

BMJ Open Testosterone therapy in hypogonadal men: a systematic review and network meta-analysis

Jesse Elliott,¹ Shannon E Kelly,¹ Adam C Millar,² Joan Peterson,³ Li Chen,¹ Amy Johnston,¹ Ahmed Kotb,⁴ Becky Skidmore,⁵ Zemin Bai,¹ Muhammad Mamdani,⁶ George A Wells¹

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

► Prepublication history and additional material for this paper are available online. To view please visit the journal online (<http://dx.doi.org/10.1136/bmjopen-2016-015284>).

Received 22 November 2016
Revised 31 July 2017
Accepted 16 August 2017



¹Cardiovascular Research Methods Centre, University of Ottawa Heart Institute, Ottawa, Ontario, Canada

²Department of Medicine, Mount Sinai Hospital, University of Toronto, Toronto, Ontario, Canada

³Clinical Epidemiology Program, Ottawa Hospital Research Institute, Ottawa, Ontario, Canada

⁴School of Medicine, Royal College of Surgeons in Ireland, Dublin, Ireland

⁵Independent Information Specialist, Ottawa, Ontario, Canada

⁶Li Ka Shing Knowledge Institute, St. Michael's Hospital, Toronto, Ontario, Canada

Correspondence to

Professor George A Wells;
gawells@ottawaheart.ca

ABSTRACT

Objective To assess the relative effects of individual testosterone products among hypogonadal men.

Design Systematic review and network meta-analysis.

Methods We searched MEDLINE, Embase, Cochrane CENTRAL, and grey literature (25 May 2017) for randomised-controlled trials (RCTs) and non-randomised studies (NRS) that involved hypogonadal men given testosterone replacement therapy (TRT) for ≥ 3 months. Comparators were placebo, another TRT, or the same product at a different dose. Outcomes were quality of life, depression, libido, erectile function, activities of daily living and testosterone levels, as well as cardiovascular death, myocardial infarction, stroke, prostate cancer, heart disease, diabetes, serious adverse events, withdrawals due to adverse events and erythrocytosis. RCT data were pooled via meta-analysis and network meta-analysis. Risk of bias was assessed using Cochrane's risk of bias tool (RCTs) and Scottish Intercollegiate Guidelines Network (SIGN)50 (NRS).

Results Eighty-seven RCTs and 51 NRS were included. Most were at high or unclear risk of bias, with short treatment duration and follow-up. When compared as a class against placebo, TRT improved quality of life (standardised mean difference (SMD) -0.26 , 95% CI -0.41 to -0.11), libido (SMD 0.33 , 95% CI 0.16 to 0.50), depression (SMD -0.23 , 95% CI -0.44 to -0.01) and erectile function (SMD 0.25 , 95% CI 0.10 to 0.41). Most individual TRTs were significantly better than placebo at improving libido (6/10). Only one TRT was better than placebo at improving quality of life, and no individual TRTs improved depression or erectile function. There was no increased risk of adverse events, with the exception of withdrawals due to adverse events with the use of some TRTs.

Conclusion Despite a class effect of improving quality of life, depression, erectile function and libido, major improvements were not observed with the use of any individual product. We observed no statistically significant increase in the risk of adverse events; however, longer-term high-quality trials are needed to fully assess the risk of harm.

PROSPERO registration number CRD42014009963.

Strengths and limitations of this study

- We performed a comprehensive systematic review of the published and grey literature to identify randomised studies and non-randomised studies (NRS) involving adult men with low testosterone levels.
- Although there is no universally agreed on level for the diagnosis of 'low' testosterone, we included only studies that enrolled men with total testosterone ≤ 12 nmol/L or free testosterone ≤ 225 pmol/L, consistent with recent guidelines.
- Data from NRS were poorly reported and were not suitable for pooling via meta-analysis.
- The included studies were generally at high or unclear risk of bias, and most studies had a relatively short treatment and follow-up duration.
- Longer-term high-quality studies are needed to more fully assess the risk of rare adverse events.

INTRODUCTION

Testosterone deficiency has well recognised negative effects on male sexuality and quality of life.¹ Recent clinical practice guidelines recommend testosterone replacement therapy (TRT) for adult men with low testosterone levels (hypogonadal men) with the goal of improving symptoms and elevating testosterone levels into the mid-normal range for young men.² However, two large observational studies have reported an increased risk of cardiovascular events with testosterone use,^{3,4} and the US Food and Drug Administration (FDA) and Health Canada both have warned of a potential increased risk of cardiovascular events among men using testosterone products.^{5,6}

Previous meta-analyses have reported positive effects of TRT compared with placebo on quality of life,⁷ depression^{8,9} and some aspects of sexual function.¹⁰ However, the results of individual trials have been mixed, and there is variation in testosterone formulations and doses.¹¹ Similarly, previous meta-analyses

of potential harms related to TRT use have reported contradictory findings.^{12–16} A 2013 meta-analysis reported an increased risk of cardiovascular-related events in a mixed population of hypogonadal and eugonadal men;¹⁵ however, others have found no increased risk of cardiovascular outcomes, including myocardial infarction, stroke or cardiovascular death.^{12–14 16}

Because there are multiple TRT products and many different dosing strategies,¹⁷ it may not be appropriate to group together all testosterone products, as in traditional meta-analyses. In this study, we performed a systematic review to identify randomised-controlled trials (RCTs) and non-randomised studies (NRS) involving hypogonadal men, and we used network meta-analysis to compare each individual product.

METHODS

This review was registered a priori (CRD42014009963) and followed the Cochrane handbook¹⁸ and the PRISMA for Network Meta-Analysis checklist.¹⁹ Our review included RCTs and NRS involving adult men with low testosterone taking any form of TRT compared with placebo, another TRT, or the same TRT at a different dose. We did not exclude studies on the basis of reported outcomes.

Patient involvement

No patients were involved in setting the research question or in developing plans for design, or implementation of the study. A patient representative was involved in selecting the outcome measures, and patient groups were given the opportunity to comment on the study protocol. No patients were asked to advise on the interpretation or writing up of results. There are no plans to disseminate the results of the research to study participants or the relevant patient community.

Search strategy

Using the OVID platform, we searched Ovid MEDLINE, Ovid MEDLINE In-Process & Other Non-Indexed Citations, and Embase Classic+Embase on 3 June 2014. We also searched CENTRAL in the Cochrane Library on Wiley on the same date. Grey literature were searched according to CADTH's Grey Matters Light.²⁰ All searches were updated on 25 May 2017.

We used controlled vocabulary, including 'Testosterone', 'Testosterone Congeners' and 'Androgens', to which we applied relevant subheadings 'administration & dosage', 'analogs & derivatives', 'adverse events' and 'therapeutic use', or combined with 'Hormone Replacement Therapy.' Our keywords included 'testosterone' or 'androgen' in combination with means of administration (buccal, cream, gel, implant, injections, oral, patch, transdermal) or function (replacement, substitute, supplement, therapy, treatment). We also searched 'TRT' and all known names for testosterone replacement products (eg, AndroGel, Bio-T-Gel, Striant). Truncation, wild cards and proximity operators were incorporated as appropriate

and terminology and syntax were adjusted according to database and platform. The search strategy is available in supplementary eAppendix 1.

Study selection

We included placebo-controlled and active-controlled RCTs and NRS involving adult men (≥ 18 years) with low testosterone (total testosterone ≤ 12 nmol/L or free testosterone ≤ 225 pmol/L) administered a testosterone product. We excluded studies that artificially suppressed endogenous testosterone, involved testosterone precursors, or had less than 10 participants. We included studies with a treatment duration of 12 weeks or longer, as follow-up is generally recommended for this time point.^{2 21} Cross-over trials were included if the initial period was at least 12 weeks. Titles and abstracts were screened in duplicate (JE, JP, ZB), and the full text of any potentially relevant record was evaluated (JE, JP, AJ, ZB). Disagreements were resolved by consensus.

Data extraction and risk of bias

Data were extracted by one reviewer using piloted standardised abstraction forms (Distiller SR) and checked by a second reviewer (JE, JP, AJ, AK, ZB). First-period data were extracted from cross-over trials. Risk of bias was assessed by two reviewers using the Cochrane Collaboration's risk of bias tool for RCTs or SIGN50 for cohort studies.²² Disagreements were resolved by discussion. Publication bias was assessed by visual inspection of funnel plots.

Outcomes

Outcomes of interest were quality of life, depression, libido, erectile function, activities of daily living and total testosterone level (at 3 months, 6 months, end of study) (continuous outcomes), as well as cardiovascular death, myocardial infarction, stroke, prostate cancer, diabetes, heart disease, serious adverse events, withdrawals due to adverse events and erythrocytosis (dichotomous outcomes). We included data for quality of life, depression, erectile function and libido that had been measured using a validated scale (supplementary eAppendix 2), and the direction of each scale was standardised before analysis. A higher effect estimate (eg, positive standardised mean difference (SMD) or mean difference (MD)) indicates improvement in libido, erectile function and testosterone level, and a lower effect estimate indicates improvement in quality of life and depression. Data from RCTs were included for quality of life, depression, libido, erectile function, and activities of daily living, and testosterone levels, and data from RCTs and NRS were included for harms outcomes.

Data analysis

The results of the included NRS are summarised narratively. Meta-analysis and network meta-analysis involving data from RCTs were performed as described below. We performed meta-analysis using RevMan (V.5.3; Cochrane Collaboration) and Bayesian network meta-analysis using

WinBUGS (V.1.4.3; MRC Biostatistics Unit). Analyses were based on mean change from baseline for quality of life, depression, erectile function, libido, and on mean after-treatment values for total testosterone level. The number randomised was used as the denominator for quality of life, depression, libido, erectile function and total testosterone level, and the number who received treatment was used for harms. Two trials^{23,24} were removed from the analyses because the data from these trials were outliers for each outcome and each had a considerable effect on heterogeneity. In an exploratory analysis, we removed trials enrolling men with major comorbidities (ie, HIV/AIDS, osteoporosis, metabolic syndrome, type 2 diabetes, angina, Alzheimer's disease, heart failure, end-stage renal disease).

In the network meta-analysis, we used a binomial likelihood model for dichotomous outcomes and a normal likelihood model for continuous outcomes, allowing for the inclusion of multiarm trials.²⁵ Network meta-analyses included all trials that reported each outcome with no restrictions based on comorbidities. Each dose of an individual testosterone product was included as a separate node in the evidence networks. A continuity correction was applied to adjust for zero events for harm outcomes. Assessment of model fit and choice of model (fixed vs random effects) was based on assessment of the deviance information criterion and comparison of residual deviance to the number of unconstrained data points.²⁵ Results are reported for the random-effects model. We derived point estimates and 95% credible intervals (CrIs) using the Markov Chain Monte Carlo method. MD, SMDs with SDs or ORs with 95% CrIs or CIs are reported for

continuous or dichotomous outcomes as appropriate. Vague priors (eg, $N(0, 100^2)$) were assigned for basic parameters throughout.²⁵ To ensure model convergence, trace plots and Brooks-Gelman-Rubin statistics were assessed.²⁶ Three chains were fit for each analysis with at least 20 000 iterations and a burn-in of at least 20 000 iterations. Inconsistency was assessed where possible by comparing the deviance, between-study variance, and deviance information criterion statistics of the consistency and inconsistency models.²⁷ The posterior mean deviance of the individual data points in the inconsistency model was plotted against the posterior mean deviance in the consistency model.²⁷ All network diagrams were constructed using NodeXL.

Role of the funding source

The funder had no role in study design, data collection, analysis, interpretation or writing. All authors had full access to the study data, and the corresponding author had final responsibility for the decision to submit for publication.

RESULTS

Search results and study characteristics

We identified 87 RCTs and 51 NRS published between 1997 and 2017, corresponding to 196 records (figure 1, supplementary eAppendix 3). Of these, 70 RCTs and 19 NRS reported an outcome of interest (supplementary eTable 1,2). The median treatment duration of the RCTs was 6 months (range: 3–36 months), with mean participant age ranging between 30 years and 78 years

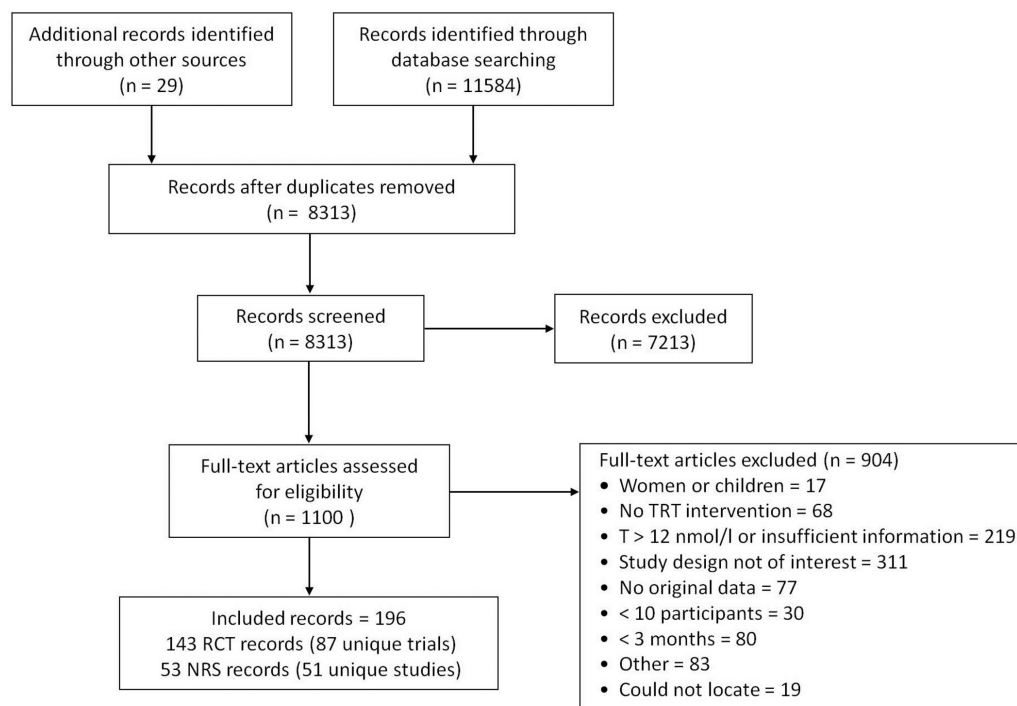


Figure 1 PRISMA flow diagram showing selection of studies. NRS, non-randomised studies; RCT, randomised-controlled trial; TRT, testosterone replacement therapy.

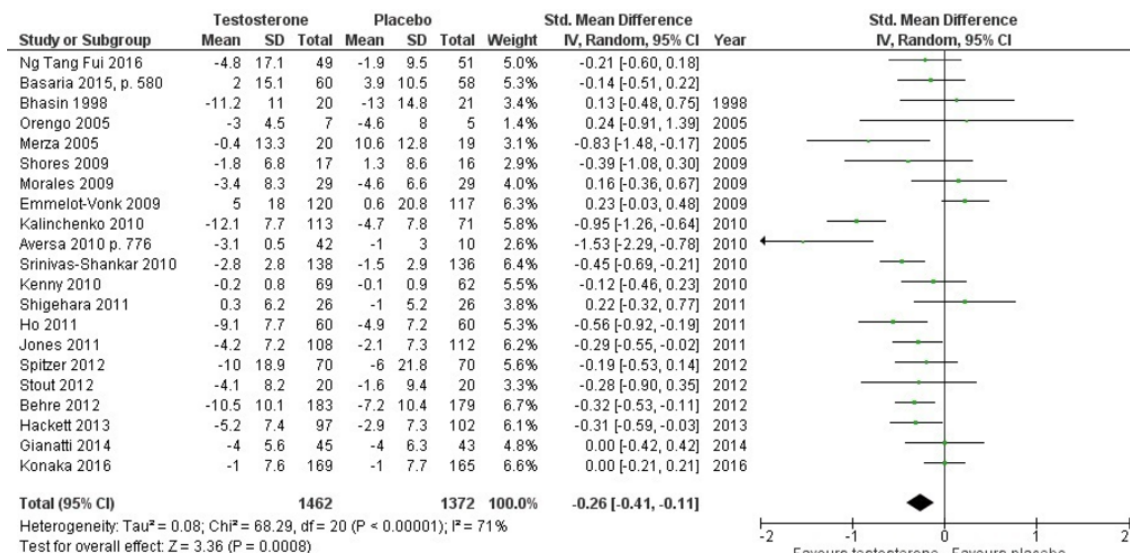


Figure 2 Meta-analysis of the effect of testosterone on quality of life.

(supplementary eAppendix 4, eTable 1). Most RCTs were placebo-controlled (87%), with two treatment groups (89%). Of the included NRS, 10 were retrospective and 9 were prospective cohorts, with a duration up to 8 years (supplementary eTable 2). Few RCTs or NRS were at low risk of bias, primarily because of a lack of details about randomisation procedures, allocation concealment and analysis populations (supplementary eTable 3,4). Based on visual inspection of funnel plots, publication bias could not be ruled out for most outcomes (supplementary eAppendix 5).

Network consistency

Comparison of the consistency and inconsistency models for libido and total testosterone levels did not show evidence of inconsistency (supplementary eAppendix 6). Consistency could not be evaluated for quality of life, depression or erectile function because of a lack of closed loops. The full network characteristics and evidence networks are presented in supplementary eAppendix 7 and supplementary eFigure 1A–E.

Outcomes

Quality of life

In total, 23 RCTs (21 placebo-controlled, two active-controlled) involving 3090 participants assessed quality of life. Compared with placebo, treatment with any TRT significantly improved quality of life (SMD -0.26 , 95% CI -0.41 to -0.11 ; $n=2834$) with substantial heterogeneity ($I^2=71%$; figure 2). To explore this heterogeneity, we excluded RCTs that involved men with major comorbidities. This had little effect on heterogeneity ($I^2=62%$) or the point estimate (SMD -0.17 , 95% CI -0.34 to -0.01 ; supplementary eFigure 2).

The evidence network for quality of life comprised 23 RCTs, representing 14 treatments in addition to placebo (figure 3). Intramuscular (IM) testosterone undecanoate (TU; 1000mg/12 weeks) significantly improved quality

of life relative to placebo (SMD -0.48 , 95% CI -0.84 to -0.10) and to oral TU (160mg/day; SMD -0.68 , 95% CI -1.32 to -0.02), with no other significant differences among the other treatments (table 1). The results were consistent when studies involving men with major comorbidities were removed from the network, despite three fewer treatments being included in the network (data not shown).

Depression

Twelve RCTs (11 placebo-controlled, 1 active-controlled) involving 852 participants randomised to nine treatments evaluated depression. Compared with placebo, treatment with any TRT improved depression (SMD -0.23 , 95% CI -0.44 to -0.01 ; $n=786$; $I^2=44%$) (supplementary eAppendix 8 (supplementary eFigure 3A)). Removal of trials involving men with major comorbidities did not reduce heterogeneity although the effect of TRT was no longer statistically significant (SMD -0.12 , 95% CI -0.49 to 0.26 ; $I^2=56%$) (supplementary eFigure 3B).

In the network meta-analysis, there were no significant differences in depression for any individual TRT compared with placebo or among the treatments (supplementary eTable 5). Removal of trials involving major comorbidities did not alter the results (two treatments removed; data not shown).

Libido

Fourteen RCTs (12 placebo-controlled, 2 active-controlled) involving 3167 patients randomised to 10 treatments investigated libido. Compared with placebo, treatment with any TRT significantly improved libido (SMD 0.33 , 95% CI 0.16 to 0.50 ; $I^2=74%$; $n=2732$) (supplementary eFigure 4A). Removal of trials involving men with major comorbidities increased heterogeneity ($I^2=80%$), and the point estimate was no longer statistically significant (SMD 0.19 , 95% CI -0.03 to 0.41) (supplementary eFigure 4B).

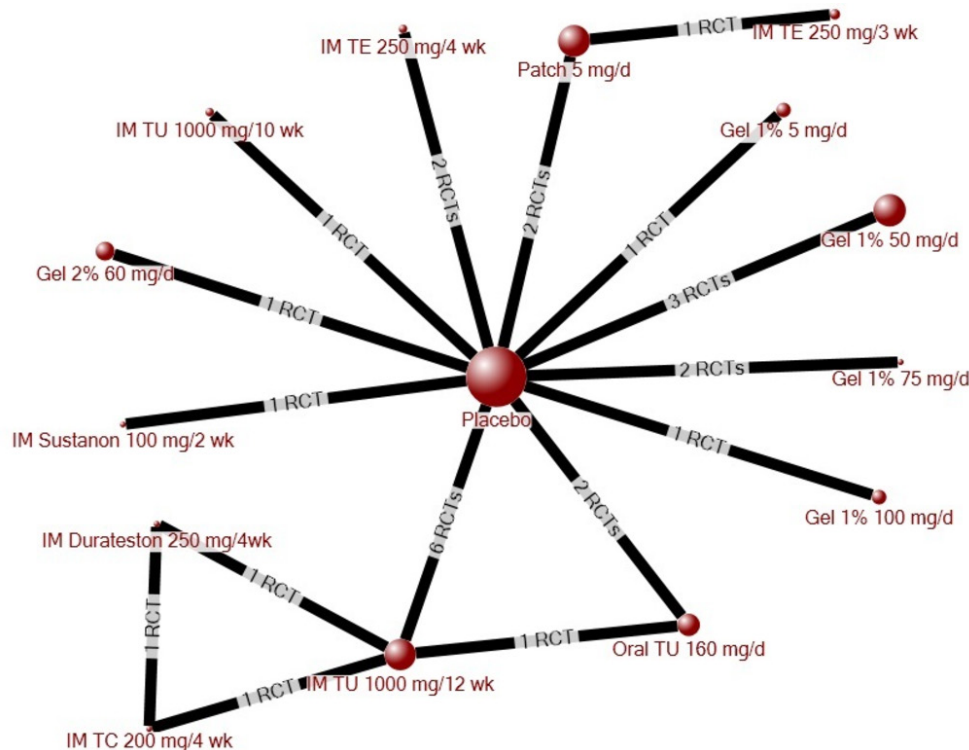


Figure 3 Evidence network for quality of life. The size of each circle (node) is proportional to the number of randomly assigned patients and indicates sample size. The number of randomised-controlled trials (RCTs) that contributed to each direct comparison is indicated on the line between nodes. IM, intramuscular injection; TC, testosterone cypionate; TE, testosterone enanthate; TU, testosterone undecanoate.

In the network meta-analysis, most individual treatments significantly improved libido compared with placebo (6/10 treatments; supplementary eTable 6). Among the treatments, 1% gel (100 mg/day) was significantly better than patch (5 mg/day) and oral TU (160 mg/day), and oral TU (120 mg/day) was significantly better than patch (5 mg/day) and 1% gel (75 mg/day). Oral TU (160 mg/day) was significantly worse than most other TRTs in the network (8/9 TRTs) (supplementary eTable 6). Removal of trials involving major comorbidities resulted in the removal of four treatments from the network; however, the results were consistent for the remaining treatments (data not shown).

Erectile function

Seventeen RCTs (all placebo-controlled) involving 3165 patients randomised to nine treatments evaluated erectile function. Compared with placebo, treatment with any TRT improved erectile function (SMD 0.25, 95% CI 0.10 to 0.41), with substantial heterogeneity ($I^2=74%$; supplementary eFigure 5A). Removing trials involving men with major comorbidities reduced heterogeneity ($I^2=58%$), with no qualitative change to the point estimate (SMD 0.36, 95% CI 0.21 to 0.51; supplementary eFigure 5B).

In the network meta-analysis, there were no significant differences in erectile function between placebo and any individual TRT or among the treatments (supplementary eTable 7). Removal of trials involving major comorbidities

did not alter the results (two treatments removed; data not shown).

Activities of daily living

One RCT reported no significant difference in activities of daily living among men with mild cognitive impairment following 6 months of TRT (testosterone gel 50–100 mg/day) compared with placebo ($p=0.31$).²⁸

Testosterone levels

In total, 26 and 23 RCTs reported total testosterone levels after 3 months or 6 months of treatment, respectively. End of treatment testosterone levels were reported in 57 RCTs. After 3 months or 6 months of treatment, about half of the treatments in each network were associated with significantly higher total testosterone levels compared with placebo (3 months: 6/15; 6 months: 11/18) (supplementary eTable 8,9). By the end of treatment (12 weeks to 36 months), most products were associated with total testosterone above 12 nmol/L (26/28 testosterone therapies), and 17 of 28 treatments had significantly higher levels relative to placebo (supplementary eTable 10).

Cardiovascular death

Ten RCTs reported the occurrence of cardiovascular death during the treatment period, while an additional nine trials reported that no cardiovascular deaths had occurred (18 placebo-controlled RCTs). Compared with placebo via pairwise meta-analysis, there was no significant

Table 1 Quality of life—indirect comparison of testosterone products

	Standardised mean difference (95% credible interval)*														
	Placebo	Patch, 5 mg/day	Gel 1%, 5 mg/day	Gel 1%, 50 mg/day	Gel 1%, 75 mg/day	Gel 1%, 100 mg/day	Gel 2%, 60 mg/day	Oral TU, 160 mg/day	IM TU, 1000 mg/10 weeks	IM TU, 1000 mg/12 weeks	IM TE, 250 mg/3 weeks	IM TE, 250 mg/4 weeks	IM TC, 200 mg/4 weeks	IM Sustanon, 100 mg/2 weeks	IM Durateston, 250 mg/4 weeks
Patch, 5 mg/day	-0.33 (-0.97 to 0.31)	-													
Gel 1%, 5 mg/day	-0.11 (-0.86 to 0.64)	0.21 (-0.78 to 1.20)	-												
Gel 1%, 50 mg/day	-0.32 (-0.74 to 0.17)	0.01 (-0.76 to 0.82)	-0.20 (-1.05 to 0.69)	-											
Gel 1%, 75 mg/day	-0.24 (-0.83 to 0.34)	0.09 (-0.78 to 0.95)	-0.12 (-1.09 to 0.82)	0.08 (-0.70 to 0.79)	-										
Gel 1%, 100 mg/day	-0.19 (-0.94 to 0.55)	0.13 (-0.84 to 1.10)	-0.08 (-1.13 to 0.97)	0.12 (-0.78 to 0.96)	0.04 (-0.89 to 0.99)	-									
Gel 2%, 60 mg/day	-0.29 (-1.01 to 0.44)	0.04 (-0.93 to 1.00)	-0.17 (-1.21 to 0.87)	0.03 (-0.85 to 0.86)	-0.05 (-0.96 to 0.89)	-0.09 (-1.12 to 0.94)	-								
Oral TU, 160 mg/day	0.20 (-0.33 to 0.74)	0.53 (-0.30 to 1.36)	0.32 (-0.61 to 1.24)	0.52 (-0.21 to 1.19)	0.44 (-0.34 to 1.24)	0.40 (-0.51 to 1.31)	0.49 (-0.41 to 1.38)	-							
IM TU, 1000 mg/10 weeks	-0.21 (-0.96 to 0.56)	0.12 (-0.87 to 1.13)	-0.09 (-1.15 to 0.98)	0.11 (-0.80 to 0.98)	0.03 (-0.92 to 1.00)	-0.01 (-1.07 to 1.06)	0.08 (-0.97 to 1.14)	-0.41 (-1.35 to 0.53)	-						
IM TU, 1000 mg/12 weeks	-0.48 (-0.84 to -0.10)	-0.15 (-0.89 to 0.60)	-0.36 (-1.19 to 0.48)	-0.16 (-0.76 to 0.41)	-0.24 (-0.91 to 0.47)	-0.28 (-1.10 to 0.55)	-0.19 (-1.00 to 0.62)	-0.68 (-1.32 to -0.02)	-0.27 (-1.12 to 0.58)	-					
IM TE, 250 mg/3 weeks	-0.05 (-1.04 to 0.94)	0.28 (-0.47 to 1.03)	0.07 (-1.17 to 1.30)	0.27 (-0.85 to 1.33)	0.19 (-1.09 to 1.34)	0.15 (-1.09 to 1.37)	0.24 (-1.00 to 1.46)	-0.25 (-1.38 to 0.88)	0.16 (-1.10 to 1.40)	0.43 (-0.63 to 1.49)	-				
IM TE, 250 mg/4 weeks	0.08 (-0.44 to 0.63)	0.41 (-0.41 to 1.26)	0.20 (-0.72 to 1.13)	0.40 (-0.32 to 1.09)	0.32 (-0.44 to 1.13)	0.28 (-0.62 to 1.21)	0.37 (-0.51 to 1.28)	-0.12 (-0.86 to 0.66)	0.29 (-0.64 to 1.22)	0.56 (-0.97 to 1.22)	0.13 (-0.97 to 1.27)	-			
IM TC, 200 mg/4 weeks	0.12 (-1.00 to 1.25)	0.45 (-0.83 to 1.74)	0.24 (-1.10 to 1.59)	0.44 (-0.78 to 1.64)	0.36 (-0.89 to 1.63)	0.32 (-1.03 to 1.66)	0.41 (-0.92 to 1.74)	-0.08 (-1.32 to 1.18)	0.33 (-1.01 to 1.68)	0.60 (-0.46 to 1.67)	0.17 (-1.32 to 1.66)	0.04 (-1.20 to 1.28)	-		
IM sustanon, 100 mg/2 weeks	-0.28 (-1.17 to 0.62)	0.05 (-1.05 to 1.14)	-0.16 (-1.32 to 1.01)	0.04 (-0.99 to 1.02)	-0.04 (-1.10 to 1.03)	-0.08 (-1.24 to 1.09)	0.01 (-1.13 to 1.16)	-0.48 (-1.52 to 0.57)	-0.07 (-1.24 to 1.11)	0.20 (-0.77 to 1.16)	-0.23 (-1.56 to 1.11)	-0.36 (-1.42 to 0.67)	-0.40 (-1.84 to 1.02)	-	
IM durateston 250 mg/4 weeks	-0.37 (-1.49 to 0.74)	-0.05 (-1.33 to 1.24)	-0.26 (-1.61 to 1.09)	-0.06 (-1.27 to 1.12)	-0.14 (-1.37 to 1.13)	-0.18 (-1.51 to 1.16)	-0.08 (-1.40 to 1.23)	-0.57 (-1.81 to 0.66)	-0.17 (-1.52 to 1.17)	0.10 (-0.95 to 1.15)	-0.33 (-1.81 to 1.17)	-0.46 (-1.70 to 0.76)	-0.50 (-1.52 to 0.52)	-	

*Random-effects model, with analysis based on mean change from baseline. A negative standardised mean difference indicates improvement in quality of life. Statistically significant changes are indicated by use of bold and colour (green indicates that the row treatment is significantly better than the column treatment), and white indicates no statistically significant difference between treatments. IM, intramuscular injection; TC, testosterone cypionate; TE, testosterone enanthate; TU, testosterone undecanoate.

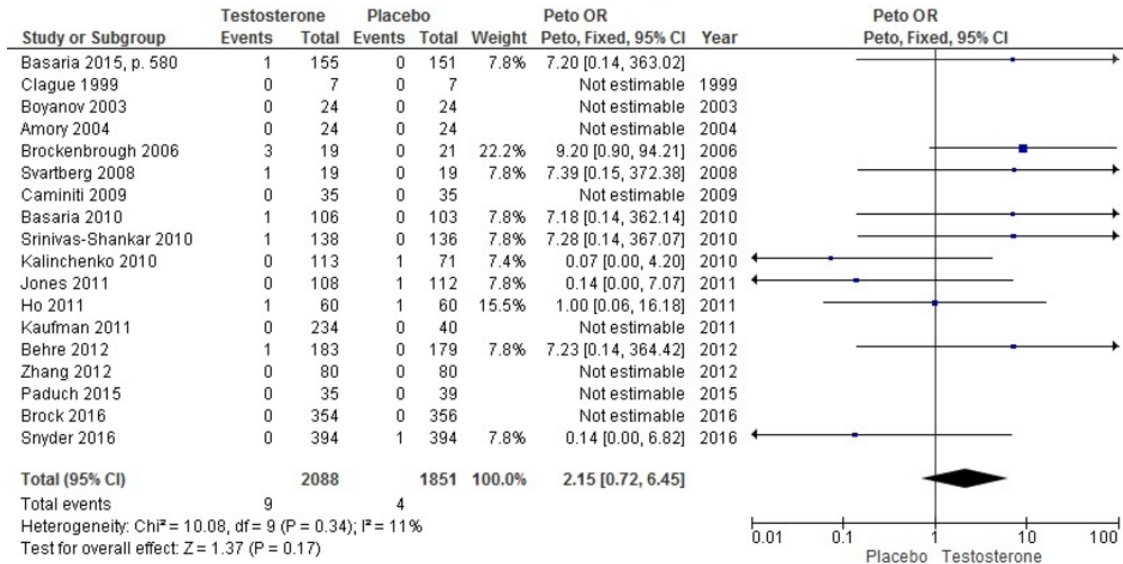


Figure 4 Odds of cardiovascular death associated with the use of any testosterone product versus placebo.

difference in the risk of cardiovascular death between placebo and any TRT (all products grouped together) (OR 2.15, 95% CI 0.72 to 6.45; $I^2=11%$) (figure 4; table 2). Because of the low event rate, network meta-analysis did not provide robust estimates for this outcome (data not reported).

Other adverse events

Compared with placebo via meta-analysis, there were no increased odds of myocardial infarction, stroke, prostate cancer, heart disease, erythrocytosis or serious adverse events, with the use of any TRT product (table 2; supplementary eFigure 6–11). Although not statistically

significant, withdrawals due to adverse events tended to be higher among TRT users compared with placebo (OR 1.31, 95% CI 0.98 to 1.73; $I^2=13%$) (supplementary eFigure 12). No RCTs reported incident diabetes during the treatment period.

Owing to the low event rates for most outcomes, network meta-analysis was only possible for serious adverse events and withdrawals due to adverse events. In the head-to-head comparison of TRTs, there was no significant difference in serious adverse events between any TRT and placebo or among the TRTs (supplementary eTable 11). Use of testosterone patch (5 mg/day) was associated with an increased odds of withdrawal due to

Table 2 Meta-analysis of adverse events associated with the use of any testosterone compared with placebo

Outcome	No. of RCTs	Treatment duration	Event/no. treated	OR (95% CI)*	I^2
Cardiovascular death	18	12 weeks to 36 months	Placebo: 4/1851 TRT: 9/2088	2.15 (0.72 to 6.45)	11%
Myocardial infarction	15	12 weeks to 36 months	Placebo: 10/1613 TRT: 6/1915	0.43 (0.15 to 1.19)	34%
Prostate cancer	13	12 weeks to 36 months	Placebo: 11/1649 TRT: 12/1877	0.97 (0.43 to 2.23)	0%
Stroke	8	12 weeks to 36 months	Placebo: 8/1103 TRT: 8/1104	0.99 (0.37 to 2.65)	29%
Heart disease	3	40 weeks to 12 months	Placebo: 5/120 TRT: 5/131	0.89 (0.24 to 3.26)	0%
Erythrocytosis	4	12 weeks to 12 months	Placebo: 0/78 TRT: 4/110	2.44 (0.26 to 22.53)	0%
Diabetes	0	–	–	–	–
Serious adverse events	18	12 weeks to 36 months	Placebo: 181/1902 TRT: 168/2138	0.88 (0.70 to 1.11)	0%
Withdrawals due to adverse events	48	12 weeks to 36 months	Placebo: 150/2551 TRT: 221/2840	1.31 (0.98 to 1.73)	13%

*Random-effects models.

RCT, randomised controlled trial; TRT, testosterone replacement therapy.

adverse events compared with placebo and many of the other TRTs (supplementary eTable 12).

Non-randomised studies

The reporting of harms in the included NRS was generally poor, primarily because of a lack of transparency around the number of patients assigned to each group, the number of events per group, and/or the type or dose of TRT (supplementary eTable 13). In the longest prospective cohort study, involving 656 men followed for up to 10 years,²⁹ men who had received IM TU (1000mg/12 weeks) were at lower risk of death (incidence of death: TRT group 0.0092, (95% CI 0.0032 to 0.0368); no TRT group 0.1145, (95% CI 0.0746 to 0.1756)), with 19 of 21 deaths in the control group attributed to cardiovascular causes and 0 of 2 deaths in the TRT group due to cardiovascular causes. Non-fatal myocardial infarction (MI) and non-fatal stroke were also more common in the untreated group than in the TRT group (TRT group: 0 MI, 0 stroke; no TRT group: 26 MI, 30 stroke). In a large retrospective cohort study in the USA, involving 8808 men dispensed a TRT product and 35527 men never dispensed TRT,³⁰ men in the TRT group were at a lower risk of sudden cardiovascular death (adjusted HR (aHR) 0.76, 95% CI 0.61 to 0.93), acute MI (aHR 0.74, 95% CI 0.63 to 0.86) and stroke (aHR 0.64, 95% CI 0.52 to 0.80) over a median follow-up time of 3.4 years. In contrast, an earlier retrospective cohort study among 8709 men who had undergone angiography reported that those who had filled a prescription for TRT were at higher risk of an adverse cardiovascular event (aHR 1.29, 95% CI 1.05 to 1.58) compared men with no TRT prescription.³ Concerning the relative safety of the different TRTs, one retrospective cohort study³¹ reported no significant difference in risk between injection or gel TRT users with low testosterone at baseline for MI (aHR 1.64, 95% CI 0.57 to 4.69) stroke (aHR 1.28, 95% CI 0.27 to 6.02) or death (aHR 5.53, 95% CI 0.98 to 31.15).

Prostate cancer was reported by 10 NRS.^{29 32–41} In a large prospective registry involving 999 newly diagnosed hypogonadal men,³⁷ there was no statistically significant difference in the risk of prostate cancer between men who used or did not use TRT (incidence rate ratio 0.52, 95% CI 0.22 to 1.26). Traish and colleagues²⁹ also reported fewer cases of prostate cancer among men who received TRT (IM TU) compared with hypogonadal men who did not receive TRT (TU: 7/360 men vs no TRT: 12/296 men) over a median follow-up of 7 years, and Yassin and colleagues³⁵ reported positive biopsies for prostate cancer among 16.7% of TRT users compared with 51.9% of non-users, as well as a lower severity in terms of staging and grading in their prospective study of prostate cancer among hypogonadal men. Among men with a previous diagnosis of prostate cancer, one small retrospective cohort study⁴¹ reported no recurrences of prostate cancer among TRT users during the follow-up period (36 months). The reporting of outcomes and treatment group in most of the remaining studies was

poor: either the group assignment of the men who experienced prostate cancer or the number of men in each treatment group was not reported in each (supplementary eTable 13).^{33 34 38}

DISCUSSION

Despite more than 70 years of clinical use, TRT remains a controversial topic. Part of the controversy may be a result of different actions of the various testosterone preparations. In an attempt to clarify the benefits and harms of individual testosterone products, we used traditional pairwise meta-analysis as well as network meta-analysis, which allows the relative comparison of products that have not been compared in head-to-head trials. Consistent with most previous meta-analyses,^{7 8 10} we found that the use of TRT improved quality of life, depression, libido and erectile function, with no increase in cardiovascular death or other major adverse events.

In the head-to-head comparison of individual testosterone treatments, we found no significant differences among products in their effect on depression or erectile function. For libido, testosterone gel (1%, 100mg/day) was significantly better than testosterone patch (5mg/day) and oral TU (160mg/day). Oral TU (160mg/day) was significantly worse than most other treatments in the network, including oral TU at 120mg/day. The finding that a lower dose of oral TU was more effective than a higher dose was unexpected and requires further investigation.

Conflicting data exist about the risk of adverse cardiovascular events among TRT users. FDA and Health Canada both issued alerts in 2015 based in part on the findings of two observational studies.^{3 4} In their retrospective cohort, Vigen and colleagues³ included men with low testosterone who had undergone angiography, reporting an increased risk of a composite cardiovascular outcome that included all-cause mortality, myocardial infarction and stroke (aHR 1.29, 95% CI 1.04 to 1.58). Most participants in this study were receiving TRT in the form of a patch (63%) and had significant medical comorbidities, which may limit generalisability. The retrospective study by Finkle and colleagues⁴ did not report the baseline testosterone level of men in their cohort nor was the cohort restricted to men with androgen deficiency; as such, this study was not eligible for inclusion in our review, but their finding that the risk of non-fatal MI is increased in the first 90 days following testosterone prescription is concerning. However, other recent large observational studies have reported a lower risk of cardiovascular death, stroke and MI among TRT users,^{29 30} supporting the findings of earlier observational studies that reported a lower risk of all-cause death among hypogonadal men using TRT.⁴⁰

Previous meta-analyses have also reported contradictory findings concerning the risk of adverse events among TRT users. One meta-analysis reported an increased risk of cardiovascular adverse events among men using TRT,¹⁵ while a number of other meta-analyses have found

no increased risk of cardiovascular events among TRT users.^{12–14 16 42 43} The meta-analysis by Xu and colleagues¹⁵ involved a broad composite outcome (cardiac disorder, cardiovascular complaints, cardiovascular events, vascular disorders) and has been criticised for use of a fixed-effects model: subsequent re-analysis using a random-effects model found no significant increase in the risk of cardiovascular events.⁴³ Our findings are consistent with previous meta-analyses that have found no increased risk of individual cardiovascular events compared with placebo.^{12–14 16 42 43}

Although we had intended to analyse the effects of individual testosterone products among men aged 65 years and older, data were limited because most RCTs included a wide age range. The Testosterone Trials were designed to address this lack of data among elderly men.⁴⁴ After 1 year of treatment with 1% testosterone gel, improved desire and erectile function were reported among men with low sexual function at baseline, with no apparent increase in the risk of adverse cardiovascular events.⁴⁴ Although these findings are encouraging, the trials were not powered to detect adverse events, and the results should not be generalised to different testosterone preparations.

Strengths and limitations

The strengths of this study include a comprehensive search of the published and grey literature without language or date restrictions. In contrast with some previous reviews, we included only studies that enrolled men with total testosterone ≤ 12 nmol/L or free testosterone ≤ 225 pmol/L. Although there is no universally agreed on threshold value for low testosterone, a recent guideline recommends TRT for men with total testosterone lower than 12 nmol/L.¹¹

Our study had several limitations. First, the included studies used a variety of assays to determine testosterone levels and a variety of cut-off values for determining 'low' testosterone. This has been noted by others.^{11 45} The US Centers for Disease Control and Prevention's hormone standardisation project⁴⁶ will help to resolve this issue. Second, the included RCTs and NRS were generally at unclear or high risk of bias, which may have an impact on the reliability of subjective data. Third, the duration of treatment and the length of follow-up may have been too short to see an effect of TRT for all outcomes, including adverse events. In keeping with the recommendation to reassess symptoms after 3 months of TRT,²¹ we included only studies with a treatment duration of 3 months or longer. The median duration of treatment was 6 months in the RCTs, but it is possible that some symptoms may take longer to resolve. Fourth, although some individual RCTs showed a positive effect of TRT on depression and erectile function compared with placebo, in the network meta-analysis no individual TRTs showed a positive effect compared with placebo. This phenomenon has been

noted previously and results in a more conservative estimate of effect.⁴⁷

CONCLUSIONS

To the best of our knowledge, this is the first study to compare the benefits and harms of individual testosterone products among hypogonadal men. Our study builds on previous meta-analyses by comparing the relative effects of individual testosterone treatments, most of which have never been compared in head-to-head trials. When considered as a class (any TRT compared with placebo), TRT improved quality of life, depression, erectile function and libido; however, when the individual products were compared head to head, there were few differences between the treatments. We found no increased risk of major harms; however, this must be viewed in light of the high risk of bias of the included studies, the rare nature of serious harms, and the short treatment duration and follow-up of most studies. Future studies need to be rigorous in design and delivery, and include comprehensive descriptions of all aspects of methodology to further enable appraisal and interpretation of results.

Acknowledgements The authors thank Wenfei Liu for assistance in screening records during the 2015 update of the literature search.

Contributors JE, SEK, MM and GAW designed the study. BS developed and executed the search strategy. JE, SEK, JP, AJ, AK and ZB selected studies for inclusion and extracted data. JE, AK and LC analysed the data. JE, SEK, LC, ACM and GAW interpreted the data. JE wrote the first draft of the manuscript, which was critically revised for intellectual content by all authors. All authors approved the final version submitted for publication and agree to be accountable for all aspects of the study.

Funding This study was supported in part by the Ontario Drug Policy Research Network, which is funded by a grant from the Ontario Ministry of Health and Long-Term Care Health System Research Fund.

Competing interests MM has received honoraria for serving on Advisory Boards for Astra Zeneca, Bristol-Myers Squibb, Eli Lilly and Company, GlaxoSmithKline, Hoffman La Roche, Novartis, Novo Nordisk and Pfizer, outside the submitted work. BS is a paid information consultant/contractor to the Ottawa Hospital Heart Institute. All other authors have declared no conflicts of interests.

Provenance and peer review Not commissioned; externally peer reviewed.

Data sharing statement Additional data are presented in supplemental files.

Open Access This is an Open Access article distributed in accordance with the Creative Commons Attribution Non Commercial (CC BY-NC 4.0) license, which permits others to distribute, remix, adapt, build upon this work non-commercially, and license their derivative works on different terms, provided the original work is properly cited and the use is non-commercial. See: <http://creativecommons.org/licenses/by-nc/4.0/>

© Article author(s) (or their employer(s) unless otherwise stated in the text of the article) 2017. All rights reserved. No commercial use is permitted unless otherwise expressly granted.

REFERENCES

1. Morgentaler A, Zitzmann M, Traish AM, *et al.* Fundamental concepts regarding Testosterone deficiency and treatment. *Mayo Clin Proc* 2016;91:881–96.
2. Morales A, Bebb RA, Manjoo P, *et al.* Diagnosis and management of Testosterone deficiency syndrome in men: clinical practice guideline. *CMAJ* 2015;187:1369–77.
3. Vigen R, O'Donnell CI, Barón AE, *et al.* Association of Testosterone therapy with mortality, myocardial infarction, and stroke in men with low testosterone levels. *JAMA* 2013;310:1829.

4. Finkle WD, Greenland S, Ridgeway GK, *et al.* Increased Risk of Non-Fatal Myocardial Infarction Following Testosterone Therapy Prescription in Men. *PLoS One* 2014;9:e85805–7.
5. Summary Safety Review – Testosterone Replacement Products Cardiovascular Risk. 2014 www.hc-sc.gc.ca/dhp-mps/medeff/reviews-examens/testosterone-eng.php (accessed 15 Aug 2016).
6. FDA Drug Safety Communication. FDA cautions about using Testosterone products for low testosterone due to aging. requires labeling change to inform of possible increased risk of heart attack and stroke with use. www.fda.gov/downloads/Drugs/DrugSafety/UCM436270.pdf (accessed 15 Aug 2015).
7. Guo C, Gu W, Liu M, *et al.* Efficacy and safety of Testosterone replacement therapy in men with hypogonadism: a meta-analysis study of placebo-controlled trials. *Exp Ther Med* 2016;11:853–63.
8. Amanatkar HR, Chibnall JT, Seo BW, *et al.* Impact of exogenous Testosterone on mood: a systematic review and meta-analysis of randomized placebo-controlled trials. *Ann Clin Psychiatry* 2014;26:19–32.
9. Zarrouf FA, Artz S, Griffith J, *et al.* Testosterone and depression: systematic review and meta-analysis. *J Psychiatr Pract* 2009;15:289–305.
10. Boloña ER, Uraga MV, Haddad RM, *et al.* Testosterone use in men with sexual dysfunction: a systematic review and meta-analysis of randomized placebo-controlled trials. *Mayo Clin Proc* 2007;82:20–8.
11. Isidori AM, Balercia G, Calogero AE, *et al.* Outcomes of androgen replacement therapy in adult male hypogonadism: recommendations from the Italian society of endocrinology. *J Endocrinol Invest* 2015;38:103–12.
12. Haddad RM, Kennedy CC, Caples SM, *et al.* Testosterone and cardiovascular risk in men: a systematic review and meta-analysis of randomized placebo-controlled trials. *Mayo Clin Proc* 2007;82:29–39.
13. Fernández-Balsells MM, Murad MH, Lane M, *et al.* Clinical review 1: Adverse effects of Testosterone therapy in adult men: a systematic review and meta-analysis. *J Clin Endocrinol Metab* 2010;95:2560–75.
14. Calof OM, Singh AB, Lee ML, *et al.* Adverse events associated with Testosterone replacement in middle-aged and older men: a meta-analysis of randomized, placebo-controlled trials. *J Gerontol A Biol Sci Med Sci* 2005;60:1451–7.
15. Xu L, Freeman G, Cowling BJ, *et al.* Testosterone therapy and cardiovascular events among men: a systematic review and meta-analysis of placebo-controlled randomized trials. *BMC Med* 2013;11:108.
16. Corona G, Maseroli E, Rastrelli G, *et al.* Cardiovascular risk associated with Testosterone-boosting medications: a systematic review and meta-analysis. *Expert Opin Drug Saf* 2014;13:1327–51.
17. McGill JJ, Shoskes DA, Sabanegh ES. Androgen deficiency in older men: indications, advantages, and pitfalls of Testosterone replacement therapy. *Cleve Clin J Med* 2012;79:797–806.
18. Higgins JP, Green S, eds. *Cochrane Handbook for Systematic Reviews of Interventions*, 2008.
19. Hutton B, Salanti G, Caldwell DM, *et al.* The PRISMA extension statement for reporting of systematic reviews incorporating network meta-analyses of health care interventions: checklist and explanations. *Ann Intern Med* 2015;162:777–84.
20. Grey matters light. https://www.cadth.ca/sites/default/files/is/cadth_Handout_greymatters_light_e.pdf (accessed 15 Aug 2016).
21. Buvat J, Maggi M, Guay A, *et al.* Testosterone deficiency in men: systematic review and standard operating procedures for diagnosis and treatment. *J Sex Med* 2013;10:245–84.
22. Sleith C. *Methodology Checklist 3: Cohort Studies*. Edinburgh, 2012 www.sign.ac.uk/checklists-and-notes.html (accessed 12 Jun 2017).
23. Cavallini G, Caracciolo S, Vitali G, *et al.* Carnitine versus androgen administration in the treatment of sexual dysfunction, depressed mood, and fatigue associated with male aging. *Urology* 2004;63:641–6 <http://www.sciencedirect.com/science/article/pii/S0090429503013013>
24. Zhang XW, Liu ZH, Hu XW, *et al.* Androgen replacement therapy improves psychological distress and health-related quality of life in late onset hypogonadism patients in Chinese population. *Chin Med J* 2012;125:3806–10.
25. Dias S, Sutton A, Welton N, *et al.* NICE DSU Technical support document 2: Generalised linear modelling framework for pairwise and network meta-analysis of randomised controlled trials. www.nicesdsu.org.uk/TSD6_Software.final.08.05.12.pdf
26. Spiegelhalter D, Thomas A, Best N, *et al.* WinBUGS User Manual. 2003 www.mrc-bsu.cam.ac.uk/wp-content/uploads/manual14.pdf
27. Dias S, Welton NJ, Sutton AJ, *et al.* Evidence synthesis for decision making 4: inconsistency in networks of evidence based on randomized controlled trials. *Med Decis Mak* 2013;33:641–56.
28. Cherrier MM, Anderson K, Shofer J, *et al.* Testosterone treatment of men with mild cognitive impairment and low testosterone levels. *Am J Alzheimers Dis Other Dement* 2015;30:421–30.
29. Traish AM, Haider A, Haider KS, *et al.* Long-term Testosterone therapy improves cardiometabolic function and reduces risk of Cardiovascular Disease in men with Hypogonadism. 2017;1074248417(United States eISSN-1940-4034 (Electronic) PT-Journal Article LG-English DC-20170419 OVID MEDLINE UP 20170421).
30. Cheetham TC, An J, Jacobsen SJ, *et al.* Association of Testosterone replacement with cardiovascular outcomes among men with androgen deficiency. *JAMA Intern Med* 2017;177:491.
31. Layton JB, Meier CR, Sharpless JL, *et al.* Comparative safety of Testosterone dosage forms. *JAMA Intern Med* 2015;175:1187.
32. Blick G. Optimal diagnostic measures and thresholds for hypogonadism in men with HIV/AIDS: comparison between 2 transdermal Testosterone replacement therapy gels. *Postgrad Med* 2013;125:30–9.
33. Wang C, Cunningham G, Dobs A, *et al.* Long-term Testosterone gel (AndroGel) treatment maintains beneficial effects on sexual function and mood, lean and fat mass, and bone mineral density in hypogonadal men. *J Clin Endocrinol Metab* 2004;89:2085–98.
34. Dean JD, Carnegie C, Rodzvilla J, *et al.* Long-term effects of testim(r) 1% testosterone gel in hypogonadal men. *Rev Urol* 2005;7:87–94 <http://www.ncbi.nlm.nih.gov/pmc/articles/>
35. Yassin A, Salman M, Talib RA, *et al.* Is there a protective role of testosterone against high-grade prostate cancer? Incidence and severity of prostate cancer in 553 patients who underwent prostate biopsy: a prospective data register. *Aging Male* 2017:1.
36. Jung HJ, Shin HS. Effect of Testosterone replacement therapy on cognitive performance and depression in men with Testosterone deficiency syndrome. *World J Mens Health* 2016;34:194.
37. Debryne FM, Behre HM, Roehrborn CG, *et al.* Testosterone treatment is not associated with increased risk of prostate cancer or worsening of lower urinary tract symptoms: prostate health outcomes in the Registry of Hypogonadism in Men. *BJU Int* 2017;119:216–24.
38. Guay AT, Perez JB, Fitaihi WA, *et al.* Testosterone treatment in hypogonadal men: prostate-specific antigen level and risk of prostate cancer. *Endocr Pract* 2000;6:132–8.
39. Rhoden EL, Morgentaler A. Influence of demographic factors and biochemical characteristics on the prostate-specific antigen (PSA) response to Testosterone replacement therapy. *Int J Impot Res* 2006;18:201–5.
40. Shores MM, Smith NL, Forsberg CW, *et al.* Testosterone treatment and mortality in men with low testosterone levels. *J Clin Endocrinol Metab* 2012;97:2050–8.
41. Pastuszak AW, Gomez LP, Scovell JM, *et al.* Comparison of the effects of Testosterone gels, injections, and pellets on serum hormones, erythrocytosis, lipids, and prostate-specific antigen. *Sex Med* 2015;3:165–73.
42. Albert SG, Morley JE. Testosterone therapy, association with age, initiation and mode of therapy with cardiovascular events: a systematic review. *Clin Endocrinol* 2016;85:436–43.
43. Borst SE, Shuster JJ, Zou B, *et al.* Cardiovascular risks and elevation of serum DHT vary by route of testosterone administration: a systematic review and meta-analysis. *BMC Med* 2014;12:211.
44. Snyder PJ, Bhasin S, Cunningham GR, *et al.* Effects of Testosterone Treatment in Older Men. *N Engl J Med* 2016;374:611–24.
45. Millar AC, Lau AN, Tomlinson G, *et al.* Predicting low testosterone in aging men: a systematic review. *CMAJ* 2016;188:E321–30.
46. Standardizing Hormone Measurements. https://www.cdc.gov/labstandards/pdf/hs/HoSt_Brochure.pdf
47. Guyatt G. Problems with Bayesian random effects in network meta-analysis. *Cochrane Colloq* 2015. Vienna (October 3–7).