CLINICAL REVIEW: Testosterone Use in Men and Its Effects on Bone Health. A Systematic Review and Meta-Analysis of Randomized Placebo-Controlled Trials

Michal J. Tracz, Kostandinos Sideras, Enrique R. Boloña, Rudy M. Haddad, Cassie C. Kennedy, Maria V. Uraga, Sean M. Caples, Patricia J. Erwin, and Victor M. Montori

Knowledge and Encounter Research Unit, Department of Medicine (M.J.T., K.S., E.R.B., R.M.H., C.C.K., M.V.U., S.M.C., P.J.E., V.M.M.); Mayo Library (P.J.E.); and Division of Endocrinology, Department of Medicine (V.M.M.), Mayo Clinic College of Medicine, Rochester, Minnesota 55905

Context: Androgen-deficient men are at increased risk of osteoporosis. The extent to which testosterone can prevent and treat osteoporosis in men remains unclear.

Objective and Design: We performed a systematic review and meta-analysis of randomized placebo-controlled trials in men to estimate the effect of testosterone use on bone health outcomes.

Data Sources: The review encompassed librarian-designed search strategies using MEDLINE (1966 to March 2005), EMBASE (1988 to March 2005), and Cochrane CENTRAL (inception to March 2005); a review of reference lists from included studies; and content expert files.

Data Collection: Independently and in duplicate, we assessed the methodological quality of the eligible trials and collected data on bone mineral density and bone fractures at the longest point of complete follow-up.

Data Synthesis: We included eight trials enrolling 365 patients. Two trials followed patients for more than 1 yr. Meta-analysis of these trials showed that, compared with placebo, intramuscular testosterone was associated with an 8% (95% confidence interval, 4%, 13%) gain in lumbar bone mineral density and transdermal testosterone had no significant impact. Testosterone use was associated with a nonsignificant 4% (95% confidence interval, −2%, 9%) gain in femoral neck bone mineral density with unexplained differences in results across trials (26% of these differences were not explained by chance alone). No trials measured or reported the effect of testosterone on fractures.

Conclusions: Intramuscular testosterone moderately increased lumbar bone density in men; the results on femoral neck bone density are inconclusive. Without bone fracture data, the available trials offer weak and indirect inferences about the clinical efficacy of testosterone on osteoporosis prevention and treatment in men. (J Clin Endocrinol Metab 91: 2011–2016, 2006)

Men with hypogonadism are at increased risk of osteoporosis (1–4). Osteoporotic fractures may be associated with loss of independence, premature mortality, and increased health care expenditures. Therefore, the identification of effective strategies to prevent osteoporotic fractures in at-risk men, such as bisphosphonates (5), is important. When faced with an identifiable cause of osteoporosis such as apparent androgen deficiency, however, clinicians and patients may be interested in primarily treating this underlying condition. Furthermore, with aging, elderly men experience a decline in testosterone levels and bone mass and an increase in the risk of osteoporotic fractures (2, 6). Chronically ill men and those exposed to glucocorticoids may have both apparent androgen deficiency and low bone mass (7, 8). Androgen deficiency may represent a key pathophysiological pathway in these situations, making testosterone an attractive intervention.

Many observational studies have found an association between testosterone use in men and important gains in bone density, favorable changes in bone turnover biomarkers, and a lower risk of osteoporotic fractures (9–12). High-quality randomized trials, however, offer the strongest inferences about the efficacy and safety of any therapeutic agent. Individual trials may not have enough events (bone fractures) to estimate the efficacy of testosterone with sufficient precision; also, different trials may reach different and even opposite conclusions. Furthermore, results from one trial may not apply well to populations that are sufficiently different in their biological or socioeconomic characteristics. However, if results are consistent across trials, clinicians in different settings would be more confident in applying this evidence to their patients.

To determine the extent to which testosterone use in men prevents osteoporotic fractures and enhances bone mass in at-risk men (hypogonadal men, elderly men with androgen decline, and men receiving glucocorticoids), we performed a systematic review and meta-analyses of randomized placebo-controlled trials conducted in these populations. These meta-analyses supported the work of The Endocrine Society Task Force on Testosterone for Men with Androgen Deficiency Syndromes in producing clinical practice guidelines and formulating evidence-based recommendations.

Materials and Methods
We prepared a review protocol (available from the authors) with extensive input from the clinical expert members of The Endocrine Society Task Force. None of the authors received funding or any other...
support from makers or distributors of testosterone preparations or have any investments in such companies. We have produced this report in adherence with the Quality of Reporting of Meta-analyses (QUOROM) standards for reporting systematic reviews of randomized trials (13).

Eligibility criteria

Eligible studies were fully published randomized trials randomly assigning patients with any degree of androgen deficiency either to testosterone use with any of the available preparations or to placebo, and measuring the impact of these interventions on the risk of bone fractures. Bone mineral density is a surrogate marker that seems to capture the bone fracture prevention effects of estrogen therapy in women; most of this effect appears mediated by testosterone (14). This provided the biological rationale for relying on bone density as a surrogate marker of the effect of testosterone use on bone health and fracture risk. Thus, trials measuring the effect of testosterone on lumbar and femoral neck bone mineral density were also eligible. Because inferences about bone fracture incidence from consideration of changes in biomarkers are generally weak, we opted for excluding trials that measured the effect of testosterone only on this outcome.

Study identification

An expert reference librarian (P.J.E.) designed and conducted the electronic search strategy with input from an endocrinologist (V.M.M.) with expertise in conducting systematic reviews. To identify eligible studies, our systematic search included electronic databases (MEDLINE, EMBASE, and CENTRAL) from their inception until October 2004; review of the reference sections of identified narrative and systematic reviews identified through a MEDLINE search in October 2004, and of each of the eligible primary studies; and contact with expert members of the Task Force. The search was updated in March 2005. Teams of two reviewers independently and with substantial reliability [change-adjusted intrarater agreement (κ = 0.71) screened all abstracts and titles as well as all resulting full-text publications for eligibility. In cases where disagreement between two reviewers existed, another member of the research team not involved in the initial assessment and with both content and methodological expertise (V.M.M.) adjudicated the study as eligible or not, after reviewing the stated reasons for the initial assessment and the full text of the report.

Data collection

Working in duplicate and using a standardized data extraction form, we abstracted the following descriptive data from every study: year and journal of publication, patient population (degree of androgen deficiency, prior exposure to testosterone, age, testosterone level), treatment (dose and route of administration of testosterone) and control interventions, and the number of patients in exposed and unexposed groups.

We classified reports by the mean testosterone level at baseline; low testosterone level was defined by total testosterone no greater than 300 ng/dl (10.4 nmol/liter). When this was not reported, we used values below the lower limit of normal for bioavailable or free testosterone levels. When laboratory values were not available, we classified studies by the type of patients enrolled (i.e., patients with previous bilateral orchiectomy). Chance-adjusted interobserver agreement for this classification was almost perfect (κ = 0.91).

We collected bone fracture rates in intervention and control at the longest point of complete follow-up after randomization. We also collected data on bone mineral density (end-of-period or change-from-baseline) at the longest duration at which follow-up was sufficiently complete and patients were still exposed to testosterone or placebo. When data seemed to have been collected but was not reported or reported only with a statement of change (e.g., ‘no change’) or significance (e.g., ‘not significant’) we contacted the authors.

Quality assessment

To ascertain the validity of eligible randomized trials, pairs of reviewers working independently and with adequate reliability (corresponding κ statistic in parentheses where appropriate) determined the adequacy of randomization (κ = 1.00) and concealment of allocation (κ = 0.82), blinding of patients (κ = 0.70), health care providers (κ = 0.70), data collectors (κ = 0.77), and outcome assessors (κ = 0.84), and extent of loss to follow-up (i.e. proportion of patients in whom the investigators were not able to ascertain outcomes).

Statistical analyses

Meta-analyses. We determined the effect size and 95%-confidence interval (CI) for the difference between arms (testosterone vs. placebo) in lumbar and femoral neck bone mineral density (BMD) by dividing the mean difference by the pooled sd between arms with adjustment for small samples (Hedges g standardized mean differences) as implemented in RevMan 4.2 (Cochrane Collaboration). We then conducted meta-analysis using the random-effects method and quantified the extent to which the variability observed corresponded to between-study differences using the I² statistic (15).

Subgroup analyses

Our a priori hypotheses to explain potential heterogeneity across studies included: study quality (particularly loss to follow-up); patient population (age greater than 60; primary or secondary prevention of osteoporosis; use of glucocorticoids; degree of androgen deficiency arbitrarily defined by baseline testosterone levels); interventions (testosterone administration route, transdermal vs. im); and outcomes (length of follow-up at time of measurement). To explore these subgroups, we tested for treatment-subgroup interactions (16).

Results

Search results

Figure 1 describes the flow of candidate and eligible articles. Two trials (11, 17) did not measure or report on BMD data, leaving eight eligible randomized trials.

Study characteristics

Methodological quality. Table 1 describes the methodological quality of included trials. Overall, the included trials had...
Table 2 describes trial characteristics. One trial (20) enrolled patients with known osteoporotic fractures (secondary prevention); we assumed participants in the other trials had not suffered osteoporotic fractures. Participants in all trials received usual testosterone doses, although trials differed in both the duration of treatment and route of testosterone administration. Two trials studied patients for 36 months (12, 21); the other trials followed patients for 1 yr or less. All but two trials (10, 12) enrolled patients with low normal and normal testosterone levels at baseline.

**Effect of testosterone on bone health.** None of the trials measured or reported on the effect of testosterone on fracture incidence. All other trials were eligible for meta-analyses of bone density (365 participants).

*Meta-analysis of these studies showed a small and significant increase in lumbar spine BMD with testosterone (effect size, 0.31; CI, 0.02, 0.61), corresponding to a 4% (CI, 0.3%, 8%) gain in lumbar BMD (Fig. 2). There was a small and non-significant increase in femoral neck BMD with testosterone (effect size, 0.17; CI, −0.11, 0.45), corresponding to a 4% (CI, −2%, 9%) gain in femoral neck BMD (Fig. 3). There was moderate heterogeneity between studies in both lumbar and femoral groups (I² = 46% and 26%, respectively). To explain these inconsistencies we conducted preplanned subgroup analyses.

**Subgroup analyses**

*Glucocorticoids.* Three studies enrolled patients taking glucocorticoids, and we assessed whether this patient subpopulation was responsible for overall heterogeneity of results. The three trials studied a total of 87 men on chronic glucocorticoid use. Two trials reported loss to follow-up in excess of 20% (10, 12, 19). The median loss to follow-up across trials was 14%; three trials reported loss to follow-up in excess of 20% (10, 12, 19).

**Clinical characteristics.** Table 2 describes trial characteristics. One trial (20) enrolled patients with known osteoporotic fractures (secondary prevention); we assumed participants in the other trials had not suffered osteoporotic fractures. Participants in all trials received usual testosterone doses, although trials differed in both the duration of treatment and route of testosterone administration. Two trials studied patients for 36 months (12, 21); the other trials followed patients for 1 yr or less. All but two trials (10, 12) enrolled patients with low normal and normal testosterone levels at baseline.

**TABLE 1. Methodological qualities of each trial**

<table>
<thead>
<tr>
<th>First author, year (Ref.)</th>
<th>Study design</th>
<th>Placebo-controlled</th>
<th>Allocation concealment</th>
<th>Loss to follow-up (%)</th>
<th>Blinding</th>
</tr>
</thead>
<tbody>
<tr>
<td>Amory, 2004 (12)</td>
<td>Parallel</td>
<td>Yes</td>
<td>Adequate</td>
<td>27</td>
<td>All groups clearly blinded.</td>
</tr>
<tr>
<td>Crawford, 2003 (19)</td>
<td>Parallel</td>
<td>Yes</td>
<td>Not reported</td>
<td>21</td>
<td>All groups presumed blinded.</td>
</tr>
<tr>
<td>Fairfield, 2001 (8)</td>
<td>Parallel, 4 arms</td>
<td>Yes</td>
<td>Not reported</td>
<td>14</td>
<td>Patients and providers clearly blinded, outcome collectors and assessors presumed blinded.</td>
</tr>
<tr>
<td>Hall, 1996 (20)</td>
<td>Parallel</td>
<td>Yes</td>
<td>Not reported</td>
<td>14</td>
<td>All groups presumed blinded, outcome assessor clearly blinded.</td>
</tr>
<tr>
<td>Howell, 2001 (29)</td>
<td>Parallel</td>
<td>Yes</td>
<td>Not reported</td>
<td>Not reported</td>
<td>Single blinded study (only patients blinded). Unknown whether data collectors or assessors were blinded.</td>
</tr>
<tr>
<td>Kenny, 2001 (10)</td>
<td>Parallel</td>
<td>Yes</td>
<td>Not reported</td>
<td>34</td>
<td>Patients and providers clearly blinded, outcome collectors and assessors presumed blinded.</td>
</tr>
<tr>
<td>Reid, 1996 (18)</td>
<td>Crossover</td>
<td>Crossover</td>
<td>Not reported</td>
<td>6</td>
<td>Clearly not blinded.</td>
</tr>
<tr>
<td>Snyder, 1999 (21)</td>
<td>Parallel</td>
<td>Yes</td>
<td>Not reported</td>
<td>12</td>
<td>All groups clearly blinded.</td>
</tr>
</tbody>
</table>

**TABLE 2. Trial characteristics**

<table>
<thead>
<tr>
<th>First author, year (Ref.)</th>
<th>Participants</th>
<th>Testosterone level at baseline [ng/dl (nmol/liter)]</th>
<th>Testosterone intervention</th>
<th>Duration (months)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Amory, 2004 (12)</td>
<td>48 men, mean age 71 yr</td>
<td>Total: 291 (10.2)</td>
<td>Testosterone enanthate 200 mg im every 2 wk vs. identical placebo.</td>
<td>36</td>
</tr>
<tr>
<td>Crawford, 2003 (19)</td>
<td>34 men, mean age 60 yr, with chronic glucocorticoid use</td>
<td>Total: 414 (14.5)</td>
<td>Testosterone mixed esters 200 mg im every 2 weeks vs. placebo.</td>
<td>12</td>
</tr>
<tr>
<td>Fairfield, 2001 (8)</td>
<td>50 men, mean age 36 yr, with AIDS wasting</td>
<td>Total: 646 (22.6)</td>
<td>Testosterone enanthate 200 mg im weekly vs. placebo.</td>
<td>3</td>
</tr>
<tr>
<td>Hall, 1996 (20)</td>
<td>30 men, mean age 61 yr, with rheumatoid arthritis</td>
<td>Total: 457 (16.0)</td>
<td>Testosterone enanthate 250 mg im monthly for 6 months then every 2 wk for 3 more months vs. placebo</td>
<td>9</td>
</tr>
<tr>
<td>Howell, 2001 (29)</td>
<td>35 men, mean age 41 yr, with Leydig cell dysfunction</td>
<td>Total: 380 (13.3)</td>
<td>Testosterone patch 2.5 mg a day</td>
<td>12</td>
</tr>
<tr>
<td>Kenny, 2001 (10)</td>
<td>67 men, mean age 75 yr</td>
<td>Bioavailable: 92 (3.2)</td>
<td>Two 2.5 mg Androderm patches (5 mg/dl) applied each evening vs. identical placebo</td>
<td>12</td>
</tr>
<tr>
<td>Reid, 1996 (18)</td>
<td>16 men, mean age 61 yr, with long-term glucocorticoid use</td>
<td>Total: 360 (12.6)</td>
<td>Testosterone mixed esters 250 mg intramuscular depot injection monthly vs. no treatment</td>
<td>12</td>
</tr>
<tr>
<td>Snyder, 1999 (21)</td>
<td>108 men, mean age 73 yr</td>
<td>Total: 363 (12.7)</td>
<td>Testosterone scrotal patch 6 mg once a day vs. identical placebo</td>
<td>36</td>
</tr>
</tbody>
</table>
corticoids. One trial investigated patients with at least 10 yr of rheumatoid arthritis (20) exposed to a mean cumulative dose of 3.5 g of prednisolone. Another trial enrolled patients with steroid-treated asthma (18) who had received a mean daily dose of 11 mg of prednisone daily for 10 yr (19). The effect of testosterone on lumbar bone density was consistent across trials (I² = 0%), significant and moderate (effect size, 0.67; CI, 0.23, 1.1), corresponding to a 9% (CI, 3%, 8%) gain in lumbar bone density. The effect on femoral neck bone density was not significant (effect size, 0.14; CI, −0.31, 0.76). The effect of testosterone on femoral neck BMD in the same subgroup was nonsignificant (effect size, 0.14; CI, −0.27, 0.55). In both cases, there was no treatment-baseline testosterone level interaction (P = 0.45 and 1.0, respectively). Test for interaction between treatment effect and on-trial testosterone levels was not possible because testosterone levels during the trial were not consistently reported across trials.

**Route of administration.** Five studies (8, 12, 18–20) administered testosterone through im injection. Analysis of this subgroup showed a significant and moderate treatment effect (effect size, 0.62; CI, 0.32, 0.92), corresponding to an 8% (CI, 4%, 13%) increase in lumbar BMD. This result is significantly different from the pooled estimate from trials of transdermal testosterone (P interaction test = 0.0009). There was no treatment-testosterone route interaction (P = 0.8) for the femoral neck site.

In addition to the above, we explored subgroups based on age (patients older than 60 yr), duration of follow-up (>1 yr), and loss to follow-up (≥20%) without identifying any other treatment-subgroup interaction.

**Discussion**

Restatement of findings

There were no randomized controlled trials (RCTs) assessing the impact of testosterone use in men and measuring the effect of testosterone on the incidence of osteoporotic bone fractures. Included RCTs followed patients for a brief period—only two of the RCTs followed patients for more than 1 yr. The pooled results suggest a beneficial effect on lumbar spine bone density and equivocal findings on femoral neck BMD (Figs. 2 and 3). Trials of im testosterone reported significantly larger effects on lumbar bone density than trials of transdermal testosterone, explaining all the observed inconsistency across trials. The inconsistency across trial results on the femoral neck bone density remains unexplained.

**Limitations and strengths**

Our systematic review has some limitations. The methodological quality of the primary studies and reliance on surrogate endpoints weaken inferences about the effect of testosterone on osteoporotic fractures. Furthermore, there is but one trial in secondary prevention population, and there are no reports of whether patients systematically implemented other effective interventions to prevent or treat osteoporosis. Despite our best efforts, we may have missed eligible studies that could contribute to publication bias, i.e. overestimating the treatment effect. Wide confidence intervals indicate that the trials were too small to determine with precision the effect of testosterone on femoral BMD.

Another limitation is the moderate inconsistency (I² of 46% for the lumbar spine and I² of 26% for the femoral neck) across study results, suggesting differential effects of testos-
Testosterone therapy across trials with different methods, patients, interventions, or outcome measures. We were able, however, to explain the inconsistency in the lumbar spine results, using a preplanned subgroup analysis, by a treatment-route interaction, observing a significantly larger pooled treatment effect in trials using im testosterone than in trials using transdermal preparations. So defined, these subgroups had no inconsistency across trials (I² = 0%).

The inferences from this review are strengthened by our focused review questions; thorough and systematic search, designed by an experienced reference librarian and enhanced by consultation with five expert andrologists; explicit and reproducible eligibility criteria; and protocol-driven and focused analyses.

Comparison with other systematic reviews

Isidori et al. (22) published a meta-analysis of randomized trials of testosterone and its effects on bone health and other outcomes in elderly men. Their findings (which differ given the difference in the research question and, therefore, include fewer trials) are consistent with those reported here: significant impact of testosterone on lumbar bone density, non-significant impact on femoral bone density, and heterogeneity explained in part by the type of testosterone preparation (im testosterone yielding a greater effect on bone density than transdermal preparations). We were not able to identify other systematic reviews on this topic, but many narrative reviews have considered the issue. For example, Allan and McLachlan (23) examined a subset of the trials included in our review and qualitatively concluded that testosterone supplementation for older men with age-related decline in testosterone levels may be beneficial in terms of bone density loss, especially in the lumbar spine. Our systematic review did not find eligible trials enrolling men with severely low testosterone levels, and we did not observe a treatment-age interaction.

Implications for practice, research, and policy

There is an evolving understanding of the biology relating testosterone use to bone health. Although testosterone seems to play an important role in bone maintenance and bone formation, it appears more likely that a complex interaction between testosterone and estrogen (through their respective receptors) is key in the regulation of the male bone skeleton such that testosterone may impact bone health directly and, perhaps to a greater extent, through aromatization to estrogen (14, 24–26). The findings summarized here support a role for testosterone use in enhancing bone health. Given the guarded inferences that can be drawn from our analysis about the efficacy of testosterone therapy in the patients studied, recommendations of the use of testosterone in men with different degrees of androgen deficiency can only be tentative. For instance, our findings may not apply to patients with profound hypogonadism who are underrepresented in the included trials.

Clinicians need to refer to the results of trials of im vs. transdermal preparations to draw stronger inferences about the apparent superiority of the former in improving lumbar spine bone density, and, more importantly, in reducing the risk of spinal fractures. Although speculative, the higher testosterone exposures achieved with im testosterone administration vs. transdermal administration (27) may partly explain this observation.

A meta-analysis of testosterone RCTs outlined the extent to which testosterone therapy is associated with adverse effects and with the burden of monitoring, particularly related to prostate cancer surveillance (28). Thus, patients and their clinicians will need to consider the testosterone choice as one associated with unclear but potential benefit, potential side effects, and treatment burden. The decision may be easier for patients with previous osteoporotic fracture, in whom several interventions could be justified [calcium, vitamin D, bisphosphonates (5)]. The extent to which adding testosterone to such a therapeutic cocktail will help patients remains unclear.

The research needs are evident: large randomized trials in patients with and without a history of osteoporotic fracture, using one or more commonly available preparations of testosterone, offering testosterone treatment for 3 yr or more; offering adequate calcium and vitamin D to the participants, and measuring bone fractures as the primary outcome. Clinicians torn about whether or not to recommend testosterone to patients should consider instead enrolling these patients in such trials. Trials comparing testosterone use with established therapies (e.g. bisphosphonates) would also be informative.

Conclusion

Testosterone use moderately increased lumbar bone density in men, particularly among patients receiving chronic glucocorticoids and im testosterone. Without bone fracture data, currently available evidence offers weak and indirect inferences about the clinical efficacy of testosterone on osteoporosis prevention and treatment. Randomized trials of testosterone vs. placebo and bisphosphonates in at-risk men with fracture outcomes would provide the necessary evidence.

Acknowledgments

We thank Gunjan Y. Gandhi, M.D.; Theophilus E. Owan, M.D.; and Laura I. Pelaez, M.D., for their assistance with study selection, data collection, and author contact. We are very grateful for the ongoing input and advice from The Endocrine Society Task Force on Androgen Use in Men. We also extend our gratitude to the research team at the Knowledge and Encounter Research Unit at Mayo Clinic College of Medicine. Received January 6, 2006. Accepted March 28, 2006.

Address all correspondence and requests for reprints to: Victor M. Montori, D.MSc., Mayo Clinic, W18A, 200 First Street SW, Rochester, Minnesota 55905. E-mail: montori.victor@mayo.edu.

The work herein was funded by a Mayo Foundation Scholarship (to V.M.M.) and supported by the Department of Medicine, Mayo Clinic College of Medicine.

References

3. Kenny AM, Prestwood KM, Marcello KM, Raisz LG 2000 Determinants of

JCEM is published monthly by The Endocrine Society (http://www.endo-society.org), the foremost professional society serving the endocrine community.