Testosterone Replacement Therapy and the Risk of Venous Thromboembolism: A Systematic Review and Meta-Analysis of Randomized Controlled Trials

Henok Tadesse Ayele PhD,1 Vanessa C. Brunetti MSc,2 Christel Renoux MD PhD,3 Vicky Tagalakis MD MSc,4 and Kristian B. Filion PhD 5

1 Department of Epidemiology, Biostatistics, Occupational Health, McGill University; and Center for Clinical Epidemiology, Lady Davis Institute, Jewish General Hospital, 3755 Cote Ste-Catherine, Montreal, Quebec, H3T 1E2, Canada; E-mail: henok.ayele@mail.mcgill.ca
2 Department of Epidemiology, Biostatistics, Occupational Health, McGill University; and Center for Clinical Epidemiology, Lady Davis Institute, Jewish General Hospital, 3755 Cote Ste-Catherine, Montreal, Quebec, H3T 1E2, Canada; E-mail: vanessa.brunetti@mail.mcgill.ca
3 Department of Epidemiology, Biostatistics, Occupational Health and Department of Neurology and Neurosurgery, McGill University; and Center for Clinical Epidemiology, Lady Davis Institute, Jewish General Hospital, 3755 Cote Ste-Catherine, Montreal, Quebec, H3T 1E2, McGill University, Montreal, Quebec, Canada. E-mail: christel.renoux@mcgill.ca
4 Division of General Internal Medicine, Department of Medicine, Jewish General Hospital, McGill University; Department of Epidemiology, Biostatistics, Occupational Health, McGill University; and Center for Clinical Epidemiology, Lady Davis Institute, Jewish General Hospital, 3755 Cote Ste-Catherine, Montreal, Quebec, H3T 1E2, Canada. E-mail: vicky.tagalakis@mcgill.ca
5 Departments of Medicine and of Epidemiology, Biostatistics, Occupational Health, McGill University; and Center for Clinical Epidemiology, Lady Davis Institute, Jewish General Hospital, 3755 Cote Ste-Catherine, Montreal, Quebec, H3T 1E2, Canada. E-mail: kristian.filion@mcgill.ca

The poster of this manuscript was presented at the 36th International Conference on Pharmacoepidemiology & Therapeutic Risk Management (ICPE 2020) on September 16th, 2020.

Body of Manuscript Word Count: 3793;
Abstract Word Count: 250

Corresponding author:
Kristian B. Filion, PhD
Associate Professor and William Dawson Scholar
Departments of Medicine and Epidemiology, Biostatistics, and Occupational Health
McGill University
3755 Cote Ste Catherine, Suite H410.1
Montreal, Quebec H3T 1E2 Canada

Phone: (514) 340-8222 x 28394
Fax: (514) 340-7564
Email: kristian.filion@mcgill.ca
ABSTRACT

Introduction: The cardiovascular safety of testosterone replacement therapy (TRT) is controversial. While several studies have investigated the association between TRT and the risk of arterial thrombosis, limited information is available regarding its risk of venous thromboembolism (VTE). We aimed to compare the risk of VTE in men randomized to TRT versus placebo or active-comparator in a systematic review.

Methods: We searched Medline, EMBASE, CINAHL, CENTRAL, and clinical trial registries to identify randomized controlled trials (RCTs) comparing TRT to placebo in men aged ≥18 years. We assessed study quality using the Cochrane Risk of Bias assessment tool and the overall quality of evidence using the Grading of Recommendations Assessment, Development, and Evaluation (GRADE) criteria. Data were pooled across RCTs using random-effects models.

Results: A total of 13 RCTs (n=5,050) were included in our meta-analysis. In all, 2,636 men were randomized to testosterone, and 2,414 men to placebo. Sample sizes ranged from 101 to 790 men, and TRT duration from 3 to 36 months. Five studies had a high risk of bias, largely driven by unclear randomization and outcome assessment. When data were pooled across RCTs, testosterone therapy was not associated with VTE compared with placebo (RR: 1.03, 95% CI: 0.49-2.14; I²: 0%; low-quality evidence). Similar estimates were obtained for deep vein thrombosis and pulmonary embolism outcomes.

Conclusions: Our systematic review suggests that TRT is not associated with an increased risk of VTE. However, estimates were accompanied by a wide 95% CIs, and a clinically important increased risk cannot be ruled out.

Keywords: Testosterone, Venous Thromboembolism, Deep Vein Thrombosis, Pulmonary Embolism, Systematic Review, Meta-analysis.
INTRODUCTION

Testosterone replacement therapy (TRT) is prescribed for the management of low testosterone levels when hypogonadism interferes with health or quality of life. TRT use has increased 10-fold in North America over the last 30 years, with a 12-fold increase in worldwide sales from $150 million to $1.8 billion during this period. However, the use of TRT declined in recent years following the publication of cardiovascular safety studies and a subsequent safety warning surrounding this issue from the US Food and Drug Administration (FDA). This warning was based on evidence that TRT can increase the risk of coronary artery dilation and potentially rupture an unstable plaque, resulting in myocardial infarction.

While much attention has focused on TRT and the risk of arterial thrombosis, the effect of TRT on the risk of venous thromboembolism (VTE) remains poorly understood. The US FDA issued a labeling change in the product information of all approved TRT products regarding the risk of VTE. However, this change was based on limited evidence, consisting only of a case series of patients with VTE. The lack of a reference group renders these data difficult to interpret. Moreover, the FDA positional statement on the benefits and safety of TRT in older men stressed the lack of conclusive data for this potential adverse drug effect.

There exists both a strong biological rationale and emerging evidence that supports a potential association between TRT and VTE. TRT is hypothesized to interact with previously undiagnosed thrombophilia-hypofibrinolysis, leading to VTE. In addition, several animal and human studies support the thrombogenic potential of TRT. A number of randomized controlled trials (RCT) reported VTE as an adverse event of TRT; however, these trials were individually underpowered to examine this safety endpoint, resulting in imprecise estimates. Three previous observational studies have examined this association, reporting heterogeneous findings. The association has also been examined in three previous systematic
reviews, which reported a disparate risk of VTE in men treated with TRT.\textsuperscript{12,26,27} However, these previous systematic reviews had several important limitations, including the exclusion of potentially relevant studies. Consequently, there remains a need to synthesize the totality of the evidence regarding the risk of VTE with TRT. The objective of this systematic review and meta-analysis of RCTs was therefore to determine if TRT, compared with placebo or an active comparator, is associated with the risk of VTE among men $\geq 18$ years.
METHODS

The protocol for this study was written following the Preferred Reporting Items for Systematic Review and Meta-Analysis Protocols (PRISMA-P) 2015 checklist (Appendix 1), and the study is reported according to the PRISMA guidelines (Appendix 2). The study involved aggregate, publicly available data and thus did not require research ethics approval.

Search Strategy

We systematically searched Medline (via Ovid; start date: 1946), Excerpta Medica database (EMBASE; start date: 1947), Cumulative Index of Nursing and Allied Health (CINAHL; start date: 1937), and the Cochrane Central Register of Controlled Trials (CENTRAL) from inception to December 18th, 2019 for RCTs of TRT that reported VTE as an outcome, with the search strategy tailored to each database. The search was restricted to RCTs in Medline using the 2005 version of the Centre for Reviews and Dissemination/Cochrane Highly Sensitive Search Strategy filter. In addition, ClinicalTrials.gov (www.clinicaltrials.gov) was searched for completed but still unpublished RCTs. Our database search was supplemented by a search of the World Health Organization trial registry and by a manual search of bibliographies of identified studies. Moreover, we used forward and backward citation searching with Scopus, and we searched the European Medical Agency (EMA) and FDA registries for unpublished trials on the safety of TRT. Finally, we reviewed the first 10 pages of Google Scholar for additional studies.

The search strategy was developed in collaboration with an experienced medical librarian. The construction of our search strategy was driven by two components of the Population, Intervention, Comparison, Outcome, and Time (PICOT) formatted research question: TRT (intervention) and VTE (outcome). VTE-related events were defined as anything
reported as such by the authors, that is, events reported as hepatic vein thrombosis, portal thrombosis, splanchnic venous thrombosis, pulmonary embolism or other venous thrombosis events. Ho et.al suggested that the use of a search strategy based on 2 elements of the PICOT question can identify a large number of relevant studies compared to 4 elements search. Since many trials focused their safety assessment on CVD rather than VTE, we included search terms for CVD to identify trials that reported VTE as a secondary outcome or a serious adverse event. To restrict our search to RCTs conducted in humans, we added the MEDLINE Cochrane RCT filter as the third concept in our search strategy. The Medline search is reproduced in full in Appendix 3.

**Eligibility and Study Selection**

RCTs that compared any commercially available TRT formulation to either placebo or an active comparator consisting of any androgenic medication (e.g., growth hormones, gonadotropin-releasing hormone antagonists), included a minimum of 50 participants in each group, and reported the risk of VTE by study arm were included in our systematic review. We included RCTs conducted in men aged 18 years or older. We excluded studies of women or children (<18 years) because they have different hormone levels than adult men. We further excluded uncontrolled trials, observational studies, previous reviews and meta-analyses, case reports and case series, letters to the editor, editorialis, and commentaries. We also excluded conference abstracts as the results of such studies are often not final, and abstracts contain insufficient information to adequately assess study quality. Finally, we excluded trials published in a language other than English or French.

Following the removal of duplicates, two independent reviewers (HTA and VCB) screened the titles and abstracts of publications identified by our electronic search. Any
publication deemed potentially relevant by either reviewer was carried forward to the full-text review. Discrepancies during the full-text review were resolved by consensus between two independent reviewers (HTA and VCB).

Data Extraction

We used an electronic form to extract data from the included studies (Appendix 4). Data extracted included: 1) study characteristics: authors’ names, journal, year of publication, sample size (total and by groups), intervention definition, duration of follow-up, duration of TRT, route of TRT administration, and outcome definition; 2) baseline demographic and clinical characteristics: age, indication for TRT, TRT dose, and pre-existing comorbidities; 3) Count data, effect estimates, and corresponding 95% confidence intervals (CI) for VTE as well as deep vein thrombosis (DVT) and pulmonary embolism (PE); and 4) study quality, as defined by the Cochrane risk of bias tool (described below). Two independent reviewers (HTA and VCB) extracted relevant information, and disagreements were resolved by consensus. The authors were contacted by email to identify missing information on VTE outcomes, particularly for studies identified in the grey literature. At least two email attempts were made for each corresponding author.

Quality Assessment

Two authors (HTA and VCB) independently assessed study quality using the Cochrane Risk of Bias assessment tool version 2.0,33 with discrepancies resolved by consensus. We also assessed funding sources as a potential source of bias; a trial was considered partially or completely supported by industry when declared by the authors or when at least one author was an employee of a pharmaceutical company, an approach used previously.34 We assessed the
overall quality of evidence using the Grading of Recommendations Assessment, Development, and Evaluation (GRADE) assessment criteria.\textsuperscript{35}

**Statistical Analyses**

The primary outcome measure was the cumulative incidence ratio for VTE among participants randomized to TRT versus placebo or an active comparator, which we estimated using reported count data. Trials with zero events were included using a 0.5 continuity correction. DerSimonian and Laird random-effects models with inverse variance weighting were used to estimate pooled relative risks (RRs) and corresponding 95\% CIs.\textsuperscript{36} We prespecified the use of a random-effects approach because of the degree of clinical and statistical heterogeneity anticipated across RCTs. In secondary analyses, we pooled risk estimates for DVT and PE separately. All analyses were conducted using R version 3.5.0,\textsuperscript{37} and we assessed the quality of evidence using the GRADEPro online platform.\textsuperscript{38}

The amount of between-study heterogeneity was assessed using the $I^2$ statistic and incorporated in the estimation of 95\% prediction intervals. The $I^2$ statistic describes the proportion of the overall variance that is due to between-study heterogeneity rather than due to chance. The prediction interval provides the range in which we can expect to find the estimated treatment effect of a subsequent RCT. We used a funnel plot and Egger’s test to evaluate the potential presence of small-study effects (typically publication bias).

**Sensitivity Analyses**

We prespecified 11 sensitivity analyses to examine the robustness of our results. First, we pooled estimates using the Hartung and Knap extension of the random-effects model, which is designed for meta-analyses in which few studies (<10) are included.\textsuperscript{39} Second, we used a fixed-effects approach with an inverse variance to determine how sensitive our results were to our
choice of modeling strategy. Third, to examine the impact of sparse data on our pooled estimates, we repeated our primary analysis using the Peto method,\textsuperscript{40} which provides the least biased estimates with good 95% CI coverage when modeling sparse outcomes. Fourth, we repeated the meta-analysis without continuity correction for treatment groups with zero events to examine the impact of using a continuity correction for RCTs with zero events. Fifth, we stratified results by funding sources to assess its potential impact on reported outcomes. Sixth, we restricted our analyses to RCTs randomized the study participants into TRT or placebo arm. Finally, while we had pre-specified sensitivity analyses across the subgroup of studies stratified by the characteristics of the participants included (such as age, total serum testosterone level at baseline), study quality, duration of follow-up, and TRT formulation, route of administration, dose, and duration, these analyses were not conducted due to sparse data.
RESULTS

Search Results

A total of 5,014 articles were identified from our search of six databases and two clinical trial registries (Supplementary Table 1). After removing duplicates, 4,019 publications underwent title/abstract screening, 418 underwent full-text review, and 13 were included in this meta-analysis (Figure 1).

Study Characteristics

Table 1 describes the characteristics of the included RCTs. The 13 RCTs included 5,050 men, 2,636 of whom were randomized to TRT and 2,414 to placebo or active comparator. The sample size of included RCTs ranged from 101 to 790 men. Nine of the trials were published in peer-reviewed scientific journals, while the remaining 4 trials were identified in ClinicalTrials.gov. Eleven trials were placebo-controlled, and two used an active comparator.41,42

The duration of TRT and follow-up ranged from 3-36 months. The characteristics of TRT varied across studies in terms of dose, duration, and route of administration. All studies were randomized by design; however, five studies had an unclear method of randomization. All trials but one (an open-label trial) blinded the study participants, researchers, or outcome assessors to treatment allocation. Most of the studies' primary outcome was total serum testosterone change from baseline. A total of 15 (0.6%) men in the TRT group and 12 (0.5%) in the placebo or active comparator group had VTE during follow-up. Most of the studies reported VTE as adverse events that were adjudicated by systematic assessment or a prespecified assessor. However, it is unclear whether the adjudication or systematic assessment included an objective assessment using diagnostic imaging. The VTE outcome classification was made following an intention-to-
treat (ITT) approach. Two studies were publicly funded, and 11 were fully or partially funded by industry.

**Risk of Bias**

Two of the studies had a low risk of bias, six studies had a moderate risk of bias, and five studies had a high risk of bias (Table 2 and Figures 2-3). The high risk of bias was largely driven by the unclear or suboptimal measurement of VTE or unclear randomization. Potentially missing outcome data and deviation from the intended intervention domains also contributed to the high risk of bias. The overall quality of evidence according to the GRADE criteria\textsuperscript{33} was low for the primary (VTE) and secondary outcomes (DVT and PE). We downgraded the quality of the evidence due to the imprecision of estimates, unclear randomization, and unclear outcome measurement (Supplementary Table 2).

**Meta-Analysis Results**

When data were pooled across RCTs, we did not find an increased risk of VTE among participants randomized to TRT compared to those randomized to placebo or active comparator (RR: 1.03, 95% CI: 0.49, 2.14; $I^2$: 0%; low-quality evidence) (Figure 4 and Supplementary Table 2). The 95% prediction interval for VTE ranged from 0.49 to 2.14. Similarly, randomization to TRT was not associated with the risk of DVT (RR: 1.14, 95% CI: 0.46 to 2.82; $I^2$: 0%; low-quality evidence) or PE (RR: 0.81, 95% CI: 0.29, 2.26; $I^2$: 0%; low-quality evidence) (Supplementary Table 2, Supplementary Figures 1-2). The 95% prediction intervals were 0.44 to 2.90 for DVT and 0.29 to 2.26 for PE. The symmetric shape of the funnel plot and Egger’s statistical test suggested that small-study effects were not present (Figure 5).

The sensitivity analyses examining different assumptions produced results that were consistent with those of our primary analysis (Supplementary Figures 3-7).
DISCUSSION

This systematic review and meta-analysis of RCTs was designed to determine the risk of VTE associated with the use of TRT among men. When study-specific estimates were pooled across 13 RCTs, we found no evidence of an increased risk of VTE with TRT. Similar results were obtained for DVT and PE and across several sensitivity analyses. We observed a high risk of bias and an overall low quality of evidence. While no evidence of an increased risk was observed, our estimates were accompanied by wide 95% CIs, and we cannot rule out a potential clinically important increased risk. Indeed, the width of prediction intervals with a Hartung and Knapp method indicates that the expected true effect of TRT on VTE could either be beneficial or detrimental in future RCTs. Consequently, while we have synthesized the totality of the evidence available from RCTs regarding this issue, there remains a need for additional large RCTs with a longer follow-up that specifically examine the risk VTE associated with TRT to further assess this potential drug safety issue.

A potential mechanism behind the purported VTE risk of TRT involves the presence of a ‘procoagulant’ state immediately after testosterone initiation.\textsuperscript{43} Moreover, the potential role of testosterone in the regulation of platelet TXA2 receptors' expression might contribute to the thrombogenicity of androgenic steroids.\textsuperscript{21} Older men with functional limitations have increased levels of factor VIII and D-dimer,\textsuperscript{44} and a high testosterone level might increase susceptibility to VTE.\textsuperscript{45}

Previous studies regarding the VTE risk of TRT have been inconclusive. Eight observational studies have examined this issue to date\textsuperscript{23-25,46-50}. Two of these studies\textsuperscript{24,25} found an increased risk of VTE in men treated with TRT, while the others did not\textsuperscript{23,46-50}. These studies also identified heterogeneity in treatment effects in some sub-populations, with VTE risk
attenuated among men with clinical hypogonadism\textsuperscript{24} and in men aged 65 years or over.\textsuperscript{25}

Previous systematic reviews of the VTE risk of TRT reported heterogeneous results.\textsuperscript{12,26,27,51} A letter by Xu et al. described an increased risk of VTE with TRT (odds ratio [OR]: 5.94, 95\% CI: 1.00-35.30) but this estimate was derived from only three RCTs.\textsuperscript{27} Two other systematic reviews did not include all relevant RCTs.\textsuperscript{12,26} Consistent with our findings, the 2020 evidence report for a Clinical Practice Guideline by the American College of Physicians presented a non-increased VTE risk with testosterone treatment in comparison to placebo.\textsuperscript{51} Given the recent increased use of TRT globally\textsuperscript{5} and the safety warnings emitted by the FDA,\textsuperscript{13} Health Canada,\textsuperscript{52} and Endocrine Society of Australia\textsuperscript{53} about the risk of VTE in testosterone use, the present study adds to the body of evidence on the safety of TRT with respect to VTE.

This study has several strengths. First, with a comprehensive and rigorous search strategy developed in consultation with an experienced medical librarian, relevant articles were identified from the published scientific and grey literature. Second, we conducted several sensitivity analyses that confirmed the robustness of our primary results. Third, we assessed the risk of bias with the Cochrane Risk of Bias 2.0 tool\textsuperscript{33} and the overall quality of evidence using GRADE\textsuperscript{35,38}. Finally, we found no evidence of small-study effects, suggesting that publication bias is unlikely to have had an important impact on our results.

This study also has several potential limitations. First, the response rate from the authors of unpublished clinical trials (grey literature) was low. Among the eight completed RCTs registered in Clinical.Trials.gov\footnote{Clinical.Trials.gov} without published results, only three authors responded to our queries. Second, VTE was rare, and treatment effects were therefore accompanied by wide 95\% CIs. In addition, we were unable to conduct all prespecified sensitivity analyses given the sparse data. However, this represents the totality of the available evidence regarding the VTE risk of
TRT. Third, we aimed to assess the risk of VTE, an outcome primarily reported as part of the adverse event reporting of these RCTs. We were limited to the outcome definition used in the included trials; the use of different outcome definitions across trials represents a potentially important source of heterogeneity. In addition, the potential under-reporting of VTE events\textsuperscript{36} might have biased the estimates toward the null. Although the proportion of lost-to-follow-up was low in these trials, it was not clear whether some of the men who were lost-to-follow-up experienced VTE events. Finally, it was not clear whether the included studies used pre-specified criteria to measure VTE-related events. Consequently, some outcome misclassification is possible.
CONCLUSIONS

Our systematic review and meta-analysis of RCTs suggests that TRT is not associated with an increased risk of VTE. However, estimates were accompanied by a wide 95% CIs, and a clinically important increased risk cannot be ruled out. Despite the limitations of the available evidence, the findings of this systematic review and meta-analysis contribute to the evolving body of evidence regarding the cardiovascular safety of TRT.

ACKNOWLEDGMENTS

We thank Andrea Quaiattini, a medical librarian at McGill University, for her help on the development of our search strategies and relevant articles identification.

SOURCES OF FUNDING

This study did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors. KBF is supported by a senior salary support award from the Fonds de recherche du Québec – Santé (FRQS; Quebec Foundation for Research - Health) and a William Dawson Scholar award from McGill University. HTA is supported by an FRQS Postdoctoral Training Award. VCB holds a Doctoral Training Award from the FRQS and a Drug Safety and Effectiveness Training (DSECT) Program Award funded by the Canadian Institutes of Health Research (CIHR). CR is the recipient of a salary award from the FRQS.

DISCLOSURES

The authors declare no competing interest.

AUTHORS’ CONTRIBUTIONS
HTA designed the study, conducted screening, extracted data, assessed quality, conducted the statistical analyses, and drafted the manuscript. VCB performed screening, data extraction, quality assessment, interpreted data, and critically reviewed the manuscript. KBF, CR, and VT contributed to study design and protocol development, interpreted data, and critically reviewed the manuscript for important intellectual content. KBF is the guarantor.
REFERENCES


Table 1. Characteristics of randomized controlled trials reporting the risk of venous thromboembolism with testosterone replacement therapy versus placebo or active comparator.

<table>
<thead>
<tr>
<th>Study population</th>
<th>Intervention</th>
<th>Outcome</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Author, Year</strong></td>
<td><strong>Sample size</strong></td>
<td><strong>Age (Mean)</strong></td>
</tr>
<tr>
<td>Behre et al., 2012</td>
<td>362</td>
<td>65.6*</td>
</tr>
<tr>
<td>Bhasin et al., 2017</td>
<td>306</td>
<td>67.6</td>
</tr>
<tr>
<td>Brock et al. (3 months), 2016</td>
<td>596</td>
<td>55.3</td>
</tr>
<tr>
<td>Brock et al. (9 months), 2016</td>
<td>596</td>
<td>54.6</td>
</tr>
<tr>
<td>Eli Lilly and company, 2015</td>
<td>715</td>
<td>55.3</td>
</tr>
<tr>
<td>Gluud et al., 1987</td>
<td>126</td>
<td>NR</td>
</tr>
<tr>
<td>Gluud et al., 1986</td>
<td>221</td>
<td>53.0</td>
</tr>
<tr>
<td>Study</td>
<td>n</td>
<td>Age</td>
</tr>
<tr>
<td>-------------------------------</td>
<td>-----</td>
<td>-----</td>
</tr>
<tr>
<td>Indevus Pharmaceuticals, 2017</td>
<td>524</td>
<td>54.4</td>
</tr>
<tr>
<td>Sinclair et al., 2016</td>
<td>101</td>
<td>55.0</td>
</tr>
<tr>
<td>Snyder et al., 2016</td>
<td>790</td>
<td>72.2</td>
</tr>
<tr>
<td>Srinivas-Shanker et al., 2010</td>
<td>262</td>
<td>73.8</td>
</tr>
<tr>
<td>Swerdloff et al., 2018</td>
<td>222</td>
<td>52.0</td>
</tr>
<tr>
<td>Tan et al., 2013</td>
<td>114</td>
<td>NR</td>
</tr>
</tbody>
</table>

*Testosterone=69.1±6.6 and Placebo=62.1±6.3; NR: Not reported; Patients and outcome assessors were blinded; Unblinding was only performed in the event of a serious adverse event; CAIMT: Carotid artery intima-media thickness; DVT: Deep vein thrombosis; FUP: Follow-Up; ITT: Intention-to-Treat; LTF: Lost-to-Followup; m: Months; PE: Pulmonary embolism; QoL: Quality of life; RoA: Route of Administration; VT: Venous thrombosis; VTE: Venous thromboembolism; TT: Testosterone; Rx: Treatment;
Table 2. The overall and domain-specific risk of bias among randomized controlled trials reporting the risk of venous thromboembolism with testosterone replacement therapy versus placebo or active comparator.

<table>
<thead>
<tr>
<th></th>
<th>Randomization process</th>
<th>Deviations from intended interventions</th>
<th>Missing outcome data</th>
<th>Measurement of the outcome</th>
<th>Selection of the reported result</th>
<th>Overall Bias</th>
</tr>
</thead>
<tbody>
<tr>
<td>Low risk</td>
<td>43.0</td>
<td>53.7</td>
<td>89.9</td>
<td>34.1</td>
<td>64.4</td>
<td>14.4</td>
</tr>
<tr>
<td>Some concerns</td>
<td>30.3</td>
<td>39.4</td>
<td>0</td>
<td>29.2</td>
<td>35.6</td>
<td>41.9</td>
</tr>
<tr>
<td>High risk</td>
<td>26.7</td>
<td>6.9</td>
<td>10.1</td>
<td>36.8</td>
<td>0</td>
<td>43.7</td>
</tr>
</tbody>
</table>

Data are reported as percentages.
* Assignment to intervention (the 'intention-to-treat' effect)
FIGURE LEGENDS

Figure 1. Flow diagram describing the selection of randomized controlled trials reporting the risk of venous thromboembolism with testosterone replacement therapy versus placebo or active comparator.

Figure 2. Percentage of the overall and domain-specific risk of bias of randomized controlled trials reporting the risk of venous thromboembolism with testosterone replacement therapy versus placebo or active comparator.

Figure 3. The risk of bias in randomized controlled trials reporting the risk of venous thromboembolism with testosterone replacement therapy versus placebo or active comparator.

Figure 4. Pooled relative risk of venous thromboembolism in men using testosterone replacement therapy versus placebo or active comparator.

Figure 5. Enhanced funnel with the risk ratios and standard error from randomized controlled trials reporting the risk of venous thromboembolism with testosterone replacement therapy versus placebo or active comparator. Regression test for funnel plot asymmetry with mixed-effects meta-regression model: \( z = 0.1207, p = 0.9039 \).
Figure 1.

Records identified through database searching (n = 4629)

Additional records identified through other sources (n = 385)

Records after duplicates removed (n = 4019)

Duplicates excluded (n = 995)

Records excluded (n = 3601)
- Animal studies (n=193)
- Children [<18 years] (n=31)
- Observational studies (n=226)
- Reviews (n=1066)
- Conference abstracts (n=703)
- Non-testosterone interventions (n=1007)
- Editorial & commentary (n=108)
- Non-randomized or quasi-experimental studies (n=141)
- Languages other than English or French (n=113)

Title & Abstract screened (n = 4019)

Full-text articles assessed for eligibility (n = 418)

Studies included in quantitative synthesis (meta-analysis) (n = 13)

Full-text articles excluded, with reasons (n = 405)
- No VTE outcome (n=211)
- Sample size < 100 (n=106)
- Women (n=72)
- No comparator (n=11)
- No response from the authors of grey literature (n=5)
Figure 2: An intention-to-treat

- Overall Bias
- Selection of the reported result
- Measurement of the outcome
- Missing outcome data
- Deviations from intended interventions
- Randomization process

Legend:
- Low risk
- Some concerns
- High risk
<table>
<thead>
<tr>
<th>Study</th>
<th>Weight</th>
<th>Randomization process</th>
<th>Deviations from intended interventions</th>
<th>Missing outcome data</th>
<th>Measurement of the outcome of the reported result</th>
<th>Overall Bias</th>
</tr>
</thead>
<tbody>
<tr>
<td>Behre <em>et al.</em>, 2012</td>
<td>5.27</td>
<td>+</td>
<td>+</td>
<td>−</td>
<td>+</td>
<td>−</td>
</tr>
<tr>
<td>Snyder <em>et al.</em>, 2016</td>
<td>31.46</td>
<td>−</td>
<td>−</td>
<td>+</td>
<td>−</td>
<td>?</td>
</tr>
<tr>
<td>Sirnivas-Shankar <em>et al.</em>, 2010</td>
<td>5.28</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>−</td>
</tr>
<tr>
<td>Swedloff <em>et al.</em>, 2018</td>
<td>5.3</td>
<td>−</td>
<td>−</td>
<td>+</td>
<td>−</td>
<td>−</td>
</tr>
<tr>
<td>Tan <em>et al.</em>, 2013</td>
<td>3.53</td>
<td>?</td>
<td>+</td>
<td>−</td>
<td>−</td>
<td>−</td>
</tr>
<tr>
<td>Brock <em>et al.</em> (9 months), 2016</td>
<td>5.87</td>
<td>?</td>
<td>?</td>
<td>+</td>
<td>?</td>
<td>!</td>
</tr>
<tr>
<td>Brock <em>et al.</em> (3 months), 2016</td>
<td>5.85</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>?</td>
<td>−</td>
</tr>
<tr>
<td>Eli Lilly, 2015</td>
<td>5.85</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>Gluud <em>et al.</em>, 1987</td>
<td>5.31</td>
<td>?</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>!</td>
</tr>
<tr>
<td>Gluud <em>et al.</em>, 1986</td>
<td>6.18</td>
<td>?</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>Bhasin <em>et al.</em>, 2017</td>
<td>5.28</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>?</td>
<td>!</td>
</tr>
</tbody>
</table>

Figure 3.
<table>
<thead>
<tr>
<th>Authors and Year</th>
<th>Testosterone VTE+</th>
<th>Testosterone VTE-</th>
<th>Placebo VTE+</th>
<th>Placebo VTE-</th>
<th>Weight (%)</th>
<th>Risk Ratio [95% CI]</th>
</tr>
</thead>
<tbody>
<tr>
<td>Behre et al., 2012</td>
<td>1</td>
<td>182</td>
<td>0</td>
<td>179</td>
<td>5.27</td>
<td>2.93 [0.12, 71.57]</td>
</tr>
<tr>
<td>Bhasin et al., 2017</td>
<td>0</td>
<td>155</td>
<td>1</td>
<td>150</td>
<td>5.28</td>
<td>0.32 [0.01, 7.91]</td>
</tr>
<tr>
<td>Brock et al. (3 months), 2016</td>
<td>0</td>
<td>358</td>
<td>2</td>
<td>355</td>
<td>5.85</td>
<td>0.20 [0.01, 4.14]</td>
</tr>
<tr>
<td>Brock et al. (9 months), 2016</td>
<td>2</td>
<td>300</td>
<td>0</td>
<td>292</td>
<td>5.85</td>
<td>4.83 [0.23, 100.28]</td>
</tr>
<tr>
<td>Eli Lilly and company, 2015</td>
<td>0</td>
<td>358</td>
<td>2</td>
<td>355</td>
<td>5.85</td>
<td>0.20 [0.01, 4.14]</td>
</tr>
<tr>
<td>Gluud et al., 1986</td>
<td>3</td>
<td>131</td>
<td>0</td>
<td>87</td>
<td>6.18</td>
<td>4.66 [0.24, 87.27]</td>
</tr>
<tr>
<td>Gluud et al., 1987</td>
<td>1</td>
<td>75</td>
<td>0</td>
<td>50</td>
<td>5.31</td>
<td>1.99 [0.08, 47.83]</td>
</tr>
<tr>
<td>Indevus Pharmace., 2017</td>
<td>1</td>
<td>271</td>
<td>0</td>
<td>252</td>
<td>5.26</td>
<td>2.78 [0.11, 67.94]</td>
</tr>
<tr>
<td>Sinclair et al., 2016</td>
<td>2</td>
<td>48</td>
<td>1</td>
<td>50</td>
<td>9.59</td>
<td>2.04 [0.19, 21.79]</td>
</tr>
<tr>
<td>Snyder et al., 2016</td>
<td>4</td>
<td>390</td>
<td>5</td>
<td>389</td>
<td>31.46</td>
<td>0.80 [0.22, 2.96]</td>
</tr>
<tr>
<td>Srinivas-Shanker et al., 2010</td>
<td>1</td>
<td>131</td>
<td>0</td>
<td>130</td>
<td>5.28</td>
<td>2.95 [0.12, 71.88]</td>
</tr>
<tr>
<td>Swerdloff et al., 2018</td>
<td>0</td>
<td>166</td>
<td>1</td>
<td>55</td>
<td>5.3</td>
<td>0.11 [0.00, 2.75]</td>
</tr>
<tr>
<td>Tan et al., 2013</td>
<td>0</td>
<td>56</td>
<td>0</td>
<td>58</td>
<td>3.53</td>
<td>1.04 [0.02, 51.29]</td>
</tr>
<tr>
<td>Total</td>
<td>15</td>
<td>2621</td>
<td>12</td>
<td>2402</td>
<td>100</td>
<td></td>
</tr>
</tbody>
</table>

Random-effects model (Q = 8.40, df = 12, p = 0.75; $i^2 = 0.0\%$)

Figure 4.