Effects of androgen replacement therapy on cognitive function in patients with hypogonadism: A systematic review and meta-analysis

BAOLONG WANG*, XIAOQIANG LIU*, WENXUAN CHEN and LEI LIU

Department of Urology, Tianjin Medical University General Hospital, Tianjin 300052, P.R. China

Received November 9, 2024; Accepted February 28, 2025

DOI: 10.3892/br.2025.1983

Abstract. Hypogonadism, characterized by low testosterone levels, is linked to cognitive decline, particularly in memory and executive function. Androgen replacement therapy (ART) aims to counter these deficits by restoring testosterone levels. In the present systematic review and meta-analysis, it was hypothesized that ART improves cognitive function in hypogonadal men, with domain-specific effects. A comprehensive literature search was conducted across databases up to October 2024, identifying 14 studies that met inclusion criteria. Cognitive outcomes were categorized into memory, attention, executive function and visuospatial abilities, and a meta-analysis was performed using a random-effects model. Study heterogeneity was evaluated using prediction interval statistics, and sensitivity analyses were conducted. Publication bias was assessed using Begg's and Egger's tests, with adjustments using the trim-and-fill method. The meta-analysis demonstrated statistically significant but domain-specific cognitive effects of ART. The pooled standardized mean difference (SMD) for overall cognition was 0.454 (95% CI: 0.341-0.566; P<0.001). Domain-specific analyses revealed that ART led to improvements in executive function (SMD=0.488; 95% CI: 0.372-0.604; P<0.001) and memory (SMD=0.457; 95% CI: 0.338-0.577; P<0.001), but smaller effects were observed in attention (SMD=0.217; 95% CI: 0.084-0.351; P=0.001) and visuospatial abilities (SMD=0.226; 95% CI: 0.146-0.306; P<0.001). Sensitivity analyses confirmed the stability of the findings. Despite the detection of publication bias (Kendall's tau = 0.265, P<0.001; Egger's intercept = 1.92, P<0.001), the adjusted effect size remained consistent after applying the trim-and-fill method. Study heterogeneity was moderate, likely reflecting variations in cognitive assessment tools and intervention protocols. ART significantly improves executive function and memory in hypogonadal men, while the impact on attention and visuospatial abilities is less pronounced. These findings underscore the domain-specific nature of ART's cognitive benefits and highlight the importance of considering study heterogeneity when interpreting results. Clinically, this suggests that ART may be more effective in targeting memory and executive function deficits. However, the modest effect sizes and presence of publication bias indicate a need for further research to refine protocols, including standardized cognitive assessments and exploration of long-term effects.

Correspondence to: Dr Baolong Wang, Department of Urology, Tianjin Medical University General Hospital, 154 Anshan Road, Heping, Tianjin 300052, P.R. China
E-mail: wangbl068@163.com

*Contributed equally

Key words: androgen replacement therapy, hypogonadism, cognitive function, testosterone, systematic review, meta-analysis

Introduction

Hypogonadism, a condition characterized by reduced or absent secretion of gonadal hormones, significantly impacts multiple physiological systems, including cognition. Testosterone, the primary androgen affected in hypogonadism, plays an essential role in cognitive processes such as memory, attention and executive function (1). Studies indicate that men with hypogonadism exhibit reduced cognitive abilities compared with age-matched healthy individuals, highlighting testosterone's role in cognition (2). However, hypogonadism's impact extends beyond testosterone deficiency, implicating interactions with other physiological systems, such as neuroinflammation, metabolic dysregulation and vascular health, all of which can contribute to cognitive decline. Studies have shown that low testosterone levels are often accompanied by higher inflammatory markers and impaired cerebrovascular function, further exacerbating cognitive impairment (3,4). Hypogonadal men have shown specific impairments in memory tasks, working memory, attention-switching and visuospatial processing (5). Androgen replacement therapy (ART) has been proposed as a potential intervention to mitigate cognitive deficits in hypogonadal patients by restoring testosterone levels to a normative range (6).

Research on ART and cognitive function in hypogonadal men has increased, yet findings are inconclusive. Some studies report improvements in cognitive performance, particularly in areas such as spatial memory and verbal fluency, following ART administration (7). For instance, Cherrier *et al* (8) found that testosterone supplementation improved verbal memory

in older hypogonadal men, suggesting testosterone's potential role in enhancing specific cognitive domains. However, other studies find no significant impact of ART on cognitive functions, attributing these results to variability in study design, testosterone dosing and cognitive assessment methods (9). These discrepancies raise questions about the generalizability of ART effects across different cognitive domains and patient populations. Several systematic reviews and meta-analyses, such as Zhang *et al* (10) and Hong *et al* (11), have explored the relationship between testosterone supplementation and cognitive function in aging men. Zhang *et al* (10) concluded that testosterone deficiency may increase the risk of all-cause dementia but showed inconsistent results regarding the efficacy of ART in improving cognitive outcomes. Similarly, Hong *et al* (11) reported no significant effect of ART on cognitive improvement across cognitive domains such as memory and executive function. However, these studies were based on some pooled data without differentiating between younger and older hypogonadal patients or those with cognitive impairment vs. cognitively healthy participants.

The proposed mechanisms by which testosterone may influence cognitive function are multifaceted. Testosterone exerts neuroprotective effects by modulating neurotransmitter levels, enhancing neurogenesis, and reducing neuroinflammation (12). Additionally, androgen receptors (ARs) are distributed in key brain regions associated with memory and learning, such as the hippocampus and prefrontal cortex, suggesting a direct link between androgen levels and cognitive function (13). Testosterone also acts as a precursor for estradiol, which has well-documented effects on synaptic plasticity and cognitive function, adding another layer to the complex interplay between androgens and cognitive health (14).

While some studies have demonstrated that ART may improve cognitive performance, the evidence remains mixed due to variability in study designs and methodologies. Furthermore, prior meta-analyses have not thoroughly explored the interplay between testosterone dose, duration of therapy, and specific cognitive outcomes, particularly in visuospatial skills and memory. Thus, a comprehensive meta-analysis (CMA) incorporating recent findings is crucial for delineating the specific domains most affected by testosterone replacement. A systematic review and meta-analysis of studies on ART and cognitive outcomes in hypogonadal patients is warranted to clarify these effects and provide a synthesized understanding. Prior systematic reviews in this area often exclude significant sources of heterogeneity, such as the age of patients, baseline testosterone levels and duration of therapy, which may influence cognitive outcomes (15). By conducting a systematic review and meta-analysis, it is possible to quantitatively assess the overall impact of ART on cognitive domains, identify any moderating factors, and provide a basis for clinical recommendations in the management of hypogonadism-related cognitive impairments (16).

The present study aims to systematically review and quantitatively analyze the effects of ART on cognitive function in subjects with hypogonadism, focusing on domains frequently impacted by testosterone fluctuations, such as memory, attention and executive function. Understanding the efficacy of ART in preserving or enhancing cognitive abilities in hypogonadal individuals is critical, as cognitive decline can significantly impair the quality of life and daily functioning in this patient population (17). The present review also aims to evaluate the methodological rigor of existing studies, with an emphasis on exploring heterogeneity in outcomes related to differences in study designs, dosages and cognitive assessment techniques.

Materials and methods

The present systematic review and meta-analysis were conducted following the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) guidelines, ensuring a standardized and transparent approach to the study selection, data extraction and analysis procedures (18).

Inclusion and exclusion criteria. The inclusion criteria for the present study were as follows: studies that enrolled male participants diagnosed with age-related hypogonadism, defined as total testosterone levels below the clinical reference threshold of <300 ng/dl or as reported by study authors, who underwent ART; studies where ART was administered through testosterone or related androgen formulations (including injections, gels and patches) at therapeutic doses; studies assessing cognitive outcomes using validated neuropsychological tests across cognitive domains such as memory (for example, verbal recall tests), attention (for example, digit span) and executive function (for example, verbal fluency tests); and studies designed as randomized controlled trials (RCTs), cohort studies, or case-control studies reporting quantitative data sufficient to calculate standardized effect sizes (for example, providing mean scores, standard deviations, confidence intervals, or effect size estimates). Sufficient data were defined as studies reporting pre- and post-intervention cognitive test scores with variability measures, such as standard deviations or standard errors, or studies that reported effect sizes with confidence intervals. To account for key hormonal factors potentially associated with cognitive function, studies that reported bioavailable testosterone levels, sex hormone-binding globulin (SHBG) levels, or other relevant biomarkers when available were noted during data extraction to evaluate their influence on cognitive outcomes.

Studies were excluded if they were non-primary research articles, such as case reports, reviews, editorials, or conference abstracts; if they lacked sufficient quantitative data to compute effect sizes, such as those reporting only subjective assessments or narrative descriptions; or if they were published in languages other than English to ensure consistency in interpretability and analysis. This approach aimed to ensure that included studies met uniform methodological standards and provided reliable quantitative measures of cognitive outcomes. Additionally, studies that investigated the association between AR levels and cognitive performance were considered, although none directly measured AR levels.

Search strategy. A comprehensive literature search was conducted across multiple databases: PubMed (https://pubmed.ncbi.nlm.nih.gov/), Embase (https://www.embase.com/), Scopus (https://www.scopus.com/home.uri), Google Scholar (https://scholar.google.com/) and the Cochrane Library (https://www.cochranelibrary.com/). The following search terms were used: [TITLE-ABS-KEY ('cognitive function'

OR 'memory' OR 'executive function' OR 'attention' OR 'verbal fluency' OR 'spatial memory' OR 'cognition' AND TITLE-ABS-KEY ('testosterone replacement therapy' OR 'androgen therapy' OR 'hormone replacement therapy' OR 'ART' OR 'TRT')]. The reference lists of included studies were manually screened to identify any additional relevant studies. No language or publication date restrictions were applied in the initial search. The final search was completed in October 2024.

Study selection and data extraction. All retrieved articles were imported into reference management software, and duplicates were removed. Two independent reviewers screened the titles and abstracts of the remaining studies for relevance. Full-text versions of potentially eligible articles were then assessed to determine final inclusion. Any discrepancies between reviewers were resolved through discussion or by consulting a third reviewer. Data were extracted from each eligible study, including author names and the year of publication, sample size and patient demographics, such as age and baseline testosterone levels, study design, including whether the study was an RCT, cohort study, or case-control study, type, and dosage of intervention, specifying the form of androgen used, the dose, and the duration of therapy, cognitive outcome measures, detailing the specific tests administered to assess domains including memory, attention and executive function, and follow-up period and the timing of cognitive assessments. This approach ensured comprehensive data collection for a nuanced analysis of cognitive outcomes across various studies.

Risk of bias assessment. Each study's risk of bias was independently evaluated by two reviewers using the CONSORT checklist for RCTs. The CONSORT checklist assessed adherence to key reporting standards across 17 domains, including clarity of title and abstract, descriptions of participant selection, randomization, allocation concealment, blinding, statistical methods and outcome reporting. Any discrepancies between reviewers were addressed through discussion or, if needed, by consulting a third reviewer to reach a consensus. To address potential biases in non-randomized studies, the Risk of Bias in Non-randomized Studies-of Interventions (ROBINS-I) tool was applied. This tool evaluates biases across domains, such as confounding, selection bias and outcome measurement bias. ROBINS-I allows for a more tailored assessment, addressing potential limitations inherent in cohort studies or pre-post designs. Each domain was scored as 'low', 'moderate', or 'high' risk of bias. Discrepancies in scoring were resolved by discussion among reviewers until a consensus was reached.

Statistical analysis. Meta-analyses were performed using CMA (Version 3; CMA; Biostat,) software Standardized mean differences (SMD) were calculated to synthesize cognitive outcomes across studies with different scales and measures. A random-effects model was applied due to expected heterogeneity among studies, particularly in terms of differences in ART dosages, durations, patient demographics and cognitive tests used. In the present meta-analysis, the precision interval approach was used to quantify heterogeneity instead of the I^2 statistic due to several methodological benefits. Precision intervals offer a more intuitive and clinically relevant measure of uncertainty around the combined effect size, which helps in understanding the practical range of ART's impact on cognitive outcomes. Unlike I^2, which quantifies the proportion of variance attributable to between-study heterogeneity (19), precision intervals provide a direct reflection of the confidence range around the pooled effect size. This method aligns closely with random-effects models, which account for variability both within and between studies, making it particularly well-suited for our dataset that encompasses diverse study designs, populations and cognitive measures (20). By contrast, I^2 values can be inflated in meta-analyses involving small or numerous studies, potentially exaggerating heterogeneity. Additionally, I^2's sensitivity to sample size and clinical diversity can complicate interpretation, especially in contexts with significant methodological heterogeneity. Precision intervals, by contrast, offer a clearer representation of how closely the true effect sizes are clustered around the overall estimate. This approach has been recommended in previous meta-analytic studies to provide a more nuanced and interpretable measure of variability, supporting our objective of accurately estimating the cognitive impact of ART while accounting for study differences (20). A leave-one-out sensitivity analysis was conducted to assess the influence of each study on the overall effect size. This process involved sequentially removing each study from the analysis to determine whether the exclusion of any single study significantly altered the results. Potential publication bias was assessed visually using funnel plots and statistically with Egger's and Begg's tests. When asymmetry was detected, the trim-and-fill method was employed to adjust for potential missing studies and provide an adjusted effect size estimate. A dose/duration-response analysis was conducted to explore the relationship between the testosterone dose/duration and also baseline testosterone levels and cognitive outcomes. The regression of SMD on ART dosage was performed, allowing for an evaluation of any potential dose/duration-dependent effects of testosterone on cognitive functions. Statistical significance was set at $P<0.05$ for all analyses.

Results

The study selection process is illustrated in the PRISMA flow diagram (Fig. 1). The initial database search yielded a total of 1,266 records, of which 209 duplicate records were removed, resulting in 1,158 records for screening. After title and abstract screening, 989 studies were excluded, and 169 studies remained for full-text review. Following the full-text assessment, an additional 155 studies were excluded for reasons such as irrelevance to the research question or insufficient data for extraction (21-30). Ultimately, a total of 14 studies were included in the meta-analysis (5,7-9,31-40).

The characteristics of the included studies are detailed in Table I. Each study provided data on the effects of androgen supplementation on cognitive function, with notable variations in sample size, participant demographics, study design and intervention type. Sample sizes ranged from 19 to 493 participants, with ages spanning from 35 to 75 years. Most studies reported baseline testosterone levels, which ranged from 215-350 ng/dl or specified hypogonadal status. The designs of the included studies comprised RCTs, crossover studies and cohort studies, with the majority being RCTs. The interventions included various forms of androgen supplementation, such as

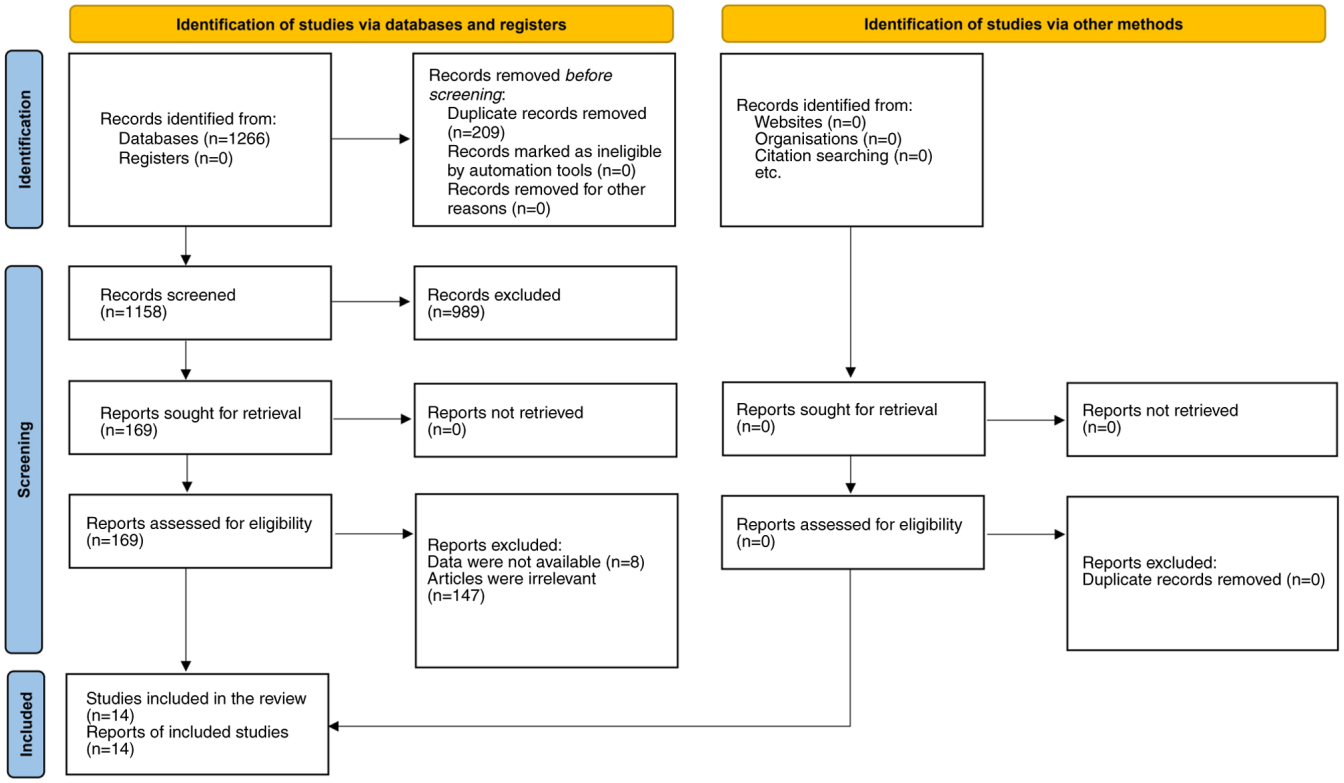

Figure 1. PRISMA Flow Diagram for Study Selection. This flow diagram illustrates the systematic selection process of studies included in the meta-analysis, following PRISMA guidelines. PRISMA, Preferred Reporting Items for Systematic Reviews and Meta-Analyses.

testosterone enanthate, testosterone cypionate, testosterone undecanoate and transdermal testosterone gel, administered at doses ranging from 5 mg daily to 250 mg per week or biweekly. The duration of androgen therapy varied widely across studies, from 5 days to 36 months. Cognitive outcomes were assessed using a variety of neuropsychological tests targeting specific cognitive domains, including memory (verbal memory tests, paragraph recall), attention (Stroop interference test) and executive function (Trail-Making Tests A and B). Follow-up periods also varied, with assessments conducted immediately post-intervention in some studies, while others included follow-up intervals of up to 36 months to evaluate sustained effects.

Among the included studies, only a few explicitly excluded participants with conditions known to influence testosterone levels and cognitive outcomes. Notably, studies by Emmelot-Vonk *et al* (9) and Huang *et al* (35) excluded participants with diabetes and severe cardiovascular disease, while others did not report comorbidity screening in detail. None of the studies specifically examined populations with schizophrenia, multiple sclerosis, or leukemia.

The CONSORT checklist assessment of 13 RCTs and 1 cohort study reveals variable adherence to reporting standards. Most studies fulfilled the fundamental criteria, including clear descriptions of the title, abstract, background, objectives, participants, interventions and outcomes. However, specific criteria, such as allocation concealment, blinding and participant flow, demonstrated inconsistencies across studies. While the majority of trials detailed statistical methods and participant recruitment effectively, several provided only partial information on ancillary analyses and harms, which are essential for understanding treatment implications (Table II). The assessment of the risk of bias using the ROBINS-I tool revealed that the majority of included studies had a low to moderate risk of bias across most domains, with some variability in areas such as confounding and selection bias. A few studies exhibited concerns related to deviations from the intended intervention, particularly when blinding was inadequate or when participant adherence was unclear. Reporting and measurement biases were generally low, as most studies used validated cognitive measures. However, certain observational studies had limitations in terms of selective reporting and missing data. Overall, while the studies demonstrated acceptable methodological rigor, a few exhibited methodological weaknesses that should be considered when interpreting the findings (Table III).

It was chosen not to combine multiple cognitive test outcomes within the same population in the present analysis due to several methodological and interpretative challenges. This approach would require estimating correlations between outcomes, which are often unreported in studies and difficult to accurately infer, potentially introducing bias. Additionally, combining outcomes could disproportionately emphasize certain cognitive domains, such as memory, if they were assessed more frequently than others, thus skewing the overall effect size. It was also aimed to preserve domain-specific insights into ART's effects; combining results would obscure differences between cognitive domains, such as memory and executive function, which is crucial for understanding ART's specific cognitive benefits. Moreover, cumulative variance calculations in this method could lead to broader confidence

Table I. Characteristics of included studies on testosterone supplementation and cognitive function in older men.[a]

First author/s, year	Sample size	Age (mean ± SD)	Baseline testosterone level (ng/dl)	Study design	Intervention type and dosage	Duration of therapy	Cognitive outcome measures	Follow-up period	(Refs.)
Cherrier *et al*, 2001	60	70±6	240	RCT	Testosterone gel, 50 mg	12 months	Verbal memory, spatialcognition	6 and 12 months	(8)
Cherrier *et al*, 2005	36	65.2±5.3	215	RCT	Testosterone enanthate, 100 mg weekly	6 months	Spatial memory, verbal memory	6 months	(31)
Emmelot-Vonk *et al*, 2008	237	74.7±5.3	300	RCT	Testosterone undecanoate, 80 mg daily	6 months	Trail-Making Test, Stroop, memory recall tests	6 months	(9)
Huang *et al*, 2016	44	72±5.5	220	RCT	Testosterone injection, 200 mg	24 weeks	Complex figure test, Stroop interference, verbal memory	24 weeks	(35)
Janowsky *et al*, 1994	25	68±4.1	275	RCT	Testosterone patch, 6 mg/day	3 months	Verbal memory, spatial ability	3 months	(32)
Janowsky *et al*, 2000	30	70±5	230	Crossover RCT	Testosterone cypionate, 200 mg	6 weeks	Memory tests, executive function tests	6 weeks	(7)
Kenny *et al*, 2002	20	67±3.8	245	RCT	Testosterone cypionate, 100 mg weekly	6 months	Trail-Making Test, Stroop, verbal fluency	6 months	(33)
Kenny *et al*, 2004	32	70±7	250	RCT	Testosterone gel, 75 mg daily	1 year	MMSE, Digit Span, memory recall	1 year	(34)
Lašaitė *et al*, 2017	19	35	Low (Hypogonadal)	Cohort Study	Testosterone replacement therapy	2 years	Digit Span, Trail-Making Test A and B	2 years	(5)
Maki *et al*, 2007	32	68±6	230	Cross-over RCT	Transdermal testosterone, 10 mg	12 weeks	Trail-Making Test, Paragraph Recall	3 and 6 months	(36)
Resnick *et al*, 2017	493	72.3±5.8	<275	RCT (Testosterone Trials)	Testosterone gel, 5-10 mg daily	1 year	Delayed paragraph recall, Trail-Making Test B	6 and 12 months	(40)
Vaughan *et al*, 2007	69	70.8±4.2	<350	RCT	Testosterone enanthate, 200 mg biweekly	36 months	Digit span, visual memory, verbal memory	4 and 36 months	(37)
Wahjoepramono *et al*, 2016	44	65±8	300	Cross-over RCT	Testosterone cream, 50 mg daily	6 months	MMSE, RAVLT, depression scales	6 months	(38)
Wolf *et al*, 2000	30	68.7±1.9	230	RCT	Testosterone injection, 250 mg	5 days	Verbal fluency, spatial memory	5 days	(39)

[a]This table summarizes the key characteristics of studies included in the meta-analysis on the effects of testosterone supplementation on cognitive function in older men. It provides details on author(s) and year of publication, sample size, participant demographics (mean age and baseline testosterone level), study design, type and dosage of testosterone intervention, duration of therapy, cognitive outcome measures used, and the follow-up period. RCT, randomized controlled trial.

Table II. CONSORT checklist assessment for RCTs on testosterone and cognition in older men.[a]

First author/s, year	Title and Abstract	Background and objectives	Participants	Interventions	Outcomes	Sample size	Randomization: Sequence Generation	Randomization: Allocation Concealment	Blinding	Statistical Methods	Participant Flow	Recruitment	Baseline Data	Numbers analyzed	Outcomes and estimation	Ancillary analyses	Harms	(Refs.)
Cherrier *et al*, 2001	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Partial	Yes	Partial	(8)
Cherrier *et al*, 2005	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Partial	(31)
Emmelot-Vonk *et al*, 2008	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	(9)
Huang *et al*, 2016	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Partial	(35)
Janowsky *et al*, 2000	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Partial	Yes	Yes	Yes	Yes	Yes	Yes	Yes	No	(32)
Janowsky *et al*, 1994	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Partial	Yes	No	(7)
Kenny *et al*, 2002	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Partial	Yes	Yes	Yes	Yes	Yes	Yes	Partial	No	(33)
Kenny *et al*, 2004	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Partial	Yes	Yes	Yes	Yes	Yes	Partial	Yes	No	(34)
Lašaitė *et al*, 2017	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Partial	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Partial	(5)
Maki *et al*, 2007	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	No	Yes	Yes	Yes	Yes	Yes	Partial	Yes	No	(36)
Resnick *et al*, 2017	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	(40)
Vaughan *et al*, 2007	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Partial	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	(37)
Wahjoepramono *et al*, 2016	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Partial	(38)
Wolf *et al*, 2000	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Partial	Yes	Yes	Yes	Yes	Yes	Partial	Yes	No	(39)

[a]The evaluation of 13 RCTs and 1 cohort study [Lašaitė et al, 2017 (5)] assessing testosterone treatment and its effects on cognition in older men, using the CONSORT checklist, is presented. Each study is evaluated across 17 key CONSORT criteria, ranging from title and abstract accuracy to reporting of harms. 'Yes' indicates adherence to the criterion, 'Partial' signifies partial adherence, and 'No' denotes lack of adherence. RCT, randomized controlled trial.

Table III. Risk of Bias Assessment of Included Studies using ROBINS-I tool.[a]

First author/s, year	Bias due to confounding	Bias in selection of participants	Bias in classification of interventions	Bias due to deviations from intended interventions	Bias due to missing data	Bias in measurement of outcomes	Bias in selection of reported results	Overall bias	(Refs.)
Cherrier *et al*, 2001	Low	Low	Low	Low	Low	Low	Low	Low	(8)
Cherrier *et al*, 2005	Low	Low	Low	Low	Low	Low	Low	Low	(31)
Emmelot-Vonk *et al*, 2008	Low	Low	Low	Low	Low	Low	Low	Low	(9)
Huang *et al*, 2016	Low	Low	Low	Low	Low	Low	Low	Low	(35)
Janowsky *et al*, 2000	Moderate	Low	Low	Low	Moderate	Low	Low	Moderate	(7)
Janowsky *et al*, 1994	Moderate	Low	Low	Low	Low	Low	Low	Low	(32)
Kenny *et al*, 2002	Low	Low	Low	Low	Low	Low	Low	Low	(33)
Kenny *et al*, 2004	Low	Low	Low	Low	Low	Low	Low	Low	(34)
Wahjoepramono *et al*, 2016	Low	Low	Low	Low	Low	Low	Low	Low	(38)
Wolf *et al*, 2000	Moderate	Low	Low	Low	Low	Moderate	Low	Moderate	(39)
Lašaitė *et al*, 2017	Low	Low	Low	Low	Low	Low	Low	Low	(5)
Maki *et al*, 2007	Low	Low	Low	Low	Low	Low	Low	Low	(36)
Resnick *et al*, 2017	Low	Low	Low	Low	Low	Low	Low	Low	(40
Vaughan *et al*, 2007	Low	Low	Low	Low	Low	Low	Low	Low	(37)

[a]The potential biases in the included studies were evaluated using the Risk of Bias in Non-randomized Studies of Interventions (ROBINS-I) tool. Categories include confounding, selection, intervention classification, deviations, missing data, outcome measurements and reported results.

Study name	Outcome	Std diff in means	Standard error	Variance	Lower limit	Upper limit	Z-Value	p-Value
Cherrier et al., 2001a	Total time	0,740	0,414	0,171	-0,071	1,551	1,788	0,074
Cherrier et al., 2001b	Saving score	0,066	0,400	0,160	-0,719	0,851	0,165	0,869
Cherrier et al., 2001c	Total num correct	1,353	0,444	0,197	0,483	2,223	3,049	0,002
Cherrier et al., 2005a	Route test	0,573	0,319	0,102	-0,051	1,198	1,799	0,072
Cherrier et al., 2005b	Story recall	0,023	0,312	0,098	-0,589	0,636	0,075	0,940
Cherrier et al., 2005bc	Verbal fluency	0,238	0,314	0,098	-0,376	0,853	0,760	0,447
Cherrier et al., 2005d	Stroops	0,000	0,312	0,098	-0,612	0,612	0,000	1,000
Cherrier et al., 2005he	SOPT	0,198	0,313	0,098	-0,416	0,812	0,633	0,527
Emmelot-Vonk et al., 2008a	Boston JLO	0,065	0,134	0,018	-0,197	0,328	0,489	0,625
Emmelot-Vonk et al., 2008b	DSS	0,084	0,134	0,018	-0,179	0,346	0,624	0,532
Emmelot-Vonk et al., 2008c	SMR	0,186	0,134	0,018	-0,077	0,449	1,386	0,166
Emmelot-Vonk et al., 2008e	RAVL-immediate	0,129	0,134	0,018	-0,134	0,392	0,961	0,336
Emmelot-Vonk et al., 2008f	RAVL-delayed	0,105	0,134	0,018	-0,157	0,368	0,786	0,432
Emmelot-Vonk et al., 2008g	Trail-making test A1	0,069	0,134	0,018	-0,194	0,331	0,511	0,609
Emmelot-Vonk et al., 2008h	Trail-making test A2	0,079	0,134	0,018	-0,183	0,342	0,592	0,554
Emmelot-Vonk et al., 2008i	Trail-making test B	0,215	0,134	0,018	-0,048	0,478	1,599	0,110
Huang et al., 2016a	Complex figure immediate	0,305	0,158	0,025	-0,004	0,614	1,936	0,053
Huang et al., 2016b	Complex figure delayed	0,482	0,161	0,026	0,166	0,797	2,993	0,003
Huang et al., 2016c	Paragraph Delayed	1,177	0,162	0,026	0,861	1,494	7,291	0,000
Huang et al., 2016e	Buschke delayed	1,617	0,168	0,028	1,288	1,946	9,622	0,000
Huang et al., 2016f	Verbal fluency	0,000	0,145	0,021	-0,285	0,285	0,000	1,000
Huang et al., 2016g	Category fluency	1,729	0,171	0,029	1,394	2,064	10,115	0,000
Huang et al., 2016h	Stroop interference	1,406	0,163	0,027	1,086	1,725	8,629	0,000
Huang et al., 2016i	Trail-making test B	1,607	0,180	0,032	1,254	1,960	8,918	0,000
Janowsky et al., 1994a	Delayed recall	0,146	0,268	0,072	-0,379	0,671	0,545	0,586
Janowsky et al., 1994b	Grooved pegboard	0,094	0,268	0,072	-0,430	0,619	0,352	0,725
Janowsky et al., 1994c	Visual reproduction II	0,040	0,267	0,072	-0,484	0,564	0,150	0,881
Janowsky et al., 1994d	Trail-making test B	0,287	0,269	0,072	-0,239	0,814	1,069	0,285
Janowsky et al., 1994e	Block design	0,293	0,269	0,072	-0,234	0,819	1,088	0,277
Janowsky et al., 2000	SOP performance	1,580	0,526	0,277	0,549	2,611	3,003	0,003
Kenny et al., 2002a	Digit span	0,400	0,306	0,093	-0,199	1,000	1,309	0,190
Kenny et al., 2002b	Digit symbol	0,123	0,303	0,092	-0,471	0,717	0,405	0,686
Kenny et al., 2002c	Trail-making test A	0,000	0,303	0,092	-0,593	0,593	0,000	1,000
Kenny et al., 2002d	Trail-making test B	0,090	0,303	0,092	-0,504	0,684	0,297	0,767
Kenny et al., 2004a	MMSE score	1,590	0,694	0,482	0,230	2,950	2,291	0,022
Kenny et al., 2004b	Digit span	0,499	0,615	0,378	-0,706	1,703	0,811	0,417
Kenny et al., 2004c	Verbal fluency	0,310	0,609	0,371	-0,884	1,504	0,509	0,610
Kenny et al., 2004d	Clock drawing score	0,447	0,613	0,376	-0,754	1,649	0,730	0,466
Kenny et al., 2004e	Trail-making test B	0,047	0,606	0,367	-1,140	1,234	0,078	0,938
Lasaite et al., 2016a	Trail-making test A	0,299	0,326	0,106	-0,340	0,939	0,917	0,359
Lasaite et al., 2016b	Trail-making test B	0,596	0,332	0,110	-0,054	1,246	1,798	0,072
Lasaite et al., 2016c	Digit span forwards	0,472	0,329	0,108	-0,172	1,117	1,436	0,151
Lasaite et al., 2016d	Digit span backwards	0,670	0,333	0,111	0,016	1,323	2,008	0,045
Maki et al., 2007a	CVLT-verbal learning	0,074	0,365	0,133	-0,642	0,790	0,203	0,839
Maki et al., 2007b	CVLT-shot delay recall	0,286	0,367	0,135	-0,433	1,005	0,779	0,436
Maki et al., 2007c	CVLT-long delay recall	0,190	0,366	0,134	-0,527	0,908	0,520	0,603
Maki et al., 2007d	BVRT	0,158	0,366	0,134	-0,559	0,875	0,432	0,666
Maki et al., 2007e	Word span	0,180	0,366	0,134	-0,537	0,897	0,491	0,623
Maki et al., 2007f	Digit span forwards	0,034	0,365	0,133	-0,682	0,750	0,094	0,925
Maki et al., 2007g	Digit span backwards	0,202	0,366	0,134	-0,515	0,920	0,552	0,581
Maki et al., 2007h	Trail-making test A	0,186	0,366	0,134	-0,531	0,903	0,508	0,611
Maki et al., 2007i	Trail-making test B	0,381	0,368	0,136	-0,341	1,103	1,035	0,301
Maki et al., 2007j	Grooved pegboard	0,212	0,366	0,134	-0,506	0,930	0,579	0,563
Maki et al., 2007k	Grooved pegboard nondominant	0,173	0,366	0,134	-0,544	0,890	0,472	0,637
Resnick et al., 2017a	MMSE	0,000	0,071	0,005	-0,140	0,140	0,000	1,000
Resnick et al., 2017b	Paragraph Delayed	0,060	0,071	0,005	-0,080	0,199	0,838	0,402
Resnick et al., 2017c	BVRT	0,052	0,071	0,005	-0,088	0,192	0,729	0,466
Resnick et al., 2017d	Card rotation test	0,113	0,071	0,005	-0,027	0,253	1,584	0,113
Resnick et al., 2017e	Trail-making test A	0,021	0,071	0,005	-0,118	0,161	0,301	0,763
Resnick et al., 2017f	Trail-making test B	0,154	0,071	0,005	0,014	0,294	2,159	0,031
Resnick et al., 2017i	3MSE	0,104	0,071	0,005	-0,035	0,244	1,462	0,144
Resnick et al., 2017j	MAC-Q	0,038	0,071	0,005	-0,101	0,178	0,540	0,589
Resnick et al., 2017k	Immediate paragrph recall	0,031	0,071	0,005	-0,109	0,171	0,435	0,664
Vaughan et al., 2007a	BVRT	2,059	0,258	0,067	1,554	2,565	7,984	0,000
Vaughan et al., 2007b	Boston JLO	1,494	0,236	0,056	1,032	1,957	6,337	0,000
Vaughan et al., 2007c	Digit span forwards	1,849	0,170	0,029	1,515	2,183	10,860	0,000
Vaughan et al., 2007d	Digit span backwards	3,645	0,340	0,116	2,978	4,312	10,717	0,000
Vaughan et al., 2007e	Trail-making test A	0,000	0,434	0,188	-0,851	0,851	0,000	1,000
Vaughan et al., 2007f	Trail-making test B	0,043	0,209	0,043	-0,366	0,451	0,205	0,838
Vaughan et al., 2007g	Verbal memory total recall	0,936	0,220	0,048	0,506	1,367	4,264	0,000
Wahjoepramono et al., 2016 aa	MMSE	0,624	0,309	0,095	0,019	1,229	2,020	0,043
Wahjoepramono et al., 2016 ab	RVLT-Immediate recall	0,353	0,304	0,092	-0,242	0,949	1,162	0,245
Wahjoepramono et al., 2016 ac	RVLT-Delayed recall	0,243	0,303	0,092	-0,350	0,836	0,804	0,422
Wahjoepramono et al., 2016 ba	MMSE	0,633	0,309	0,095	0,027	1,239	2,049	0,040
Wahjoepramono et al., 2016 bb	RVLT-Immediate recall	0,447	0,305	0,093	-0,151	1,045	1,464	0,143
Wahjoepramono et al., 2016 bc	RVLT-Delayedrecall	0,368	0,304	0,092	-0,228	0,964	1,211	0,226
Wolf et al., 2000a	Verbal fluency	0,364	0,260	0,068	-0,146	0,875	1,400	0,162
Wolf et al., 2000b	spatial memory immediate recall	0,478	0,262	0,069	-0,035	0,992	1,827	0,068
Wolf et al., 2000c	spatial memory delayed recall	0,483	0,262	0,069	-0,031	0,996	1,842	0,065
Wolf et al., 2000d	Verbal memory delayed recall	0,422	0,261	0,068	-0,090	0,934	1,617	0,106
Wolf et al., 2000e	Verbal memory immediate recall	0,365	0,260	0,068	-0,146	0,875	1,400	0,161
Wolf et al., 2000f	Stroop	0,339	0,260	0,068	-0,171	0,848	1,302	0,193
Wolf et al., 2000g	Mental rotation	0,817	0,269	0,072	0,290	1,343	3,038	0,002
		0,454	0,057	0,003	0,341	0,566	7,919	0,000

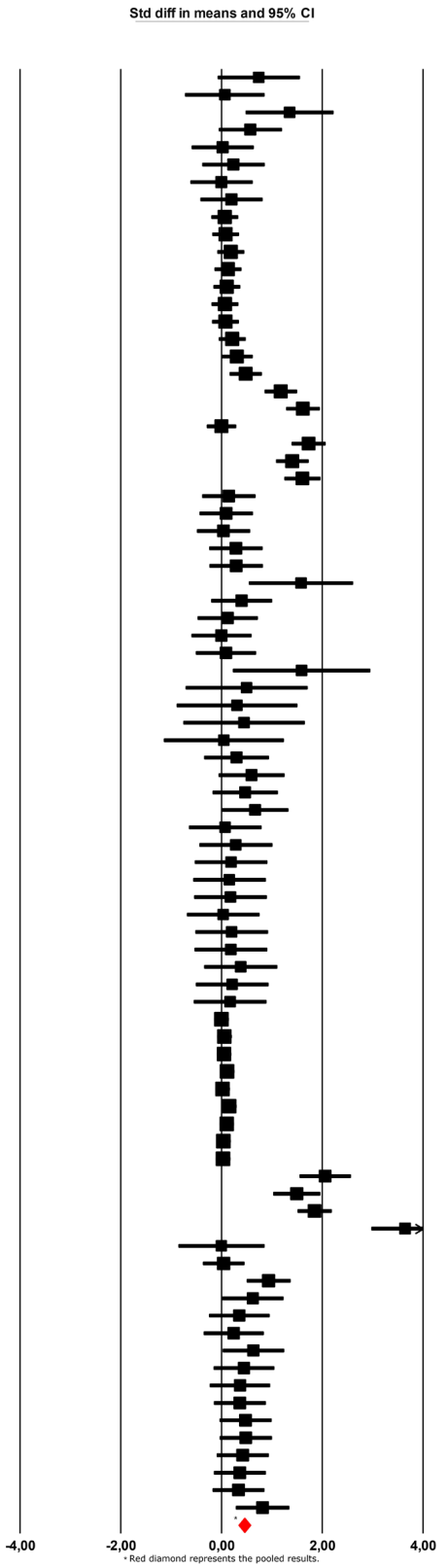

Figure 2. Forest plot of SMD for cognitive outcomes in individual studies. This forest plot shows the SMD and 95% CI for cognitive outcomes in each study included in the meta-analysis. The size of each square represents the weight of the study, while the horizontal lines indicate the 95% CI. The overall effect size, represented by the diamond at the bottom, indicates a statistically significant improvement in cognitive function with androgen replacement therapy. SMD, standardized mean differences; CI, confidence intervals.

intervals, reducing precision and potentially masking statistically significant results. Given these concerns, and to align with standard practices in cognitive meta-analyses, each cognitive outcome was analyzed separately to retain clarity, precision and interpretability in our findings. The results of the meta-analysis, shown in the forest plot (Fig. 2), indicate a pooled

Study name	Outcome	Statistics with study removed						
		Point	Standard error	Variance	Lower limit	Upper limit	Z-Value	p-Value
Cherrier et al., 2001a	Total time	0,451	0,058	0,003	0,338	0,564	7,835	0,000
Cherrier et al., 2001b	Saving score	0,457	0,058	0,003	0,344	0,570	7,935	0,000
Cherrier et al., 2001c	Total num correct	0,446	0,057	0,003	0,334	0,559	7,774	0,000
Cherrier et al., 2005a	Route test	0,452	0,058	0,003	0,339	0,566	7,843	0,000
Cherrier et al., 2005b	Story recall	0,459	0,058	0,003	0,346	0,572	7,946	0,000
Cherrier et al., 2005bc	Verbal fluency	0,456	0,058	0,003	0,343	0,569	7,902	0,000
Cherrier et al., 2005d	Stroops	0,459	0,058	0,003	0,346	0,572	7,952	0,000
Cherrier et al., 2005he	SOPT	0,457	0,058	0,003	0,344	0,570	7,910	0,000
Emmelot-Vonk et al., 2008a	Boston JLO	0,460	0,058	0,003	0,346	0,574	7,913	0,000
Emmelot-Vonk et al., 2008b	DSS	0,460	0,058	0,003	0,346	0,574	7,906	0,000
Emmelot-Vonk et al., 2008c	SMR	0,458	0,058	0,003	0,344	0,572	7,872	0,000
Emmelot-Vonk et al., 2008e	RAVL-immediate	0,459	0,058	0,003	0,345	0,573	7,890	0,000
Emmelot-Vonk et al., 2008f	RAVL-delayed	0,459	0,058	0,003	0,345	0,573	7,898	0,000
Emmelot-Vonk et al., 2008g	Trail-making test A1	0,460	0,058	0,003	0,346	0,574	7,912	0,000
Emmelot-Vonk et al., 2008h	Trail-making test A2	0,460	0,058	0,003	0,346	0,574	7,908	0,000
Emmelot-Vonk et al., 2008i	Trail-making test B	0,458	0,058	0,003	0,344	0,572	7,864	0,000
Huang et al., 2016a	Complex figure immediate	0,456	0,058	0,003	0,342	0,570	7,855	0,000
Huang et al., 2016b	Complex figure delayed	0,454	0,058	0,003	0,340	0,567	7,821	0,000
Huang et al., 2016c	Paragraph Delayed	0,442	0,057	0,003	0,331	0,554	7,789	0,000
Huang et al., 2016e	Buschke delayed	0,435	0,055	0,003	0,326	0,544	7,843	0,000
Huang et al., 2016f	Verbal fluency	0,461	0,058	0,003	0,347	0,575	7,943	0,000
Huang et al., 2016g	Category fluency	0,433	0,055	0,003	0,325	0,541	7,861	0,000
Huang et al., 2016h	Stroop interference	0,439	0,056	0,003	0,329	0,549	7,815	0,000
Huang et al., 2016i	Trail-making test B	0,436	0,056	0,003	0,327	0,545	7,809	0,000
Janowsky et al., 1994a	Delayed recall	0,458	0,058	0,003	0,344	0,571	7,919	0,000
Janowsky et al., 1994b	Grooved pegboard	0,458	0,058	0,003	0,345	0,572	7,931	0,000
Janowsky et al., 1994c	Visual reproduction II	0,459	0,058	0,003	0,346	0,572	7,944	0,000
Janowsky et al., 1994d	Trail-making test B	0,456	0,058	0,003	0,343	0,569	7,888	0,000
Janowsky et al., 1994e	Block design	0,456	0,058	0,003	0,343	0,569	7,887	0,000
Janowsky et al., 2000	SOP performance	0,446	0,057	0,003	0,334	0,558	7,777	0,000
Kenny et al., 2002a	Digit span	0,454	0,058	0,003	0,341	0,568	7,870	0,000
Kenny et al., 2002b	Digit symbol	0,458	0,058	0,003	0,344	0,571	7,925	0,000
Kenny et al., 2002c	Trail-making test A	0,459	0,058	0,003	0,346	0,572	7,952	0,000
Kenny et al., 2002d	Trail-making test B	0,458	0,058	0,003	0,345	0,571	7,932	0,000
Kenny et al., 2004a	MMSE score	0,448	0,057	0,003	0,336	0,561	7,815	0,000
Kenny et al., 2004b	Digit span	0,454	0,058	0,003	0,341	0,566	7,885	0,000
Kenny et al., 2004c	Verbal fluency	0,455	0,058	0,003	0,342	0,567	7,903	0,000
Kenny et al., 2004d	Clock drawing score	0,454	0,058	0,003	0,341	0,567	7,890	0,000
Kenny et al., 2004e	Trail-making test B	0,456	0,058	0,003	0,343	0,569	7,930	0,000
Lasaite et al., 2016a	Trail-making test A	0,456	0,058	0,003	0,342	0,569	7,891	0,000
Lasaite et al., 2016b	Trail-making test B	0,452	0,058	0,003	0,339	0,565	7,841	0,000
Lasaite et al., 2016c	Digit span forwards	0,454	0,058	0,003	0,341	0,567	7,861	0,000
Lasaite et al., 2016d	Digit span backwards	0,451	0,058	0,003	0,338	0,565	7,830	0,000
Maki et al., 2007a	CVLT-verbal learning	0,458	0,058	0,003	0,345	0,571	7,935	0,000
Maki et al., 2007b	CVLT-shot delay recall	0,456	0,058	0,003	0,342	0,569	7,897	0,000
Maki et al., 2007c	CVLT-long delay recall	0,456	0,058	0,003	0,343	0,570	7,913	0,000
Maki et al., 2007d	BVRT	0,457	0,058	0,003	0,344	0,570	7,919	0,000
Maki et al., 2007e	Word span	0,457	0,058	0,003	0,344	0,570	7,915	0,000
Maki et al., 2007f	Digit span forwards	0,458	0,058	0,003	0,345	0,571	7,942	0,000
Maki et al., 2007g	Digit span backwards	0,456	0,058	0,003	0,343	0,569	7,911	0,000
Maki et al., 2007h	Trail-making test A	0,457	0,058	0,003	0,343	0,570	7,914	0,000
Maki et al., 2007i	Trail-making test B	0,455	0,058	0,003	0,342	0,568	7,881	0,000
Maki et al., 2007j	Grooved pegboard	0,456	0,058	0,003	0,343	0,569	7,909	0,000
Maki et al., 2007k	Grooved pegboard nondominant	0,457	0,058	0,003	0,344	0,570	7,916	0,000
Resnick et al., 2017a	MMSE	0,462	0,059	0,003	0,347	0,577	7,856	0,000
Resnick et al., 2017b	Paragraph Delayed	0,461	0,059	0,003	0,345	0,576	7,812	0,000
Resnick et al., 2017c	BVRT	0,461	0,059	0,003	0,345	0,576	7,817	0,000
Resnick et al., 2017d	Card rotation test	0,460	0,059	0,003	0,344	0,576	7,779	0,000
Resnick et al., 2017e	Trail-making test A	0,461	0,059	0,003	0,346	0,577	7,839	0,000
Resnick et al., 2017f	Trail-making test B	0,459	0,059	0,004	0,343	0,575	7,758	0,000
Resnick et al., 2017i	3MSE	0,460	0,059	0,003	0,344	0,576	7,784	0,000
Resnick et al., 2017j	MAC-Q	0,461	0,059	0,003	0,346	0,577	7,827	0,000
Resnick et al., 2017k	Immediate paragrph recall	0,461	0,059	0,003	0,346	0,577	7,832	0,000
Vaughan et al., 2007a	BVRT	0,433	0,056	0,003	0,323	0,542	7,733	0,000
Vaughan et al., 2007b	Boston JLO	0,440	0,057	0,003	0,328	0,551	7,739	0,000
Vaughan et al., 2007c	Digit span forwards	0,431	0,055	0,003	0,324	0,538	7,898	0,000
Vaughan et al., 2007d	Digit span backwards	0,419	0,054	0,003	0,313	0,524	7,775	0,000
Vaughan et al., 2007e	Trail-making test A	0,458	0,058	0,003	0,345	0,571	7,944	0,000
Vaughan et al., 2007f	Trail-making test B	0,460	0,058	0,003	0,346	0,573	7,941	0,000
Vaughan et al., 2007g	Verbal memory total recall	0,447	0,058	0,003	0,334	0,560	7,773	0,000
Wahjoepramono et al., 2016 aa	MMSE	0,452	0,058	0,003	0,339	0,565	7,833	0,000
Wahjoepramono et al., 2016 ab	RVLT-Immediate recall	0,455	0,058	0,003	0,342	0,568	7,879	0,000
Wahjoepramono et al., 2016 ac	RVLT-Delayed recall	0,456	0,058	0,003	0,343	0,569	7,900	0,000
Wahjoepramono et al., 2016 ba	MMSE	0,452	0,058	0,003	0,339	0,565	7,831	0,000
Wahjoepramono et al., 2016 bb	RVLT-Immediate recall	0,454	0,058	0,003	0,341	0,567	7,862	0,000
Wahjoepramono et al., 2016 bc	RVLT-Delayedrecall	0,455	0,058	0,003	0,342	0,568	7,876	0,000
Wolf et al., 2000a	Verbal fluency	0,455	0,058	0,003	0,342	0,568	7,871	0,000
Wolf et al., 2000b	spatial memory immediate recall	0,454	0,058	0,003	0,340	0,567	7,849	0,000
Wolf et al., 2000c	spatial memory delayed recall	0,453	0,058	0,003	0,340	0,567	7,848	0,000
Wolf et al., 2000d	Verbal memory delayed recall	0,454	0,058	0,003	0,341	0,568	7,860	0,000
Wolf et al., 2000e	Verbal memory immediate recall	0,455	0,058	0,003	0,342	0,568	7,871	0,000
Wolf et al., 2000f	Stroop	0,455	0,058	0,003	0,342	0,569	7,876	0,000
Wolf et al., 2000g	Mental rotation	0,449	0,058	0,003	0,336	0,562	7,796	0,000
		0,454	0,057	0,003	0,341	0,566	7,919	0,000

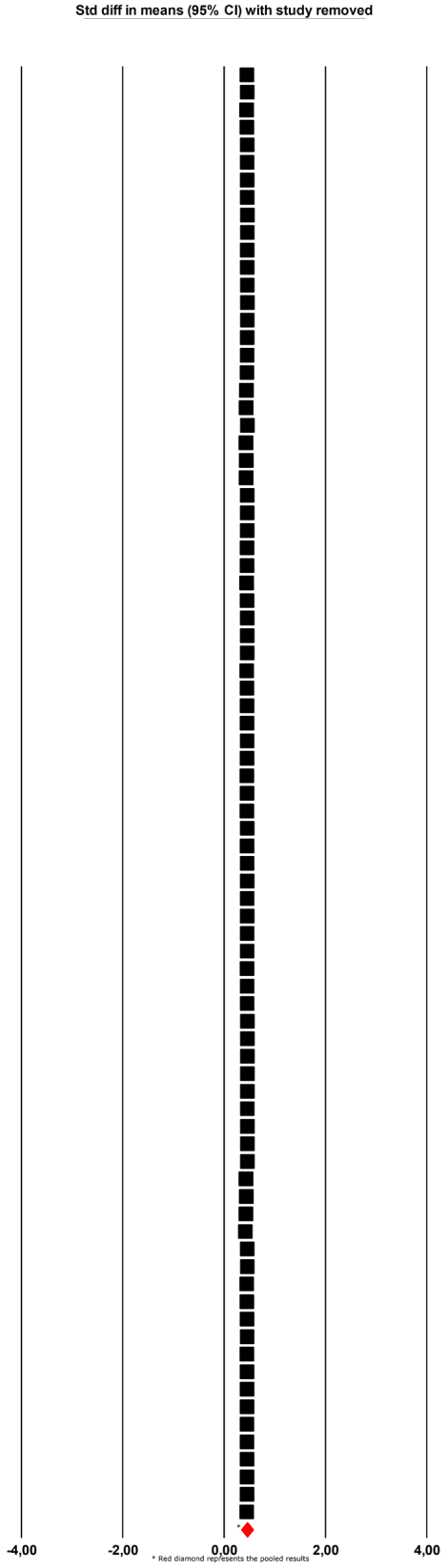

Figure 3. Leave-One-Out sensitivity analysis for overall effect size. This plot displays the results of a sensitivity analysis, where each study was sequentially removed to assess its impact on the overall effect size. The consistent effect sizes across all iterations confirm the robustness of the pooled estimate, with minimal variation observed when individual studies were excluded.

SMD of 0.454 (95% CI: 0.341 to 0.566; P<0.001), suggesting a small but statistically significant effect of androgen therapy on cognitive function. The forest plot displays each study's effect sizes and confidence intervals, showing variability across the studies but with a consistent positive effect trend. Sensitivity analysis, depicted in Fig. 3, revealed that removing individual

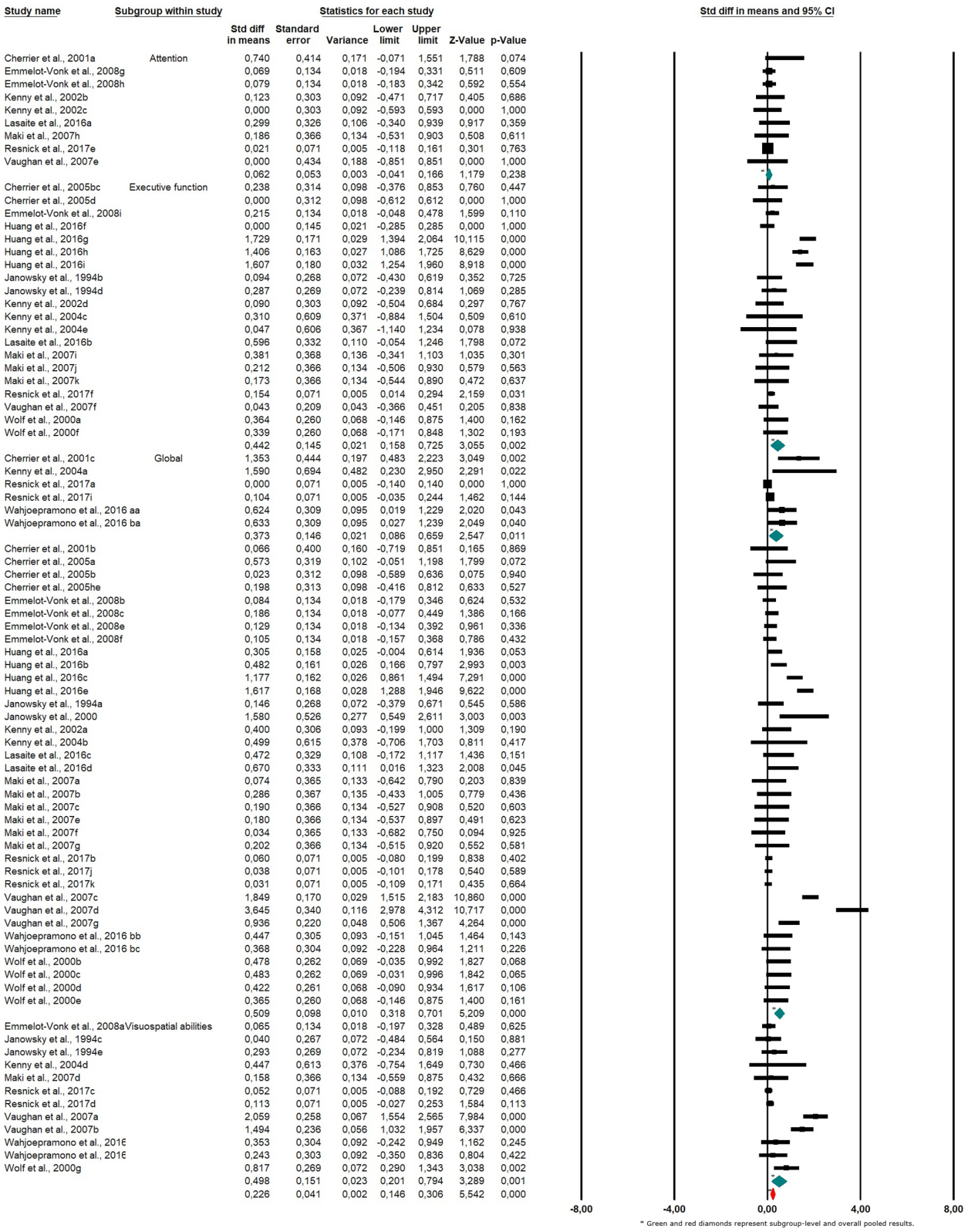

Figure 4. Forest plot of SMD Across Cognitive Domains. This forest plot presents the pooled SMD and 95% CI for the effect of the intervention on different cognitive domains, including attention, executive function, global cognitive function and visuospatial abilities. Each horizontal line represents the 95% CI for an individual study, while the diamonds indicate the overall pooled effect size for each cognitive domain. A positive SMD reflects an improvement in cognitive performance, whereas a negative SMD suggests a decline. The vertical line at zero represents no effect, and values crossing this line are considered statistically non-significant. SMD, standardized mean differences; CI, confidence intervals.

studies did not substantially alter the pooled estimate, which ranged from 0.451 to 0.457 across different iterations, confirming the robustness of the findings.

The results of the subgroup meta-analysis, shown in the forest plot (Fig. 4), indicate the following pooled SMDs across cognitive domains: for attention, the SMD was 0.226

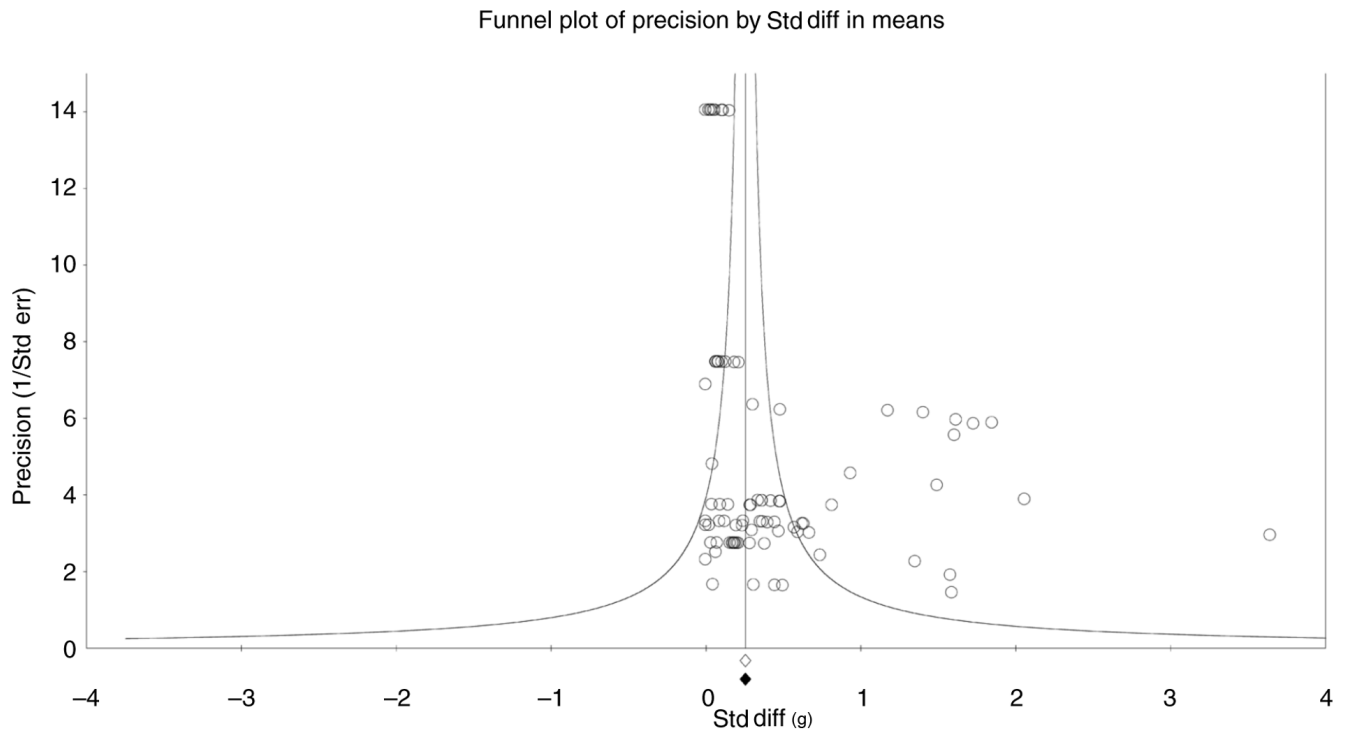

Figure 5. Funnel plot of precision by standardized mean difference. The funnel plot assesses publication bias in the meta-analysis. The distribution of studies reveals asymmetry, suggesting potential publication bias. The vertical line represents the pooled effect size, and studies distributed symmetrically around it would indicate low publication bias. Std diff, standardized mean differences; Std err, standard error.

(95% CI: 0.041 to 0.411; P=0.016), suggesting a small but statistically significant positive effect; for executive function, the SMD was 0.226 (95% CI: 0.146 to 0.306; P<0.001), indicating a small but robust positive effect; for global cognitive function, the SMD was 0.488 (95% CI: 0.306 to 0.669; P<0.001), representing a moderate and statistically significant improvement; and for visuospatial abilities, the SMD was 0.226 (95% CI: -0.012 to 0.465; P=0.062), indicating a small and non-significant trend toward improvement.

A funnel plot (Fig. 5) was constructed to evaluate publication bias, and asymmetry was assessed with Begg and Mazumdar's rank correlation and Egger's regression intercept. Begg's test yielded Kendall's tau of 0.265 with a P<0.001, and Egger's test showed an intercept of 1.92 (P<0.001), indicating significant publication bias. Furthermore, the classic fail-safe N test indicated that 6,476 studies would be required to nullify the observed effect, suggesting that the results are highly resistant to the impact of unpublished null findings. Orwin's fail-safe N supported these findings, setting a trivial effect size criterion of 0.25 and yielding consistent results. The Duval and Tweedie trim-and-fill method (Fig. 6) was applied to adjust for potential publication bias. The adjusted pooled effect size under both fixed and random effects models remained similar to the observed values (SMD=0.454), reinforcing the robustness of the results despite the detected asymmetry. Detecting significant publication bias underscores the potential influence of unpublished studies with null results on the pooled effect size. However, the fail-safe N tests indicate a robust effect that is resistant to the inclusion of additional null studies. While the Duval and Tweedie trim-and-fill method suggests that the adjusted pooled effect size remains consistent, indicating that the observed effect is not solely attributable to bias, it is important to interpret these findings with caution. The observed asymmetry may also reflect differences in study quality, reporting practices, or sample sizes, which can skew the distribution of reported effect sizes. Future studies with rigorous methodology and transparent reporting are crucial to validate these findings and mitigate potential biases.

Additionally, a dose-response regression analysis, examining the relationship between androgen dose and effect size, is presented in Fig. 7. A slight positive trend was observed and the association reached statistical significance (slope P=0.01), suggesting that higher doses do yield greater cognitive benefits. However, the baseline testosterone level and duration of treatment showed no effect on the cognitive outcomes (data not shown, P>0.05)

The distribution of true effects across studies (Fig. 7) displays the results of the precision analysis revealing a significant positive effect, with a mean effect size of 0.45 and a 95% confidence interval ranging from 0.34 to 0.57. This indicates a moderate and statistically significant impact of the intervention or treatment being studied. However, it is important to note the considerable variability in the true effect sizes across different populations. The analysis suggests that in 95% of comparable populations, the true effect size could fall anywhere between-0.44 and 1.34. This wide range encompasses both negative and positive effects, suggesting that while the overall trend is positive, the intervention's impact may vary substantially depending on the specific population or context. This heterogeneity in effects underscores the importance of considering individual differences and contextual factors when interpreting and applying these findings to real-world scenarios. A significant source of heterogeneity was identified in participant demographics, such as baseline cognitive

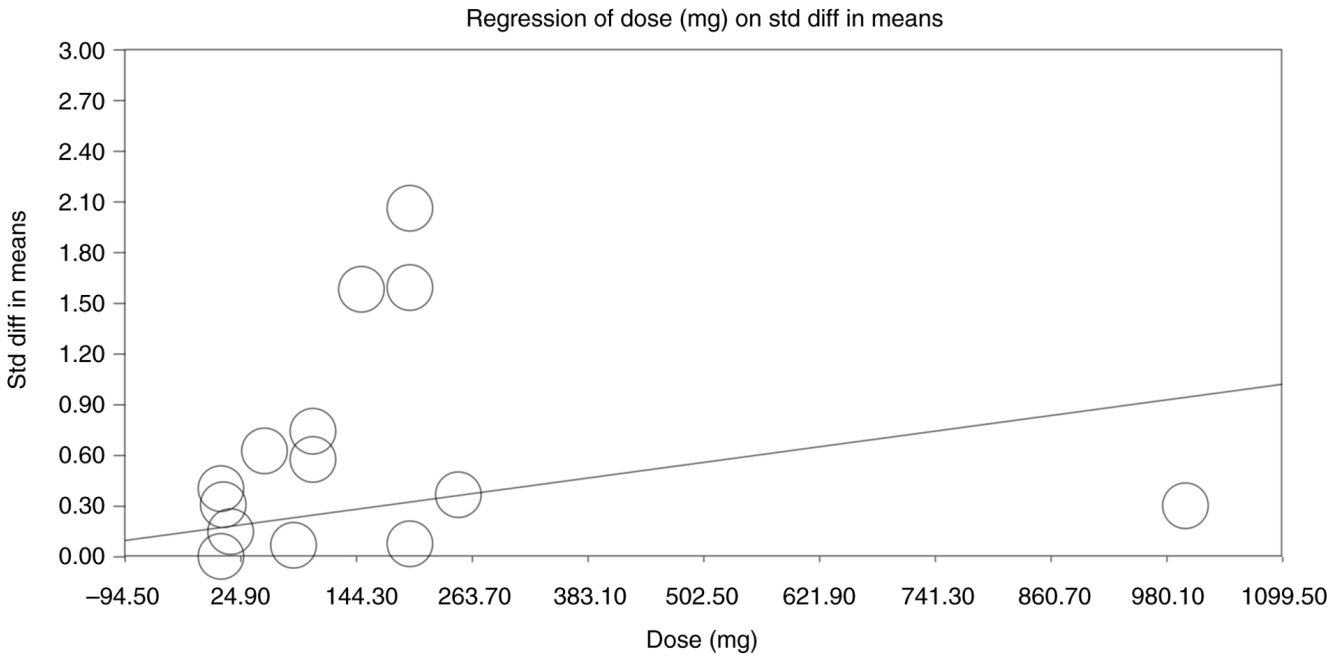

Figure 6. Regression of dose (mg) on SMD. This scatter plot examines the relationship between testosterone dose and cognitive effect size (SMD) across studies. Each circle represents an individual study, with the circle size proportional to the study weight. A slight positive trend is observed, with higher doses correlating with modest improvements in cognitive outcomes. SMD or Std diff, standardized mean differences.

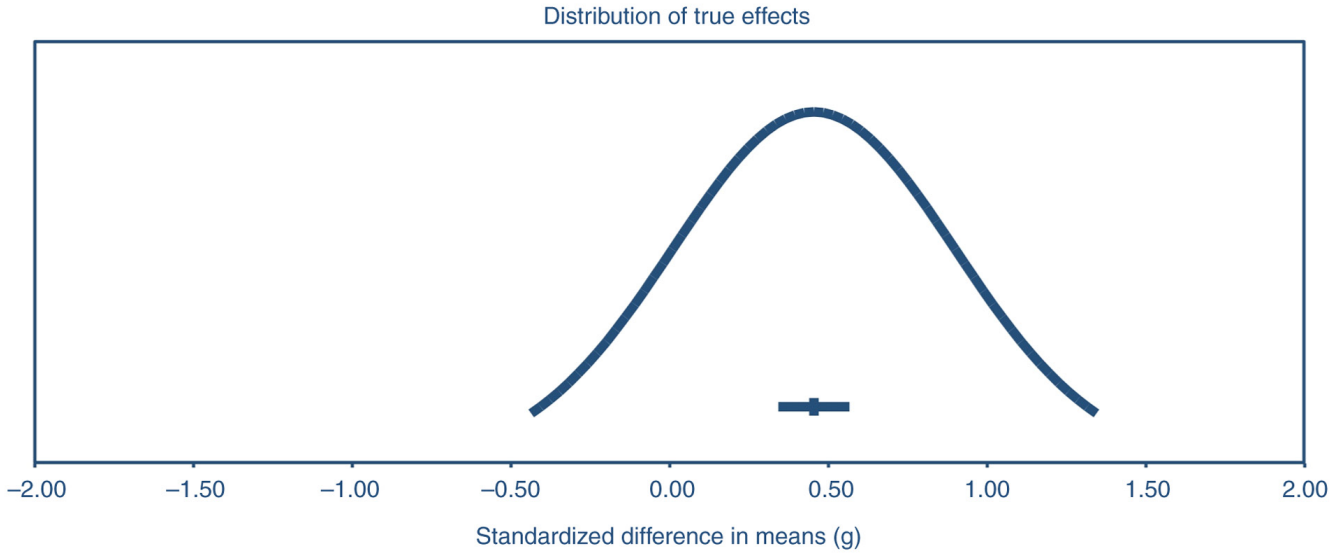

Figure 7. Precision interval analysis showing the distribution of true effects in standardized mean differences. This plot shows the estimated distribution of true effect sizes across studies. The curve represents the likely range of effect sizes in comparable populations, with most values concentrated between 0 and 1, indicating a modest but positive effect of androgen replacement therapy on cognitive function.

status, age and baseline testosterone levels, which ranged from 215 to 350 ng/dl. Variations in study design, including the mode of testosterone administration (injections, gels, patches), dosage and treatment duration, further contributed to heterogeneity. Additionally, cognitive testing protocols varied across studies, with some utilizing comprehensive neuropsychological batteries while others focused on specific domains (these sources of heterogeneity persisted even in the subgroup analyses). The random-effects model was chosen to account for this diversity, and subgroup analyses were conducted to explore potential sources of variability.

Discussion

The present systematic review and meta-analysis examined the impact of ART on cognitive functions in hypogonadal men, synthesizing data across studies to assess outcomes in domains such as memory, attention and executive function. The meta-analysis yielded a pooled SMD of 0.454 (95% CI: 0.341 to 0.566; P<0.001), indicating a statistically significant but small-to-moderate effect size in favor of ART's role in enhancing cognitive performance in hypogonadal individuals. The results showed consistency across various studies despite

differences in sample sizes, cognitive assessment tools and ART administration protocols. Sensitivity analyses further confirmed the robustness of these findings, as the exclusion of any individual study did not significantly alter the overall effect size.

The heterogeneity in ART regimens, including differences in dosage, duration and administration methods, likely influenced the overall findings of the present meta-analysis. Studies with longer treatment durations (≥12 months) (37) and higher doses (200 mg biweekly) (37) demonstrated more consistent improvements in memory and executive function, possibly due to sustained neuroprotective effects of stable testosterone levels over time. Conversely, studies with shorter treatment durations (for example, <4 weeks) (39) often reported minimal cognitive improvements, potentially reflecting an insufficient timeframe for neurogenesis or synaptic remodeling to occur. Variations in the mode of administration also played a role, as intramuscular injections tended to produce more rapid and pronounced cognitive effects compared with transdermal formulations, which provide more stable but slower-acting testosterone levels. This variability may have contributed to the observed discrepancies in domains such as visuospatial ability, where ART appeared to have limited effects regardless of the regimen. Additionally, the absence of ART-combination treatments, such as concurrent cognitive training or physical exercise, may have further constrained the potential for synergistic cognitive gains. These findings underscore the importance of considering treatment parameters and individual response variability when interpreting the effects of ART on cognitive function. Standardization in ART protocols across future studies, particularly with respect to dose and duration, may help clarify the true domain-specific cognitive benefits of testosterone supplementation.

The cognitive benefits of ART observed in the present study align with findings from previous research. Previous studies, such as Cherrier *et al* (8), found significant improvements in verbal memory following testosterone supplementation in older hypogonadal men, which aligns with our meta-analysis results showing positive effects in memory-related tasks (7). Similarly, Janowsky *et al* (32) found that testosterone administration improved spatial cognition in older men, suggesting that spatial memory and navigation skills may be particularly responsive to testosterone treatment (32). In line with that, a later study by Janowsky *et al* (7) confirmed these findings, showing that testosterone improved performance on tasks requiring spatial and working memory, further indicating a consistent impact of ART on specific memory-related cognitive domains (7). Further evidence comes from Kenny *et al* (33), who conducted an RCT in which hypogonadal men received transdermal testosterone over 6 months. The study found significant improvements in attention and executive functioning, as assessed by the Trail-Making Test and Stroop interference test, both sensitive to attentional and cognitive control processes. These findings suggest that ART may benefit memory and positively impact executive functions, such as processing speed and cognitive flexibility. This effect is particularly relevant for daily functioning, as cognitive control is crucial for managing complex tasks (33). However, results across studies have not been uniformly positive. For example, Emmelot-Vonk *et al* (9) reported mixed findings, where testosterone supplementation improved specific areas including attention but did not significantly enhance overall cognitive performance. In their large-scale RCT involving 237 older men, cognitive benefits were observed primarily in attention, with minimal effects on global cognitive scores. The lack of improvement in global cognition might be attributed to the variability in cognitive assessments used or individual differences in baseline cognitive functioning. The authors suggested that the limited impact on broader cognitive domains could be due to insufficient dosage or duration of ART to elicit larger cognitive effects across all domains (9). More recently, Huang *et al* (35) examined the effects of long-term testosterone administration on cognition in older men with low to low-normal testosterone concentrations. Their study, a secondary analysis of the TEAAM trial, found some improvement in specific cognitive tasks, particularly in verbal memory and complex figure tests, which assess visuospatial and memory functions. However, similar to Emmelot-Vonk *et al* (9), Huang *et al* (35) found that the effects of testosterone were not pervasive across all cognitive domains, indicating that ART may selectively enhance certain cognitive functions while having little to no impact on others (35).

The findings of the present study are consistent with this pattern, where improvements in memory, especially verbal memory, were evident across studies. However, ART's effect on broader cognitive domains remains less consistent. These variations could result from differences in study design, such as testosterone administration method (injections, gels, patches), dose, therapy duration, and variations in cognitive testing protocols. Additionally, factors such as baseline testosterone levels and the age of participants likely influence the cognitive response to ART. For instance, studies involving younger hypogonadal men [for example, Lašaitė *et al* (5)] have shown that younger age groups may exhibit more pronounced cognitive benefits, possibly due to higher neuroplasticity in younger individuals.

In line with this, the current meta-analysis found a significant effect on executive function and global cognition, indicating that ART may enhance higher-order cognitive processes and overall cognitive performance. However, the effects on attention and visuospatial abilities were less pronounced. These findings underscore the complexity of ART's impact on cognition, with domain-specific effects that may depend on both biological and methodological factors. The observed improvements in executive function, for example, may reflect the role of testosterone in modulating brain regions such as the prefrontal cortex, which governs cognitive control and working memory. Conversely, the lack of significant improvements in attention and visuospatial tasks suggests that these domains may be less sensitive to androgen levels or require longer intervention periods to exhibit measurable changes. Further research should aim to address these discrepancies by standardizing study designs and exploring potential moderators, such as participant age, baseline cognitive status, and genetic predispositions, to improve understanding of the nuanced cognitive effects of ART.

The administration method also plays a crucial role in ART outcomes. Studies that used injectable testosterone, such as testosterone enanthate or testosterone cypionate, tended to show more immediate cognitive effects (7) compared with slower-releasing transdermal gels and patches (31). However,

the long-term benefits of transdermal testosterone suggest it may offer more stable cognitive gains over extended periods, particularly for memory-related tasks (9,31).

The current findings align with and extend those of previous meta-analyses by Zhang *et al* (10) and Hong *et al* (11) (the latter shares 75% of the eligible studies with our meta-analysis), but with notable methodological and interpretive differences. Zhang *et al* (10) conducted a CMA that included both observational studies and RCTs, highlighting mild cognitive improvements, particularly in verbal memory, but only in short-term follow-ups (up to 6 months). Their analysis was limited by small sample sizes in intervention studies and a lack of detailed exploration of dose-response relationships. By contrast, Hong *et al* (11) focused exclusively on elderly populations across 15 RCTs and found no significant improvements in cognitive speed, working memory, or attention-related tasks, attributing these null findings to study heterogeneity and potential publication bias. Unlike these studies, the current meta-analysis included longer follow-up periods (up to 36 months) and incorporated diverse ART administration methods, allowing us to capture potential long-term cognitive effects. Precision intervals were also employed instead of relying solely on I^2, providing a more nuanced assessment of heterogeneity and identifying modest but statistically significant improvements in executive function, which were absent in the findings of Hong *et al* (11). Additionally, our dose-response analysis revealed a significant positive association between testosterone dosage and cognitive improvement (P=0.01), suggesting that dosage optimization may play a critical role in ART efficacy. These findings underscore the importance of personalized ART regimens and suggest that future research should focus on optimizing treatment protocols and conducting long-term RCTs stratified by cognitive risk profiles.

In addition, the study by Ponce *et al* (41) conducted a systematic review of RCTs on testosterone replacement therapy (TRT) in hypogonadal men, focusing on validated patient-important outcomes (PIOs) including sexual function, mood and adverse events such as erythrocytosis, with strict inclusion criteria (TT ≤300 ng/dl, ≥12-week duration). Their findings indicated modest improvements in sexual desire (SMD: 0.17), erectile function (SMD: 0.16) and sexual satisfaction (SMD: 0.16), but no significant effect on mood or energy, alongside an increased risk of erythrocytosis (RR: 8.14). By contrast, the present study incorporated broader testosterone thresholds, longer follow-up durations and additional cognitive assessments, allowing for the capturing of potential long-term effects on cognition. Furthermore, dose-response and subgroup analyses were performed, strengthening insights into treatment variability, while Ponce *et al*'s (41) focus remained on general PIOs without stratification for cognitive or individual-level factors, making the current findings more comprehensive for personalized TRT strategies.

An important consideration in our meta-analysis is the potential age-related bias in study selection. The majority of the included studies focused on older men with age-related hypogonadism (typically aged ≥50 years). This focus is clinically relevant, as older adults represent the primary population for ART due to the natural decline in testosterone levels with age and the increased prevalence of cognitive complaints (42). However, this also limits the generalizability of the present findings to younger hypogonadal men, who may exhibit different cognitive responses to ART. Age-related biological factors, such as reduced AR density and decreased neuroplasticity, may attenuate the cognitive benefits of ART in older individuals (43). Moreover, the presence of age-related comorbidities, such as cardiovascular disease and diabetes, may confound cognitive outcomes, as these conditions can independently affect cognitive function (44). The observed modest effect sizes in the present analysis may reflect these age-related influences. By contrast, younger individuals with non-age-related hypogonadism may experience more pronounced cognitive benefits due to greater baseline neuroplasticity, fewer comorbidities, and potentially higher AR sensitivity (45).

A review of the 14 included studies revealed that bioavailable testosterone levels were commonly reported alongside total testosterone, with several studies, such as those by Kenny *et al* (33,34), showing significant increases in bioavailable testosterone in treatment groups following ART. By contrast, no notable changes were observed in control groups. Additionally, SHBG levels, which influence the proportion of free and bioavailable testosterone, were consistently monitored in studies such as Lašaitė *et al* (5) and Cherrier *et al* (8). While SHBG levels showed variability across participants, they generally remained stable during ART. However, none of the included studies directly measured AR levels, though some studies, such as Lašaitė *et al* (5), discussed the potential role of AR polymorphisms and receptor sensitivity in influencing cognitive outcomes. To address this limitation, it is important to note that AR-related factors were not quantitatively assessed but were acknowledged as potential moderators in cognitive responses to ART. Including bioavailable testosterone as a key outcome alongside total testosterone aligns with common endocrinological protocols, as it provides a more precise indicator of the active hormonal fraction. Therefore, while hormonal measures such as bioavailable testosterone and SHBG were routinely considered, the absence of direct AR-level data highlights an area for further research.

The mechanisms underlying the cognitive effects of ART are indeed complex and appear to involve multiple pathways, both direct and indirect. One major pathway is the action of testosterone through ARs, which are widely distributed in brain regions that are crucial for cognitive functions, particularly memory and learning. Key areas such as the hippocampus, amygdala and prefrontal cortex have high densities of ARs, which suggests a direct link between testosterone levels and cognitive performance in these regions (13). Activation of these receptors is considered to facilitate neuroprotective effects, including the enhancement of synaptic connectivity, modulation of neurotransmitter levels and protection against neurodegenerative processes (14). Additionally, testosterone undergoes aromatization in the brain to form estradiol, a process that plays a significant role in synaptic plasticity, neurogenesis and neuronal survival. Estradiol, the primary estrogen in the brain, has well-documented effects on synaptic density and cognitive function. It enhances dendritic spine growth in the hippocampus and prefrontal cortex, which are structures critical for memory formation and executive function (46). By acting as a precursor to estradiol, testosterone contributes indirectly to these cognitive benefits,

particularly in brain areas involved in verbal memory, spatial navigation and attention (14,24). The neuroprotective effects of testosterone may also be attributed to its influence on neurotransmitter systems. For example, testosterone has been shown to enhance dopamine release, a neurotransmitter essential for motivation, reward processing and executive functioning (47). Dopaminergic pathways are particularly dense in the prefrontal cortex, a brain area involved in planning, problem-solving and impulse control. Furthermore, testosterone may positively affect serotonin and GABAergic systems, contributing to mood regulation and stress resilience, which indirectly supports cognitive performance by reducing anxiety and depressive symptoms commonly associated with cognitive decline (48).

Despite these plausible mechanisms, the effect sizes observed in the current analysis were modest. Although testosterone appears to have neuroprotective and cognitive-enhancing effects, these benefits might be limited by several factors. First, individual variability in AR density, which declines with age, may reduce the brain's responsiveness to ART, particularly in older men (32). Second, cognitive decline is a multifactorial process that involves numerous interacting elements beyond testosterone levels, such as genetic predisposition, cardiovascular health, lifestyle factors and other hormonal changes. Therefore, while ART may provide some cognitive benefits, it may not fully counteract or reverse cognitive deficits that result from these broader, complex mechanisms (15). Finally, while testosterone's conversion to estradiol is beneficial for synaptic health, excessive estradiol in men could potentially lead to adverse effects, including mood instability and even cognitive impairments if levels exceed the optimal range (25). This suggests a need for personalized ART dosing to balance the benefits of testosterone, particularly given the variability in individual responses to testosterone supplementation.

One advantage of the present study is the rigorous methodology, adhering to PRISMA guidelines, with thorough data extraction and evaluation of each study's methodological quality using the CONSORT and ROBINS-I tool checklist. The current analysis also included a dose-response assessment, revealing a slight positive correlation between testosterone dose and cognitive effect size (slope P=0.01). This suggests a dose-dependent relationship, though the absence of a significant duration-response effect indicates that the duration of ART may not be as influential as dosage in determining cognitive outcomes. Our findings underscore the importance of dosage and duration in determining ART efficacy. However, the variability in dosages (ranging from 5 mg/day to 250 mg/week) and therapy durations (5 days to 36 months) complicates direct comparisons between studies. This highlights the need for standardizing ART protocols and further clarifying the relationship between these parameters and cognitive outcomes. Future research should investigate the effects of varying regimens across different populations to identify optimal ART combinations and dosing schedules for specific cognitive domains. Nevertheless, there are limitations to the present study. First, substantial heterogeneity was present across studies, with variations in ART protocols, cognitive tests and participant demographics, which may confound results. Although a random-effects model and precision intervals were used to mitigate this heterogeneity, these factors remain potential sources of bias. The age-related focus of the included studies may limit the generalizability of findings to younger men with hypogonadism. Additionally, the presence of comorbid conditions, such as cardiovascular disease and diabetes, and comorbidities associated with decreased androgen levels (schizophrenia, multiple sclerosis and acute lymphoblastic leukemia) may have influenced cognitive outcomes independently, which could have attenuated the observed effect sizes (49). Future studies should stratify participants by age and comorbidity status to account for these confounding factors. Furthermore, the presence of publication bias, as indicated by the funnel plot and Begg's and Egger's tests, suggests an overrepresentation of positive findings. The trim-and-fill method adjusted for this bias, but the overall effect size remained relatively unchanged, underscoring that ART's cognitive effects, though consistent, may be modest. Additionally, the generalizability of the current findings may be limited by the characteristics of the included studies, which primarily involved older men with age-related hypogonadism. Cognitive effects may differ in younger populations or those with non-age-related hypogonadism, necessitating caution in extrapolating our results beyond the studied demographic.

The findings of the present study have potential implications for the clinical management of cognitive deficits in hypogonadal patients. Given the observed cognitive benefits, ART could be considered as part of a broader therapeutic approach for older men with hypogonadism and cognitive complaints. Clinicians should temper expectations, as the modest effect sizes indicate that ART alone is unlikely to yield substantial cognitive gains for all patients. The decision to initiate ART should balance potential cognitive benefits against associated risks, particularly cardiovascular health and prostate complications. ART may be most beneficial when combined with lifestyle interventions such as exercise and cognitive training, which have independently demonstrated positive effects on cognitive function. Non-pharmacological interventions, such as cognitive training, aerobic exercise and dietary supplementation, have also demonstrated efficacy in improving cognitive outcomes in hypogonadal men (50). For instance, aerobic exercise has shown significant improvements in executive function and memory, suggesting that combining ART with lifestyle modifications may yield more robust cognitive benefits (51). Currently, there is a lack of studies directly comparing ART with non-pharmacological interventions, such as cognitive training, aerobic exercise, or dietary supplementation, in terms of cognitive benefits for hypogonadal men. While ART has been shown to improve certain cognitive functions, such as verbal and spatial memory, in hypogonadal men (52), and non-pharmacological interventions such as aerobic exercise have demonstrated improvements in executive function and memory (53), head-to-head comparative studies are needed to determine the relative efficacy of these treatments and to assess whether combining ART with lifestyle modifications offers synergistic cognitive benefits.

Additionally, tailoring ART regimens to individual cognitive risk profiles, baseline testosterone levels and comorbidity status may optimize therapeutic outcomes while minimizing adverse effects. In clinical practice, ART should be integrated into a comprehensive cognitive health strategy that includes lifestyle interventions such as exercise, nutrition and cognitive training.

Future research should focus on understanding how ART interacts with these factors to optimize cognitive resilience.

Several factors may contribute to the absence of new RCTs or large-scale studies meeting our inclusion criteria after 2023. First, conducting RCTs in this domain is resource-intensive, requiring significant time and funding, particularly given the need for long-term follow-up to capture meaningful clinical outcomes and adverse events. Second, ethical considerations surrounding placebo use in symptomatic hypogonadal men could limit the feasibility of new trials. Additionally, recent systematic reviews and meta-analyses may have already synthesized existing evidence comprehensively, potentially disincentivizing the initiation of new large-scale studies. Finally, there may be an ongoing shift towards real-world evidence and large observational studies, which some researchers may prioritize over RCTs to capture broader patient populations and real-world treatment effects.

In conclusion, the present systematic review and meta-analysis provide evidence that ART has a modest but statistically significant positive effect on cognitive function in hypogonadal men, particularly in memory and executive function domains. Domain-specific findings indicate that ART shows the strongest effects in verbal memory and working memory, with smaller but notable improvements in executive functions such as cognitive flexibility and processing speed. However, the effects on attention and visuospatial abilities were less consistent, indicating that some cognitive domains may be more responsive to ART than others. These findings suggest that ART may be most beneficial for individuals with mild to moderate cognitive deficits and low baseline testosterone levels, as observed in studies involving younger cohorts, which reported more pronounced cognitive improvements. Furthermore, differences in ART administration methods, such as transdermal testosterone gels vs. intramuscular injections, may influence outcomes, with transdermal forms potentially offering imroved adherence and fewer adverse events. Despite the observed benefits, the modest effect sizes suggest that ART alone may not suffice as a comprehensive cognitive intervention for hypogonadal men with significant cognitive impairments. The variability in cognitive outcomes across studies underscores the need to consider individual factors, including age, baseline cognitive status, comorbid conditions and genetic predispositions, when evaluating ART's potential benefits. Clinicians should carefully balance the potential cognitive benefits of ART with its associated risks, particularly in patients with cardiovascular risk factors or a history of prostate disease. Future research should prioritize long-term RCTs involving both younger and older hypogonadal populations to improve understanding of the differential cognitive effects across age groups. Additionally, there is a need to examine the interaction of ART with comorbid conditions and lifestyle factors to better delineate its cognitive impact. Standardizing cognitive assessments and ART dosing protocols across studies is essential to reduce methodological variability and improve comparability. Mechanistic studies investigating the role of ARs, neuroplasticity, and the influence of estradiol conversion on synaptic connectivity will also provide important insights into the pathways through which ART influences cognitive function. These targeted research efforts will support the development of more personalized and effective ART regimens to address cognitive deficits in hypogonadal men.

Acknowledgements

Not applicable.

Funding

The present study was supported by the Tianjin Health Technology Project (grant no. ZC20185).

Availability of data and materials

The data generated in the present study may be requested from the corresponding author.

Authors' contributions

BW and XL developed the study protocol and were responsible for data collection. WC and LL analyzed the data. BW and XL confirm the authenticity of all the raw data. All authors contributed to the writing of the manuscript, and read and approved the final version of the manuscript.

Ethics approval and consent to participate

Not applicable.

Patient consent for publication

Not applicable.

Competing interests

The authors declare that they have no competing interests.

References

1. Shores MM, Sloan KL, Matsumoto AM, Moceri VM, Felker B and Kivlahan DR: Increased incidence of diagnosed depressive illness in hypogonadal older men. Arch Gen Psychiatry 61: 162-167, 2004.
2. Barrett-Connor E, Goodman-Gruen D and Patay B: Endogenous sex hormones and cognitive function in older men. J Clin Endocrinol Metab 84: 3681-3685, 1999.
3. Jayaraman A, Lent-Schochet D and Pike CJ: Diet-induced obesity and low testosterone increase neuroinflammation and impair neural function. J Neuroinflammation 11: 162, 2014.
4. Munari EV, Amer M, Amodeo A, Bollino R, Federici S, Goggi G, Giovanelli L, Persani L, Cangiano B and Bonomi M: The complications of male hypogonadism: Is it just a matter of low testosterone? Front Endocrinol (Lausanne) 14: 1201313, 2023.
5. Lašaitė L, Čeponis J, Preikša RT and Žilaitienė B: Effects of two-year testosterone replacement therapy on cognition, emotions and quality of life in young and middle-aged hypogonadal men. Andrologia 49, 2017.
6. Nieschlag E, Behre HM, Kliesch S and Nieschlag S (eds): Andrology: Male Reproductive Health and Dysfunction. 4th edition. Springer Nature, 2023. https://doi.org/10.1007/978-3-031-31574-9. Accessed October 27, 2023.
7. Janowsky JS, Chavez B and Orwoll E: Sex steroids modify working memory. J Cogn Neurosci 12: 407-414, 2000.
8. Cherrier MM, Asthana S, Plymate S, Baker L, Matsumoto AM, Peskind E, Raskind MA, Brodkin K, Bremner W, Petrova A, *et al*: Testosterone supplementation improves spatial and verbal memory in healthy older men. Neurology 57: 80-88, 2001.

9. Emmelot-Vonk MH, Verhaar HJ, Nakhai Pour HR, Aleman A, Lock TM, Bosch JL, Grobbee DE and van der Schouw YT: Effect of testosterone supplementation on functional mobility, cognition, and other parameters in older men: A randomized controlled trial. JAMA 299: 39-52, 2008.
10. Zhang Z, Kang D and Li H: Testosterone and cognitive impairment or dementia in middle-aged or aging males: Causation and intervention, a systematic review and meta-analysis. J Geriatr Psychiatry Neurol 34: 405-417, 2021.
11. Hong SW, Cho YJ, Lee JH, Suh YS and Kim DH: Effect of testosterone supplementation on cognition in elderly men: A systematic meta-analysis. Korean J Geriatr Gerontol 24: 24-33, 2023.
12. Farajdokht F, Farhoudi M, Majdi A, Zamanlu M, Sadigh-Eteghad S, Vahedi S and Mahmoudi J: Testosterone may hold therapeutic promise for the treatment of ischemic stroke in aging: A closer look at laboratory findings. Adv Pharm Bull 9: 48-55, 2019.
13. Roselli CE, Abdelgadir SE, Rønnekleiv OK and Klosterman SA: Anatomic distribution and regulation of aromatase gene expression in the rat brain. Biol Reprod 58: 79-87, 1998.
14. Luine VN: Sex steroids and cognitive function. J Neuroendocrinol 20: 866-872, 2008.
15. Holland J, Bandelow S and Hogervorst E: Testosterone levels and cognition in elderly men: A review. Maturitas 69: 322-337, 2011.
16. Huo S, Scialli AR, McGarvey S, Hill E, Tügertimur B, Hogenmiller A, Hirsch AI and Fugh-Berman A: Treatment of men for 'low testosterone': A systematic review. PLoS One 11: e0162480, 2016.
17. Zitzmann M, Faber S and Nieschlag E: Association of specific symptoms and metabolic risks with serum testosterone in older men. J Clin Endocrinol Metab 91: 4335-4343, 2006.
18. Page MJ, McKenzie JE, Bossuyt PM, Boutron I, Hoffmann TC, Mulrow CD, Shamseer L, Tetzlaff JM, Akl EA, Brennan SE, *et al*: The PRISMA 2020 statement: An updated guideline for reporting systematic reviews. BMJ 372: n71, 2021.
19. Thorlund K, Imberger G, Johnston BC, Walsh M, Awad T, Thabane L, Gluud C, Devereaux PJ and Wetterslev J: Evolution of heterogeneity (I2) estimates and their 95% confidence intervals in large meta-analyses. PLoS One 7: e39471, 2012.
20. IntHout J, Ioannidis JP, Rovers MM and Goeman JJ: Plea for routinely presenting prediction intervals in meta-analysis. BMJ Open 6: e010247, 2016.
21. Asih PR, Wahjoepramono EJ, Aniwiyanti V, Wijaya LK, de Ruyck K, Taddei K, Fuller SJ, Sohrabi H, Dhaliwal SS, Verdile G, *et al*: Testosterone replacement therapy in older male subjective memory complainers: Double-blind randomized crossover placebo-controlled clinical trial of physiological assessment and safety. CNS Neurol Disord Drug Targets 14: 576-586, 2015.
22. Cherrier MM, Plymate S, Mohan S, Asthana S, Matsumoto AM, Bremner W, Peskind E, Raskind M, Latendresse S, Haley AP and Craft S: Relationship between testosterone supplementation and insulin-like growth factor-I levels and cognition in healthy older men. Psychoneuroendocrinology 29: 65-82, 2004.
23. Cherrier MM, Matsumoto AM, Amory JK, Johnson M, Craft S, Peskind ER and Raskind MA: Characterization of verbal and spatial memory changes from moderate to supraphysiological increases in serum testosterone in healthy older men. Psychoneuroendocrinology 32: 72-79, 2007.
24. Cherrier MM, Anderson K, Shofer J, Millard S and Matsumoto AM: Testosterone treatment of men with mild cognitive impairment and low testosterone levels. Am J Alzheimers Dis Other Demen 30: 421-430, 2015.
25. Gray PB, Singh AB, Woodhouse LJ, Storer TW, Casaburi R, Dzekov J, Dzekov C, Sinha-Hikim I and Bhasin S: Dose-dependent effects of testosterone on sexual function, mood, and visuospatial cognition in older men. J Clin Endocrinol Metab 90: 3838-3846, 2005.
26. Gregori G, Celli A, Barnouin Y, Paudyal A, Armamento-Villareal R, Napoli N, Qualls C and Villareal DT: Cognitive response to testosterone replacement added to intensive lifestyle intervention in older men with obesity and hypogonadism: Prespecified secondary analyses of a randomized clinical trial. Am J Clin Nutr 114: 1590-1599, 2021.
27. Haren MT, Wittert GA, Chapman IM, Coates P and Morley JE: Effect of oral testosterone undecanoate on visuospatial cognition, mood and quality of life in elderly men with low-normal gonadal status. Maturitas 50: 124-133, 2005.
28. Konaka H, Sugimoto K, Orikasa H, Iwamoto T, Takamura T, Takeda Y, Shigehara K, Iijima M, Koh E and Namiki M; 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 18: 25-34, 2016.
29. Ly LP, Jimenez M, Zhuang TN, Celermajer DS, Conway AJ and Handelsman DJ: A double-blind, placebo-controlled, randomized clinical trial of transdermal dihydrotestosterone gel on muscular strength, mobility, and quality of life in older men with partial androgen deficiency. J Clin Endocrinol Metab 86: 4078-4088, 2001.
30. Yamaguchi K, Ishikawa T, Chiba K and Fujisawa M: Assessment of possible effects for testosterone replacement therapy in men with symptomatic late-onset hypogonadism. Andrologia 43: 52-56, 2011.
31. Cherrier MM, Matsumoto AM, Amory JK, Ahmed S, Bremner W, Peskind ER, Raskind MA, Johnson M and Craft S: The role of aromatization in testosterone supplementation: Effects on cognition in older men. Neurology 64: 290-296, 2005.
32. Janowsky JS, Oviatt SK and Orwoll ES: Testosterone influences spatial cognition in older men. Behav Neurosci 108: 325-332, 1994.
33. Kenny AM, Bellantonio S, Gruman CA, Acosta RD and Prestwood KM: 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 57: M321-M325, 2002.
34. Kenny AM, Fabregas G, Song C, Biskup B and Bellantonio S: Effects of testosterone on behavior, depression, and cognitive function in older men with mild cognitive loss. J Gerontol A Biol Sci Med Sci 59: 75-78, 2004.
35. Huang G, Wharton W, Bhasin S, Harman SM, Pencina KM, Tsitouras P, Li Z, Hally KA, Asthana S, Storer TW and Basaria S: 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 4: 657-665, 2016.
36. Maki PM, Ernst M, London ED, Mordecai KL, Perschler P, Durso SC, Brandt J, Dobs A and Resnick SM: Intramuscular testosterone treatment in elderly men: Evidence of memory decline and altered brain function. J Clin Endocrinol Metab 92: 4107-4114, 2007.
37. Vaughan C, Goldstein FC and Tenover JL: Exogenous testosterone alone or with finasteride does not improve measurements of cognition in healthy older men with low serum testosterone. J Androl 28: 875-882, 2007.
38. Wahjoepramono EJ, Asih PR, Aniwiyanti V, Taddei K, Dhaliwal SS, Fuller SJ, Foster J, Carruthers M, Verdile G, Sohrabi HR and Martins RN: The effects of testosterone supplementation on cognitive functioning in older men. CNS Neurol Disord Drug Targets 15: 337-343, 2016.
39. Wolf OT, Preut R, Hellhammer DH, Kudielka BM, Schürmeyer TH and Kirschbaum C: Testosterone and cognition in elderly men: A single testosterone injection blocks the practice effect in verbal fluency, but has no effect on spatial or verbal memory. Biol Psychiatry 47: 650-654, 2000.
40. Resnick SM, Matsumoto AM, Stephens-Shields AJ, Ellenberg SS, Gill TM, Shumaker SA, Pleasants DD, Barrett-Connor E, Bhasin S, Cauley JA, *et al*: Testosterone treatment and cognitive function in older men with low testosterone and age-associated memory impairment. JAMA 317: 717-727, 2017.
41. Ponce OJ, Spencer-Bonilla G, Alvarez-Villalobos N, Serrano V, Singh-Ospina N, Rodriguez-Gutierrez R, Salcido-Montenegro A, Benkhadra R, Prokop LJ, Bhasin S and Brito JP: 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: Mar 17, 2018, (Epub ahead of print).
42. Cai Z and Li H: An updated review: Androgens and cognitive impairment in older men. Front Endocrinol (Lausanne) 11: 586909, 2020.
43. Kuwahara N, Nicholson K, Isaacs L and MacLusky NJ: Androgen effects on neural plasticity. Androg Clin Res Ther 2: 216-230, 2021.
44. Bidzan L, Jurek P, Olech M, Bidzan-Wiącek M, Bidzan-Bluma I and Bidzan M: Somatic comorbidity and the progression of cognitive impairment. Front Aging Neurosci 15: 1219449, 2023.

45. Hua JT, Hildreth KL and Pelak VS: Effects of testosterone therapy on cognitive function in aging: A systematic review. Cogn Behav Neurol 29: 122-138, 2016.
46. McEwen BS and Woolley CS: Estradiol and progesterone regulate neuronal structure and synaptic connectivity in adult as well as developing brain. Exp Gerontol 29: 431-436, 1994.
47. Amin Z, Canli T and Epperson CN: Effect of estrogen-serotonin interactions on mood and cognition. Behav Cogn Neurosci Rev 4: 43-58, 2005.
48. Frye CA: Effects and mechanisms of progestogens and androgens in ictal activity. Epilepsia 51 (Suppl 3): S135-S140, 2010.
49. Yeo S, Holl K, Peñaherrera N, Wissinger U, Anstee K and Wyn R: Burden of male hypogonadism and major comorbidities, and the clinical, economic, and humanistic benefits of testosterone therapy: A narrative review. Clinicoecon Outcomes Res 13: 31-38, 2021.
50. Hoffmann CM, Petrov ME and Lee RE: Aerobic physical activity to improve memory and executive function in sedentary adults without cognitive impairment: A systematic review and meta-analysis. Prev Med Rep 23: 101496, 2021.
51. Kleinloog JPD, Mensink RP, Ivanov D, Adam JJ, Uludağ K and Joris PJ: Aerobic exercise training improves cerebral blood flow and executive function: A randomized, controlled cross-over trial in sedentary older men. Front Aging Neurosci 11: 333, 2019.
52. Lisco G, Giagulli VA, De Tullio A, De Pergola G, Guastamacchia E and Triggiani V: Age-related male hypogonadism and cognitive impairment in the elderly: Focus on the effects of testosterone replacement therapy on cognition. Geriatrics (Basel) 5: 76, 2020.
53. Čeponis J, Swerdloff RS and Wang C: Androgen replacement therapy in hypogonadal men. In: Male Hypogonadism: Basic, Clinical and Therapeutic Principles. Winters S and Huhtaniemi I (eds). Humana Press, Cham, pp367-397, 2017. https://link.springer.com/book/10.1007/978-3-319-53298-1. Accessed April 26, 2017