Author's response to reviews

Title: Cardiovascular Risks and Elevation of Serum DHT Vary by Route of Testosterone Administration: a Systematic Review and Meta-analysis

Authors:

Stephen E Borst (seborst@ufl.edu)
Jonathan J Shuster (shusterj@ufl.edu)
Baiming Zou (bzou@ufl.edu)
Fan Ye (evan1979@ufl.edu)
Huanguang Jia (huanguang.jia@va.gov)
Anita Wokhlu (anita.wokhlu@va.gov)
Joshua F Yarrow (joshua.yarrow@va.gov)

Version: 2 Date: 7 October 2014

Author's response to reviews: see over
We are pleased to receive comments from two knowledgeable reviewers. While they pointed out the strengths and novelty of our meta-analysis, they also pointed out a number of issues that need to be addressed. Our responses are listed below and appropriate changes have been made in the ms.

**Reviewer 1**

1. The Copenhagen study used an oral formulation of micronized testosterone. The resulting serum T concentrations were very high. This should be factored into the conclusion that oral T has definite cardiovascular (CV) risks.

   *The very high T values in the treatment group may be due to cirrhosis and reduced metabolism of T. This is now listed as a limitation in the Discussion. This study was not used to assess elevation of T/DHT because DHT was not measured.*

2. On page 8, we stated that “oral TRT appeared to produce a post-treatment serum T that was similar to other administration routes.” The reviewer points out that oral TRT produced highly variable serum T concentrations.

   *We agree and we also think that since serum T levels are not sustained following oral administration, the time of blood acquisition may be a factor in the variability. We have added a statement on page 8, that although the mean is similar to other routes, the variance is high.*

3. Reviewer suggests discussion of the whether compliance varies by mode of administration and cites 2 references.

   *We thank the reviewer for pointing out these articles. The Schoenfeld paper reports that that gel TRT adherence was 34.7% at 6 months. The Donatucci paper reports that at 3 months, adherence to topical TRT was 52% and adherence to injected TRT was 32%. We did not think these papers show clear differences in adherence to different routes of administration for TRT. However, we did think these studies demonstrate that adherence is low in the clinic setting compared to clinical trials. We have added this point to the Discussion section on limitations and cited the two references. Regardless of adherence, the gold standard for analysis is by intent-to-treat, and hence even if we had patient level data, it would not be appropriate to consider adherence in the primary analysis.*

**Reviewer 2**

1. Authors conclude that oral T produces significant risk despite the fact that the effects were mixed – 2 studies showed increased risk and 2 did not.

   *The 2 studies that did not show risk had only 3 events between them including both the treatment and placebo arms. However we have added information in the Discussion section on limitations. Both the other studies were close to statistically significant and even 2/4 with P<0.10 would occur with about 5% probability when the null of 0 effect is true.*

2. Avoid use of the term non-significant direction trends

   *We agree and have substituted the following language. “While no significant effects on CV risk were observed with either injected or transdermal TRT, the point estimates suggest that further research is needed to establish whether administration by these routes is protective or detrimental, respectively.”*

3. The included studies did not have CV events as a pre-specified endpoint.

   *We agree and have added a statement to this effect in the Discussion section on limitations. However, we would like to point out that this is not a negative and that the same is true for all other meta-analyses of serious adverse side events.*
4. The authors stated that this meta-analysis includes less publication bias. Reviewer suggests that we present a funnel-plot to check for publication bias and present the results using trim-and-fill analysis. We agree that we cannot show whether we have less publication bias and have removed the statement. However, we are reluctant to include a funnel-plot for the following reason. Co-author Dr. Shuster is an expert in meta-analysis and he states in his peer reviewed review of the Cochrane Hand Book for Systematic Reviews for Interventions (2011), published in 2012 in Research Synthesis Methods that funnel plots should not be used to assess publication bias, although they often are. What funnel plots show is whether there is an association between the precision of the estimate and the estimate. This can happen in many ways totally unrelated to selection bias. Correlation is not evidence of cause-effect.

5. Was sensitivity analysis performed?

Thank you for the comment. A part of the methods on page 6 was mislabeled a sensitivity analysis. In fact we did not perform sensitivity analysis. We have moved this paragraph to the Discussion section on limitations and we now state that there was no sensitivity analysis.

6. Discussion is weak and needs discussion of strength and limitations
We have now added a section to the Discussion on study strengths and limitations. The latter include those pointed out by the reviewer in comments 1, 3, 4, 5, 7, and 9.

7. Two trials were stopped early. Could this have affected your results?

Basaria et al. (gel TRT) was stopped early because of CV events. The Copenhagen study (oral TRT) was a very long study and was eventually stopped for lack of efficacy. The estimates derived when studies are terminated early due to differences in the endpoint under study do yield slightly biased estimates but this cannot be adjusted for, and no one to our knowledge has taken this into account. If the studies were stopped early for issues unrelated to cardiovascular endpoints, then the issue is not relevant. (Copenhagan). Even if they were related (Basaria), note that as noted above, the bias is slightly against the null, so this actually reinforces the overall null conclusion. As a result we have added this to the Discussion section on limitations.

8. In Table 3, presentation of pre-post fold increases in T/DHT is confusing. For example, an increase from 0.99 to 3.43 is listed as a 5.5-fold increase.

The values are not arithmetic means, but sample-weighted means. A sample calculation is listed below. We have added information in the statistical section to clarify this. But this was an astute observation by the reviewer.

<table>
<thead>
<tr>
<th>Obs</th>
<th>Author</th>
<th>Mode</th>
<th>Tnum</th>
<th>DHTfold</th>
<th>Tnum</th>
<th>fold_times_Tnum</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Bals-Pratch</td>
<td>Patch</td>
<td>7</td>
<td>4.6</td>
<td>7</td>
<td>32.20</td>
</tr>
<tr>
<td>2</td>
<td>Mazer</td>
<td>Patch</td>
<td>28</td>
<td>3.45</td>
<td>28</td>
<td>96.60</td>
</tr>
<tr>
<td>3</td>
<td>Behre</td>
<td>Patch</td>
<td>11</td>
<td>4</td>
<td>11</td>
<td>44.00</td>
</tr>
<tr>
<td>4</td>
<td>Ahmed</td>
<td>Patch</td>
<td>5</td>
<td>26</td>
<td>5</td>
<td>130.00</td>
</tr>
<tr>
<td>5</td>
<td>Cunningham</td>
<td>Patch</td>
<td>12</td>
<td>27.33</td>
<td>12</td>
<td>327.96</td>
</tr>
<tr>
<td>6</td>
<td>Meikle</td>
<td>Patch</td>
<td>6</td>
<td>5.95</td>
<td>6</td>
<td>35.70</td>
</tr>
<tr>
<td>7</td>
<td>Raynaud</td>
<td>Patch</td>
<td>131</td>
<td>5</td>
<td>131</td>
<td>655.00</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Variable</th>
<th>Sum</th>
</tr>
</thead>
<tbody>
<tr>
<td>fold_times_Tnum</td>
<td>1321.46</td>
</tr>
<tr>
<td>Tnum</td>
<td>200.00000000</td>
</tr>
</tbody>
</table>

Point Estimate=1321/200=6.61
9. Given the discrepancies of DHT assay methods (LC-MS/MS, GC-MS, RIA, Fluoro-immuno assay and others) comparisons may not be valid.

We thank the reviewer for bringing up a good point that we were aware of. LC-MS/MS and GC/MS are the gold standard measures and are highly accurate. The EIA assay method for measuring DHT is invalid as shown by our recent paper (Yarrow JF, Beck DT, Conover CF, Beggs LA, Goldberger BA, Borst, SE. (2013) Invalidation of a commercially available human 5α-dihydrotestosterone immunoassay. Steroids 78(12-13):1220-1225) and studies using EIA were excluded. In contrast the RIA and other methods are valid, but the concentrations reported are a little higher than with MS-based assays (Wang et al. Steroids 73:1345-1352, 2008). RIA has good specificity for DHT and has been used to demonstrate the reduced serum DHT concentration following finasteride administration (Stanczyk et al. j Steroid Biochem & Molec Biol, 138:10-16, 2013). The fact that some assays produced DHT concentrations that are somewhat higher than those produced with MS-based assay should not present any bias when examining the ratio of DHT pre-treatment/post-treatment, as shown in table 3. However, it may introduce some bias in the post-treatment DHT values shown in Figure 4 and we have now addressed this in the discussion section on potential limitations. This is one of the beauties of random effects methods. You can incorporate different methods (as long as the units match) and different follow-up times. This increases between study variation over completely controlling this from study to study, but it does not invalidate the inference.

Yours sincerely -

Steve Borst, Ph.D.

Assoc. Director for Research, Geriatric Research
Malcom Randall VA Medical Center
1601 SW Archer Rd. Gainesville FL 32608-1197
Professor, Department of Applied Physiology & Kinesiology
University of Florida
seborst@ufl.edu phone: 352-374-6114 FAX: 352-374-6142