The Benefits and Harms of Systemic Testosterone Therapy in Postmenopausal Women With Normal Adrenal Function: A Systematic Review and Meta-analysis

Tarig Elraiyah, Mohamad Bassam Sonbol, Zhen Wang, Tagwa Khairalseed, Noor Asi, Chaitanya Undavalli, Mohammad Nabhan, Belal Firwana, Osama Altayar, Larry Prokop, Victor M. Montori, and Mohammad Hassan Murad

Knowledge and Evaluation Research Unit (T.E., M.B.S., Z.W., T.K., N.A., C.U., M.N., B.F., O.A., V.M.M., M.H.M.), and Center for the Science of Healthcare Delivery (T.E., Z.W., N.A., M.H.M.), Mayo Clinic, Rochester, Minnesota 55905; Internal Medicine Department (M.B.S.), Georgia Regents University, Augusta, Georgia 30901; and Mayo Clinic Libraries (L.P.), Division of Endocrinology, Diabetes, Metabolism and Nutrition (V.M.M.), and Division of Preventive, Occupational, and Aerospace Medicine (M.H.M.), Mayo Clinic, Rochester, Minnesota 55905

Context: The use of T has been suggested to improve women’s health during the postmenopausal period.

Objective: We conducted a systematic review and meta-analysis of randomized trials to summarize the best available evidence regarding the benefits and harms of systemic T in postmenopausal women with normal adrenal function.

Methods: A comprehensive search of MEDLINE, EMBASE, PsycInfo, Cochrane Central Register of Controlled Trials, Cochrane Database of Systematic Reviews, EBSCO CINAHL, and Scopus was conducted through January 2014. We conducted study selection, data extraction, and appraisal in duplicate. Random-effects meta-analysis was used to pool results.

Results: We identified 35 randomized trials (n = 5053) at a moderate risk of bias. T use was associated with statistically significant improvement in various domains of sexual function and personal distress in postmenopausal women, although the majority of the trials did not have specific or contemporary diagnostic criteria for androgen deficiency in women. T use was also associated with a reduction in total cholesterol, triglyceride, and high-density lipoprotein and an increase in low-density lipoprotein and in the incidence of acne and hirsutism. No significant effect was noted on anthropometric measures and bone density. Long-term safety data were sparse, and the quality of such evidence was low.

Conclusion: Despite the improvement in sexual function associated with T use in postmenopausal women, long-term safety data are lacking. (*J Clin Endocrinol Metab* 99: 3543–3550, 2014)

Testosterone in women exerts its effect through androgen receptors located in several organs of the body, including breast, brain, ovaries, bone, muscle, fat, liver, and skin (1, 2). T plays an important role in female sexual function, is considered a major driver for sexual desire, and modulates the response to sexual stimuli (3). Androgens also affect specific receptors in the brain to reduce the level of depression and anxiety and improve the sense of well-being (4, 5). Furthermore, T in women is known to have multiple anabolic effects on muscles, body fat, and bone mineral content (1).
During the years of menopause, a decline in androgen level takes place due to age-related reduction in secretion by ovaries and adrenal glands (6, 7). In addition, bilaterally oophorectomized women experience a sharp drop in T level of up to 50% after surgery (8). Whether this decline is associated with natural or surgical menopause, it manifests itself clinically by a decrease in sexual function, an increase in personal distress and anxiety levels (9), and a decline in bone density with increased tendency for osteoporotic fractures (10).

In both surgically and naturally menopausal women, T therapy, alone or combined with hormonal replacement therapy (HRT), was suggested to be associated with improvement in sexual function, energy, and quality of life (11–13); a reduction in personal distress and mood (14, 15); and an increase in bone mineral density (BMD) (16, 17).

However, safety concerns regarding the use of T in women need to be considered. Some of the most common side effects that are associated with the use of T are the androgenic side effects, especially hirsutism and acne (7). In addition, it has been found to be associated with changes in lipid profile as reported by many studies (7, 15, 18).

Given this, The Endocrine Society commissioned a task force to develop clinical practice guidelines for the use of androgen in postmenopausal women. Subsequently, the task force requested a systematic review and meta-analysis summarizing the best available evidence on this topic. Utilizing their input, we conducted this review to bring the knowledge base up to date in regard to beneficial effects as well as the risks that are associated with T use.

**Materials and Methods**

A panel of experts from The Endocrine Society provided the content expertise and developed a priori protocol for this systematic review and meta-analysis. The population criteria, description of interventions, comparisons, and the outcomes of interest were predefined in the protocol. This report is consistent with recommendations from the Preferred Reporting Items for Systematic Reviews and Meta-analyses (PRISMA) statement (19).

**Eligibility criteria**

We included randomized controlled trials (RCTs) that enrolled postmenopausal women with normal adrenal function who were assigned to T therapy with or without estrogen and compared to either estrogens alone or placebo. The outcomes included sexual function, psychological symptoms, bone health, body measurements, lipid profile, and incidence of hirsutism and acne.

**Literature search**

The search strategy was developed by a reference librarian (L.P.), who received input from the study’s principal investigator (M.H.M.) in compliance with the protocol and the predefined criteria. Several databases were searched, including Medline In-Process & Other Non-Indexed Citations, MEDLINE, EMBASE, PsycInfo, Cochrane Central Register of Controlled Trials, EBSCO CINAHL, and Scopus. These databases were searched from July 2008 to January 2014. Eligible studies prior to July 2008 were identified from previous Endocrine Society guidelines (20), from a previously published systematic review addressing the same question by the Cochrane collaboration (21), and from a review about T without HRT (7). The first author (T.E.) searched the reference sections of the included studies to identify other eligible studies that could have been missed by the electronic search. Experts from The Endocrine Society were contacted to verify the final list and identify any missing studies. The full search strategy is provided in the Supplemental Data.

**Study selection**

Relevant studies were screened and selected using DistillerSR (Evidence Partners Inc), which is a web-based software specifically designed to conduct systematic reviews. References were screened in duplicate, and conflicts were included for full-text review. Full-text screening was also conducted in duplicate, and conflicts were resolved by consensus. We calculated the chance-adjusted inter-reviewer agreement using kappa statistics (22).

**Data extraction**

Reviewers, working independently and in pairs, used a piloted and comprehensive data form to extract the following information from each study: patient demographic data, eligibility criteria, and details about intervention and the outcomes of interest. The outcomes extracted included: 1) sexual function, namely the number of satisfying sexual episodes, frequency of sexual activity, libido, orgasm, arousal, pleasure or enjoyment of sex, sexual responsiveness, sexual self-image, sexual or relationship satisfaction, and sexual concern; 2) psychological symptoms including personal distress, anxiety or fear, and depressed mood; 3) body measurements including body weight and body mass index (BMI); 4) bone health as measured by spine, hip, and total body BMD; 5) lipid profile data, including total cholesterol, low-density lipoprotein (LDL) cholesterol, high-density lipoprotein (HDL) cholesterol, and triglycerides; and 6) androgenic side effects, including acne and hirsutism.

**Author contact**

When the data were not available in the published articles or a further clarification was needed, we contacted the authors by e-mail. If there was no response, a repeat e-mail was sent in 2 weeks.

**Risk of bias assessment**

To assess the quality of the included study, we used Cochrane Collaboration’s Risk of Bias assessment tool (23). Two reviewers independently assessed the quality of randomization methods, allocation concealment, baseline imbalances, whether blinding was done adequately and who was blinded, monitoring of adherence to the medication use, and whether analysis was done by intention-to-treat. The quality of evidence was evaluated using
the Grading of Recommendations Assessment, Development and Evaluation (GRADE) methods (24).

**Statistical analysis**

Meta-analysis was done using the random effect described by DerSimonian and Laird (25). We pooled relative risk (RR) and 95% confidence interval (CI) for dichotomous outcomes. For continuous outcomes, we calculated the weighted mean difference (WMD) along with the 95% CI. When the studies used different scales to assess the outcomes, we calculated standardized mean difference (SMD). Heterogeneity was evaluated by the $I^2$ statistics. Whenever appropriate (more than 10–20 studies and low between-study heterogeneity), we assessed publication bias using the Egger regression asymmetry test and visual inspection of funnel plots (26, 27). To avoid the potential false-positive differences due to multiple outcomes and multiple testing, we adopted the false discovery rate-controlled procedures proposed by Benjamini and Hochberg (28) and set the $P$ value for significance at .015 instead of the usual .05.

All analyses were conducted using STATA, version 12.1 (StataCorp LP).  

**Results**

**Search results and study description**

A total of 35 trials were included (Figure 1). Average weighted kappa for study selection was 0.85. The RCTs enrolled a total of 5053 postmenopausal women who received T with or without HRT. Patient follow-up ranged from 1.5 to 24 months (median, 6 mo). One author was contacted to verify data about lipids. The characteristics of the included studies and the interventions are summarized in Supplemental Table 1.

**Methodological quality**

The included RCTs were of low to moderate quality overall. Randomization method was adequate in 35%, with adequate allocation concealment reported in 75% of all; more than 90% masked patients and study investigators. One-fifth of the studies reported treatment adherence

---

**Subgroup analysis**

Predefined subgroups were established to explore explanations for inconsistency between studies. These subgroups were defined based on the type of menopause (natural vs surgical), type of T (T vs methyltestosterone), route of T administration (oral vs non-oral), whether HRT was given with T (HRT-containing vs HRT-free regimens), HRT routes of administration (oral vs non-oral), and different estrogen preparations (conjugated equine estrogen, esterified estrogen, and estradiol). Supplemental Tables 3 and 4 show the results of subgroup analyses.
assessments, and only 25% analyzed participants as randomized (intention-to-treat analysis). Figure 2 depicts the overall quality, and the detailed assessment is provided in Supplemental Table 2.

**Meta-analysis**

The meta-analytic estimates of the effect sizes for benefit outcomes are presented in Table 1, and those for the adverse effect are in Table 2.

**Sexual function outcomes**

Compared with the T-free regimen (TFR), the T-containing regimen (TCR) was associated with statistically significant improvement in frequency of sexual activity (SMD, 0.29; 95% CI, 0.16 to 0.41; $I^2 = 41%$; $n = 9$), the number of satisfying sexual episodes (WMD, 1.20; 95% CI, 0.88 to 1.51; $I^2 = 0%$; $n = 8$), and interest in sex (SMD, 0.35; 95% CI, 0.19 to 0.52; $I^2 = 79%$; $n = 17$). The quality of encounters was also found to be improved by TCR, including improvements in orgasm (SMD, 0.20; 95% CI, 0.09 to 0.31; $I^2 = 52%$; $n = 15$), arousal (SMD, 0.25; 95% CI, 0.09 to 0.40; $I^2 = 77%$; $n = 12$), and enjoyment of sex (SMD, 0.31; 95% CI, 0.12 to 0.51; $I^2 = 79%$; $n = 11$). Concern about sexual function impairment in postmenopausal women was also found to be significantly less when these women were treated with TCRs (SMD, 0.24; 95% CI, 0.03 to 0.44; $I^2 = 85%$; $n = 8$). Sexual self-image (SMD, 0.20; 95% CI, 0.05 to 0.35; $I^2 = 70%$; $n = 8$), sexual responsiveness (SMD, 0.29; 95% CI, 0.19 to 0.43; $I^2 = 0%$; $n = 8$), sexual or relationship satisfaction (SMD, 0.83; 95% CI, 0.28 to 1.37; $I^2 = 87%$; $n = 7$), and sexual concern (SMD, 0.24; 95% CI, 0.03 to 0.44; $I^2 = 85%$; $n = 8$).

### Table 1. Meta-analysis of Sexual, Psychological, and Body Composition Outcomes

<table>
<thead>
<tr>
<th>Outcome</th>
<th>No. of Studies</th>
<th>No. of Patients</th>
<th>SMD/WMD</th>
<th>LL 95% CI</th>
<th>UL 95% CI</th>
<th>P Value</th>
<th>$I^2$, %</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Sexual outcomes</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>No. of satisfying sexual episodes$^b$</td>
<td>8</td>
<td>2684</td>
<td>SMD;1.20</td>
<td>0.88</td>
<td>1.51</td>
<td>&lt;.001</td>
<td>0</td>
</tr>
<tr>
<td>Frequency of sexual activity</td>
<td>9</td>
<td>2047</td>
<td>SMD;0.29</td>
<td>0.16</td>
<td>0.41</td>
<td>&lt;.001</td>
<td>41</td>
</tr>
<tr>
<td>Libido/desire/interest in sex</td>
<td>17</td>
<td>3288</td>
<td>SMD;0.35</td>
<td>0.19</td>
<td>0.52</td>
<td>&lt;.001</td>
<td>79</td>
</tr>
<tr>
<td>Orgasm</td>
<td>15</td>
<td>3019</td>
<td>SMD;0.20</td>
<td>0.09</td>
<td>0.31</td>
<td>&lt;.001</td>
<td>52</td>
</tr>
<tr>
<td>Arousal</td>
<td>12</td>
<td>2901</td>
<td>SMD;0.25</td>
<td>0.09</td>
<td>0.41</td>
<td>.002</td>
<td>77</td>
</tr>
<tr>
<td>Pleasure/enjoyment of sex</td>
<td>11</td>
<td>2596</td>
<td>SMD;0.31</td>
<td>0.12</td>
<td>0.51</td>
<td>.002</td>
<td>79</td>
</tr>
<tr>
<td>Sexual responsiveness</td>
<td>12</td>
<td>3027</td>
<td>SMD;0.29</td>
<td>0.17</td>
<td>0.41</td>
<td>&lt;.001</td>
<td>57</td>
</tr>
<tr>
<td>Sexual self-image</td>
<td>8</td>
<td>2955</td>
<td>SMD;0.20</td>
<td>0.05</td>
<td>0.35</td>
<td>.008</td>
<td>70</td>
</tr>
<tr>
<td>Sexual or relationship satisfaction</td>
<td>7</td>
<td>303</td>
<td>SMD;0.83</td>
<td>0.28</td>
<td>1.37</td>
<td>.003</td>
<td>87</td>
</tr>
<tr>
<td>Sexual concern</td>
<td>8</td>
<td>2688</td>
<td>SMD;0.24</td>
<td>0.03</td>
<td>0.44</td>
<td>.023</td>
<td>85</td>
</tr>
<tr>
<td><strong>Psychological outcomes</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Personal distress</td>
<td>9</td>
<td>2669</td>
<td>SMD;−0.34</td>
<td>−0.43</td>
<td>−0.24</td>
<td>&lt;.001</td>
<td>30</td>
</tr>
<tr>
<td>Anxiety or fear</td>
<td>4</td>
<td>249</td>
<td>SMD;−0.03</td>
<td>−0.23</td>
<td>0.16</td>
<td>.739</td>
<td>0</td>
</tr>
<tr>
<td>Depressed mood</td>
<td>5</td>
<td>282</td>
<td>SMD;0.13</td>
<td>−0.06</td>
<td>0.32</td>
<td>.18</td>
<td>0</td>
</tr>
<tr>
<td><strong>Bone health outcomes</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Spine BMD, g/cm²</td>
<td>4</td>
<td>177</td>
<td>WMD;−0.00</td>
<td>−0.02</td>
<td>0.01</td>
<td>.636</td>
<td>0</td>
</tr>
<tr>
<td>Hip BMD, g/cm²</td>
<td>3</td>
<td>134</td>
<td>WMD;0.01</td>
<td>−0.00</td>
<td>0.03</td>
<td>.094</td>
<td>0</td>
</tr>
<tr>
<td>Total body BMD, g/cm²</td>
<td>3</td>
<td>120</td>
<td>WMD;0.00</td>
<td>−0.03</td>
<td>0.04</td>
<td>.922</td>
<td>0</td>
</tr>
<tr>
<td><strong>Body composition outcomes</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Body weight, kg</td>
<td>4</td>
<td>149</td>
<td>WMD;1.17</td>
<td>−0.20</td>
<td>2.54</td>
<td>.095</td>
<td>0</td>
</tr>
<tr>
<td>BMI, kg/m²</td>
<td>5</td>
<td>210</td>
<td>WMD;0.25</td>
<td>−0.81</td>
<td>1.30</td>
<td>.647</td>
<td>0</td>
</tr>
</tbody>
</table>

**Abbreviations:** LL 95% CI, lower limit of 95% CI; UL 95% CI, upper limit of 95% CI.

$^a$ The difference in means is calculated as TCR minus TFR.

$^b$ Results are presented in WMD.

### Table 2. Meta-analysis of Lipids and Adverse Effect Outcomes

<table>
<thead>
<tr>
<th>Outcome</th>
<th>No. of Studies</th>
<th>No. of Patients</th>
<th>WMD/RR$^a$</th>
<th>LL 95% CI</th>
<th>UL 95% CI</th>
<th>P Value</th>
<th>$I^2$, %</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Lipid profile</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total cholesterol, mg/dL</td>
<td>24</td>
<td>2817</td>
<td>WMD;−8.51</td>
<td>−13.30</td>
<td>−3.71</td>
<td>.001</td>
<td>73</td>
</tr>
<tr>
<td>Triglyceride, mg/dL</td>
<td>23</td>
<td>2701</td>
<td>WMD;−15.29</td>
<td>−22.90</td>
<td>−7.67</td>
<td>&lt;.001</td>
<td>68</td>
</tr>
<tr>
<td>LDL, mg/dL</td>
<td>21</td>
<td>2697</td>
<td>WMD;2.62</td>
<td>0.24</td>
<td>5.01</td>
<td>.031</td>
<td>19</td>
</tr>
<tr>
<td>HDL, mg/dL</td>
<td>22</td>
<td>2719</td>
<td>WMD;−9.11</td>
<td>−12.89</td>
<td>−5.32</td>
<td>&lt;.001</td>
<td>94</td>
</tr>
<tr>
<td>Acne and hirsutism</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Acne</td>
<td>10</td>
<td>3161</td>
<td>RR;1.62</td>
<td>1.28</td>
<td>2.06</td>
<td>&lt;.001</td>
<td>0</td>
</tr>
<tr>
<td>Hirsutism</td>
<td>10</td>
<td>2983</td>
<td>RR;1.45</td>
<td>1.09</td>
<td>1.93</td>
<td>.011</td>
<td>0</td>
</tr>
</tbody>
</table>

**Abbreviations:** LL 95% CI, lower limit of 95% CI; UL 95% CI, upper limit of 95% CI.

$^a$ The difference in means is calculated as TCR minus TFR.
0.17 to 0.41; \( I^2 = 57\%; n = 12 \) ), and satisfaction with sex and relationship (SMD, 0.83; 95% CI, 0.28 to 1.37; \( I^2 = 87\%; n = 7 \)) were also found to be significantly better with TCRs. The quality of evidence for these outcomes ranged from low to moderate, due to high heterogeneity and low-to-moderate risk of bias.

**Psychological outcomes**

There was statistically significant improvement in personal distress with TCRs (SMD, \(-0.34\); 95% CI, \(-0.43\) to \(-0.24\); \( I^2 = 30\%; n = 9 \)), but not in anxiety (SMD, \(-0.03\); 95% CI, \(-0.23\) to 0.16; \( I^2 = 0\%; n = 4 \)) or depression (SMD, 0.13; 95% CI, \(-0.06\) to 0.32; \( I^2 = 0\%; n = 5 \)). The quality of evidence is low, due to increased risk of bias and imprecision.

**Bone health**

We found no trials that reported fractures as an outcome. Only four RCTs assessed its effect on BMD, which is considered a surrogate for fracture risk. No significant improvement was detected in BMD, measured in grams per square centimeter, in any of the sites, including spine (WMD, \(-0.00\); 95% CI, \(-0.02\) to 0.01; \( I^2 = 0\%; n = 4 \)), hip (WMD, 0.01; 95% CI, \(-0.00\) to 0.03; \( I^2 = 0\%; n = 3 \)), and total body BMD (WMD, 0.00; 95% CI, \(-0.03\) to 0.04; \( I^2 = 3\%; n = 3 \)). Quality of evidence is very low, rated down due to indirectness, imprecision, and moderate risk of bias.

**Body composition**

No change was detected in body weight (WMD, 1.17 kg; 95% CI, \(-0.20\) to 2.54; \( I^2 = 0\%; n = 4 \)) or BMI (WMD, 0.25 kg/m²; 95% CI, \(-0.81\) to 1.30; \( I^2 = 0\%; n = 5 \)) when TCR was compared to TFR in four and five studies, respectively. Quality of evidence is low, rated down due to imprecision and the risk of bias.

**Lipid profile**

Based on a meta-analysis of 24 and 23 trials, respectively, significant decreases in total cholesterol (\(-8.51\) mg/dL; 95% CI, \(-13.30\) to \(-3.71\); \( I^2 = 73\%; n = 24 \)) and triglyceride (\(-15.29\) mg/dL; 95% CI, \(-22.90\) to \(-7.67\); \( I^2 = 68\%; n = 23 \)) were noted with TCR. The quality of evidence for both of these outcomes is low to moderate, downgraded due to high inconsistency and moderate risk of bias in the included studies. On the other hand, T was found to be associated with a significant increase in LDL (2.62 mg/dL; 95% CI, 0.24 to 5.01; \( I^2 = 19\%; n = 21 \)) and reduction in HDL (\(-9.11\) mg/dL; 95% CI, \(-12.89\) to \(-5.32\); \( I^2 = 94\%; n = 22 \)). The quality of evidence was moderate for LDL and low for HDL, rated down due to moderate risk of bias in the included studies and significant heterogeneity for studies that assessed HDL. All lipid outcomes are surrogate; therefore, rating down due to indirectness is also plausible and can cause all these outcomes to warrant low confidence.

**Acne and hirsutism**

TCR had an increased risk of developing acne of 7.0% vs 4.7% (RR, 1.62; 95% CI, 1.28 to 2.06; \( I^2 = 0\%; n = 10 \)) and hirsutism of 10.7% vs 6.6% (RR, 1.45; 95% CI, 1.09 to 1.93; \( I^2 = 0\%; n = 10 \)). The quality of evidence is moderate for both due to moderate risk of bias of the included studies.

**Publication bias**

No evidence of publication bias was detected in any of the outcomes that were tested. Nevertheless, it is worth mentioning that the methods that are used to detect the publication bias are of limited yield in the presence of significant heterogeneity and the small number of studies pooled per outcome (26, 27).

**Subgroup analysis**

**T benefits**

We detected a significant difference between groups that used T vs those that used methyltestosterone or T undecanoate in sexual or relationship satisfaction outcome; the \( P \) value for the subgroup interaction (P-interaction) was .01. Patients who used T achieved better improvement than those who used the other two types. Surgically menopausal women achieved significantly higher scores in pleasure and enjoyment on sex scales than those who had natural menopause (P-interaction, .01).

No statistically significant differences were detected for other subgroups that evaluated the outcomes based on the route of T, whether or not the regimen used contained HRT, type of estrogen preparation, and route of HRT. The results of subgroup analysis are presented in Supplemental Table 3.

**T adverse effects**

The addition of HRT to T was found to be associated with a more pronounced reduction in total cholesterol as compared to the studies that used T alone (P-interaction, .003). The reduction in total cholesterol was also found to be more noticeable with methyltestosterone (P-interaction, .001) and with oral route (P-interaction, <.001). Triglyceride reduction was also found to be significant with the oral route, but not with non-oral route (P-interaction, .001).

Both HRT-containing regimens and HRT-free regimens showed a significant decline in HDL, but the decline...
was more prominent in trials that included HRT in their regimen. Furthermore, methyltestosterone and the oral route revealed more significant reduction in HDL when compared to other groups; interaction \( P \) values for both were less than .001.

No significant subgroup differences were detected for acne and hirsutism outcomes. The subgroup analysis results are presented in more detail in Supplemental Table 4.

Other androgenic side effects were also reported in some trials (weight gain, alopecia, and voice deepening); however, data were insufficient for meta-analysis.

Discussion

Main findings

This systematic review and meta-analysis demonstrates that the use of T either alone or with HRT has statistically significant beneficial effects on multiple domains of sexual function in postmenopausal women. These outcomes include the number of satisfying sexual episodes, frequency of sexual activity, libido, orgasm, arousal, pleasure or enjoyment of sex, sexual responsiveness, sexual self-image, and sexual or relationship satisfaction. Moreover, TCR was found to be associated with a reduction in sexual function concern and personal distress.

TCR was not associated with any significant benefit when compared to TFR regarding anxiety, depressed mood, BMD, body weight, or BMI. This inconclusive evidence could be due in part to the small number of studies that assessed each one of these outcomes.

It was also noted that TCR is significantly associated with a reduction in total cholesterol, triglyceride, and HDL cholesterol, whereas it increases LDL cholesterol. However, changes in lipids were small in magnitude and have unknown effect on future development of cardiac events. As expected, T increased the risk of acne and hirsutism. We found several significant subgroup analyses; however, such findings should only be considered as non-randomized observations that are exploratory and hypothesis generating.

Limitations and strengths

The strength of this systematic review stems from the fact that it was protocol-driven, done in a systematic manner, involved a comprehensive search of several databases, and was performed in collaboration with content experts. In addition, several measures that protect from bias were undertaken to ensure high quality of article selection and data extraction. The risk of bias assessment of the included studies was done in duplicate to ensure adequate determination of the quality of each study and the whole body of evidence.

The limitations of this systematic review include the possibility of publication bias, which could not be ruled out reliably, and the possibility of reporting bias (particularly for outcomes reported by a small proportion of the trials). For example, the outcome of bone health was only reported in three RCTs that were not designed to study bone health. Heterogeneity and imprecision affected several outcomes. Indirectness (use of surrogate outcomes) affected lipids and bone density outcomes.

In terms of sexual function measured by a variety of scales, interpretation of the finding is also challenging. We standardized some scales (reporting the results in standard deviation units). Communication of these results to patients requires knowledge of the minimally important difference of each scale (recognized by patients) and the use of shared decision-making tools.

Although we conducted six different subgroup analyses, data on T dose and T level achieved were insufficient for meta-analysis. Such analyses cannot reliably be done in a trial-level meta-analysis such as this one and will require individual patient data meta-analysis (ie, data on every participant). Lack of data on T dose and T level achieved add another layer of uncertainty to the clinical implications of our findings.

Comparison with previous reviews

Our systematic review reflects the current state of the best available evidence through updating previous reviews. A Cochrane systematic review by Somboonporn et al (21) was conducted in 2005 and updated in 2009 and focused on T therapy only when added to existing HRT. The authors concluded that there was good evidence that adding T to HRT has a beneficial effect on sexual function in postmenopausal women. However, the combined therapy was associated with a higher incidence of hair growth and acne and a reduction in HDL cholesterol. Our systematic review included trials in which T was used with or without HRT and evaluated the effect in each of these settings. Reviews by Hubayter and Simon (29), Braunstein (30), and Davis and Braunstein (7) did not provide meta-analytic estimates; however, they narratively concluded that T was an effective therapy for postmenopausal women with hypoactive sexual desire disorder.

Implications for practice

A major challenge in using T by postmenopausal women is establishing the diagnosis. There is disagreement about the normal serum androgen level for age and sex and about which assay to use. The Endocrine Society supports...
efforts to standardize assays, and several certified assays for T exist. The length of treatment and monitoring procedures are also not well established. Products approved by the US Food and Drug Administration are not available. Nevertheless, symptomatology and related distress cannot be ignored. Patients’ values and preferences should be the main driver for treatment decisions in this context. The Endocrine Society task force will integrate the available evidence summarized in this systematic review with patients’ context to develop clinical practice recommendations.

Implications for research

Very few RCTs studied the effect of T vs placebo without the use of HRT. The existing RCTs had short follow-up duration, which decreases confidence in benefits and harms when treatment for longer duration is contemplated. Data on treatment monitoring are also needed. Rigorously designed prospective observational studies could have a major role in evaluating the safety profile of T therapy. These studies have the advantage of enrolling large numbers of participants and following them up for extended periods of time, which allows investigators to assess patient-important safety outcomes.

Conclusion

The available evidence indicates that T is associated with significant improvement in sexual function in postmenopausal women with increased risk of acne and hirsutism. Changes in serum lipids are statistically significant but of unknown overall effect.

Acknowledgments

Address all correspondence and requests for reprints to: M. Hassan Murad, MD, MPH, Mayo Clinic, The Knowledge and Evaluation Research Unit, 200 First Street SW, Rochester, MN 55905. E-mail: Murad.mohammad@mayo.edu.

This review was partially funded by a contract from The Endocrine Society.


References


