Review Article

pISSN: 2287-4208 / eISSN: 2287-4690
World J Mens Health 2018 May 36(2): 110-122
https://doi.org/10.5534/wjmh.182001

The World Journal of Men's Health

Testosterone Replacement, Muscle Strength, and Physical Function

You-Seon Nam[1], Gyeongsil Lee[2,*], Jae Moon Yun[2], Belong Cho[1,2,3,4]

[1]Department of Family Medicine, Seoul National University College of Medicine, [2]Department of Family Medicine, Seoul National University Hospital, Seoul, [3]Advanced Institutes of Convergence Technology, Seoul National University, Suwon, [4]Institute on Aging, Seoul National University College of Medicine, Seoul, Korea

Muscle strength and physical function decrease in older men, as do testosterone levels. Nonetheless, the effects of testosterone replacement therapy on muscle strength and physical function remain inconclusive and equivocal. We conducted a rapid systematic review, the results of which showed that testosterone replacement does not affect muscle strength (measured by hand grip strength and leg muscle strength), although it may increase physical function (measured by the 6-minute walk test, Physical Activity Scale for the Elderly score, and other physical performance tests). However, most of the studies were conducted in the United States or Europe and did not include participants from Asian or other ethnic backgrounds; therefore, further studies are needed to evaluate the effects of testosterone replacement in a broader population.

Keywords: Aged; Muscle strength; Physical activity; Testosterone

INTRODUCTION

Decreased physical function is a representative sign of frailty in elderly people [1], which is usually accompanied by a decrease in testosterone levels in elderly men [2]. Lower testosterone levels are associated with decreased physical function and increased mortality [3], and frailty is associated with decreased quality of life [4]. However, there is no effective clinical treatment to restore physical function in the elderly. One of the mechanisms for the decreased physical function in the elderly is sarcopenia [5], and especially in elderly men, muscle mass is associated with testosterone levels [6].

Although many randomized controlled trials (RCTs) of testosterone replacement therapy (TRT) have been performed, these studies generally had small sample sizes and a variety of study designs, and the effects of TRT on improving physical function are equivocal [7-9]. While in one study [7] testosterone supplementation improved strength and was suggested to have a role in the treatment of frailty in hypogonadic males, another study [8] did not observe any increase in muscle strength, but only an increase in muscle mass.

Some systematic reviews (SRs) [10-13] investigating the effects of TRT on the body have been published. However, one study [10] included middle-aged subjects and did not include recent studies. Although another study [13] performed a comprehensive SR including

Received: Nov 29, 2017 Revised: Dec 19, 2017 Accepted: Dec 26, 2017 Published online Mar 22, 2018
Correspondence to: Belong Cho https://orcid.org/0000-0001-9558-689X
Department of Family Medicine, Seoul National University Hospital, 101 Daehak-ro, Jongno-gu, Seoul 03080, Korea.
Tel: +82-2-2072-2195, Fax: +82-2-766-3276, E-mail: belong@snu.ac.kr
*Current affiliation: Health Promotion Center, Chung-Ang University Hospital, Seoul, Korea

elderly men, the heterogeneity in the target groups and outcomes was large, and no significant conclusions were suggested. Accordingly, a quantitative analysis including recently published high-quality RCTs assessing physical function [2,14] could yield additional clinical significance. Therefore, we conducted a rapid SRs, including elderly men receiving TRT, to assess the effects of TRT on muscle strength and physical function. We also performed quantitative analyses of the outcomes of some RCTs in order to determine their clinical significance. The search strategies and study selection criteria are explained at the end.

TESTOSTERONE REPLACEMENT THERAPY AND MUSCLE STRENGTH

Muscle strength is commonly measured with hand grip strength and the 1-repetition maximum (1-RM) for exercises performed as part of strength testing. Sixteen studies examined the effects of testosterone supplementation on muscle strength. Fourteen studies assessed muscle strength as a primary outcome and 2 as a secondary outcome. The other primary and secondary outcomes investigated in each of the studies are listed in Table 1 [2,7-9,14-27]. The most common measurements for muscle strength were hand grip strength and leg muscle strength (knee extension and flexion). Other measurements included single or double leg press, chest press, and triceps extension.

1. Hand grip

Ten studies included in the analysis measured muscle strength in terms of hand grip strength. Five of these studies were conducted in the United States, 2 in the United Kingdom, 1 in the Netherlands, 1 in Australia, and 1 in New Zealand. A total of 1002 participants were included in the analysis. The intervention period varied from 8 weeks to 1 year. All hand grip strength measurements were expressed in kilograms, except for one study [15], for which a pound-to-kilogram conversion was done for the purposes of our analysis. The overall results (Fig. 1) showed that TRT does not increase hand grip strength. However, individual studies, such as a randomized, placebo-controlled, double-blind study [15], have found that hand grip strength improved after testosterone administration. Other studies [7,26] that showed improved hand grip strength after testosterone supplementation could not be included in the analysis because they only presented data in figures.

2. Leg muscle strength

Seven studies included in the analysis measured leg muscle strength. All 7 studies included information for the knee extensor muscles, consisting of a total of 780 participants. Three studies were conducted in the United States, 2 in the United Kingdom, 1 in New Zealand, and 1 in the Netherlands. The study conducted in the Netherlands [20] did not include measurements for the knee flexor muscles, so only 6 of the 7 studies included for the knee extensor analysis were included in the analysis of knee flexor strength. This resulted in a total of 543 participants for the knee flexor analysis. The intervention period varied from 12 weeks to 1 year. Lower muscle strength measurements were expressed as 1-RM in kilograms, newton-meters per second, or watts. Although marginal effects were shown for increased knee extensor strength (Fig. 2), when combined with the outcomes for knee flexor strength, the overall results showed no significant difference.

TESTOSTERONE REPLACEMENT THERAPY AND PHYSICAL FUNCTION

Physical function was commonly measured by the 6-minute walk test [28], the Physical Activity Scale for the Elderly (PASE) questionnaire [29], the physical-function domain (PF-10) of the Medical Outcomes Study 36-Item Short-Form Health Survey [30], a physical performance test (PPT) [31], and the Short Physical Performance Battery (ability to rise from a chair, static balance, and 8-foot walk) [32]. Other tests performed in some studies that were not included in our analyses included the supine-to-stand test [33], and the Get-Up-and-Go test [34].

1. 6-minute walk test

Three studies included in the analysis showed results of the 6-minute walk test. Two of these studies were conducted in the United States, and 1 in New Zealand. A total of 733 participants were included in the analysis. The intervention period varied from 6 months to 1 year. All TRT interventions were done using transdermal testosterone gel, although the dosages differed between studies. The overall results (Fig. 3) showed that TRT improved the 6-minute walking distance by 9.35

https://doi.org/10.5534/wjmh.182001

Table 1. Effects of testosterone on muscle strength and physical function

Study (year)	Country	Type of intervention	Intervention period	Primary outcome	Secondary outcome
Bakhshi et al [15] (2000)	US	Testosterone enanthate (100 mg weekly, intramuscular injections)	8 wk	Functional independence measure, hand grip strength (dominant)	Geriatric Depression Scale - Short Form
Borst et al [16] (2014)	US	Testosterone and placebo	12 mo	Prostate digital rectal examination and transrectal ultrasonography, 1-repetition maximum strength testing (leg press, knee flexion, knee extension, chest press, and triceps extension), grip strength, urinary symptoms	N/A
Brill et al [17] (2002)	US	Transdermal placebo or T patches (Androderm, two 2.5-mg patches applied at bedtime) and evening subcutaneous injections of saline or rhGH (Genotropin, 6.25 g/kg per day)	1–3 mo wash-out	Body composition, performance, mood, sexual function, bone turnover, and muscle gene expression	
Clague et al [18] (1999)	UK	Testosterone enanthate (200 mg intramuscular injections, at 2-wk intervals)	12 wk	Muscle function	
Dias et al [19] (2016)	US	5 g of transdermal testosterone gel	12 mo	BMD	Body composition; muscle strength; gait speed; sex hormone levels
Emmelot-Vonk et al [20] (2008)	Netherlands	80 mg of oral testosterone undecanoate twice daily	6 mo	Functional mobility (Stanford Health Assessment Questionnaire, timed get-up-and-go test, isometric handgrip strength, isometric leg extensor strength), cognitive function (8 different cognitive instruments), BMD of the hip and lumbar spine (dual-energy x-ray absorptiometry scanning), body composition (total body dual-energy x-ray absorptiometry and abdominal ultrasound of fat mass), metabolic risk factors (fasting plasma lipids, glucose, and insulin), quality of life (Short-Form Health 36 Survey and the Questions on Life Satisfaction Modules), and safety parameters (serum prostate-specific antigen level, ultrasonographic prostate volume, International Prostate Symptom score, serum levels of creatinine, aspartate aminotransferase, alanine aminotransferase, gamma-glutamyltransferase, hemoglobin, and hematocrit)	N/A

Table 1. Continued

Study (year)	Country	Type of intervention	Intervention period	Primary outcome	Secondary outcome
Giannoulis et al [21] (2006)	UK	Fixed dose of 5 mg of testosterone through transdermal patches (testosterone and placebo rhGH)	6 mo	Lean body mass, total body fat, mid-thigh muscle cross-section area, muscle strength, aerobic capacity, condition-specific quality of life (Age-Related Hormone Deficiency-Dependent Quality of Life questionnaire), and generic health status (36-Item Short-Form Health Survey)	
Hildreth et al [22] (2013)	US	Transdermal testosterone gel (2 doses targeting either a lower [400–550 ng/dL] or higher [600–1,000 ng/dL] testosterone range)	12 mo	Functional performance	Strength and body composition
Kenny et al [23] (2001)	US	Transdermal testosterone supplementation	12 mo	Sex hormone levels, calcium-regulating hormone levels, BMD, bone markers, frailty evaluation, strength measures, multiple physical performance measures, and body composition were assessed; a prostate examination was performed; and hemoglobin and cholesterol levels were checked	
Kenny et al [24] (2010)	US	5 mg/d of testosterone gel	12 mo	BMD of hip, lumbar spine, and mid-radius; body composition; sex hormones, calcium-regulating hormones; bone turnover markers; strength; physical performance; and safety parameters	N/A
Ly et al [9] (2001)	Australia	Application of 70 mg of dihydrotestosterone with a derma gel daily	3 mo	Muscular strength	Muscular function (gait, balance, mobility), body composition, reproductive hormones, hematopoiesis, prostate size and prostate-specific antigen, and vascular reactivity
Nair et al [25] (2006)	US	Transdermal testosterone patch (5 mg per day; D-TRANS, Alza)	24 mo	Physical performance, peak aerobic capacity, body composition, BMD, and levels of plasma insulin and glucose after an overnight fast	Body weight, the proportion of body fat, the insulin-sensitivity index, quality of life, levels of various hormones, and levels of alkaline phosphatase, alanine aminotransferase, aspartate transferase, and hemoglobin, and adverse events
Page et al [26] (2005)	US	Testosterone enanthate, 200 mg intramuscularly every 2 wk, with placebo pills daily (testosterone-only)	36 mo	Timed physical performance, grip strength, lower extremity strength, body composition (by dual-energy x-ray absorptiometry), fasting cholesterol profiles, and hormones	

https://doi.org/10.5534/wjmh.182001

Table 1. Continued

Study (year)	Country	Type of intervention	Intervention period	Primary outcome	Secondary outcome
Sih et al [7] (1997)	US	200 mg (1 mL) of testosterone cypionate intramuscularly every 14–17 d (n=17)	12 mo	Strength (hand grip), hemoglobin, prostate-specific antigen, leptin, cognition	Alanine aminotransferase, gamma-glutamyltransferase, lactate dehydrogenase, blood pressure, cholesterol, 1,25 vitamin D, body mass index, body fat
Snyder et al [2] (2016)	US	5 g of a transdermal testosterone gel	12 mo	The percentage of men who increased the distance walked in the 6-min walk test by at least 50 m	The percentage of men whose score on the physical-function domain (PF-10) of the Medical Outcomes Study 36-Item Short-Form Health Survey (SF-36) increased by at least 8 points and changes from baseline in the 6-min walking distance and PF-10 score.
Srinivas-Shankar et al [27] (2010)	New Zealand	Transdermal testosterone (50 mg/d)	6 mo	Isometric knee extension peak torque and isokinetic knee extension peak torque	Isometric knee flexion peak torque, isokinetic knee flexion peak torque, physical function tests, body composition, and quality of life. All outcome assessments were carried out by a single assessor at baseline and at 6 months (end of treatment).
Storer et al [14] (2017)	US	7.5 g of 1% testosterone	36 mo	Muscle strength, power, and fatigability, muscle strength (1-RM), muscle power (leg-press, chest-press), muscle fatigability, stair-climb power, lean body mass, hormone assays	N/A
Wittert et al [8] (2003)	Australia	Standard dose (80 mg twice daily) of testosterone undecanoate (Andriol; Organon, Oss, The Netherlands)	1 y	Body composition, muscle strength, hormones, and safety parameters (hematocrit, prostate-specific antigen, urine flow, systolic and diastolic blood pressure, high-density lipoprotein, low-density lipoprotein, total cholesterol, triglycerides)	

rhGH: recombinant human growth hormone, BMD: bone mineral density, N/A: not applicable.

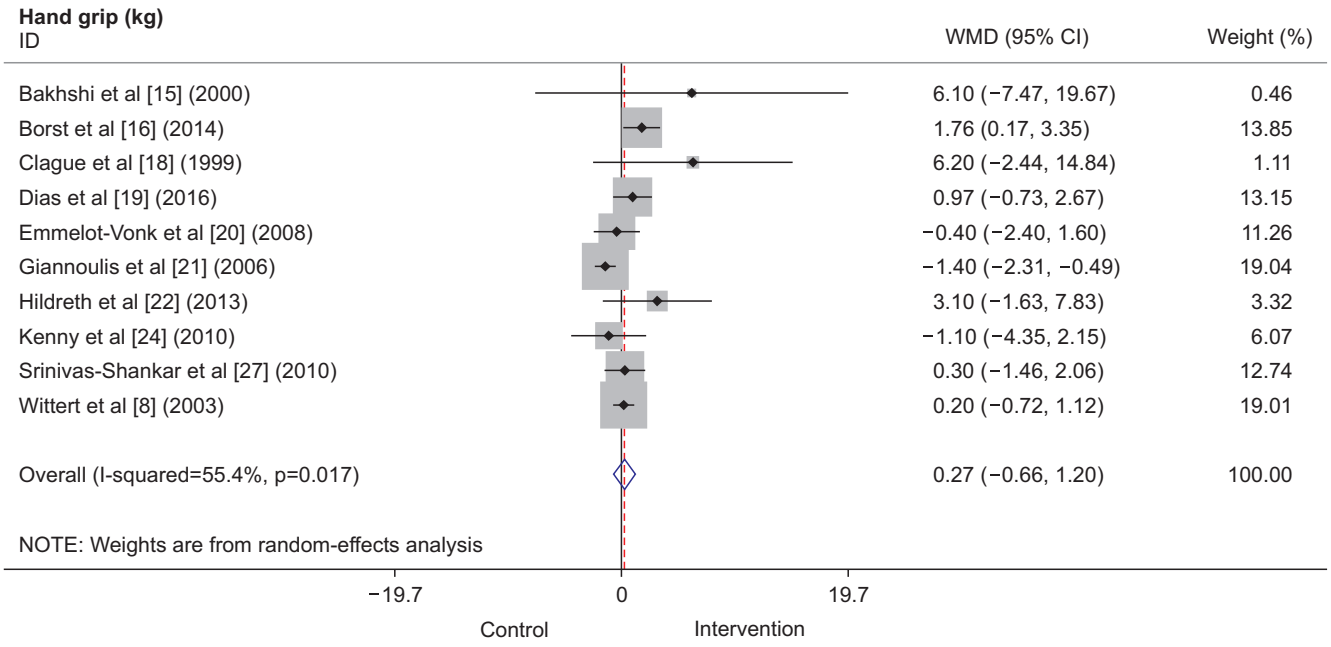

Fig. 1. Forest plot showing the weighted mean differences (WMDs) and 95% confidence intervals (CIs) for hand grip strength in kilograms as derived from available randomized controlled trials on the effect of testosterone replacement therapy *vs*. placebo.

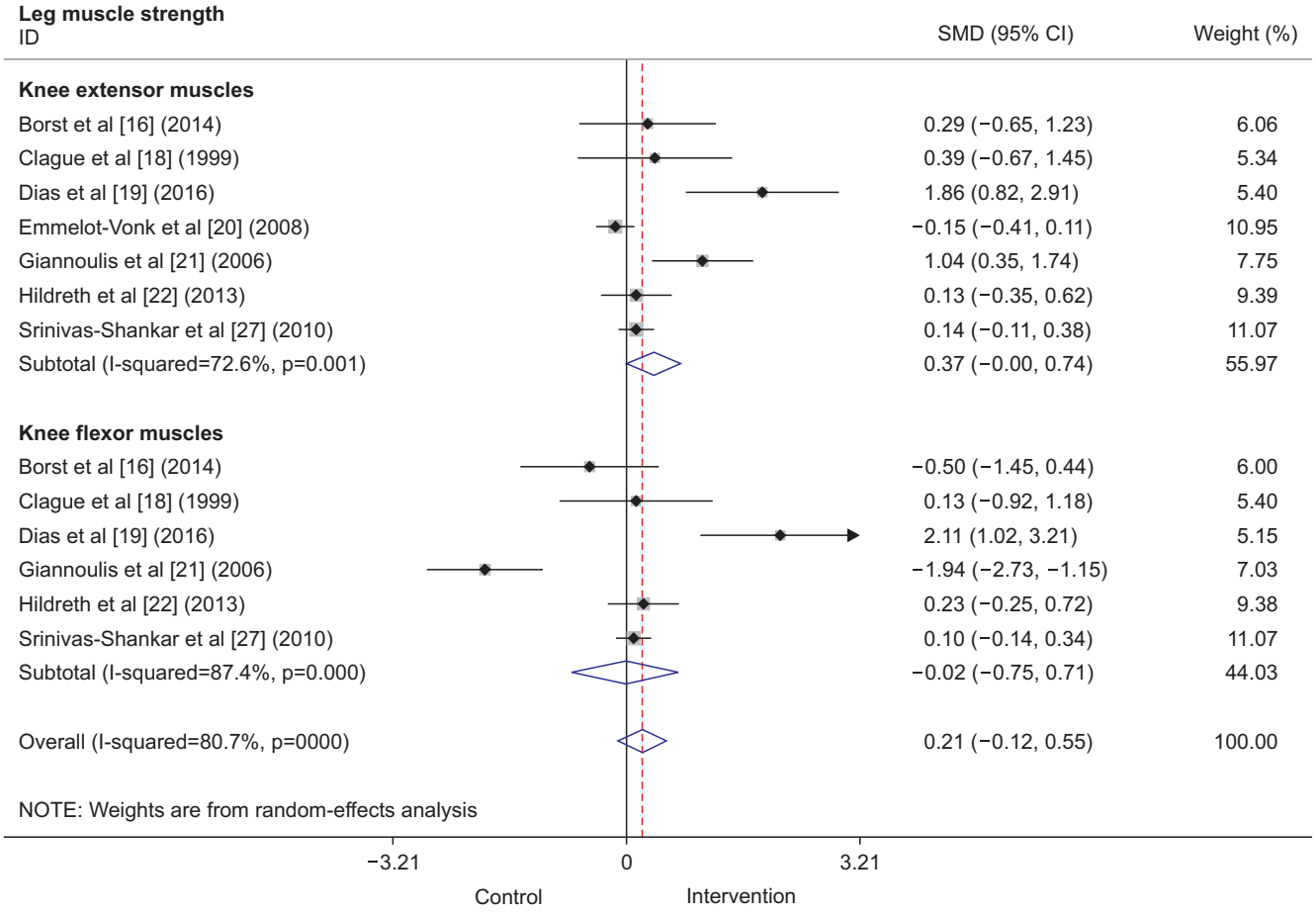

Fig. 2. Forest plot showing standardized mean differences (SMDs) and 95% confidence intervals (CIs) for leg muscle strength as derived from available randomized controlled trials on the effect of testosterone replacement therapy *vs*. placebo.

m (95% confidence interval [CI], 0.64–18.07 m).

2. Physical Activity Scale for the Elderly score

Three studies included results from the PASE questionnaire [23,24,27], which is known to be a reliable instrument for assessing physical activity in older people [29]. Two studies were conducted in the United States, and 1 was conducted in New Zealand. A total of 504 participants were included in the analysis. The intervention period varied from 6 months to 1 year, and the participants in all 3 studies received transdermal testosterone supplementation. Two studies [23,24] administered 5 mg per day for 12 months, while another [27] administered 50 mg per day for 6 months. Even though each of the individual studies did not show statistically significant improvements in the PASE score, when the individual results were combined, the overall results (Fig. 4) showed that there was an increase of 18.22 points (95% CI, 1.27–35.18 points) in the PASE score with TRT.

3. Physical performance test

Three studies were combined to analyze the effects of TRT on physical performance. Two studies were conducted in the United States, and 1 was conducted in New Zealand. A total of 733 participants were included.

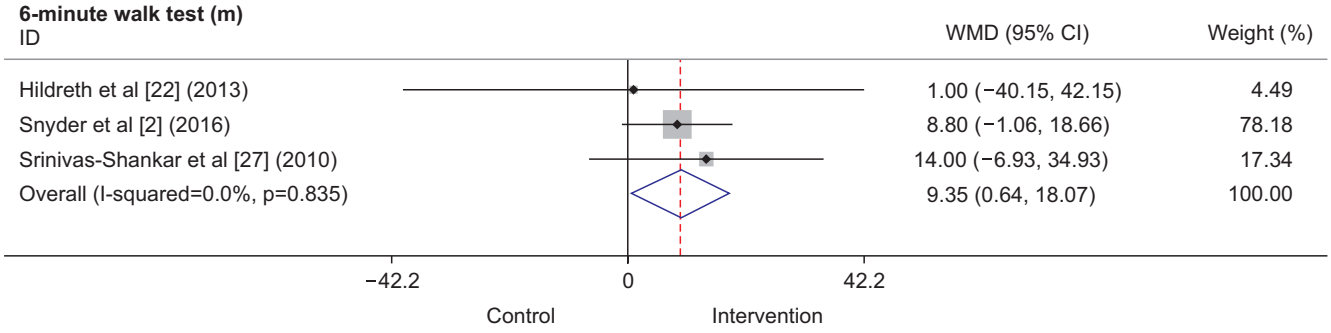

Fig. 3. Forest plot showing weighted mean differences (WMDs) and 95% confidence intervals (CIs) for the 6-minute walk test in meters as derived from available randomized controlled trials on the effect of testosterone replacement therapy *vs*. placebo.

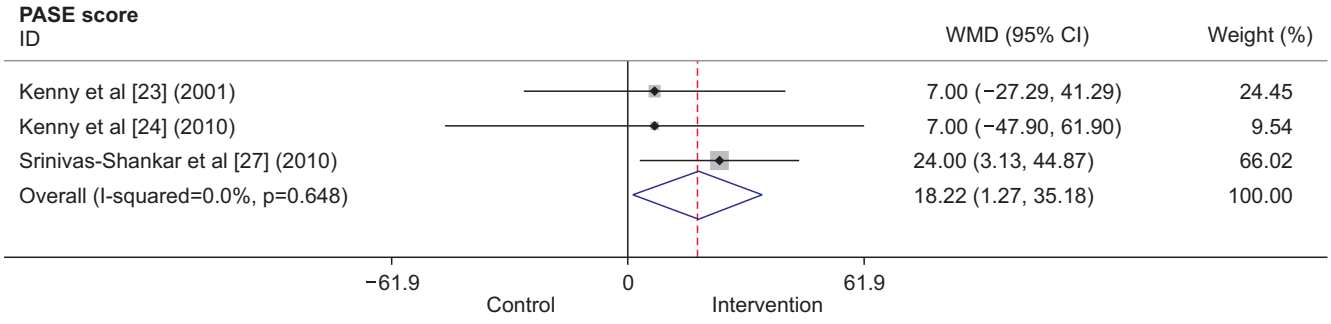

Fig. 4. Forest plot showing weighted mean differences (WMDs) and 95% confidence intervals (CIs) for physical activity scale for the elderly (PASE) scale as derived from available randomized controlled trials on the effect of testosterone replacement therapy *vs*. placebo.

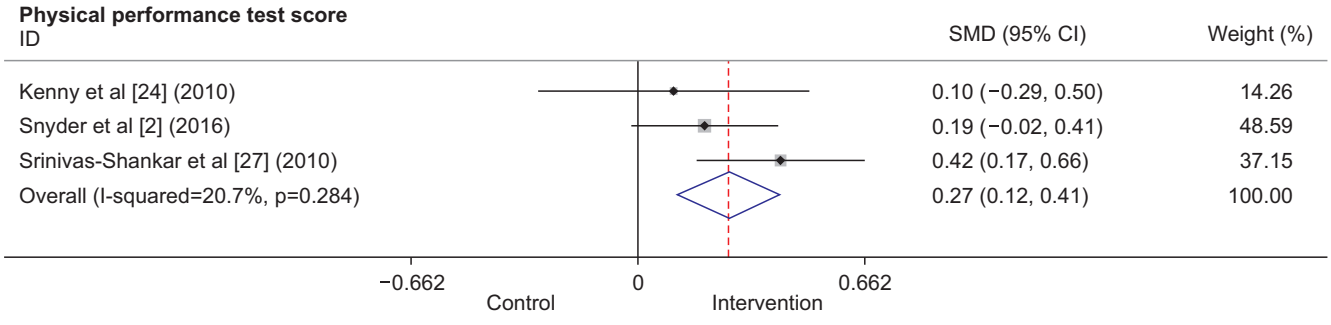

Fig. 5. Forest plot showing standardized mean differences (SMDs) and 95% confidence intervals (CIs) for physical performance test scores as derived from available randomized controlled trials on the effect of testosterone replacement therapy *vs*. placebo.

All interventions consisted of transdermal testosterone, and the intervention period varied from 6 months to 1 year. The overall results (Fig. 5) showed improved physical performance after TRT. Other studies that also showed improved physical performance after testosterone supplementation could not be included in the analysis because they only presented data in figures for a timed PPT [26] or gait speed [19] and did not provide exact measurements. Another study did not report outcomes for the functional assessment tests that were conducted, such as maximal reach, standing balance, fast walk, and chair rise [9], and hence, could not be included.

SUMMARY

The overall analyses showed that TRT is not associated with increased muscle strength, whereas it does increase physical function. This is in accordance with previous findings [7-9]. However, as the number of studies included in some analyses was limited, these results are not generalizable to elderly men in general as a whole. Additionally, most of the reviewed RCTs were conducted in the United States or Europe, and there was no mention of the inclusion of participants from other ethnic backgrounds, such as Asians. Therefore, further studies are needed to investigate the effects of TRT on muscle strength and physical function in populations with a broader ethnic background.

Some difficulties were encountered during the review. For instance, when comparing results for muscle strength, some studies included exercise training and some did not. Moreover, some studies measured isokinetic muscle strength and some measured isometric muscle strength with different dynamometers, which made them difficult to compare. Additionally, the muscle strength evaluation protocols differed for each study. The lack of a standardized, validated exercise protocol to evaluate muscle strength made it difficult to compare the results and suggests that further studies are needed to develop a standardized protocol to be disseminated and widely used.

SEARCH STRATEGIES

We performed 2 sets of searches, one as a rapid SR and an additional traditional SR, because the latest published SR included studies published up to April 9, 2016. The results for the rapid and traditional SRs and meta-analyses were reported following the criteria of the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) statement [35] (Fig. 6).

For the rapid SR, the PubMed Central, MEDLINE, and Embase databases were searched for studies published up to October 30, 2017. The initial search was conducted using the following keywords: elderly men aged at least 60 years (for participants), testosterone replacement (for intervention), placebo (for comparison), physical performance (for outcome), and SR (for study design) (see Supplementary Table 1 for the complete list of the search terms). We only included articles published in English. Studies were considered eligible for the rapid SR if they met the following inclusion criteria: 1) being a SR; 2) including elderly men at least 60 years of age; 3) reporting physical function, physical performance, and/or muscle strength; 4) having TRT as the major intervention; and 5) including a comparison with a placebo. Although we searched for healthy elderly men, we did not exclude studies if they included participants with hypogonadism, but studies of other specific diseases, such as diabetes mellitus and Parkinson disease, were excluded from the analysis. We included interventions that compared testosterone supplementation *vs.* placebo or other substances, such as growth hormone. As a result, 1 SR [13] was retrieved for assessment. The other 3 SRs were excluded because 2 [10,11] included inappropriate populations and one [12] had inappropriate outcomes. We then applied AMSTAR [36], a measurement tool for assessing the methodological quality of the retrieved SR, which gave a score of 4 out of 11 (see Appendix 1 for details).

For the additional traditional SR, as the retrieved SR used for the rapid SR included studies between January 1, 1950, and April 9, 2016, we searched for additional RCTs from January 1, 2015, to October 30, 2017 through the PubMed Central, MEDLINE, and Embase databases. We used the same search terms as were used for the rapid SR except for the study design, which was RCTs (see Supplementary Table 2 for a complete list of the search terms). After excluding studies that did not match our inclusion criteria, we included a total of 17 studies and added the 4 RCTs retrieved from the additional search. Subsequently, a total of 21 full-text articles were reviewed for the analysis and 3 studies were further excluded because their population [37,38] and outcome measures [39] did not

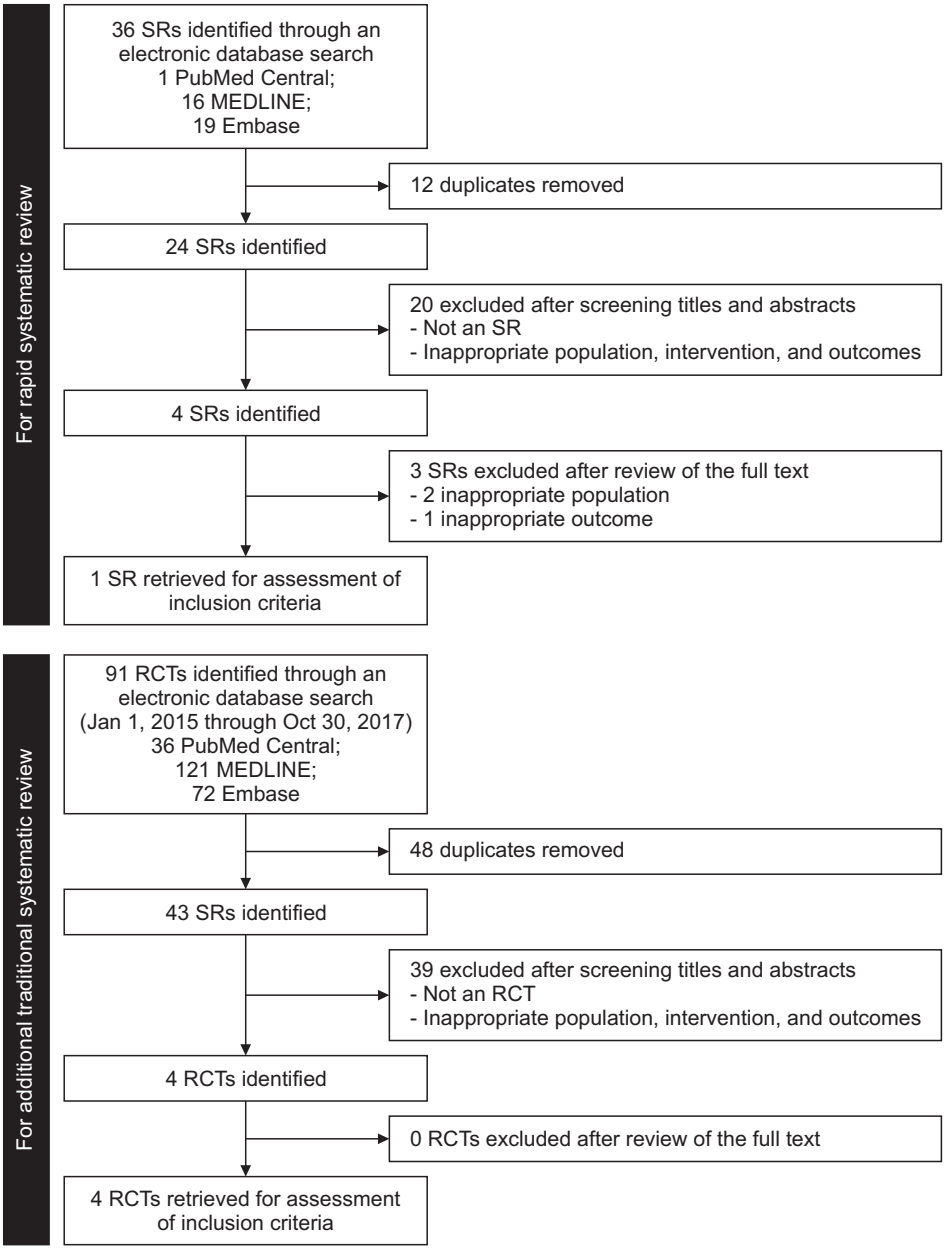

Fig. 6. Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) statement. SR: systematic review, RCT: randomized controlled trial.

meet our criteria. A total of 18 studies were included in the final analysis.

STATISTICAL ANALYSIS

Statistical analyses were performed using STATA ver. 15.0 (StataCorp, College Station, TX, USA) using the *metan* command with a random effect. When outcome measurements were given in the same units or the units could be converted, weighted mean differences were calculated, whereas if outcome measurements were given in different, non-convertible units, standardized mean differences were calculated instead.

Disclosure

The authors have no potential conflicts of interest to disclose.

Author Contribution

Drafting of the manuscript: Nam YS. Data acquisition: Lee G, Yun JM. Data analysis and interpretation: Lee G, Nam YS, Yun JM. Statistical analysis: Lee G. Critical revision of the manuscript: Yun JM, Cho B. Approval of final manuscript: all authors.

Supplementary Materials

Supplementary materials can be found via https://doi.org/10.5534/wjmh.182001.

REFERENCES

1. Buckinx F, Reginster JY, Petermans J, Croisier JL, Beaudart C, Brunois T, et al. Relationship between frailty, physical performance and quality of life among nursing home residents: the SENIOR cohort. Aging Clin Exp Res 2016;28:1149-57.
2. Snyder PJ, Bhasin S, Cunningham GR, Matsumoto AM, Stephens-Shields AJ, Cauley JA, et al. Effects of Testosterone Treatment in Older Men. N Engl J Med 2016;374:611-24.
3. Shores MM, Moceri VM, Gruenewald DA, Brodkin KI, Matsumoto AM, Kivlahan DR. Low testosterone is associated with decreased function and increased mortality risk: a preliminary study of men in a geriatric rehabilitation unit. J Am Geriatr Soc 2004;52:2077-81.
4. Chou CH, Hwang CL, Wu YT. Effect of exercise on physical function, daily living activities, and quality of life in the frail older adults: a meta-analysis. Arch Phys Med Rehabil 2012;93:237-44.
5. Narici MV, Maffulli N. Sarcopenia: characteristics, mechanisms and functional significance. Br Med Bull 2010;95:139-59.
6. Baumgartner RN, Waters DL, Gallagher D, Morley JE, Garry PJ. Predictors of skeletal muscle mass in elderly men and women. Mech Ageing Dev 1999;107:123-36.
7. Sih R, Morley JE, Kaiser FE, Perry HM 3rd, Patrick P, Ross C. Testosterone replacement in older hypogonadal men: a 12-month randomized controlled trial. J Clin Endocrinol Metab 1997;82:1661-7.
8. Wittert GA, Chapman IM, Haren MT, Mackintosh S, Coates P, Morley JE. Oral testosterone supplementation increases muscle and decreases fat mass in healthy elderly males with low-normal gonadal status. J Gerontol A Biol Sci Med Sci 2003;58:618-25.
9. Ly LP, Jimenez M, Zhuang TN, Celermajer DS, Conway AJ, 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 2001;86:4078-88.
10. Isidori AM, Giannetta E, Greco EA, Gianfrilli D, Bonifacio V, Isidori A, et al. Effects of testosterone on body composition, bone metabolism and serum lipid profile in middle-aged men: a meta-analysis. Clin Endocrinol (Oxf) 2005;63:280-93.
11. Vigano A, Piccioni M, Trutschnigg B, Hornby L, Chaudhury P, Kilgour R. Male hypogonadism associated with advanced cancer: a systematic review. Lancet Oncol 2010;11:679-84.
12. Corona G, Giagulli VA, Maseroli E, Vignozzi L, Aversa A, Zitzmann M, et al. Therapy of endocrine disease: testosterone supplementation and body composition: results from a meta-analysis study. Eur J Endocrinol 2016;174:R99-116.
13. Huo S, Scialli AR, McGarvey S, Hill E, Tügertimur B, Hogenmiller A, et al. Treatment of men for "Low Testosterone": a systematic review. PLoS One 2016;11:e0162480.
14. Storer TW, Basaria S, Traustadottir T, Harman SM, Pencina K, Li Z, et al. Effects of testosterone supplementation for 3 years on muscle performance and physical function in older men. J Clin Endocrinol Metab 2017;102:583-93.
15. Bakhshi V, Elliott M, Gentili A, Godschalk M, Mulligan T. Testosterone improves rehabilitation outcomes in ill older men. J Am Geriatr Soc 2000;48:550-3.
16. Borst SE, Yarrow JF, Conover CF, Nseyo U, Meuleman JR, Lipinska JA, et al. Musculoskeletal and prostate effects of combined testosterone and finasteride administration in older hypogonadal men: a randomized, controlled trial. Am J Physiol Endocrinol Metab 2014;306:E433-42.
17. Brill KT, Weltman AL, Gentili A, Patrie JT, Fryburg DA, Hanks JB, et al. Single and combined effects of growth hormone and testosterone administration on measures of body composition, physical performance, mood, sexual function, bone turnover, and muscle gene expression in healthy older men. J Clin Endocrinol Metab 2002;87:5649-57.
18. Clague JE, Wu FC, Horan MA. Difficulties in measuring the effect of testosterone replacement therapy on muscle function in older men. Int J Androl 1999;22:261-5.
19. Dias JP, Melvin D, Simonsick EM, Carlson O, Shardell MD, Ferrucci L, et al. Effects of aromatase inhibition vs. testosterone in older men with low testosterone: randomized-controlled trial. Andrology 2016;4:33-40.
20. Emmelot-Vonk MH, Verhaar HJ, Nakhai Pour HR, Aleman A, Lock TM, Bosch JL, 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.
21. Giannoulis MG, Sonksen PH, Umpleby M, Breen L, Pentecost C, Whyte M, et al. The effects of growth hormone and/or testosterone in healthy elderly men: a randomized controlled trial. J Clin Endocrinol Metab 2006;91:477-84.
22. Hildreth KL, Barry DW, Moreau KL, Vande Griend J, Meacham RB, Nakamura T, et al. Effects of testosterone and progressive resistance exercise in healthy, highly functioning older men with low-normal testosterone levels. J Clin Endo-

crinol Metab 2013;98:1891-900.
23. Kenny AM, Prestwood KM, Gruman CA, Marcello KM, Raisz LG. Effects of transdermal testosterone on bone and muscle in older men with low bioavailable testosterone levels. J Gerontol A Biol Sci Med Sci 2001;56:M266-72.
24. Kenny AM, Kleppinger A, Annis K, Rathier M, Browner B, Judge JO, et al. Effects of transdermal testosterone on bone and muscle in older men with low bioavailable testosterone levels, low bone mass, and physical frailty. J Am Geriatr Soc 2010;58:1134-43.
25. Nair KS, Rizza RA, O'Brien P, Dhatariya K, Short KR, Nehra A, et al. DHEA in elderly women and DHEA or testosterone in elderly men. N Engl J Med 2006;355:1647-59.
26. Page ST, Amory JK, Bowman FD, Anawalt BD, Matsumoto AM, Bremner WJ, et al. Exogenous testosterone (T) alone or with finasteride increases physical performance, grip strength, and lean body mass in older men with low serum T. J Clin Endocrinol Metab 2005;90:1502-10.
27. Srinivas-Shankar U, Roberts SA, Connolly MJ, O'Connell MD, Adams JE, Oldham JA, et al. Effects of testosterone on muscle strength, physical function, body composition, and quality of life in intermediate-frail and frail elderly men: a randomized, double-blind, placebo-controlled study. J Clin Endocrinol Metab 2010;95:639-50.
28. ATS Committee on Proficiency Standards for Clinical Pulmonary Function Laboratories. ATS statement: guidelines for the six-minute walk test. Am J Respir Crit Care Med 2002;166:111-7.
29. Washburn RA, Smith KW, Jette AM, Janney CA. The Physical Activity Scale for the Elderly (PASE): development and evaluation. J Clin Epidemiol 1993;46:153-62.
30. Ware JE, Kosinski M, Gandek B. SF-36 health survey: manual and interpretation guide. Lincoln (RI): QualityMetric Inc; 2000.
31. Reuben DB, Siu AL. An objective measure of physical function of elderly outpatients. The Physical Performance Test. J Am Geriatr Soc 1990;38:1105-12.
32. Guralnik JM, Branch LG, Cummings SR, Curb JD. Physical performance measures in aging research. J Gerontol 1989;44:M141-6.
33. Alexander NB, Ulbrich J, Raheja A, Channer D. Rising from the floor in older adults. J Am Geriatr Soc 1997;45:564-9.
34. Podsiadlo D, Richardson S. The timed "Up & Go": a test of basic functional mobility for frail elderly persons. J Am Geriatr Soc 1991;39:142-8.
35. Moher D, Liberati A, Tetzlaff J, Altman DG; PRISMA Group. Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. PLoS Med 2009;6:e1000097.
36. Shea BJ, Grimshaw JM, Wells GA, Boers M, Andersson N, Hamel C, et al. Development of AMSTAR: a measurement tool to assess the methodological quality of systematic reviews. BMC Med Res Methodol 2007;7:10.
37. Travison TG, Basaria S, Storer TW, Jette AM, Miciek R, Farwell WR, et al. Clinical meaningfulness of the changes in muscle performance and physical function associated with testosterone administration in older men with mobility limitation. J Gerontol A Biol Sci Med Sci 2011;66:1090-9.
38. Magnussen LV. Testosterone therapy of men with type 2 diabetes mellitus: a randomized, double-blinded, placebo-controlled study. Dan Med J 2017;64:B5396.
39. Sullivan DH, Roberson PK, Johnson LE, Bishara O, Evans WJ, Smith ES, et al. Effects of muscle strength training and testosterone in frail elderly males. Med Sci Sports Exerc 2005;37:1664-72.

Appendix 1. AMSTAR – a measurement tool to assess the methodological quality of systematic reviews

Huo, 2016	
1. Was an 'a priori' design provided? The research question and inclusion criteria should be established before the conduct of the review. *Note: Need to refer to a protocol, ethics approval, or pre-determined/a priori published research objectives to score a "yes."*	□ Yes ■ No □ Can't answer □ Not applicable
Comments: Not stated.	
2. Was there duplicate study selection and data extraction? There should be at least two independent data extractors and a consensus procedure for disagreements should be in place. *Note: 2 people do study selection, 2 people do data extraction, consensus process or one person checks the other's work.*	□ Yes □ No ■ Can't answer □ Not applicable
Quote: "Data were extracted into tables by 4 independent reviewers according to the presence of information on cardiovascular health, sexual function, muscle weakness/wasting, mood and behavior, or cognition." However, there is no information 2 people do study selection, 2 people do data extraction, consensus process or one person checks the other's work."	
3. Was a comprehensive literature search performed? At least two electronic sources should be searched. The report must include years and databases used (e.g., Central, EMBASE, and MEDLINE). Key words and/or MESH terms must be stated and where feasible the search strategy should be provided. All searches should be supplemented by consulting current contents, reviews, textbooks, specialized registers, or experts in the particular field of study, and by reviewing the references in the studies found. *Note: If at least 2 sources + one supplementary strategy used, select "yes" (Cochrane register/Central counts as 2 sources; a grey literature search counts as supplementary).*	■ Yes □ No □ Can't answer □ Not applicable
Quote: "~ literature searches were conducted in PubMed, Embase, and APA PsycNET."	
4. Was the status of publication (i.e. grey literature) used as an inclusion criterion? The authors should state that they searched for reports regardless of their publication type. The authors should state whether or not they excluded any reports (from the systematic review), based on their publication status, language etc. *Note: If review indicates that there was a search for "grey literature" or "unpublished literature," indicate "yes." SIGLE database, dissertations, conference proceedings, and trial registries are all considered grey for this purpose. If searching a source that contains both grey and non-grey, must specify that they were searching for grey/unpublished lit.*	□ Yes ■ No □ Can't answer □ Not applicable
Comments: Not stated.	
5. Was a list of studies (included and excluded) provided? A list of included and excluded studies should be provided. *Note: Acceptable if the excluded studies are referenced. If there is an electronic link to the list but the link is dead, select "no."*	□ Yes ■ No □ Can't answer □ Not applicable
Comments: Not stated.	
6. Were the characteristics of the included studies provided? In an aggregated form such as a table, data from the original studies should be provided on the participants, interventions and outcomes. The ranges of characteristics in all the studies analyzed e.g., age, race, sex, relevant socioeconomic data, disease status, duration, severity, or other diseases should be reported. *Note: Acceptable if not in table format as long as they are described as above.*	■ Yes □ No □ Can't answer □ Not applicable
Comments: Reported in Table 3.	
7. Was the scientific quality of the included studies assessed and documented? 'A priori' methods of assessment should be provided (e.g., for effectiveness studies if the author(s) chose to include only randomized, double-blind, placebo controlled studies, or allocation concealment as inclusion criteria); for other types of studies alternative items will be relevant. *Note: Can include use of a quality scoring tool or checklist, e.g., Jadad scale, risk of bias, sensitivity analysis, etc., or a description of quality items, with some kind of result for EACH study ("low" or "high" is fine, as long as it is clear which studies scored "low" and which scored "high"; a summary score/range for all studies is not acceptable).*	■ Yes □ No □ Can't answer □ Not applicable
Quote: "We assessed quality of studies by a 5-point Jadad score. In order to be as inclusive as possible, we included all studies identified regardless of Jadad score. For clinical endpoints only (angina/ ischemia, congestive heart failure, and erectile dysfunction) we also included an analysis of studies restricted to Jadad scores of 4 or 5. We accepted whatever criteria were used by individual study authors to define low testosterone."	

8. Was the scientific quality of the included studies used appropriately in formulating conclusions?
The results of the methodological rigor and scientific quality should be considered in the analysis and the conclusions of the review, and explicitly stated in formulating recommendations.
Note: Might say something such as "the results should be interpreted with caution due to poor quality of included studies." Cannot score "yes" for this question if scored "no" for question 7.

■ Yes
□ No
□ Can't answer
□ Not applicable

Quote: In summary, the majority of studies show increased muscle mass but no effect of testosterone on muscle strength or function.

9. Were the methods used to combine the findings of studies appropriate?
For the pooled results, a test should be done to ensure the studies were combinable, to assess their homogeneity (i.e., Chi-squared test for homogeneity, I^2). If heterogeneity exists a random effects model should be used and/or the clinical appropriateness of combining should be taken into consideration (i.e., is it sensible to combine?).
Note: Indicate "yes" if they mention or describe heterogeneity, i.e., if they explain that they cannot pool because of heterogeneity/variability between interventions.

□ Yes
■ No
□ Can't answer
□ Not applicable

Comments: There's no quantitative comparison; Not stated about heterogeneity of including studies.

10. Was the likelihood of publication bias assessed?
An assessment of publication bias should include a combination of graphical aids (e.g.,funnel plot, other available tests) and/or statistical tests (e.g., Egger regression test, Hedges-Olken).
Note: If no test values or funnel plot included, score "no". Score "yes" if mentions that publication bias could not be assessed because there were fewer than 10 included studies.

□ Yes
■ No
□ Can't answer
□ Not applicable

Comments: Not stated.

11. Was the conflict of interest included?
Potential sources of support should be clearly acknowledged in both the systematic review and the included studies.
Note: To get a "yes," must indicate source of funding or support for the systematic review AND for each of the included studies.

□ Yes
■ No
□ Can't answer
□ Not applicable

Comments: No review on each of the included studies.

Total 4/11

Shea et al. BMC Medical Research Methodology 2007;7:10. doi:10.1186/1471-2288-7-10. Additional notes (in italics) made by Michelle Weir, Julia Worswick, and Carolyn Wayne based on conversations with Bev Shea and/or Jeremy Grimshaw in June and October 2008 and July and September 2010.

Supplementary Table 1. Search Queries for Systematic Reviews (SR)

CENTRAL_SR		
1	testosterone near/3 (replacement* or supplement* or treatment*):ti,ab,kw	962
2	exogenous testosterone.mp.	160
3	#1 or #2	1,087
4	physical near/3 (function* or performance*):ti,ab,kw	8,375
5	body composition:ti,ab,kw	8,245
6	#5 or #6	16,155
7	#3 and #6	152
8	#7; filter: review	1
MEDLINE_SR		
1	(testosterone adj3 (replacement* or supplement* or treatment*)).mp.	6,641
2	exogen* testosteron*.mp.	796
3	1 or 2	7,200
4	(physical adj3 (function* or performance*)).mp.	37,679
5	body composition.mp.	52,047
6	4 or 5	88,337
7	(review or review,tutorial or review, academic).pt.	2,431,078
8	(medline or medlars or embase or pubmed or cochrane).tw,sh.	161,299
9	(scisearch or psychinfo or psycinfo).tw,sh.	20,174
10	(psychlit or psyclit).tw,sh.	947
11	cinahl.tw,sh.	19,053
12	((hand adj2 search*) or (manual* adj2 search*)).tw,sh.	10,922
13	(electronic database* or bibliographic database* or computeri?ed database* or online database*).tw,sh.	26,467
14	(pooling or pooled or mantel haenszel).tw,sh.	85,046
15	(peto or dersimonian or der simonian or fixed effect).tw,sh.	5,876
16	(retraction of publication or retracted publication).pt.	11,356
17	or/8-16	260,062
18	7 and 17	130,870
19	meta-analysis.pt.	92,088
20	meta-analysis.sh.	92,088
21	(meta-analys* or meta analys* or metaanalys*).tw,sh.	153,085
22	(systematic* adj5 (review* or overview*)).tw,sh.	125,789
23	(quantitativ* adj5 (review* or overview* or synthes?s)).tw,sh.	8,816
24	(methodologic* adj5 (review* or overview*)).tw,sh.	5,467
25	(integrative research review* or research integration).tw.	127
26	or/19-25	236,137
27	18 and 26	287,625
28	3 and 6 and 27	16

The World Journal of
Men's Health

Supplementary Table 1. Continued

EMBASE_SR		
1	(testosterone NEAR/3 (replacement* or supplement* or treatment*)):ti,ab	8,402
2	exogenous testosterone':ti,ab	910
3	1 or 2	9,025
4	(physical NEAR/3 (function* or performance*)):ti,ab	49,591
5	body composition':ti,ab	38,383
6	4 or 5	86,598
7	review/exp	2,324,693
8	(literature NEAR/3 review*):ti,ab	285,910
9	meta analysis'/exp	131,402
10	systematic review'/exp	145,768
11	or/7-10	2,554,435
12	(medline or medlars or embase or pubmed or cinahl or amed or psychlit or psyclit or psychinfo or psycinfo or scisearch or cochrane):ti,ab	190,166
13	retracted article'	8,815
14	12 or 13	198,805
15	11 and 14	148,368
16	(systematic* NEAR/2 (review* or overview)):ti,ab.	143,829
17	(meta?anal* or 'meta anal*' or meta-anal* or metaanal$ or metanal*):ti,ab.	148,047
18	or/15-17	293,480
19	3 and 6 and 18	19

Supplementary Table 2. Search Queries for Randomized Control Trials (RCTs)

CENTRAL_RCT		
1	testosterone near/3 (replacement* or supplement* or treatment*):ti,ab,kw	962
2	exogenous testosterone.mp.	160
3	#1 or #2	1,087
4	physical near/3 (function* or performance*):ti,ab,kw	8,375
5	body composition:ti,ab,kw	8,245
6	#5 or #6	16,155
7	#3 and #6	152
8	#7; filter: trials	151
MEDLINE_SR		
1	(testosterone adj3 (replacement* or supplement* or treatment*)).mp.	6,641
2	exogen* testosteron*.mp.	796
3	1 or 2	7,200
4	(physical adj3 (function* or performance*)).mp.	37,679
5	body composition.mp.	52,047
6	4 or 5	88,337
7	"randomized controlled trial".pt.	497,816
8	(random* or placebo* or single blind* or double blind* or triple blind*).ti,ab.	1,090,333
9	(retraction of publication or retracted publication).pt.	11,356
10	or/7-10	1,201,185
11	(animals not humans).sh.	4,647,954
12	((comment or editorial or meta-analysis or practice-guideline or review or letter or journal correspondence) not "randomized controlled trial").pt.	4,118,531
13	(random sampl* or random digit* or random effect* or random survey or random regression).ti,ab. not "randomized controlled trial".pt.	72,108
14	or/11-13	8,603,216
15	10 not 14	886,830
16	3 and 6 and 15	121
EMBASE_RCT		
1	(testosterone NEAR/3 (replacement* or supplement* or treatment*)):ti,ab	8,402
2	exogenous testosterone':ti,ab	910
3	1 or 2	9,025
4	(physical NEAR/3 (function* or performance*)):ti,ab	49,591
5	body composition':ti,ab	38,383
6	4 or 5	86,598
7	3 and 6; filter: randomized controlled trial	72