

# Clinical Meaningfulness of the Changes in Muscle Performance and Physical Function Associated With Testosterone Administration in Older Men With Mobility Limitation

Thomas G. Travison,<sup>1,2</sup> Shehzad Basaria,<sup>\*,1</sup> Thomas W. Storer,<sup>\*,1</sup> Alan M. Jette,<sup>3</sup> Renee Miciek,<sup>1</sup> Wildon R. Farwell,<sup>4,5</sup> Karen Choong,<sup>1</sup> Kishore Lakshman,<sup>1</sup> Norman A. Mazer,<sup>1,6</sup> Andrea D. Coviello,<sup>1</sup> Philip E. Knapp,<sup>1</sup> Jagadish Ulloor,<sup>1</sup> Anqi Zhang,<sup>1</sup> Brad Brooks,<sup>1</sup> Ahn-Hoa Nguyen,<sup>1</sup> Richard Eder,<sup>1</sup> Nathan LeBrasseur,<sup>1</sup> Ayan Elmi,<sup>1</sup> Erica Appleman,<sup>1</sup> Leife Hede-Brierley,<sup>1</sup> Geeta Bhasin,<sup>1</sup> Ashmeet Bhatia,<sup>1</sup> Antonio Lazzari,<sup>4</sup> Samuel Davis,<sup>4</sup> Pengsheng Ni,<sup>3</sup> Lauren Collins,<sup>1</sup> and Shalender Bhasin<sup>1</sup>

<sup>1</sup>Section of Endocrinology, Diabetes, and Nutrition, Department of Medicine, Boston University School of Medicine and Boston Medical Center, Massachusetts.

<sup>2</sup>Department of Biostatistics and <sup>3</sup>Health and Disability Research Institute, Boston University School of Public Health, Massachusetts.

<sup>4</sup>Department of Medicine, VA Boston Healthcare System, Jamaica Plains, Massachusetts.

<sup>5</sup>Division of Aging, Department of Medicine, Brigham and Women's Hospital, Harvard Medical School, Boston, Massachusetts.

<sup>6</sup>Present address: Department of Clinical Modeling and Simulation, Hoffmann LaRoche, Ltd., Basel, Switzerland.

\*These authors contributed equally to this work.

Clinical Trials Government Identifier: NCT00240981

Address correspondence to Shalender Bhasin, MD, Section of Endocrinology, Diabetes, and Nutrition, Boston University School of Medicine and Boston Medical Center, 670 Albany Street, Boston, MA 02118. Email: [bhasin@bu.edu](mailto:bhasin@bu.edu)

**Context.** Testosterone in Older Men with Mobility Limitations Trial determined the effects of testosterone on muscle performance and physical function in older men with mobility limitation. Trial's Data and Safety Monitoring Board recommended enrollment cessation due to increased frequency of adverse events in testosterone arm. The changes in muscle performance and physical function were evaluated in relation to participant's perception of change.

**Methods.** Men aged 65 years and older, with mobility limitation, total testosterone 100–350 ng/dL, or free testosterone less than 50 pg/mL, were randomized to placebo or 10 g testosterone gel daily for 6 months. Primary outcome was leg-press strength. Secondary outcomes included chest-press strength, stair-climb, 40-m walk, muscle mass, physical activity, self-reported function, and fatigue. Proportions of participants exceeding minimally important difference in study arms were compared.

**Results.** Of 209 randomized participants, 165 had follow-up efficacy measures. Mean (*SD*) age was 74 (5.4) years and short physical performance battery score 7.7 (1.4). Testosterone arm exhibited greater improvements in leg-press strength, chest-press strength and power, and loaded stair-climb than placebo. Compared with placebo, significantly greater proportion of men receiving testosterone improved their leg-press and chest-press strengths (43% vs 18%,  $p = .01$ ) and stair-climbing power (28% vs 10%,  $p = .03$ ) more than minimally important difference. Increases in leg-press strength and stair-climbing power were associated with changes in testosterone levels and muscle mass. Physical activity, walking speed, self-reported function, and fatigue did not change.

**Conclusions.** Testosterone administration in older men with mobility limitation was associated with patient-important improvements in muscle strength and stair-climbing power. Improvements in muscle strength and only some physical function measures should be weighed against the risk of adverse events in this population.

**Key Words:** Testosterone—Minimally important difference—Mobility limitation—Older men—Function promoting therapies.

Received March 12, 2011; Accepted May 11, 2011

Decision Editor: Luigi Ferrucci, MD, PhD

**M**OBILITY limitation, a common geriatric condition, affects an individual's quality of life, health care costs, and resource utilization while being predictive of disability,

risk of hospitalization, and mortality (1–5). Among the function promoting therapies that are being developed for the treatment of functional limitations, androgens are the

farthest in development. Testosterone levels are associated with skeletal muscle mass, lower extremity strength, and physical function (6–12). Men with low free testosterone are at increased risk of mobility limitation and its progression (8). Testosterone supplementation increases skeletal muscle mass (13–21), but its effects on muscle strength and physical function have been inconsistent across trials (13–21). Most testosterone trials have been conducted in healthy older men without functional limitations; the safety and efficacy of testosterone in improving muscle performance and physical function have not been demonstrated in older individuals with mobility limitation.

The *Testosterone in Older Men with Mobility Limitations (TOM) Trial* was a placebo-controlled randomized trial (22,23), whose aim was to determine whether testosterone therapy in older men with mobility limitation and low total or free testosterone levels improves lower extremity muscle strength and physical function (22). As reported recently (23), because of a significantly higher incidence of adverse cardiovascular events in men assigned to the testosterone arm, the trial's Data and Safety Monitoring Board recommended that further enrollment and administration of study medications to participants be discontinued.

Here, we evaluated the clinical meaningfulness of the effects of intervention on performance-based as well as self-reported measures of muscle performance and physical function in relation to participant's perception of change. To determine whether changes in the measures of muscle performance, skeletal muscle mass, and physical function were patient important and thus clinically meaningful, we categorized the participants as "improved" or "not improved" based on whether the change in outcome exceeded the minimally important difference (MID), determined using an anchor-based method within this trial. We analyzed the factors associated with improvements in outcomes and explored the relation between change in efficacy outcomes and adverse events.

## METHODS

The trial's design and safety results have been published (22,23).

### *Study Design*

The TOM Trial was a parallel group, placebo-controlled, double-blind randomized trial, approved by the Institutional Review Board of Boston University Medical Center (BUMC), New England Research Institutes, Watertown, MA, and the Boston Veterans Administration Health Care System (BVAHCS). Participant recruitment took place at BUMC, New England Research Institutes, and BVAHCS, but outcome assessments were performed only at BUMC. All participants provided written informed consent.

### *Eligibility Criteria*

The participants were community-dwelling men, aged 65 years and older, with total testosterone between 100 and 350 ng/dL or free testosterone less than 50 pg/mL, and mobility limitation. The participants were deemed to have mobility limitation if they reported difficulty walking two blocks on a level surface or climbing 10 steps and had a summary score between 4 and 9 on the Short Physical Performance Battery (3), which reflects moderate-to-mild degree of physical dysfunction.

We excluded men who had prostate cancer, lower urinary tract symptom score greater than 21, prostrate specific antigen greater than 4 ng/mL, unstable angina, congestive heart failure, myocardial infarction within 3 months, uncontrolled hypertension, neuromuscular diseases that limited mobility, alanine or aspartate aminotransferase concentrations more than three times the upper limit of normal, creatinine more than 3.5 mg/dL, hemoglobin A1c more than 8.5%, hematocrit more than 48%, untreated severe obstructive sleep apnea, or body mass index more than 40 kg/m<sup>2</sup>. Men using testosterone, growth hormone, or any anabolic therapy or drugs that affect gonadal function were excluded.

### *Randomization and Blinding*

Eligible participants were randomized to either placebo or testosterone gel using a concealed computer-generated randomization table and a block size of 6. Participants were stratified by age (65–75 and >75 years). The participants and outcome assessors were blinded to intervention.

### *Study Intervention*

The participants applied daily transdermal gel containing either placebo or 100 mg testosterone (Testim 1%; Auxilium Pharmaceuticals, Norristown, PA (24)) for 6 months. This regimen of testosterone gel raises total testosterone concentration into the mid-to-high normal range in hypogonadal men (24). To maintain blinding, all participants applied daily three tubes of the gel that were identical in appearance; those assigned to testosterone group applied two tubes each containing 5 g testosterone gel (containing 50 mg testosterone) plus one tube containing placebo gel; and those assigned to placebo group received three tubes containing placebo that were identical in appearance to the tubes containing testosterone. Testosterone was measured 2 weeks after randomization in blood samples drawn 2–4 hours after gel application. If the average of the two testosterone concentrations was less than 500 ng/dL or more than 1,000 ng/dL, the unblinded physician either increased the daily dose to 15 g or decreased it to 5 g.

### *Outcomes*

The primary outcome was change in maximal voluntary strength in the leg-press exercise assessed by the one-repetition maximum (1-RM) method (16,22,25) using pneumatic

resistance machines (Keiser Sport, Fresno, CA). Leg-press strength is important for activities of daily living such as walking and climbing stairs and decreases with advancing age (26,27). Following a 5-minute warm-up, loads were progressively increased until 1-RM, the maximum amount of weight a participant could lift, was achieved. This procedure was also performed for the seated chest-press exercise. Both tests were repeated within 7 days of the initial test.

Physical function was assessed using a 12-step stair-climb, 40-m walk, and a lift-and-lower task (22–25). The stair-climb and walk tests were performed with and without a load equal to 20% body weight. Participants completed two trials of the stair-climb and walk as quickly as possible without running. Walking speed was recorded using a switch mat and infrared timing system (22–25). Stair-climb test required participants to complete two trials of the 12-step staircase ascent as quickly as possible (22–25). Stair-climbing power was calculated as the product of body weight plus weight carried, total stair-rise, divided by ascent time. In the lift-and-lower task, participants completed two trials of lifting and lowering a basket holding a weight equivalent to 15% body weight. The number of shelves completed was recorded.

Lean body mass (LBM) and appendicular lean soft tissue (ALST) mass were determined using dual-energy x-ray absorptiometer (Hologic QDR 4500A), calibrated using a soft tissue phantom (28). ALST was calculated as the sum of the lean masses of the upper and lower extremities.

Self-reported function and disability were measured using Late-Life Function and Disability Index (29). Fatigue was measured using Chalder Fatigue Scale (30).

Measures of muscle performance and physical function were assessed at baseline and during Week 24. Body composition and self-reported measures were assessed at baseline and Weeks 12 and 24. For participants whose study medication was discontinued because of trial's cessation, the efficacy outcomes were assessed at intervention discontinuation, if possible.

Participants with limited exercise tolerance or those at high risk of cardiovascular events underwent a modified testing protocol, which used outcomes that do not elicit greater cardiovascular stress than activities of daily living (22). All others underwent a cardiopulmonary exercise tolerance test to assess evidence for ischemia or exercise intolerance and underwent exercise testing consistent with exercise tolerance test (Figure 1).

#### *Minimally Important Difference*

The anchor-based MID was estimated for each outcome by comparing it against a global rating provided by each participant at the 6-month visit (31; A.M. Jette, N. Latham, S. M. Haley, et al., unpublished data, 2011). Participants rated their perception of change in each outcome over the previous 6 months. Participants were grouped into two cat-

egories based on response to global rating on each outcome: better (those who responded as completely recovered, recovered, much better, or better) and no change or worse (those who responded no better, worse, or much worse). A receiver operating curve compared participants who rated their strength and physical function as “better” at 6 months to those reporting “no change or worse” by examining the mean raw score change on each outcome measure from baseline to 6 months to the global measure of change in function (31; A.M. Jette, N. Latham, S. M. Haley, et al., unpublished data, 2011). We computed the mean MID and percentage of participants whose change from baseline exceeded the MID. The MID was calculated as the mean raw score change in “better” group, and we computed the percentage of participants whose change from baseline exceeded the MID.

#### *Hormone Assays*

Total testosterone level was measured at Quest Diagnostics, San Juan Capistrano, CA, using a Bayer-Advia-Centaur immunoassay with sensitivity 10 ng/dL (32). Sex hormone binding globulin levels were measured using an immunofluorometric assay with sensitivity 2.5 nmol/L (DELFLIA-Wallac, Turku, Finland) (8,13,23). Free testosterone was calculated using a published law of mass action equation (33).

#### *Statistical Analyses*

The planned sample size of 252 participants was based on Type I error of 0.05, 90% power, and assumption of 25 kg difference in the change in leg-press strength (*SD* 55 kg) (23,34). When the data and safety monitoring board recommended enrollment cessation, 209 men had been randomized. Because of participant withdrawal and the early stopping of the trial, a number of participants had missing outcomes data, and in some participants, outcomes data were obtained at an earlier time than the planned 6 months of intervention. The primary analysis included all participants with a baseline and at least one postrandomization measurement (“intention-to-treat [ITT] sample”). This was supplemented by a “per-protocol” analysis restricted to participants who had completed 6 months of study medication (“completion sample”).

Exploratory analyses were conducted using tabular and graphical summaries obtained via generalized additive models. Change in each outcome measure was calculated for each participant and compared across study arms via Fisher's exact and Student's *t* tests. We categorized participants as “improved” or “not improved” depending on whether their change from baseline equaled or exceeded the MID estimates. The proportion of participants “improved” in the two arms was compared using Fisher's exact tests. The associations between baseline levels and changes in LBM and ASLT and total and free testosterone levels and changes in muscle performance and physical function measures were evaluated using generalized additive models and Spearman's rank correlation statistics. Sensitivity analyses

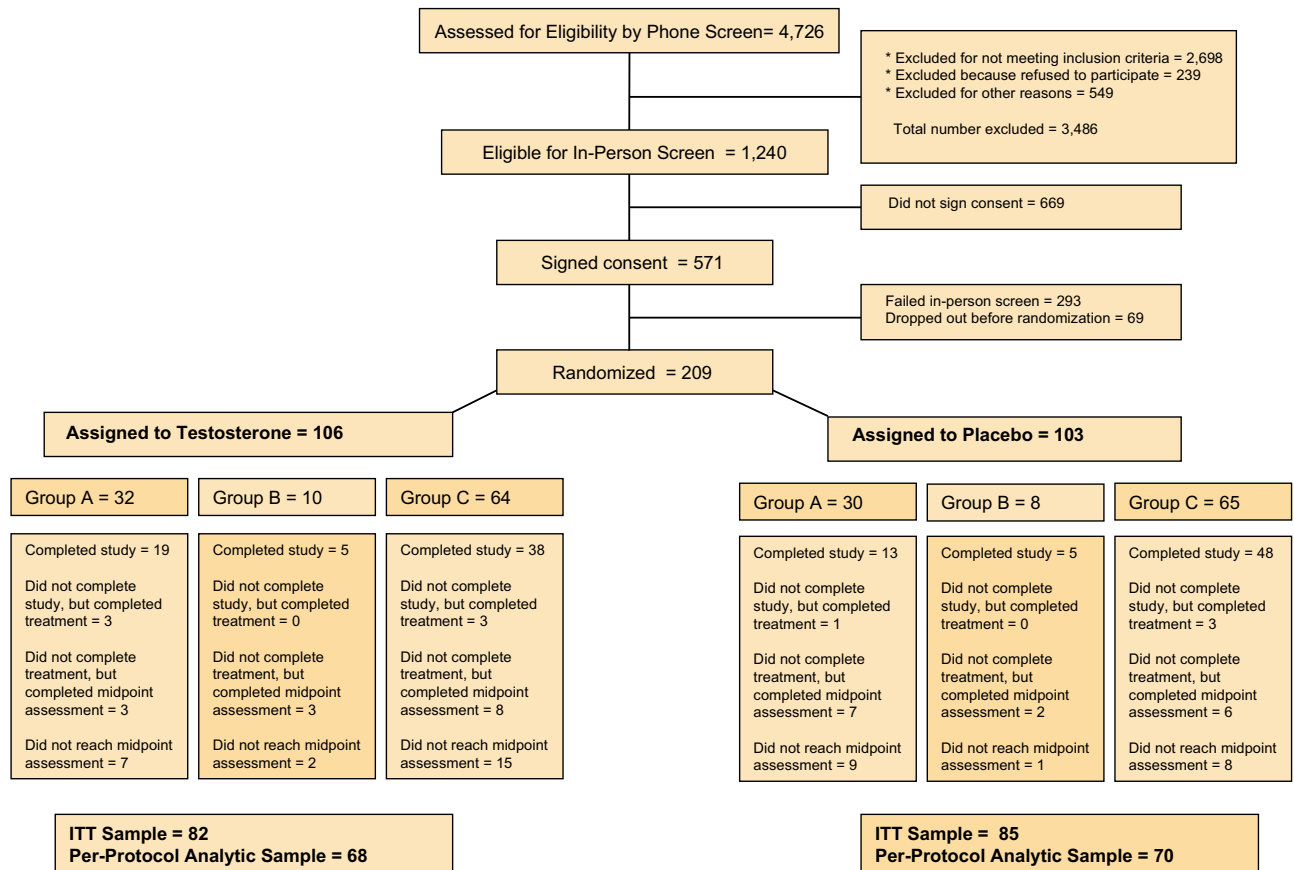


Figure 1. The flow of participants through the TOM Trial is depicted. A two-stage screening process was used. Of 4,726 participants who underwent telephone screening, 1,240 were deemed eligible for further screening, 571 signed consent and were screened in person, 278 met eligibility criteria, and 209 were randomized. Participants with limited exercise tolerance or those at high risk of cardiovascular events underwent a modified testing protocol, which included unloaded 40-m walk test, Bassey's leg power, physical activity, and self-reported measures (Group A). All others underwent a cardiopulmonary exercise tolerance test. Participants who demonstrated evidence of ischemia or exercise intolerance during exercise tolerance test were assigned to Group B and underwent loaded and unloaded walk and stair-climb, leg-press strength and power, and self-reported measures. Participants who underwent an exercise tolerance test without demonstrating ischemic changes underwent complete battery of tests, including loaded and unloaded walk and stair-climb, lift and reach, leg press and chest press, physical activity, and self-reported measures (Group C).

were conducted using multiple linear and logistic regression models.

As the trial was designed with a primary outcome in mind and was stopped early reducing statistical power, no adjustment for multiple comparisons was performed. Results therefore deemphasized hypothesis testing in favor of estimation of effects and comparison of differences vis-à-vis the calculated MID.

## RESULTS

### Flow of Participants Through the Study

The details of the study have been published (23). Briefly, when data and safety monitoring board recommended cessation of further enrollment, 4,726 men had been screened, 278 had met eligibility criteria, and 209 had been randomized, 106 to testosterone, and 103 to placebo. Hundred and twenty-eight men completed all phases of the study, and 10 men had completed the 6-month intervention phase; these

138 participants constituted the completion sample. An additional 27 men had undergone at least one postrandomization assessment (Figure 1). These 165 men constituted the ITT sample.

### Baseline Characteristics

Table 1 shows the baseline characteristics of the ITT and completion samples. The mean (*SD*) age of the ITT sample was 74 (5.4) years. The mean (*SD*) Short Physical Performance Battery score of 7.7 (1.4) indicates a moderate level of physical dysfunction. Fifty-four percent of men had walking speed less than 1.0 m/s.

### Compliance

Participants using more than 90% of the prescribed gel tubes were deemed compliant. The proportion of men compliant by this criterion exceeded 90% in both groups (23).

Table 1. Baseline Characteristics of the Participants Included in the Efficacy Analyses

	ITT Sample (N = 165)		Completer Sample (N = 138)	
	Testosterone (n = 82)	Placebo (n = 83)	Testosterone (n = 69)	Placebo (n = 69)
Age, y	73.6 ± 5.8	74.1 ± 5	73.8 ± 5.8	73.9 ± 4.9
Body mass index, kg/m <sup>2</sup>	29.6 ± 3.9	29.9 ± 4.2	29.2 ± 4	30.1 ± 4.4
Total testosterone, nmol/L*	8.6 ± 2.1	8.2 ± 2.3	8.7 ± 2.2	8.0 ± 2.3
Free testosterone, pmol/L†	165.4 ± 40.2			141.8 ± 44.0
Lean body mass, kg‡	55.1 ± 6.9	56.5 ± 6.5 (81)	54.5 ± 7 (69)	57 ± 6.8 (67)
ASLT, kg‡	23.7 ± 3.3	24.1 ± 3 (81)	23.4 ± 3.4 (69)	24.3 ± 3.2 (67)
SPPB score	7.7 ± 1.4	7.7 ± 1.4	7.6 ± 1.5 (69)	7.6 ± 1.5 (69)
Leg-press strength, N‡	1913.9 ± 431.4 (59)	1968.8 ± 346.9 (60)	1908.9 ± 405.6 (51)	1971.7 ± 336.9 (49)
Chest-press strength, N‡	417.8 ± 100 (52)	421.2 ± 87.9 (53)	418.1 ± 95.4 (45)	428.3 ± 90.3 (45)
Leg-press power, W‡	499.5 ± 160.3 (58)	507.7 ± 133.5 (59)	491.5 ± 150.3 (50)	513.1 ± 123.9 (49)
Chest-press power, W‡	161.9 ± 43.7 (52)	163.7 ± 43.2 (52)	159.4 ± 41.3 (45)	167.3 ± 44.8 (45)
Dominant hand grip strength, kg‡	27.6 ± 6.9 (75)	26.5 ± 7.6 (75)	27.6 ± 6.7 (62)	26.2 ± 8 (61)
Unloaded 40-m walking speed, m/s‡	1.7 ± 0.4 (72)	1.7 ± 0.4 (76)	1.6 ± 0.4 (60)	1.7 ± 0.4 (63)
Loaded 40-m walking speed, m/s‡	1.6 ± 0.4 (55)	1.6 ± 0.4 (57)	1.6 ± 0.4 (49)	1.7 ± 0.4 (46)
Unloaded stair-climb power, W‡	322.9 ± 114.6 (57)	322 ± 92.1 (58)	312.8 ± 112.9 (50)	324.1 ± 86.9 (47)
Loaded stair-climb power, W‡	353.3 ± 138.2 (55)	348.8 ± 116.3 (56)	346.7 ± 134.9 (49)	357.7 ± 109.7 (46)
Lift-lower score‡	25.2 ± 9 (50)	23.6 ± 10.1 (51)	25.2 ± 8.7 (44)	23.1 ± 10.2 (43)
Late Life Functional Disability Index	60.5 ± 9.6 (67)	62.6 ± 10.2 (68)	61.2 ± 9.9 (56)	61.7 ± 9.1 (55)
Fatigue score	4.9 ± 3.1 (77)	4.2 ± 3.2 (77)	4.8 ± 3.1 (64)	4.2 ± 3.3 (65)

Notes: ALST = appendicular lean soft tissue; ITT = intention-to-treat; and SPPB = Short Physical Performance Battery.

\* Measured in a single morning sample drawn between 7 and 11 AM.

† Calculated via mass action equation.

‡ Number of nonmissing records shown in parentheses.

### Testosterone Levels

Total and free testosterone levels increased significantly more in testosterone arm than in placebo arm (Supplementary Table 1).

### Efficacy Outcomes

The increase in leg-press strength was greater in men assigned to testosterone arm than in those assigned to placebo arm, whether the change is expressed on a proportionate or an absolute scale (Figure 2); results were similar in the completion sample (not shown). The percent of men whose leg-press strength improved more than the MID was significantly greater in the testosterone group (43%) than in the placebo group (18%,  $p = .01$ ; Figure 2). The men whose leg-press strength improved more than the MID did not differ significantly in their baseline body composition, muscle performance, or physical function from those who improved less than the MID (not shown). Among participants assigned to testosterone arm, increases in total and free testosterone were associated with increased leg-press strength, appendicular skeletal muscle mass, and loaded stair-climb power (Supplementary Table 2). Lean mass gains and fat mass losses were significantly greater in the testosterone than in the placebo arm (Table 2).

The proportion of men whose chest-press strength improved more than the MID was greater in the testosterone group than in the placebo group (Figure 2). Changes in chest-press strength were associated with changes in total ( $r = .34$ ,  $p = .002$ ) and free testosterone ( $r = .36$ ,  $p = .001$ ) and changes in LBM ( $r = .42$ ,  $p = .0001$ ) and ALST ( $r = .36$ ,  $p = .001$ ).

A greater proportion of men randomized to testosterone arm improved more than the MID in their loaded stair-climbing power than those randomized to placebo ( $p = .03$ ); those achieving the MID exhibited significantly greater increases in total and free testosterone, LBM and ALST, and leg-press and chest-press strengths but did not differ in baseline measures (not shown). Changes in loaded stair-climbing power were significantly related to changes in total and free testosterone, LBM and ALST, and leg-press strength (Supplementary Table 2).

Changes in unloaded walking speed and unloaded stair-climbing did not differ significantly between the two groups. However, changes in loaded walking speed and loaded stair-climbing power tracked with changes in leg-press strength (Supplementary Table 2).

The self-reported measures of functional disability and fatigue and physical activity counts did not differ between arms (Figure 2).

### Treatment Effect Sizes

The differences between testosterone and placebo groups for leg and chest press, lean mass and ASLT, and physical function measures ranged from 0.25 to 0.35 *SD* units (Figure 3).

### Relation of Cardiovascular Adverse Events to Treatment Responsiveness

As reported previously (23), cardiovascular-related adverse events occurred in 28 participants (23 in testosterone and 5 in placebo arm); of these, 22 (17 testosterone and

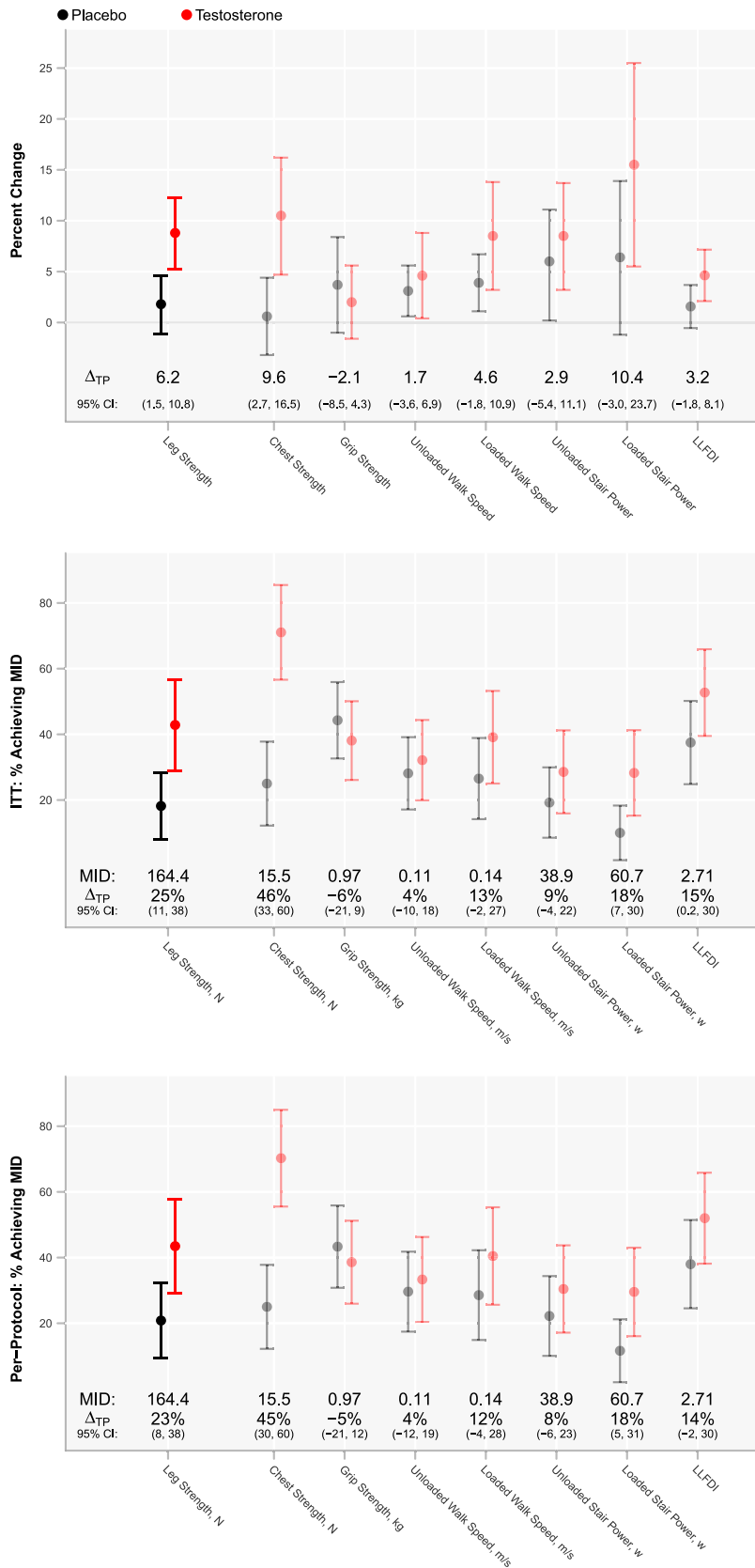


Figure 2. The percent change from baseline in selected measures of muscle performance and physical function in testosterone and placebo arms of the trial. Data are mean and 95% confidence intervals. Estimates of mean treatment difference attributable to testosterone administration ( $\Delta_{TP}$ ) are presented along with 95% confidence intervals. These may be interpreted as estimates of difference (testosterone minus placebo) in proportionate change achieved over 6 months.

Table 2. Changes in Measures of Skeletal Muscle Mass and Fat Mass by Dual-Energy x-Ray Absorptiometer

Body Composition Measurement (kg)	Testosterone, Mean $\pm$ SD			Placebo, Mean $\pm$ SD			Treatment Difference, Midtreatment		Treatment Difference, Posttreatment	
	Baseline	Midtreatment	Posttreatment	Baseline	Midtreatment	Posttreatment	Mean(CI)	<i>p</i>	Mean(CI)	<i>p</i>
ITT sample	(N = 82)	(N = 73)	(N = 77)	(N = 81)	(N = 75)	(N = 80)				
Total lean mass	55.1 $\pm$ 6.9	57.2 $\pm$ 7.1	56.1 $\pm$ 6.9	56.5 $\pm$ 6.5	56.7 $\pm$ 6.2	56.2 $\pm$ 6.3	1.8 (1.2, 2.5)	.2	1.5 (0.4, 2.2)	<.0001
Total fat mass	26.2 $\pm$ 7.5	24.9 $\pm$ 7.1	24.0 $\pm$ 6.8	28 $\pm$ 8.4	28 $\pm$ 8.1	27.5 $\pm$ 7.9	-1.7 (-2.7, -0.6)	.002	-2.0 (-2.8, -1.2)	<.0001
ASMM	23.7 $\pm$ 3.3	24.7 $\pm$ 3.3	24.2 $\pm$ 3.3	24.1 $\pm$ 3	24.1 $\pm$ 3	23.9 $\pm$ 3.1	1.1 (0.7, 1.5)	<.0001	0.9 (0.4, 1.4)	.0009
Appendicular fat mass	9.7 $\pm$ 2.8	9.3 $\pm$ 2.8	8.9 $\pm$ 2.6	10.7 $\pm$ 3.5	10.6 $\pm$ 3	10.5 $\pm$ 3	-0.3 (-0.6, -0.01)	.04	-0.8 (-1.2, -0.4)	.0003
Completion sample	(N = 69)	(N = 62)	(N = 69)	(N = 67)	(N = 63)	(N = 68)				
Total lean mass	54.5 $\pm$ 7	56.7 $\pm$ 7.2	55.7 $\pm$ 7.1	57 $\pm$ 6.8	57.1 $\pm$ 6.5	56.5 $\pm$ 6.6	2.0 (1.3, 2.8)	<.0001	1.6 (0.8, 2.4)	<.0001
Total fat mass	25.6 $\pm$ 7.4	24.1 $\pm$ 6.9	23.5 $\pm$ 6.6	28.5 $\pm$ 8.6	28.7 $\pm$ 8.1	28.1 $\pm$ 8	-1.9 (-3.1, -0.6)	<.0001	-2.2 (-1.4, -3.1)	<.0001
ASMM	23.4 $\pm$ 3.4	24.5 $\pm$ 3.4	24.2 $\pm$ 3.4	24.3 $\pm$ 3.2	24.3 $\pm$ 3.1	24 $\pm$ 3.3	1.2 (0.8, 1.6)	<.0001	1.0 (0.4, 1.5)	.0004
Appendicular fat mass	9.5 $\pm$ 2.9	9.1 $\pm$ 2.8	8.7 $\pm$ 2.6	10.9 $\pm$ 3.6	10.8 $\pm$ 3.1	10.7 $\pm$ 3	-0.4 (-0.7, -0.04)	.02	-0.8 (-1.2, -0.4)	.0003

Notes: ASMM = appendicular skeletal muscle mass; CI = confidence interval; ITT = intention to treat.

5 placebo) were included in the ITT sample. There were no significant differences between these participants and others for any outcome, although the statistical power for these comparisons was limited due to the small number of participants with adverse events.

## DISCUSSION

In this population of older men with mobility limitations and high burden of chronic diseases, the men assigned to the testosterone arm experienced greater gains in leg-press strength, chest-press strength and power, skeletal muscle

mass, and loaded stair-climbing power than those assigned to placebo. Because of early trial cessation, the preplanned enrollment target of 252 randomized men was not met; furthermore, a substantial fraction of randomized participants did not complete the planned 6 months of intervention. In spite of the early cessation of enrollment and earlier than planned discontinuation of intervention in many participants, changes in muscle strength and stair-climbing power associated with randomization to testosterone arm of the study were statistically significant. As a greater proportion of men experienced improvements in leg-press and chest-press strengths and stair-climbing power that exceeded the MID, these treatment effects are patient important and therefore clinically meaningful. Walking speed, a key determinant of mobility, did not change significantly. The improvements in muscle mass and strength without significant improvements in walking speed should be weighed against the greater risk of adverse cardiovascular events observed among men assigned to testosterone arm.

The changes in maximal voluntary strength and loaded stair-climbing power were related to changes in testosterone concentrations. The muscle strength gains were related to gains in skeletal muscle mass, which were related to changes in total and free testosterone concentrations. Changes in stair-climbing power and walking speed were related to changes in leg-press strength, which is an important determinant of stair-climbing power and walking speed. These correlational analyses are consistent with the following mechanistic directionality: increases in testosterone levels  $\rightarrow$  gains in skeletal muscle mass  $\rightarrow$  increase in muscle strength  $\rightarrow$  improved physical function.

Meta-analyses of testosterone trials have reported significant gains in LBM but inconsistent changes in muscle strength and physical function (35,36). The trials in these meta-analyses were limited by the small number of participants, shorter treatment durations than the 6-month intervention period used in the TOM Trial, or the failure to include comprehensive assessments of muscle strength and

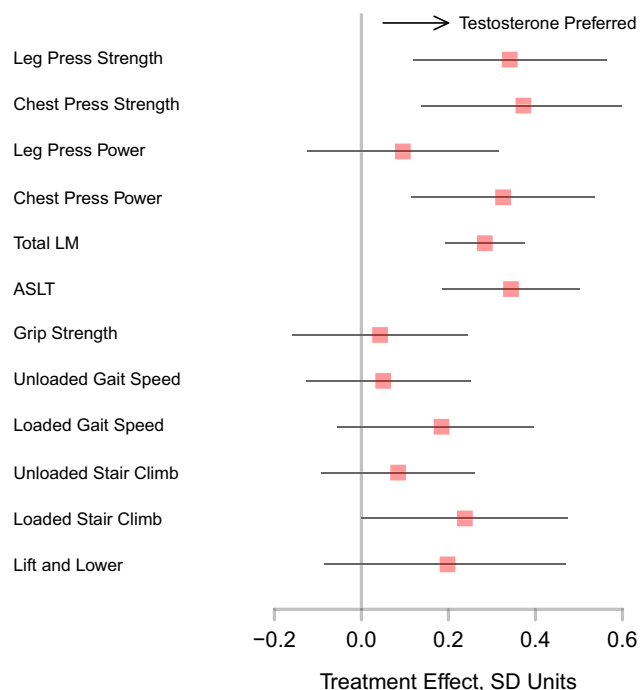


Figure 3. Absolute treatment differences (testosterone vs placebo arms) are plotted for the primary and secondary outcomes in units normalized to the baseline standard deviation of measurement; point estimates (red) are accompanied by 95% confidence intervals.

physical function. Some trials used small doses of testosterone that resulted in significantly smaller increments in testosterone levels than were achieved in the TOM Trial (18,19). A unique aspect of this trial is the use of MID estimates derived using an anchor within this trial to determine the clinical meaningfulness of the observed treatment effects. Additionally, the trial included a comprehensive assessment of upper and lower extremity function and performance-based as well as self-reported measures.

Most testosterone trials have enrolled healthy older men; this is the first randomized testosterone trial in older men with mobility limitations. A recent important testosterone trial in older men with frailty reported improvements in LBM and some self-reported measures of physical function, but the trial found no significant differences in muscle strength or performance-based measures of physical function between placebo and testosterone arms (14). A significant fraction of participants in that trial were prefrail. Another trial in frail elderly men also failed to find significant improvement in muscle strength or physical function measures (37); the testosterone dose in that trial was lower than that used in the TOM Trial. Small testosterone trials in men with congestive heart failure have suggested improvements in exercise capacity with inconsistent changes in muscle mass and strength (38).

In spite of substantial gains in muscle mass and maximal voluntary strength in men assigned to testosterone arm, several measures of physical function did not improve significantly beyond the improvements seen in the placebo arm. Other factors, such as neuromuscular integration and functional training, might be required to optimize the translation of muscle mass and strength gains into improved physical function. It is possible that neuromuscular adaptations that result in improved function may require more time than the 6-month duration of our trial. Page and colleagues reported significant improvement in a continuous timed physical performance test after 12 months of testosterone administration (20), but other studies of 1- to 3-year duration (17–19) have failed to note improvements in functional measures.

In our trial, treatment effects for most measures of physical function favored testosterone administration; in general, the measures having the highest ceiling (loaded stair-climb and loaded gait speed) showed greater testosterone effect than unloaded tests with lower ceilings. As substantial numbers of randomized participants did not complete the trial due to early study cessation, this may have reduced statistical power. However, close examination of proportional gains in functional outcomes indicates that variation in the unloaded functional measures was comparable to that in the strength indices. Thus, there were indeed gains in the functional measures among participants in the testosterone arm, but there were comparable gains in function among participants assigned to placebo (whereas participants assigned to placebo evinced little to no gains in strength; Figure 2). Thus,

the estimated treatment effects (testosterone vs placebo) for unloaded walk, unloaded stair-climb, and late life function and disability index were of modest magnitude. By contrast, proportionate variation in the loaded functional measures—particularly stair-climb power—was somewhat greater than that in the strength measures, so that sizeable estimates of treatment effect (testosterone vs placebo) did not achieve statistical significance due to the large variation. One may postulate that had the trial achieved full enrollment, these differences in the loaded measures would have achieved statistical significance, whereas the available evidence suggests that the unloaded measures would have failed to demonstrate efficacy even under full enrollment.

Several factors merit consideration in weighing the observed improvements in muscle strength and physical function against the increased risk of cardiovascular events in men assigned to the testosterone arm. The number needed to treat has been used to describe the efficacy of health care interventions. In the TOM Trial, the numbers needed to treat to achieve clinically meaningful improvements in leg-press strength and stair-climbing power were 4.0 and 5.5, respectively. However, the gains in skeletal muscle mass and muscle strength were not associated with significant improvements in walking speed. Adjunctive strategies, such as physical activity or other interventions (39), may be needed to induce neuromuscular, cognitive, and behavioral adaptations that are necessary for translating muscle mass and strength gains induced by testosterone into functional improvements. Also, the testosterone administration at the dose used in the trial was associated with adverse events. Therefore, the clinical application of testosterone as a function promoting anabolic therapy might be predicated upon strategies, which augment the anabolic effects of testosterone and facilitate translation of testosterone-induced gains in muscle mass and strength into functional improvements at lower testosterone concentrations that can be safely administered. Such strategies, which might include physical activity interventions (39), resistance exercise training, cognitive and behavioral training, or combined administration of testosterone with other anabolic agents, such as recombinant human growth hormone, should be investigated.

#### FUNDING

This study was supported primarily by a grant from the National Institutes on Aging administered under a cooperative agreement (1U01AG14369). Additional support was provided by Boston Claude D. Pepper Older Americans Independence Center grant (5P30AG031679) and a BU Clinical and Translational Science Institute grant (1UL1RR025771). Testosterone and placebo gel for the study were provided by Auxilium Pharmaceuticals, Inc., Norristown, PA. The MID Substudy was supported in part by a research grant from Merck and Co. to Boston University. A part of the work was supported by the resources and facilities of the VA Boston Healthcare System.

#### SUPPLEMENTARY MATERIAL

Supplementary material can be found at: <http://biomedgerontology.oxfordjournals.org/>



## ACKNOWLEDGMENTS

Data and Safety Monitoring Board members: Eric Orwoll (Chair), Anne Newman, Kenneth Schechtman, Wayne Meikle, and Mark Litwin. We thank Drs. Sergei Romashkan and Evan Hadley of the National Institute in Aging for their oversight and guidance throughout the duration of the trial; Lindsay Cloutier, Daniela Ciccolini, Newsha Lajevardi, Dr. Sharon Tennstedt, and the staff of the General Clinical Research Unit of the Boston University's Clinical and Translational Science Institute for their help in the conduct of these studies; and the study participants for their commitment and generosity.

## REFERENCES

- Gardener EA, Huppert FA, Guralnik JM, Melzer D. Middle-aged and mobility-limited: prevalence of disability and symptom attributions in a national survey. *J Gen Intern Med*. 2006;21:1091–1096.
- Kelly-Hayes M, Jette AM, Wolf PA, D'Agostino RB, Odell PM. Functional limitations and disability among elders in the Framingham Study. *Am J Public Health*. 1992;82:841–845.
- Guralnik JM, Ferrucci L, Simonsick EM, Salive ME, Wallace RB. Lower-extremity function in persons over the age of 70 years as a predictor of subsequent disability. *N Engl J Med*. 1995;332:556–561.
- Studenski S, Perera S, Patel K, et al. Gait speed and survival in older adults. *JAMA*. 2011;305:50–58.
- Newman AB, Simonsick EM, Naydeck BL, et al. Association of long-distance corridor walk performance with mortality, cardiovascular disease, mobility limitation, and disability. *JAMA*. 2006;295:2018–2026.
- Roy TA, Blackman MR, Harman SM, Tobin JD, Schragger M, Metter EJ. Interrelationships of serum testosterone and free testosterone index with FFM and strength in aging men. *Am J Physiol Endocrinol Metab*. 2002;283:E284–E294.
- Baumgartner RN, Waters DL, Gallagher D, Morley JE, Garry PJ. Predictors of skeletal muscle mass in elderly men and women. *Mech Ageing Dev*. 1999;107:123–136.
- Schaap LA, Pluijm SM, Smit JH, et al. The association of sex hormone levels with poor mobility, low muscle strength and incidence of falls among older men and women. *Clin Endocrinol (Oxf)*. 2005;63:152–160.
- Cawthon PM, Ensrud KE, Laughlin GA, et al. Sex hormones and frailty in older men: the osteoporotic fractures in men (MrOS) study. *J Clin Endocrinol Metab*. 2009;94:3806–3815.
- Krasnoff JB, Basaria S, Pencina MJ, et al. Free testosterone levels are associated with mobility limitation and physical performance in community-dwelling men: the Framingham Offspring Study. *J Clin Endocrinol Metab*. 2010;95:2790–2799.
- Orwoll E, Lambert LC, Marshall LM, et al. Endogenous testosterone levels, physical performance, and fall risk in older men. *Arch Intern Med*. 2006;166:2124–2131.
- Araujo AB, Travison TG, Bhasin S, et al. Association between testosterone and estradiol and age-related decline in physical function in a diverse sample of men. *J Am Geriatr Soc*. 2008;56:2000–2008.
- Bhasin S, Woodhouse L, Casaburi R, et al. Older men are as responsive as young men to the anabolic effects of graded doses of testosterone on the skeletal muscle. *J Clin Endocrinol Metab*. 2005;90:678–688.
- Srinivas-Shankar U, Roberts SA, Connolly MJ, et al. Effects of testosterone on muscle strength, physical function, body composition, and quality of life in intermediate-frail and frail elderly men: a randomized, double-blind, placebo-controlled study. *J Clin Endocrinol Metab*. 2010;95:639–650.
- Sih R, Morley JE, Kaiser FE, Perry HM 3rd, Patrick P, Ross C. Testosterone replacement in older hypogonadal men: a 12-month randomized controlled trial. *J Clin Endocrinol Metab*. 1997;82:1661–1667.
- Storer TW, Magliano L, Woodhouse L, et al. Testosterone dose-dependently increases maximal voluntary strength and leg power, but does not affect fatigability or specific tension. *J Clin Endocrinol Metab*. 2003;88:1478–1485.
- Snyder PJ, Peachey H, Hannoush P, et al. Effect of testosterone treatment on body composition and muscle strength in men over 65 years of age. *J Clin Endocrinol Metab*. 1999;84:2647–2653.
- Nair KS, Rizza RA, O'Brien P, et al. DHEA in elderly women and DHEA or testosterone in elderly men. *N Engl J Med*. 2006;355:1647–1659.
- Emmelot-Vonk MH, Verhaar HJ, Nakhai Pour HR, et al. Effect of testosterone supplementation on functional mobility, cognition, and other parameters in older men: a randomized controlled trial. *JAMA*. 2008;299:39–52.
- Page ST, Amory JK, Bowman FD, et al. Exogenous testosterone (T) alone or with finasteride increases physical performance, grip strength, and lean body mass in older men with low serum T. *J Clin Endocrinol Metab*. 2005;90:1502–1510.
- Ferrando AA, Sheffield-Moore M, Yeckel CW, et al. Testosterone administration to older men improves muscle function: molecular and physiological mechanisms. *Am J Physiol Endocrinol Metab*. 2002;282:E601–E607.
- LeBrasseur NK, Lajevardi N, Miciek R, Mazer N, Storer TW, Bhasin S. Effects of testosterone therapy on muscle performance and physical function in older men with mobility limitations (The TOM Trial): design and methods. *Contemp Clin Trials*. 2009;30:133–140.
- Basaria S, Coviello AD, Travison TG, et al. Adverse events associated with testosterone administration. *N Engl J Med*. 2010;363:109–122.
- Steidle C, Schwartz S, Jacoby K, Seebree T, Smith T, Bachand R. AA2500 testosterone gel normalizes androgen levels in aging males with improvements in body composition and sexual function. *J Clin Endocrinol Metab*. 2003;88:2673–2681.
- LeBrasseur NK, Bhasin S, Miciek R, Storer TW. Tests of muscle strength and physical function: reliability and discrimination of performance in younger and older men and older men with mobility limitations. *J Am Geriatr Soc*. 2008;56:2118–2123.
- Danneskiold-Samsøe B, Kofod V, Munter J, Grimby G, Schnohr P, Jensen G. Muscle strength and functional capacity in 78–81-year-old men and women. *Eur J Appl Physiol Occup Physiol*. 1984;52:310–314.
- Wolfson L, Judge J, Whipple R, King M. Strength is a major factor in balance, gait, and the occurrence of falls. *J Gerontol A Biol Sci Med Sci*. 1995;50:64–67.
- Kim J, Wang Z, Heymsfield SB, Baumgartner RN, Gallagher D. Total-body skeletal muscle mass: estimation by a dual-energy X-ray absorptiometry method. *Am J Clin Nutr*. 2002;76:378–383.
- Haley SM, Jette AM, Coster WJ, et al. Late life function and disability instrument: II. Development and evaluation of the function component. *J Gerontol A Biol Sci Med Sci*. 2002;57:M217–M222.
- Chalder T, Berelowitz G, Pawlikowska T, et al. Development of fatigue scale. *J Psychosom Res*. 1993;37:147–153.
- Revicki D, Hays RD, Cella D, Sloan J. Recommended methods for determining responsiveness and minimally important differences for patient-reported outcomes. *J Clin Epidemiol*. 2008;61:102–109.
- Salameh WA, Redon-Goldman MM, Clarke NJ, Reitz RE, Caulfield MP. Validation of a total testosterone assay using high turbulence liquid chromatography tandem mass spectrometry: total and free testosterone reference ranges. *Steroids*. 2010;75:165–175.
- Mazer NA. A novel spreadsheet method for calculating the free serum concentrations of testosterone, dihydrotestosterone, estradiol, estrone and cortisol: with illustrative examples from male and female populations. *Steroids*. 2009;74:512–519.
- Bhasin S, Storer TW, Javanbakht M, et al. Testosterone replacement and resistance exercise in HIV-infected men with weight loss and low testosterone levels. *JAMA*. 2000;283:763–770.
- Bhasin S, Calof OM, Storer TW, et al. Testosterone and selective androgen receptor modulators as anabolic therapies for chronic illness and aging. *Nat Clin Pract Endocrinol Metab*. 2006;2:146–159.

36. Isidori AM, Giannetta E, Greco EA, et al. Effects of testosterone on body composition, bone metabolism and serum lipid profile in middle-aged men: a meta-analysis. *Clin Endocrinol (Oxf)*. 2005; 63:280–293.
37. Kenny AM, Kleppinger A, Annis K, et al. Effects of transdermal testosterone on bone and muscle in older men with low bioavailable testosterone levels, low bone mass, and physical frailty. *J Am Geriatr Soc*. 2010;58:1134–1143.
38. Malkin CJ, Pugh PJ, West JN, van Beek EJ, Jones TH, Channer KS. Testosterone therapy in men with moderate severity heart failure: a double-blind randomized placebo controlled trial. *Eur Heart J*. 2006; 27:57–64.
39. LIFE Study Investigators, Pahor M, Blair SN, et al. Effects of a physical activity intervention on measures of physical performance: results of the lifestyle interventions and independence for Elders Pilot (LIFE-P) study. *J Gerontol A Biol Sci Med Sci*. 2006;61:1157–1165.