Androgen treatment and muscle strength in elderly men: a meta-analysis

Ottenbacher K J, Ottenbacher M E, Ottenbacher A J, Acha A A, Ostir G V

CRD summary
The authors concluded that testosterone moderately increased muscle strength in the older men included in the studies, but the results were influenced by one study. The authors’ conclusion appears to reflect the review findings. However, a more cautious conclusion might have been appropriate given the considerable influence of one small study and other potential differences between the studies.

Authors’ objectives
To evaluate the effects of androgen treatment on muscle strength in older men.

Searching
MEDLINE, EMBASE, CINAHL and the Cochrane CENTRAL Register were searched for studies published in English in peer-reviewed journals between 1980 and 2005; the search terms were reported. In addition, the reference lists of identified studies were screened and citations of authors of studies in the field were tracked.

Study selection
Study designs of evaluations included in the review
Double-blind randomised controlled trials (RCTs) were eligible for inclusion. Both parallel group and crossover trials were eligible. The duration of the included studies ranged from 4 to 156 weeks.

Specific interventions included in the review
Studies that compared testosterone or dihydrotestosterone (DHT) with placebo were eligible for inclusion. Studies had to describe the specific method and dose of androgen. Most of the included studies evaluated testosterone given by injection (ranging generally from 100 to 200 mg biweekly) or patch (2.5 to 6 mg daily); one study each evaluated testosterone administered orally (80 mg twice daily) and as a gel (70 mg daily).

Participants included in the review
Studies of older men with a mean age of 65 years or older were eligible for inclusion. The mean age of participants in the review was 69 years and baseline testosterone levels ranged from 2.9 to 5.9 mg/mL.

Outcomes assessed in the review
Studies that reported an operationally defined measure of upper and/or lower muscle strength were eligible for inclusion. Studies had to report sufficient results data to permit calculation of an effect size. It is noted, however, that an effect size of 0 was assumed for some comparisons (e.g. when a non-statistical result was reported in a study). The included studies assessed muscle strength at various sites and in various positions of the upper and lower limb; one study assessed total body strength. The review also assessed adverse events.

How were decisions on the relevance of primary studies made?
One reviewer screened identified abstracts and two reviewers screened full reports. No other details of the study selection process were described.

Assessment of study quality
Two reviewers independently assessed, coded and scored validity using 13 items in relation to study design. Studies scoring less than 10 points out of 13 were classified as low quality. Studies were also assessed for intention-to-treat (ITT) analysis and level of attrition.
Data extraction
Two reviewers independently extracted and coded the data. Where data were reported at multiple time points, the data for the time closest to the end of treatment were extracted. Where possible, for each study, standardised mean effect sizes (g-index) were extracted or calculated from reported statistics for each measure of muscle strength. Where required, authors were contacted for raw data to enable the calculation of effect sizes. An effect size of 0.00 was assumed where adequate data to calculate an effect size were not available. In addition, all effect sizes within one study were averaged to obtain one effect size per study.

Methods of synthesis
How were the studies combined?
Pooled effect sizes (Hedges’ g) with 95% confidence intervals (CIs) were calculated using a random-effects model weighted by sample size. Pooled effect sizes were calculated using each of the 38 effect sizes. In the review, effect sizes were considered small if between 0.2 and 0.49, medium if between 0.5 and 0.79, and large if 0.8 or greater. The Shapiro-Wilks test of normality was used to assess the use of multiple effect sizes from single studies. The potential for publication bias was assessed using a funnel plot.

How were differences between studies investigated?
Heterogeneity was examined using Galbraith plots at the homogeneity statistic. The influence of each study was assessed by repeating the analysis after omitting each study in turn. Subgroup analyses were used to examine the influence on treatment effect size of the type of muscle strength (upper or lower extremity or lower body), the method of testosterone administration, attrition rate, ITT analysis, design quality, publication date and study duration.

Results of the review
Eleven RCTs were included (n=474 according to the text and n=455 according to the data extraction tables). These studies provided 38 effect sizes.

The mean unadjusted effect size using one effect size per study was 0.58 (95% CI: 0.22, 0.93). The mean effect size using all 38 effect sizes was 0.53 (95% CI: 0.21, 0.86).

There appeared to be evidence of heterogeneity but it was not clear to which analysis this applied (all 38 effect sizes, or the 1 effect size per study analysis). Other analyses were based on all 38 effect sizes.

Effect sizes were larger for measures of lower extremity muscle strength (0.63, 95% CI: 0.03, 1.28; based on 20 effect sizes) compared with upper extremity strength (0.47, 95% CI: 0.12, 0.84; based on 17 effect sizes).

Effect sizes were larger for studies using injected testosterone (0.95, 95% CI: 0.33, 1.58; based on 18 effect sizes) compared with topical (0.26, 95% CI: 0.08, 0.42; based on 16 effect sizes) or oral testosterone (-0.21, 95% CI: -1.40, 1.02; based on 4 effect sizes. Effect sizes from studies with no attrition were larger than those from studies with attrition levels greater than 10% (1.27 versus 0.10). Effect sizes from studies using ITT analysis were smaller than those from studies that did not use ITT analysis (0.15 versus 0.89).

Effect sizes from higher quality studies were smaller than those from lower quality studies (0.30 versus 0.64).

The removal of 1 study (n=12, provided 5 effect sizes) substantially decreased the treatment effect size from 0.53 to 0.23 (95% CI: 0.09, 0.38). This study administered an individually determined injected dose of testosterone.

Adverse events.
All 11 studies monitored adverse events but definitions varied widely. Three studies reported high prostate-specific antigen or prostate disease. Four studies reported no adverse events.

Authors’ conclusions
Testosterone or DHT treatment moderately increased muscle strength in the older men included in the review, but the
results were influenced by one study.

**CRD commentary**
The review addressed a clear question that was defined in terms of the participants, intervention, outcomes and study design. Several relevant sources were searched but no attempts were made to minimise publication or language bias. Some attempts were made to reduce reviewer error and bias in the study selection, data extraction and validity assessment processes, but the methods were not completely described (methods used to identify potentially relevant abstracts or to resolve disagreements were not reported). The validity of the included studies was assessed systematically and the influence of study quality on treatment effects was examined.

Little information was provided about the participants, making it difficult to determine to what population these results may be applicable. The studies were pooled using meta-analysis and statistical heterogeneity was assessed, but the results of the assessment were not clearly reported. In the absence of meta-analysis graphs or effect sizes from individual studies it was not possible to determine the extent of the heterogeneity. The authors' conclusion appears to reflect the review findings but a more cautious conclusion about the effects of testosterone might have been appropriate in view of the considerable influence of one small study with 12 participants, and other potential differences between the studies.

**Implications of the review for practice and research**
Practice: The authors did not state any implications for practice.

Research: The authors stated that future research to evaluate the efficacy of testosterone in older men should take into account factors such as the method of administration, treatment duration, individually adjusted dose and drop-outs.

**Bibliographic details**

**PubMedID**
17087692

**DOI**
10.1111/j.1532-5415.2006.00938.x

**Indexing Status**
Subject indexing assigned by NLM

**MeSH**
Aged; Dihydrotestosterone /therapeutic use; Hormone Replacement Therapy; Humans; Male; Muscle Strength /drug effects; Randomized Controlled Trials as Topic; Testosterone /therapeutic use

**AccessionNumber**
12006007877

**Date bibliographic record published**
31/03/2008

**Date abstract record published**
31/03/2008

**Record Status**
This is a critical abstract of a systematic review that meets the criteria for inclusion on DARE. Each critical abstract
contains a brief summary of the review methods, results and conclusions followed by a detailed critical assessment on the reliability of the review and the conclusions drawn.