supplementation on cognition can be quite difficult (Valenti and Schwartz, 2008).

The ideal way to conduct any androgen-related research would be to monitor the levels of T precursors, byproducts, and interacting proteins in order to infer causality and fully understand the direct versus indirect influences that T and T supplementation have on cognition. While some studies have tried to evaluate T therapy by measuring or influencing these complex pathways, such as investigating the difference between T and DHT supplementation (Cherrier et al, 2003) or adding a 5-alpha-reductase inhibitor (Vaughan et al, 2007), the complex endocrine environment of T supplementation has not been completely explained.

Assessment of Testosterone Levels

Like the many ways that T can be supplemented, there are many ways that T levels can be measured and interpreted. TT levels consist of both *unbound* and *bound* T, both of which can be determined by a serum test. The best measure of biologically functional T is not TT, but rather the combination of free T (FT) and bioavailable T (BT). When T enters the bloodstream, a small percent of it stays unbound as FT while the rest either interacts weakly

with albumin or is inactivated through binding with SHBG. Both FT and T bound to albumin constitute BT (Martin et al, 2007). Most T studies measure TT from serum and then calculate either FT or BT using a formula.

Measurement of serum TT has become more convenient with the development of commercial radioimmunoassay kits and automated chemiluminescence assays. Although most of these assays are reasonably accurate within the normal adult male reference range established by the individual laboratory, they might be significantly less precise and accurate in populations with low T (Wang et al, 2004). Radioimmunoassay is often used to measure free T directly, but this technique has been criticized for being overly influenced by TT and SHBG and for discrepancies between radioimmunoassay and the gold standard method of equilibrium dialysis (Fritz et al, 2008). Radioimmunoassay for FT should be performed by a reliable reference laboratory for the most accurate results (Bhasin et al, 2010). Salivary measurement of FT is convenient, but binding of FT to salivary proteins may limit accuracy (Fiers et al, 2014).

Time of testing matters in evaluating T levels. Serum total T concentration declines progressively after age 40, and about 30% of men over age 70 have levels below the normal range for younger men (Feldman et al, 2002; Kaufman and Vermeulen, 1997; Matsumoto, 1993). The time of day the sample is taken is also critical. T has diurnal fluctuation, with levels peaking in the morning and falling throughout the day and into the evening; as noted, however, these fluctuations are attenuated in older men (Bremner et al, 1983). Most studies take blood samples in the morning. However, some data suggest that sleep/wake time has a stronger influence on T levels than circadian rhythms (Axelsson et al, 2005). Both the quality and quantity of sleep may affect T levels (Penev, 2007). Such detailed information about sample collection is rarely available in published studies, further evidence of the difficulty of interpreting study results given the variability in assessment of T levels.

Measurement of Cognitive Function

Many of the tests available to measure cognition focus on global cognitive function and/or the cognitive domains of attention, executive function, memory, visuospatial and visuoperceptual ability, and/or language. Many tests overlap several domains, but whether a test reflects function in more than one domain, and in which domains, is not always agreed upon. Unless the same tests and subtests are used to measure cognition, it is difficult to compare cognitive outcomes between studies and to make domain-specific assertions about T effects. This is especially true across age groups and in older men since some tests become less sensitive to small, but meaningful, cognitive changes that may occur with the decline in T in older age.

In Table 1 we index most of the cognitive tasks used in the studies of T and cognition that we review in this paper. (A few visuospatial function tests mentioned in Table 4 are not indexed because they are not often used in clinical practice.) Because the studies used different versions of the cognitive tasks, Table 1 cites only the best or most recent reference for each task. For the same reason, we do not cite references for individual tasks within the text. Interested readers should check the individual studies to learn which version of each task was given.

METHODS

Eligibility Criteria for Studies in This Review

We used PRISMA as a guide in conducting this work. We modeled our methodology on other systematic T reviews, such as those by Fernandez-Balsells et al (2010) and Oskui et al (2013).

In this systematic review, we defined our outcome of interest as cognitive function, as measured by neurocognitive tasks of memory, attention, executive function, visuospatial function, and global function. We considered only studies published in the English language.

Eligible studies enrolled older men (average age > 50 years), evaluated our outcome of interest, and met one of these four distinct criteria:

- Observational studies that measured T level
- Randomized controlled trials that enrolled men with *low* T level and *no* known cognitive deficits, and treated the men with T for at least 4 weeks
- Randomized controlled trials that enrolled men with *normal* T level and *with* known cognitive deficits, and treated the men with T for at least 4 weeks
- Randomized controlled trials that enrolled men with *low* T level and *with* known cognitive deficits, and treated the men with T for at least 4 weeks

We excluded studies if they did not report the outcome of interest, if they treated participants with androgens other than T, or if they did not include a non-T control group.

Data Collection

Authors J.T.H. and V.S.P. conducted the search with the MEDLINE and Embase® electronic databases. We searched for articles published between January 1995 and February 2016, and examined each paper and its reference list. The details of our strategy are available on request. We extracted the following data from each study:

- Description of participants: age, presence or absence of known cognitive deficits
- T measurement method: TT, FT, or BT
- T treatment regimen: formulation, dose, frequency, and duration

We contacted study authors to request missing data or seek clarification.

Quality Assessment

We conducted a critical appraisal of included studies to determine the quality of their methods. We considered the following standards:

- Pre-study exclusion of candidates currently receiving treatments that could affect T level and interactions

- Method of T measurement, particularly whether serum T levels were obtained in the morning
- Adequate participant follow-up for the outcome measure of documented cognitive function
- Adequate duration of follow-up
- Randomization of participant groups
- Blinding of both participants and study personnel to treatments
- Funding source for the study

RESULTS

Search Results

Our systematic literature search, depicted in Figure 1, identified 865 potential studies, of which we excluded 804 after screening the abstracts. We reviewed the full text of the 61 remaining articles and excluded 32 for not meeting our full eligibility criteria. We disqualified many of the studies because they failed to meet all of our quality standards. For example, we excluded randomized controlled studies of T supplementation in healthy older men if the participants were not explicitly screened for a low T state.

We analyzed a total of 29 studies in total for this review.

Cognition and Testosterone Levels in Healthy Aging Men

Table 2 presents results on 20 studies that evaluate the relationship between T and cognition in healthy older men. The table makes clear the differences in both methodology and findings among the studies.

Most studies on the relationship between T levels and cognition have included a mix of eugonadal men and men with a low T level, while very few studies have looked solely at cognition in men with a low T or hypogonadal state. Thus, it is difficult to generalize about gonadal state and cognition; however, our review of studies with mixed populations shows evidence for associations between cognitive performance and endogenous T levels.

Global cognition has been evaluated with the Mini-Mental State Examination (MMSE) in a number of studies with mixed results (Barrett-Connor et al, 1999; Hogervorst et al, 2010; LeBlanc et al, 2010; Lessov-Schlaggar et al, 2005; Moffat et al, 2002; Perry et al, 2001; Yaffe et al, 2007; Yeap et al, 2008). Although the MMSE has been widely used to assess global cognition, the measure is not sensitive to subtle changes in cognition, particularly in healthy aging and community-dwelling people who are not approaching dementia (Gluhm et al, 2013). Nonetheless, several studies report on the relationship between the MMSE and T. One of the largest cross-sectional studies of T and global cognition (Yeap et al, 2008) evaluated 2932 community-dwelling older men and identified a weak association between higher FT level and higher MMSE score (Spearman rho 0.06, $P = 0.001$); the men scoring in the highest MMSE quintile had higher FT levels than those in the lowest quintile (FT 278 versus 262 pmol/L, $P = 0.003$).

Studies with longitudinal analyses provide insight into cognitive change over time. Hogervorst et al (2010) gave 257 cognitively intact older men a baseline MMSE and tested their TT levels. The authors found a non-linear relationship, with lower MMSE scores in the participants who had TT levels above or below an optimal level. The authors followed the men for 2 years and then repeated the MMSE. After correcting for age and SHBG level, the authors found less cognitive decline (with decline defined as a drop of ≥ 4 points on the MMSE) in the men who had had a higher baseline TT (odds ratio = 0.93; 95% confidence interval = 0.87, 0.99), suggesting perhaps a more nuanced relationship.

By contrast, large longitudinal studies from LeBlanc et al (2010), Moffat et al (2002), and Lessov-Schlaggar et al (2005), with follow-up of 4.5, 9.7, and 10 to 16 years, respectively, showed no significant association between T levels and change over time in MMSE scores.

A positive relationship between FT and visuospatial or visuoperceptual function has been noted on cognitive tests like the Block Design Task (Thilers et al, 2006), Backward Masking Task (Van Strien et al, 2009), and Card Rotation Test (Moffat et al, 2002). In a large cross-sectional analysis, Thilers et al (2006) recruited 1107 men from the Betula study, a population-based longitudinal study of aging and health in Sweden (Nilsson et al, 1997, 2004). The men performed the Block Design Task, a well-established measure of spatial ability. The authors found a weak positive relationship between task scores and FT (β = 0.091, $P < 0.001$), and a stronger relationship between task and education level.

Interestingly, Yonker et al (2006) performed a similar study of participants recruited from the same Betula Study, also using the Block Design Task, and found a negative correlation with FT. The primary difference in analysis between the studies was a dichotomous comparison between the men with low versus high T by Yonker et al (2006), while Thilers et al (2006) used a more in-depth hierarchical regression model.

Inverse correlations have also been found using the Pattern and Letter Comparison Test (Hogervorst et al, 2004) and, again, the Block Design Task (Aleman et al, 2001), though in much smaller populations. In a cross-sectional observational study of 96 men, Martin et al (2008) found no significant correlation between FT and the Mental Rotation Test. These findings suggest the need to explore a nonlinear or task-specific relationship between T level and the visuospatial and visuoperceptual cognitive domains.

A similar pattern has been observed in investigations of endogenous T and memory. A large cross-sectional Australian study compared TT and FT levels against memory performance in 1046 participants using the Fuld Object Memory Evaluation (Martin et al, 2007). Both TT and FT levels revealed nonlinear, quadratic moderation effects on the relationship between age and scores on the Evaluation, suggesting that higher T levels “amplify” the negative effects of aging on cognition. Similar nonlinear associations with memory were reported by other authors, including Barrett-Connor et al (1999) in an early longitudinal study using the Visual Reproduction Test from the Wechsler Memory Scale. This test assesses short- and longer-term memory of geometric forms. In 547 men aged 59 to 89 years, a multiple regression model showed BT levels to be significantly related to the Visual Reproduction

Test score in a nonlinear, quadratic fashion (βs = 0.399, $P < 0.05$) after adjustment for the covariates age and education level.

Muller et al (2005) studied 395 older men with no known cognitive impairment and found quadratic significance ($P = 0.05$) for TT and scores on the Rey Auditory Verbal Learning Test, a test of immediate and delayed word recall. A positive association has been reported between BT, FT, and TT levels and verbal memory (Barrett-Connor et al, 1999; Moffat et al, 2002), working memory (Fonda et al, 2005; Thilers et al, 2006), and visual memory (Moffat et al, 2002).

Other studies of verbal memory have not been able to reproduce the same positive results (Aleman et al, 2001; Hogervorst et al, 2010; Lessov-Schlaggar et al, 2005; Matousek and Sherwin, 2010; Wolf and Kirschbaum, 2002; Yaffe et al, 2007). Likewise, some groups have found no significant association with working memory (Yonker et al, 2006) or visual memory (Lessov-Schlaggar et al, 2005; Wolf and Kirschbaum, 2002), or negative associations between TT and measures of memory and attention (Martin, 2008).

Perhaps the mixed results on research into the effects of T level on memory are explained by differences in methods and the inherent differences between memory tests. Studies that have explored nonlinear relationships suggest a nonlinear relationship similar to that observed with global cognition.

The theory of nonlinearity and an optimal T level also appears in studies of attention and executive function. In the same large study that showed a nonlinear relationship between BT and the Visual Reproduction Test, Barrett-Connor et al (1999) also observed a quadratic association between TT and attention when they asked participants to spell the word *world* backward (β = − 0.107). A trial of 54 older men by Matousek and Sherwin (2010) suggested a weak quadratic association between TT and BT and the Letter-Number Sequencing task, although the sample size and power of effect were small.

Cognition and Testosterone Supplementation in Aging Men

Cognitively Normal Older Men with Low Testosterone—Given the association between cognitive loss and low T, cognition has been evaluated during T therapy in low T and hypogonadal men of all ages, but older men are of particular interest because T levels and cognition both decline with age. Table 3 summarizes the characteristics and conclusions of studies on T supplementation in older men with low endogenous T levels.

One of the earliest of these studies showed no significant improvement in cognitive function after therapy (Sih et al, 1997). In this randomized controlled trial, 17 of 32 men (BT < 60 ng/dL) were injected intramuscularly with 200 mg of testosterone cypionate every 14 to 17 days for a year. Post-treatment evaluations showed no significant change in memory, recall, or verbal fluency between the treatment and placebo groups (Tables 3 and 4). Notably, the study design made no mention of excluding participants who were receiving other hormonal therapies that might influence their T metabolism, nor did the study draw all serum samples in the morning; both of these factors can unpredictably influence results.

Other studies have demonstrated modest improvement in cognition with T supplementation. Kenny et al (2002) randomized 67 participants with low BT (< 128 ng/dL) to transdermal T or placebo. The treatment group, but not the placebo group, had markedly higher scores on the Trail Making Test Part B than at baseline, indicating improvement in processing speed and executive function with T therapy. Of note, neither group had changes in their Digit Scan, Digit Symbol, or Trail Making Test Part A scores compared to baseline.

Vaughan et al (2007) performed a similar study with 69 healthy older men with low T (TT < 350 ng/dL) over 3 years. All the men received injected T, and a subset also received daily doses of finasteride, a 5-alpha-reductase inhibitor that slows conversion of T to DHT. Cognitive testing evaluated the participants' attention, executive function, visuospatial skills, visual memory, and verbal memory. The T-only group improved significantly in one test of attention, the Digit Span Test, and the T-plus-finasteride group improved in one test of verbal memory, the Selective Reminding Test.

Cherrier et al (2003) evaluated the effects of both T and DHT supplementation in a trial of 12 older men with TT < 300 ng/dL, and found that T supplementation improved verbal memory on the Proactive Interference test and the Wechsler Memory Scale Revised Story Recall. Interestingly, the DHT supplementation alone correlated with improved spatial memory as measured by the Route Test, which measures the ability to navigate within a room. Neither T nor DHT improved attention and executive function on the Stroop Color Word Interference Task. While this was a smaller and shorter trial than Vaughan et al (2007), with 12 versus 69 participants and 180 days versus 3 years of follow-up, it is intriguing that Cherrier et al found memory test improvements similar to Vaughan's, and suggests a more sustained memory response to T supplementation.

Haren et al's (2005) study of T supplementation in men with FT levels in the low-normal range (FT index 0.3 to 0.5), rather than unequivocal hypogonadism, found no differences between the treatment and placebo groups on the MMSE, Block Design Test, or Trail Making Test Part B after 1 year. Because Haren et al did not test their participants' memory, it is difficult to determine whether their results correlate with the findings of Vaughan et al (2007), Cherrier et al (2003), and Kenny et al (2002) in the other cognitive domains or whether participants must have low T levels to obtain significant clinical benefit from T therapy.

Visuospatial ability is one domain that often shows a significant positive association with endogenous T levels in men across a wide age range (Martin et al, 2007; Moffat et al, 2002; Thilers et al, 2006; Van Strien et al, 2009). Thus, many studies of T supplementation have focused on evaluating changes in visuospatial performance (Cherrier et al, 2004; Emmelot-Vonk et al, 2008; Young et al, 2010), as summarized in Table 4.

Emmelot-Vonk et al (2008) performed a large prospective randomized trial of T supplementation in older men with T levels in the lower half of the normal range, but found no correlation between treatment and the Mental Rotation Test. By contrast, smaller trials by Cherrier et al (2001, 2004) showed better performance on the Block Design Test with T supplementation. Since results of T supplementation on visuospatial function have been

mixed, some authors have again suggested a nonlinear association to explain the findings (Barrett-Connor et al, 1999; Hogervorst et al, 2010). We review this proposed association further below.

Cognitively Impaired Older Men with Normal Testosterone—In older men with cognitive impairment, T supplementation has been proposed as a possible treatment for declining cognitive function, particularly Alzheimer disease and mild cognitive impairment. To explore this possibility, a handful of clinical trials have tested whether T therapy can slow or even reverse cognitive dysfunction. Table 5 summarizes three studies of T supplementation in cognitively impaired older men with normal T levels.

In 2005, Cherrier et al reported a randomized controlled trial of 32 eugonadal, cognitively impaired older men, 15 with Alzheimer disease and 17 with mild cognitive impairment. Six weekly injections of T enanthate 100 mg were given to 19 of the 32 men. This group improved more over baseline than did the untreated group on the Route Test of visuospatial memory and the Block Design test of visuospatial function, but had no significant improvement on the verbal fluency test of language, the Stroop Color Test of selective attention, or the Trail Making Test Part B of divided attention.

Cherrier and colleagues (2005) also noted an interesting trend in the verbal memory tasks of Proactive Interference and Story Recall. While the T-treated group did not have significant *improvement* over the 6 weeks, they performed significantly better than the placebo group, which had a *decline.* The authors interpreted this finding as a therapeutic effect of T supplementation.

Lu et al (2006) also noted this potential “protective effect” when they randomized 47 men (18 with Alzheimer disease and 29 healthy controls) to 24 weeks of transdermal T gel or placebo and then tested the participants’ cognition with the California Verbal Learning Test, Block Design, Judgment of Line Orientation, and Visual-Motor Integration. After the 24 weeks, the Alzheimer disease group receiving T had greater improvement, or less decline, on Visual-Motor Integration than the placebo group, as well as a similar though nonsignificant trend in Judgment of Line Orientation.

Fukai et al (2010) showed a more dramatic improvement in global cognition after 6 months of T therapy than placebo in cognitively impaired men. The measures were the MMSE (2.4 versus 0.1 point increase, respectively) and the revised Hasegawa Dementia Scale (3.0 increase versus 0.8 point decrease, respectively). However, our ability to interpret the study’s results is limited by the small sample size and the lack of a statement that the trial excluded men who were already on treatments that could affect T.

Cognitively Impaired Older Men with Low Testosterone—Table 6 summarizes the only two trials of T therapy in cognitively impaired men with a low T. Tan and Pu (2003) performed a pilot study on 10 men (TT < 250 ng/dL) with Alzheimer disease. Five received T enanthate intramuscularly weekly for a year, and the other five were controls. The average TT level in the treatment group rose from 126 to 341 ng/dL. In terms of cognition, the treatment group had a non-statistically significant improvement ($P < 0.07$) in their

visuospatial function over baseline on the Clock-Drawing Task. The treated group also scored significantly better than the controls on the MMSE and Alzheimer's Disease Assessment Scale–Cognitive subscale. This study showed promising results in a small group. As far as we know, no larger follow-up trial has yet been implemented.

In the other trial, Kenny et al (2004) studied 11 men with low BT and mild-to-moderate cognitive decline. Six of the men were randomized to receive T enanthate injections of 200 mg every 3 weeks for 12 weeks, and the other five men received a placebo. All participants took the Trail Making Test Part B, Clock Drawing Test, Clock Face Perception, and Digit Span at baseline, 4 weeks, and 10 weeks. While the authors concluded that T treatment appeared to be safe, they found that neither group had significant change in attention, visuospatial function, problem solving, or executive function, or in behavior, depression, or daily activity. However, like Tan and Pu (2003), the Kenny et al (2004) study was a pilot study with a small sample, from which generalizations are difficult.

Adverse Effects of Testosterone Therapy

All nine controlled studies listed in Table 7 commented on adverse events of T therapy in their participants (the adverse effects are not shown in the table). The only significant adverse effect was an elevated hemoglobin concentration, reported by Sih et al (1997). Four men in this study developed a hematocrit above 52%, and a few men reported skin irritation at the injection site. No study reported significant changes in participants' aggression level, mood, prostate-specific antigen levels, prostate size on rectal exam, leg or breast swelling, serum cholesterol levels, or liver function tests, and no study reported cardiovascular events.

A thorough review of the safety profile of T therapy is beyond the scope of this paper. Multiple other groups, however, have evaluated T's adverse effects in some detail. For example, Fernandez-Balsells et al (2010) reported in a systematic review and meta-analysis that T therapy correlated with an increase in hemoglobin concentration and a small decrease in high density lipoprotein. Both associations are of unknown clinical significance.

Methodological Quality of Controlled Studies in the Review

As shown in Table 7, each of the nine controlled trials of T supplementation had methodologic limitations, and many of the trials were pilot studies. The table lists factors that we considered important for experimental quality and generalizability of results. The higher-quality studies are toward the top of the table.

In terms of limitations, three of the nine studies had sample sizes of only 10 or 11 men. Only five of the nine studies had what we considered an adequate length of follow-up, ≥1 year. A different group of five studies did not measure (or did not report measuring) T levels at the preferred time in the morning. Four studies did not report whether the researchers excluded participants who were taking medications that could influence their T levels. Three studies had a >30% dropout rate, and two studies did not report the dropout rate at all.

Among the nine studies, we considered Vaughan et al (2007) to meet the highest number of our quality criteria. Vaughan et al (2007) had the most participants (69 men), the longest follow-up (36 months), morning blood sampling for T level, and a randomized double-

blinded protocol; however, the high reported 33% dropout rate limited our ability to generalize from the results.

None of the other studies offered quality equal to Vaughan et al (2007). Although Lu et al (2006) followed 47 participants for 24 months in a randomized double-blinded study with a lower dropout rate of 19%, the investigators did not take morning blood samples.

In summary, the overall quality of controlled studies published as of February 2016 and included in our review is not adequate to provide recommendations and clinical guidelines for T supplementation to preserve or improve cognitive performance.

DISCUSSION

Suggested Modifications to Future Study Designs

Our systematic review of the current state of knowledge about cognition and T in older men and the potential role of T supplementation in age-related cognitive disorders has led us to make several recommendations for the directions of research.

Future studies should use a standardized method of T delivery, serum T assessments, and the cognitive tests and subtests used to assess cognition. The first step toward standardizing protocols could be the use of transdermal T formulations and measurement of TT, FT and BT levels at several specific points during the treatment period, in order to determine an average level during the study, rather than just at the beginning and end of the study. Further, despite receiving the same treatment in a study, men may have very different final T levels as a result of individual variations in T absorption, metabolism, and excretion. For this reason, we suggest that researchers consider either evaluating outcomes based on T levels measured at multiple points during the study, or varying the supplementation to enable participants to reach pre-specified serum levels.

Whenever possible, we believe that serum T should be measured within 1 to 2 hours of each man’s usual waking time, and men should be asked how long and how well they slept the night before the collection. Adherence to such a protocol would more accurately reflect the physiologic fluctuations of T levels (Axelsson et al, 2005; Bremner et al, 1983).

A major challenge in studying T and cognition is that the underlying physiology is both complex and dynamic, involving the entire hypothalamic-pituitary-adrenal-gonadal axis as well as multiple additional internal and external factors (Ulubaev et al, 2009). Thus, the notion of finding a simple correlation between a single hormone and an outcome of interest is unrealistic, especially in older populations in whom mild dysregulation across multiple systems is the norm. Indeed, the term *syncrinology* has been proposed to describe the mechanisms by which multiple hormones act simultaneously in similar or opposing ways to affect a single target. (Valenti and Schwartz, 2008).

While appealing, this more holistic approach incorporating multiple moving parts collides with the realities of conducting clinical studies and the practical considerations of sample size, cost, and burden on the participant. Controlling for estradiol, DHT, dihydroepiandrosterone, and/or SHBG may be the most practical solution until we have a

better understanding of the impact that these hormones have on T and cognition. Carefully designed, adequately powered, hypothesis-driven studies are needed to understand better how T affects cognition, but even high-quality studies will be limited in their ability to provide unequivocal answers and therefore will require cautious interpretation.

Principal Findings

Studies of the association between endogenous T level and cognitive performance have shown a possible inverted U-shaped relationship. In particular, there may be an optimal T level at which men improve in their global cognitive function, specifically memory, attention, and executive function. Indirect evidence suggests that there is an optimal T level for visuospatial and visuoperceptual function. T supplementation may enhance cognitive function in older men with low T levels, but improvement varies with the cognitive domain assessed and the test given.

In men with cognitive deficits such as mild cognitive impairment or Alzheimer disease, T may have a “protective effect” by slowing the rate of cognitive decline in those who are eugonadal at baseline. Data are insufficient to draw any conclusions about the value of T supplements for cognitively impaired men who are hypogonadal or in a low T state.

As for clinical implications, this review illustrates the potential for T as a therapy for improving cognitive abilities and function, but also shows that more study is essential before standardized clinical recommendations can be made.

The strengths of this review include the broad scope of studies reviewed and the analyses of subpopulations by T levels and cognitive status. To help the reader put the reviewed studies in context, our tables summarize the methodology of each: longitudinal versus cross-sectional, number of participants, T levels, specific cognitive tests used as outcome measures, and results. We analyzed the quality of the studies by evaluating potential methodological biases (Table 7).

One limitation is that we may have failed to identify all relevant studies within our chosen time frame. We included observational studies, which broaden the scope of the review but are subject to inherent biases. Finally, we were limited by the limitations of the studies themselves in sample size, duration, and variability of methods; this limitation precluded our performing quantitative analyses and drawing more definitive conclusions about the association between T and cognition in older men.

In summary, this relationship has been studied extensively, with mixed results. The variability in results is likely related to and dependent on study design and methods and the study populations’ age, gonadal state, and baseline cognitive status. Despite these limitations, trends and promising results have been reported for T supplementation in aging men with both normal and low T states, with and without baseline cognitive dysfunction. Larger studies with a more standardized approach to assessment will be needed to understand fully and realize sustained benefits from T supplementation in the elderly male population. Given that people with Alzheimer disease can suffer marked cognitive decline in just 6 to 12 months, short studies in these patients have the potential to show treatment

effects. Further research is greatly needed to establish a therapeutic relationship between cognition and T before clinical recommendations can be made, and a multi-center trial that accommodates standardization of protocols and a large participant population should be the goal.

Acknowledgments

Supported in part by NIH Grant T35 EY021455.

Glossary

BT	bioavailable testosterone
DHT	dihydrotestosterone
FT	free testosterone
MMSE	Mini-Mental State Examination
PRISMA	Preferred Reporting Items for Systematic Reviews and Meta-Analyses
SHBG	sex hormone binding globulin
T	testosterone
TT	total testosterone

References

Aleman A, de Vries WR, Koppeschaar HP, et al. Relationship between circulating levels of sex hormones and insulin-like growth factor-1 and fluid intelligence in older men. Exp Aging Res. 2001; 27:283–291. [PubMed: 11441649]

Alexander GM, Swerdloff RS, Wang C, et al. Androgen-behavior correlations in hypogonadal men and eugonadal men. II. Cognitive abilities. Horm Behav. 1998; 33:85–94. [PubMed: 9647934]

Axelsson J, Ingre M, Akerstedt T, et al. Effects of acutely displaced sleep on testosterone. J Clin Endocrinol Metab. 2005; 90:4530–4535. [PubMed: 15914523]

Barrash J, Damasio H, Adolphs R, et al. The neuroanatomical correlates of route learning impairment. Neuropsychologia. 2000; 38:820–836. [PubMed: 10689057]

Barrett-Connor E, Goodman-Gruen D, Patay B. Endogenous sex hormones and cognitive function in older men. J Clin Endocrinol Metab. 1999; 84:3681–3685. [PubMed: 10523014]

Bäumler, G. Lern- und Gedächtnistest. LGT-3 [Learning and Memory Test LGT-3]. Göttingen, Germany: Hogrefe; 1974.

Beery, KE. The Beery-Buktenica Developmental Test of Visual-Motor Integration. 4th. Parsippany, New Jersey: Modern Curriculum Press; 1997.

Benton AL. A visual retention test for clinical use. Arch Neurol Psychiatry. 1945; 54:212–216.

Benton AL, Varney NR, Hamsher KS. Visuospatial judgment. Arch Neurol. 1978; 35:364–367. [PubMed: 655909]

Berg EA. A simple objective technique for measuring flexibility in thinking. J Gen Psychol. 1948; 39:15–22. [PubMed: 18889466]

Bhasin S, Cunningham GR, Hayes FJ, et al. Testosterone therapy in men with androgen deficiency syndromes: an Endocrine Society clinical practice guideline. J Clin Endocrinol Metab. 2010; 95:2536–2559. [PubMed: 20525905]

Bimonte-Nelson HA, Singleton RS, Nelson ME, et al. Testosterone, but not nonaromatizable dihydrotestosterone, improves working memory and alters nerve growth factor levels in aged male rats. Exp Neurol. 2003; 181:301–312. [PubMed: 12782002]

Bremner WJ, Vitiello MV, Prinz PN. Loss of circadian rhythmicity in blood testosterone levels with aging in normal men. J Clin Endocrinol Metab. 1983; 56:1278–1281. [PubMed: 6841562]

Buschke H, Fuld PA. Evaluating storage, retention, and retrieval in disordered memory and learning. Neurology. 1974; 24:1019–1025. [PubMed: 4473151]

Cappa SF, Guariglia C, Papagno C, et al. Patterns of lateralization and performance levels for verbal and spatial tasks in congenital androgen deficiency. Behav Brain Res. 1988; 31:177–183. [PubMed: 3202949]

Cherrier MM, Asthana S, Plymate S, et al. Testosterone supplementation improves spatial and verbal memory in healthy older men. Neurology. 2001; 57:80–88. [PubMed: 11445632]

Cherrier MM, Craft S, Matsumoto AH. Cognitive changes associated with supplementation of testosterone or dihydrotestosterone in mildly hypogonadal men: a preliminary report. J Androl. 2003; 24:568–576. [PubMed: 12826696]

Cherrier MM, Matsumoto AM, Amory JK, et al. Testosterone improves spatial memory in men with Alzheimer disease and mild cognitive impairment. Neurology. 2005; 64:2063–2068. [PubMed: 15985573]

Cherrier MM, Plymate S, Mohan S, et al. Relationship between testosterone supplementation and insulin-like growth factor-I levels and cognition in healthy older men. Psychoneuroendocrinology. 2004; 29:65–82. [PubMed: 14575730]

Delis, DC.; Kramer, JH.; Kaplan, E.; Ober, RA. California Verbal Learning Test. San Antonio, Texas: The Psychological Corporation; 1987.

Ekstrom, RB.; French, JW.; Harman, HH. Manual for Kit of Factor-Referenced Cognitive Tests. Princeton, New Jersey: Educational Testing Service; 1976.

Emmelot-Vonk MH, Verhaar HJ, Nakhai Pour HR, et al. Effect of testosterone supplementation on functional mobility, cognition, and other parameters in older men: a randomized controlled trial. JAMA. 2008; 299:39–52. [PubMed: 18167405]

Feldman HA, Longcope C, Derby CA, et al. Age trends in the level of serum testosterone and other hormones in middle-aged men: longitudinal results from the Massachusetts Male Aging Study. J Clin Endocrinol Metab. 2002; 87:589–598. [PubMed: 11836290]

Fernandez-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; 95:2560–2575. [PubMed: 20525906]

Fiers T, Delanghe J, T'Sjoen G, et al. A critical evaluation of salivary testosterone as a method for the assessment of serum testosterone. Steroids. 2014; 86:5–9. [PubMed: 24793565]

Folstein MF, Folstein SE, McHugh PR. "Mini-mental state:" a practical method for grading the cognitive state of patients for the clinician. J Psychiatr Res. 1975; 12:189–198. [PubMed: 1202204]

Fonda SJ, Bertrand R, O'Donnell A, et al. Age, hormones, and cognitive functioning among middle-aged and elderly men: cross-sectional evidence from the Massachusetts Male Aging Study. J Gerontol A Biol Sci Med Sci. 2005; 60:385–390. [PubMed: 15860479]

Fritz KS, McKean AJ, Nelson JC, et al. Analog-based free testosterone test results linked to total testosterone concentrations, not free testosterone concentrations. Clin Chem. 2008; 54:512–516. [PubMed: 18171714]

Fukai S, Akishita M, Yamada S, et al. Effects of testosterone in older men with mild-to-moderate cognitive impairment. J Am Geriatr Soc. 2010; 58:1419–1421. [PubMed: 20672465]

Fuld PA, Masur DM, Blau AD, et al. Object-memory evaluation for prospective detection of dementia in normal functioning elderly: predictive and normative data. J Clin Exp Neuropsychol. 1990; 12:520–528. [PubMed: 2211974]

Gluhm S, Goldstein J, Loc K, et al. Cognitive performance on the mini-mental state examination and the Montreal Cognitive Assessment across the healthy adult lifespan. Cogn Behav Neurol. 2013; 26:1–5. [PubMed: 23538566]

Gray PB, Singh AB, Woodhouse LJ, et al. Dose-dependent effects of testosterone on sexual function, mood, and visuospatial cognition in older men. J Clin Endocrinol Metab. 2005; 90:3838–3846. [PubMed: 15827094]

Haren MT, Wittert GA, Chapman IM, et al. Effect of oral testosterone undecanoate on visuospatial cognition, mood and quality of life in elderly men with low-normal gonadal status. Maturitas. 2005; 50:124–133. [PubMed: 15653010]

Hasegawa K. Psychology of the aged patient and the nursing approach [in Japanese]. Kangogaku Zasshi. 1974; 38:893–900. [PubMed: 4216699]

Hier DB, Crowley WF Jr. Spatial ability in androgen-deficient men. N Engl J Med. 1982; 306:1202–1205. [PubMed: 7070432]

Hogervorst E. Effects of gonadal hormones on cognitive behaviour in elderly men and women. J Neuroendocrinol. 2013; 25:1182–1195. [PubMed: 23895362]

Hogervorst E, De Jager C, Budge M, et al. Serum levels of estradiol and testosterone and performance in different cognitive domains in healthy elderly men and women. Psychoneuroendocrinology. 2004; 29:405–421. [PubMed: 14644069]

Hogervorst E, Matthews FE, Brayne C. Are optimal levels of testosterone associated with better cognitive function in healthy older women and men? Biochim Biophys Acta. 2010; 1800:1145–1152. [PubMed: 20060437]

Holland J, Bandelow S, Hogervorst E. Testosterone levels and cognition in elderly men: a review. Maturitas. 2011; 69:322–337. [PubMed: 21696899]

Janowsky JS, Chavez B, Orwoll E. Sex steroids modify working memory. J Cogn Neurosci. 2000; 12:407–414. [PubMed: 10931767]

Janowsky JS, Oviatt SK, Orwoll ES. Testosterone influences spatial cognition in older men. Behav Neurosci. 1994; 108:325–332. [PubMed: 8037876]

Kaufman JM, Vermeulen A. Declining gonadal function in elderly men. Baillieres Clin Endocrinol Metab. 1997; 11:289–309. [PubMed: 9403124]

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–M325. [PubMed: 11983727]

Kenny AM, Fabregas G, Song C, et al. Effects of testosterone on behavior, depression, and cognitive function in older men with mild cognitive loss. J Gerontol A Biol Sci Med Sci. 2004; 59:75–78. [PubMed: 14718489]

Kluger A, Ferris SH, Golomb J, et al. Neuropsychological prediction of decline to dementia in nondemented elderly. J Geriatr Psychiatry Neurol. 1999; 12:168–179. [PubMed: 10616864]

Law DJ, Morrin KA, Pellegrino JW. Training effects and working memory contributions to skill acquisition in a complex coordination task. Learn Individ Differ. 1995; 7:207–234.

LeBlanc ES, Wang PY, Janowsky JS, et al. Association between sex steroids and cognition in elderly men. Clin Endocrinol. 2010; 72:393–403.

Lessov-Schlaggar CN, Reed T, Swan GE, et al. Association of sex steroid hormones with brain morphology and cognition in healthy elderly men. Neurology. 2005; 65:1591–1596. [PubMed: 16301487]

Lezak, MD. Neuropsychological Assessment. 3rd. New York, New York: Oxford University Press; 1995.

Lu PH, Masterman DA, Mulnard R, et al. Effects of testosterone on cognition and mood in male patients with mild Alzheimer disease and healthy elderly men. Arch Neurol. 2006; 63:177–185. [PubMed: 16344336]

Malec JF, Ivnik RJ, Smith GE, et al. Visual Spatial Learning Test: normative data and further validation. Psychol Assess. 1992; 4:433–441.

Martin DM, Wittert G, Burns NR, et al. Testosterone and cognitive function in ageing men: data from the Florey Adelaide Male Ageing Study (FAMAS). Maturitas. 2007; 57:182–194. [PubMed: 17287097]

Martin DM, Wittert G, Burns NR, et al. Endogenous testosterone levels, mental rotation performance, and constituent abilities in middle-to-older aged men. Horm Behav. 2008; 53:431–441. [PubMed: 18190916]

Matousek RH, Sherwin BB. Sex steroid hormones and cognitive functioning in healthy, older men. Horm Behav. 2010; 57:352–359. [PubMed: 20079740]

Matsumoto, AM. Aging and human male reproductive function. In: Haseltine, F.; Paulsen, CA.; Wang, C., editors. Reproductive Issues and the Aging Male. Washington, DC: American Association for the Advancement of Science; 1993. p. 1-14.

McKenna P, Warrington EK. Testing for nominal dysphasia. J Neurol Neurosurg Psychiatry. 1980; 43:781–788. [PubMed: 7420102]

Moffat SD, Zonderman AB, Metter EJ, et al. Longitudinal assessment of serum free testosterone concentration predicts memory performance and cognitive status in elderly men. J Clin Endocrinol Metab. 2002; 87:5001–5007. [PubMed: 12414864]

Moher D, Liberati A, Tetzlaff J, et al. Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. Ann Intern Med. 2009; 151:264–269. [PubMed: 19622511]

Morris JC, Heyman A, Mohs RC, et al. The Consortium to Establish a Registry for Alzheimer's Disease (CERAD). Part I. Clinical and neuropsychological assessment of Alzheimer's disease. Neurology. 1989; 39:1159–1165. [PubMed: 2771064]

Moscovitch M. Cognitive resources and dual-task interference effects on retrieval in normal people: the role of the frontal lobes and medial temporal cortex. Neuropsychol. 1994; 8:524–534.

Muller M, Aleman A, Grobbee DE, et al. Endogenous sex hormone levels and cognitive function in aging men: is there an optimal level? Neurology. 2005; 64:866–871. [PubMed: 15753424]

Nilsson LG, Bäckman L, Erngrund K, et al. The Betula prospective cohort study: memory, health, and aging. Aging Neuropsychol Cogn. 1997; 4:1–32.

Nilsson LG, Adolfsson R, Bäckman L, et al. Betula: a prospective cohort study on memory, health and aging. Aging Neuropsychol Cogn. 2004; 11:134–148.

Oskui PM, French WJ, Herring MJ, et al. Testosterone and the cardiovascular system: a comprehensive review of the clinical literature. J Am Heart Assoc. 2013; 2:e000272. [PubMed: 24242682]

Penev PD. Association between sleep and morning testosterone levels in older men. Sleep. 2007; 30:427–432. [PubMed: 17520786]

Perry PJ, Lund BC, Arndt S, et al. Bioavailable testosterone as a correlate of cognition, psychological status, quality of life, and sexual function in aging males: implications for testosterone replacement therapy. Ann Clin Psychiatry. 2001; 13:75–80. [PubMed: 11534928]

Petrides M, Milner B. Deficits on subject-ordered tasks after frontal- and temporal-lobe lesions in man. Neuropsychologia. 1982; 20:249–262. [PubMed: 7121793]

Piaget, J.; Inhelder, B. The Child's Conception of Space. New York, New York: W.W. Norton; 1956.

Pike CJ, Carroll JC, Rosario ER, et al. Protective actions of sex steroid hormones in Alzheimer's disease. Front Neuroendocrinol. 2009; 30:239–258. [PubMed: 19427328]

Raab DH. Backward masking. Psychological Bulletin. 1963; 60:118–129. [PubMed: 13972804]

Reitan RM. Validity of the Trail Making Test as an Indicator of Organic Brain Damage. Percept Mot Skills. 1958; 8:271–276.

Rey A. L'examen psychologique dans les cas d'encephalopathie traumatique [The psychological examination in cases of traumatic encephalopathy]. Archives de Psychologie. 1941; 28:286–340.

Rey, A. Epreuves mnésiques et d'apprentissage. Actualités Pedagogiques et Psychologiques [Memory Tests and Learning. Recent Developments in Pedagogy and Psychology]. Neuchatel, Switzerland: Delachaux & Neistle; 1968.

Robbins TW, James M, Owen AM, et al. Cambridge Neuropsychological Test Automated Battery (CANTAB): a factor analytic study of a large sample of normal elderly volunteers. Dementia. 1994; 5:266–281. [PubMed: 7951684]

Robertson IH, Manly T, Andrade J, et al. 'Oops!': performance correlates of everyday attentional failures in traumatic brain injured and normal subjects. Neuropsychologia. 1997; 35:747–758. [PubMed: 9204482]

Rosen WG, Mohs RC, Davis KL. A new rating scale for Alzheimer's disease. Am J Psychiatry. 1984; 141:1356–1364. [PubMed: 6496779]

Royall DR, Cordes JA, Polk M. CLOX: an executive clock drawing task. J Neurol Neurosurg Psychiatry. 1998; 64:588–594. [PubMed: 9598672]

Royall DR, Mahurin RK, Gray KF. Bedside assessment of executive cognitive impairment: the executive interview. J Am Geriatr Soc. 1992; 40:1221–1226. [PubMed: 1447438]

Salthouse TA, Babcock RL. Decomposing adult age differences in working memory. Dev Psychol. 1991; 27:763–776.

Schacter DL, Harbluk JL, Maclachlan DR. Retrieval without recollection: an experimental analysis of source amnesia. J Verbal Learning Verbal Behav. 1984; 23:593–611.

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; 82:1661–1667. [PubMed: 9177359]

Spreen, O.; Strauss, E. A Compendium of Neuropsychological Tests: Administration, Norms, and Commentary. 2nd. New York, New York: Oxford University Press; 1998.

Stankov L. Complexity, metacognition, and fluid intelligence. Intelligence. 2000; 28:121–143.

Stroop JR. Studies of interference in serial verbal reactions. J Exper Psychol. 1935; 18:643–662.

Tan RS, Pu SJ. A pilot study on the effects of testosterone in hypogonadal aging male patients with Alzheimer's disease. Aging Male. 2003; 6:13–17. [PubMed: 12809076]

Teng EL, Chui HC. The Modified Mini-Mental State (3MS) examination. J Clin Psychiatry. 1987; 48:314–318. [PubMed: 3611032]

Thilers PP, Macdonald SW, Herlitz A. The association between endogenous free testosterone and cognitive performance: a population-based study in 35 to 90 year-old men and women. Psychoneuroendocrinology. 2006; 31:565–576. [PubMed: 16487665]

Ulubaev A, Lee DM, Purandare N, et al. Activational effects of sex hormones on cognition in men. Clin Endocrinol. 2009; 71:607–623.

Valenti G, Schwartz RS. Anabolic decline in the aging male: a situation of unbalanced syncrinology. Aging Male. 2008; 11:153–156. [PubMed: 19172545]

Vandenberg SG, Kuse AR. Mental rotations, a group test of three-dimensional spatial visualization. Percept Mot Skills. 1978; 47:599–604. [PubMed: 724398]

Van Strien JW, Weber RF, Burdorf A, et al. Higher free testosterone level is associated with faster visual processing and more flanker interference in older men. Psychoneuroendocrinology. 2009; 34:546–554. [PubMed: 19042092]

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; 28:875–882. [PubMed: 17609296]

Vink M, Jolles J. A new version of the Trail Making Test as an information processing task. J Clin Neuropsychol. 1985; 7:162.

Wang C, Catlin DH, Demers LM, et al. Measurement of total serum testosterone in adult men: comparison of current laboratory methods versus liquid chromatography-tandem mass spectrometry. J Clin Endocrinol Metab. 2004; 89:534–543. [PubMed: 14764758]

Wechsler, D. Wechsler Adult Intelligence Scale–Fourth Edition. San Antonio, Texas: Pearson; 2008.

Wechsler, D. Wechsler Memory Scale–Fourth Edition. San Antonio, Texas: Pearson; 2009.

Wolf OT, Kirschbaum C. Endogenous estradiol and testosterone levels are associated with cognitive performance in older women and men. Horm Behav. 2002; 41:259–266. [PubMed: 11971659]

Woodcock, R.; Johnson, M.; Mather, N. Woodcock-Johnson Psycho-Educational Battery-Revised. Allen, Texas: DLM Teaching Resources; 1989.

Yaffe K, Lui LY, Zmuda J, et al. Sex hormones and cognitive function in older men. J Am Geriatr Soc. 2002; 50:707–712. [PubMed: 11982672]

Yaffe K, Barnes D, Lindquist K, et al. Endogenous sex hormone levels and risk of cognitive decline in an older biracial cohort. Neurobiol Aging. 2007; 28:171–178. [PubMed: 17097195]

Yeap BB, Almeida OP, Hyde Z, et al. Higher serum free testosterone is associated with better cognitive function in older men, while total testosterone is not. The Health In Men Study. Clin Endocrinol. 2008; 68:404–412.

Yonker JE, Eriksson E, Nilsson LG, et al. Negative association of testosterone on spatial visualization in 35 to 80 year old men. Cortex. 2006; 42:376–386. [PubMed: 16771043]

Young LA, Neiss MB, Samuels MH, et al. Cognition is not modified by large but temporary changes in sex hormones in men. J Clin Endocrinol Metab. 2010; 95:280–288. [PubMed: 19880786]

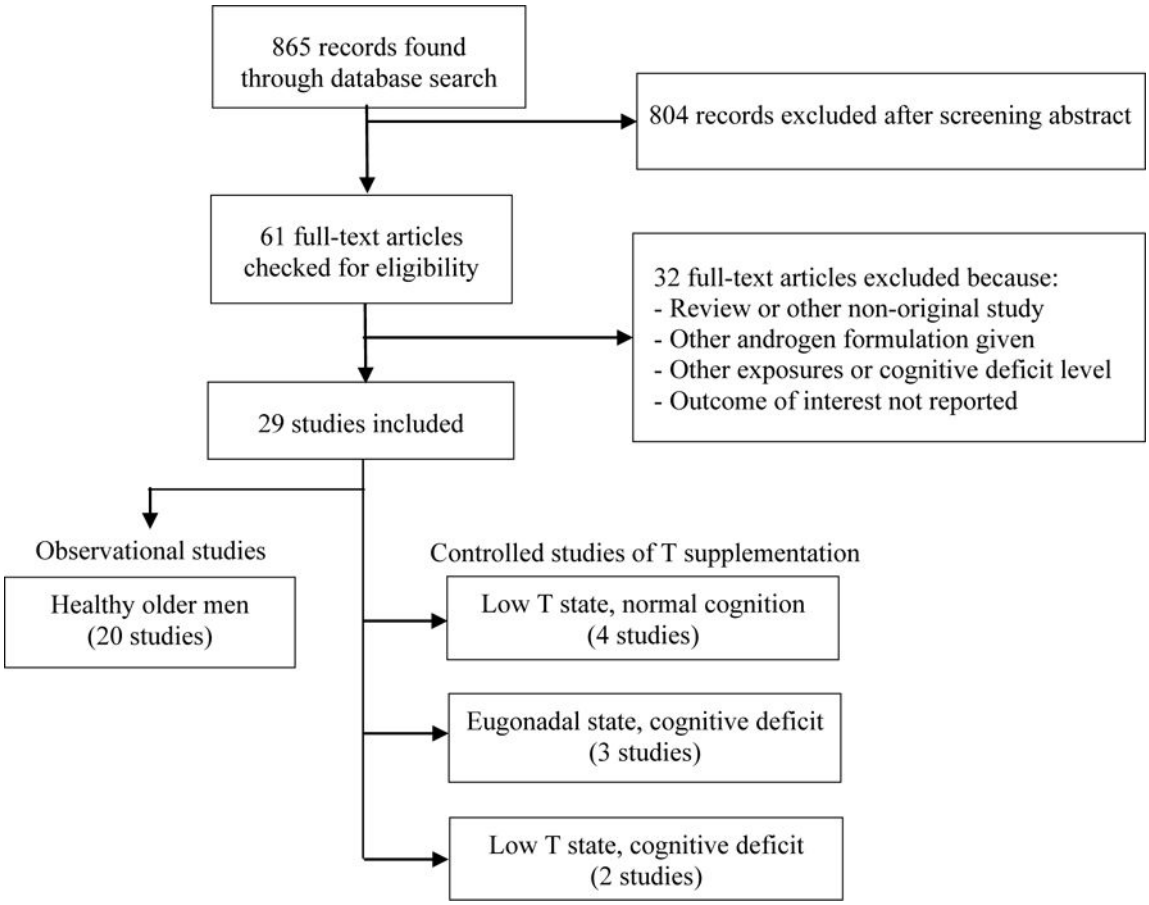

FIGURE 1.

This review's systematic search of the MEDLINE and Embase® electronic databases of publications from January 1995 to February 2016. We used the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) guidelines (Moher et al, 2009) to seek the current state of knowledge about cognition and testosterone (T) in older men and the potential role of T supplementation in age-related cognitive disorders.

TABLE 1

Cognitive Tasks Given to Men in Testosterone Studies Cited in Tables 2–6

Cognitive Domain	Task	Best and/or Most Recent Reference*
Global (GLOB)	1. Mini-Mental State Examination (MMSE)	Folstein et al, 1975
	2. Modified Mini-Mental State Examination (3MS)	Teng and Chui, 1987
	3. Hasegawa Dementia Scale	Hasegawa, 1974
	4. Alzheimer's Disease Assessment Scale–Cognitive (ADAScog)	Rosen et al, 1984
Memory (MEM)	1. Digit Span–Wechsler Adult Intelligence Scale (WAIS)	Wechsler, 2008
	2. Digit Symbol Substitution–Wechsler Adult Intelligence Scale (WAIS)	Wechsler, 2008
	3. Paragraph Recall	Kluger et al, 1999
	4. Story Recall–Wechsler Memory Scale (WMS)	Wechsler, 2009
	5. Verbal Paired Associates Test–Wechsler Memory Scale (WMS)	Wechsler, 2009
	6. Graded Naming Test	McKenna and Warrington, 1980
	7. California Verbal Learning Test (CVLT)	Delis et al, 1987
	8. Rey Auditory Verbal Learning Test (RAVLT)	Rey, 1941
	9. Rey Auditory Verbal Learning Test (RAVLT) Recall	Rey, 1941
	10. Rey Visual Design Learning Test (RVDLT)	Rey, 1968
	11. Rey Visual Design Learning Test (RVDLT) Recall	Rey, 1968
	12. Benton Visual Retention Test (BVRT)	Benton, 1945
	13. Visual Reproduction Test–Wechsler Memory Scale (WMS)	Wechsler, 2009
	14. Selective Reminding Test	Buschke and Fuld, 1974
	15. Cambridge Neuropsychological Test Automated Battery–Spatial Span Test	Robbins et al, 1994
	16. Consortium to Establish a Registry for Alzheimer's Disease (CERAD) 10-Word List	Morris et al, 1989
	17. Dot Matrix Task	Law et al, 1995
	18. Fuld Object Memory Evaluation	Fuld et al, 1990
	19. Proactive Interference	Moscovitch, 1994
	20. Picture Swap	Stankov, 2000
Visuospatial and visuoperceptual performance (VS)	1. Mental Rotation Test	Vandenberg and Kuse, 1978
	2. Judgment of Line Orientation	Benton et al, 1978
	3. Block Design–Wechsler Adult Intelligence Scale (WAIS)	Wechsler, 2008
	4. Clock Drawing Test	Royall et al, 1998
	5. Clock Face Perception	Kenny et al, 2004
	6. City Map Task	Bäumler, 1974
	7. Route Test	Barrash et al, 2000
	8. Backward Masking Task	Raab, 1963
	9. Cambridge Neuropsychological Test Automated Battery–Spatial Recognition Test	Robbins et al, 1994

Cognitive Domain	Task	Best and/or Most Recent Reference*
	10. Paper Folding Test	Ekstrom et al, 1976
	11. Pattern and Letter Comparison Test	Salthouse and Babcock, 1991
	12. Card Rotation Test	Ekstrom et al, 1976
	13. Spatial Array Learning Test (SALT)	Malec et al, 1992
	14. Developmental Test of Visual-Motor Integration	Beery, 1997
	15. Water-Level Test	Piaget and Inhelder, 1956
Attention and executive function (ATT)	1. Digit Span–Wechsler Adult Intelligence Scale (WAIS)	Wechsler, 2008
	2. Digit Symbol Substitution–Wechsler Adult Intelligence Scale (WAIS)	Wechsler, 2008
	3. Stroop Color Word Interference Task	Stroop, 1935
	4. Trail Making Part A	Reitan, 1958
	5. Trail Making Part B	Reitan, 1958
	6. Tower of Hanoi	Lezak, 1995
	7. Cross Out Test	Woodcock and Johnson, 1989
	8. “World” Backwards or Serial Sevens (MMSE)	Folstein et al, 1975
	9. Executive Interview	Royall et al, 1992
	10. Wisconsin Card Sort	Berg, 1948
	11. Animal Naming	Spreen and Strauss, 1998
	12. Sustained Attention to Response Task	Robertson et al, 1997
	13. Self-Ordered Pointing Task	Petrides and Milner, 1982
	14. Controlled Oral Word Association Task (COWAT)	Spreen and Strauss, 1998
	15. Semantic Memory	Schacter et al, 1984
	16. Vocabulary–Wechsler Adult Intelligence Scale (WAIS)	Wechsler, 2008
	17. Concept Shifting Task	Vink and Jolles, 1985
	18. Letter-Number Sequencing–Wechsler Adult Intelligence Scale (WAIS)	Wechsler, 2008

*These references are not necessarily the source for the version of the task used in each of the studies listed in Tables 2–6. Please check each study to learn which version of the task was given.

TABLE 2

Studies of Endogenous Testosterone and Cognition in Healthy Older Men

Authors, Year (Study Type)	N	Mean Age (Range) If Reported	Method of T Measurement	Mean T Level	Cognitive Measures*	Significant Conclusions
Aleman et al, 2001 (cross-sectional)	25	69.1 (65–76)	TT, FT	TT 21 nmol/L FT 398 pmol/L	MEM 8 VS 2, 3 ATT 2, 16, 17	Negative association between TT and VS 3, ATT 2
Barrett-Connor et al, 1999 (longitudinal)	547	70.2 (55–89)	Morning TT, BT	TT 10.8 nmol/L BT 3.4 nmol/L	GLOB 1 MEM 13, 14 ATT 5, 8	Positive association between BT and MEM 14 Quadratic association between TT and ATT 8, and between BT and MEM 13
Fonda et al, 2005 (cross-sectional)	981	62.7 (48–80)	FT, TT	TT 4.52 ng/mL FT 0.07 ng/mL	ATT 1, 2	Positive association between TT and ATT 1
Hogervorst et al, 2004 (cross-sectional)	79	74	Non-fasting TT	TT 16.33 nmol/L	MEM 5, 6, 15, 16 VS 9, 11	Negative association between TT and VS 11
Hogervorst et al, 2010 (longitudinal)	257	74 (64–94)	TT	TT 12.91 nmol/L	GLOB 1 MEM 4	Curvilinear association between TT and GLOB 1
LeBlanc et al, 2010 (longitudinal)	1397	> 65	Morning fasting TT, FT	FT 0.26 nmol/L	GLOB 2 ATT 5	No significant associations between FT and any tests
Lessov-Schlaggar et al, 2005 (longitudinal)	514 pairs of twins	63.1 (59–70)	TT	Not reported	GLOB 1 MEM 2, 7, 12 ATT 3, 4, 5, 14	No significant associations between TT and any tests
Martin et al, 2007 (cross-sectional)	1046	54.3 (35–80)	TT, BT, FT	TT 14.1 nmol/L BT 5.1 nmol/L FT 212 pmol/L	MEM 18 ATT 4, 5	Negative association between TT and MEM 18 Positive association between TT and ATT 4 Quadratic moderation of TT and FT on age and MEM 18
Martin et al, 2008 (cross-sectional)	96	38–49 (n = 29), 50–59 (n = 37), 60–69 (n = 30) (38–69)	TT, FT	TT 14 - 14.7 nmol/L FT 216 - 245 pmol/L	MEM 2, 4, 17, 20 VS 1 ATT 7, 12, 13	Negative association between TT and MEM 2 and between TT and ATT 7, 12, 13
Matousek and Sherwin, 2010 (cross-sectional)	54	68.6 (61–77)	TT, FT, BT	TT 9.7 nmol/L BT 3.9 nmol/L FT 190 pmol/L	MEM 2, 5, 14 VS 1, 3, 10, 15 ATT 18	Nonlinear association between FT and BT and ATT 18
Moffat et al, 2002 (longitudinal)	407	64.1 (50–91)	Morning fasting TT, FT	TT 421 ng/dL FT 5.1 ng/dL	GLOB 1 MEM 1, 7, 12 VS 1 ATT 4, 5	Positive association between FT and MEM 7, 12, VS 1, and ATT 4, 5
Muller et al, 2005 (cross-sectional)	395	60.2 (40–80)	TT, BT	TT 18.6 nmol/L BT 8.2 nmol/L	GLOB 1 MEM 1, 8 ATT 4, 5	Curvilinear association between TT and MEM 8 Positive association between TT and ATT 5

Authors, Year (Study Type)	N	Mean Age (Range) If Reported	Method of T Measurement	Mean T Level	Cognitive Measures*	Significant Conclusions
Perry et al, 2001 (cross-sectional)	78	65.6 (55–75)	Morning BT	BT 91 ng/dL	GLOB 1 ATT 4, 5, 9, 10	Nonsignificant negative association between BT and ATT 9
Thilers et al, 2006 (cross-sectional)	1107	(35–90)	TT, FT	TT 17 nmol/L FT 7.7 nmol/L	VS 3 ATT 15	Positive association between FT and VS 3 and ATT 15
Van Strien et al, 2009 (cross-sectional)	72	67.2 (57–79)	TT, FT	TT 13.5 nmol/L FT 0.25 nmol/L	VS 1, 8	Positive association between FT and VS 1,8
Wolf et al, 2002 (cross-sectional)	30	69	TT, FT	TT 3.6 ng/mL FT 13.3 pg/mL	MEM 5 VS 1, 6 ATT 3	No significant associations between TT/FT and any tests
Yaffe et al, 2002 (cross-sectional)	310	73	Morning fasting TT, BT	TT 426 ng/dL BT 127 ng/dL	GLOB 1 MEM 2 ATT 5	Positive association between BT and GLOB 1, MEM 2, ATT 5
Yaffe et al, 2007 (longitudinal)	439	75.4 (70–79)	FT	FT 6.1 - 6.4 pg/mL	GLOB 2 MEM 14 VS 4	No significant associations between FT and any tests
Yeap et al, 2008 (cross-sectional)	2932	(70–89)	Morning TT, FT	TT 14.8 - 15.2 nmol/L FT 262 - 278 pmol/L	GLOB 1	Positive association between TT and FT and GLOB 1
Yonker et al, 2006 (cross-sectional)	450	54.3 (35–80)	FT	FT 49.5 pg/mL	GLOB 1 MEM 3 VS 3 ATT 6, 15	Negative association between FT and VS 3

*Table 1 explains the abbreviated cognitive domains and the numbered cognitive tests.

T = testosterone. **TT** = total testosterone. **FT** = free testosterone. **BT** = bioavailable testosterone.

TABLE 3

Controlled Studies of Testosterone Supplementation in Cognitively Normal Older Men with Low Testosterone

Authors, Year	N	Mean Age (Range)	Method of T Measurement	Baseline T Level	Cognitive Measures*	Supplementation Method	Significant Conclusions
Cherrier et al, 2003	12	57 (34–70)	Morning TT	TT < 300 ng/dL	MEM 4, 19 VS 7, 13 ATT 3	T gel 50 or 100 mg/day for 180 days	Improvement in MEM 4 with T
Kenny et al, 2002	67	76 (65–87)	TT, BT	BT < 128 ng/dL	MEM 1, 2 ATT 4, 5	T transdermally 5 mg/day for 1 year	Improvement in ATT 5 with T
Sih et al, 1997	32	68 (51–79)	TT, BT	BT < 60 ng/dL	MEM 8, 9, 10, 11	T cypionate intramuscularly 200 mg every 14–17 days for 1 year	No significant differences in any tests with T
Vaughan et al, 2007	69	70.8 (65–83)	Morning TT, BT, DHT	TT < 350 ng/dL	MEM 1, 12, 14 VS 2 ATT 4, 5	T enanthate intramuscularly 200 mg every 2 weeks (± finasteride) for 3 years	Improvement in MEM 1 with T, and in MEM 14 with T + finasteride

* Table 1 explains the abbreviated cognitive domains and the numbered cognitive tests.

T = testosterone. **TT** = total testosterone. **BT** = bioavailable testosterone. **DHT** = dihydrotestosterone.

TABLE 4

Studies of Testosterone Supplementation and Visuospatial Function in Older Men

Visuospatial Task*	Low Testosterone State		Eugonadal State	
	No Change	Improvement	No Change	Improvement
3D Spatial Memory Test				Cherrier et al, 2001
Block Design (VS 3)	Haren et al, 2005			Cherrier et al, 2001, 2004
Figure Discrimination Task			Young et al, 2010	
Judgment of Line Orientation (VS 2)	Vaughan et al, 2007			
Mental Rotation Task			Emmelot-Vonk et al, 2008	Young et al, 2010
Route Test		Cherrier et al, 2003 (with dihydrotestosterone)		
Rey Visual Design Learning Test (MEM 10)	Sih et al, 1997			
Rey Visual Design Learning Test Recall (MEM 11)	Sih et al, 1997			
Spatial Array Learning Test (VS 13)		Cherrier et al, 2003 (with dihydrotestosterone)	Cherrier et al, 2001	
Visuospatial Memory				Gray et al, 2005

*Table 1 explains the abbreviated cognitive domains and the numbered cognitive tests. If an entry does not list a domain and test number, the task is not often used in clinical care and the reader is referred to the original source for more information.

TABLE 5

Controlled Studies of Testosterone Supplementation in Cognitively Impaired Older Men with Normal Testosterone

Authors, Year	N	Mean Age (Range) If Reported	Method of T Measurement	Baseline T Level (Cognitive Impairment)	Cognitive Measures*	Supplementation Method	Significant Conclusions
Cherrier et al, 2005	32	76 (63–85)	TT	"Low-normal" (Alzheimer disease or mild cognitive impairment)	MEM 4, 19 VS 3, 7 ATT 3, 5	T enanthate intramuscularly 100 mg weekly for 6 weeks	Improvement in MEM 3, 19 and VS 3, 7 with T
Fukai et al, 2010	11	81	Fasting morning TT	Not reported (mild to moderate cognitive impairment)	GLOB 1, 3	T undecanoate orally 40 mg/day for 6 months	Improvement in GLOB 1, 3 with T
Lu et al, 2006	47	70	TT, DHT, FT	Not reported (Alzheimer disease)	MEM 7 VS 2, 3, 14	T gel 75 mg/day for 24 weeks	Less decline in T group versus placebo in VS 14

*Table 1 explains the abbreviated cognitive domains and the numbered cognitive tests.

T = testosterone. **TT** = total testosterone. **AD** = Alzheimer disease. **DHT** = dihydrotestosterone. **FT** = free testosterone.

TABLE 6

Controlled Studies of Testosterone Supplementation in Cognitively Impaired Older Men with Low Testosterone

Authors, Year	N	Mean Age (Range)	Method of T Measurement	Baseline T Level (Cognitive Impairment)	Cognitive Measures*	Supplementation Method	Significant Conclusions
Kenny et al, 2004	11	80 (73–87)	Fasting TT, BT	BT < 128 ng/dL (early cognitive decline)	MEM 1 VS 4, 5 ATT 5, 6	T enanthate intramuscularly 200 mg every 3 weeks for 12 weeks	No significant differences in any tests with T
Tan and Pu, 2003	10	72.4 (68–80)	TT, BT	TT < 250 ng/dL (Alzheimer disease)	GLOB 1, 4 VS 4	T enanthate intramuscularly 200 mg every 2 weeks for 1 year	Nonsignificant improvement in VS 4 with T

*Table 1 explains the abbreviated cognitive domains and the numbered cognitive tests.

T = testosterone. **TT** = total testosterone. **BT** = bioavailable testosterone.

TABLE 7

Factors Affecting the Methodological Quality of Controlled Studies of Testosterone Supplementation

Authors, Year	Participants Lost to Follow-Up	Months of Follow-Up	Morning Blood Sampling	Participants Free of Other Treatments That Could Profoundly Affect Testosterone Levels	Randomized Participant Groups	Double-Blinded Study	For-Profit Funding Source
Vaughan et al, 2007	23/69 (33%)	36	Yes	Yes	Yes	Yes	No
Lu et al, 2006	9/47 (19%)	24	No	Yes	Yes	Yes	Partial
Sih et al, 1997	10/32 (31%)	12	No	No	Yes	Yes	No
Cherrier et al, 2003	0/12 (0%)	6	Yes	Yes	Yes	Yes	Partial
Tan and Pu, 2003	Not reported	12	No	No	Yes	Not reported	No
Kenny et al, 2002	23/67 (33%)	12	No	Yes	Yes	Not reported	No
Cherrier et al, 2005	3/32 (9%)	3	No	Yes	Yes	Yes	No
Kenny et al, 2004	0/11 (0%)	2.5	Yes	No	Yes	Yes	No
Fukai et al, 2010	Not reported	6	Yes	No	No	Not reported	No