Supplementary Appendix

This appendix has been provided by the authors to give readers additional information about their work.

Supplement to: Snyder PJ, Bhasin S, Cunningham GR, et al. Effects of testosterone treatment in older men. N Engl J Med 2016;374:611-24. DOI: 10.1056/NEJMoa1506119

EFFECTS OF TESTOSTERONE TREATMENT OF OLDER MEN

SUPPLEMENTARY APPENDIX TABLE OF CONTENTS

ROLES OF INVESTIGATORS AND OTHERS IN THE TESTOSTERONE TRIALS

Steering Committee

Peter J. Snyder, MD, Chair	University of Pennsylvania
Elizabeth Barrett-Connor, MD	University of California, San Diego
Shalendar Bhasin, MD	Harvard Medical School
Jane A Cauley, DrPH	University of Pittsburgh
David Cella, PhD	Northwestern University
Jill P Crandall, MD	Albert Einstein College of Medicine
Glenn Cunningham, MD	Baylor Medical School
Susan S Ellenberg, PhD	University of Pennsylvania
Kristine E Ensrud, MD, MPH	University of Minnesota
John T Farrar, MD, PhD	University of Pennsylvania
Thomas M Gill, MD	Yale University
Cora E Lewis, MD, MSPH	University of Alabama at Birmingham
Alvin M Matsumoto, MD	University of Washington
Mark E Molitch, MD	Northwestern University
Marco Pahor, MD	University of Florida
Susan Resnick, PhD	National Institute of Aging
Raymond C Rosen, PhD	New England Research Institute
Ronald S Swerdloff, MD	University of California, Los Angeles

Statisticians (Department of Biostatistics and Epidemiology, University of Pennsylvania)

Susan S Ellenberg, PhD	
Xiaoling Hou, MS	Alisa Stephens-Shields, PhD
J Richard Landis, PhD	Bret Zeldow, MS
Liyi Cen, MD	Renee H Moore, PhD

Data Coordinating Center (Clinical Research Computing Unit, Department of Biostatistics and Epidemiology, University of Pennsylvania)

Denise Cifelli, MS	Darlene Dougar, MPH
Laura Fluharty, MPH	Laura Gallagher, MPH, CCRP
Shawn Ballard, MS	Tracy Chai, MS
James Dattilo, BS	Trina Brown
Sandra Smith, AS	Fran Chicchi, BS

Adjudicators of Cardiovascular and Cerebrovascular Adverse Events

Scott E Kasner, MD	University of Pennsylvania
Cora E Lewis, MD, MSPH	University of Alabama at Birmingham
Steven R Messe, MD	University of Pennsylvania
Emile R Mohler III, MD	University of Pennsylvania

Clinical Trial Sites

Site		PI	Co-Is and Sub-Is	Research Coordinator(s)
11	Albert Einstein College of Medicine	Jill P. Crandall, MD	Vafa Tabatabaie, MD Eric Epstein, MD Uriel Barzel, MD	Gilda Trandafirescu
12	Baylor College of Medicine	Glenn R. Cunningham, MD	N/A	Emilia Cordero, MS, RN, ANP-C Patti Marino
13	Brigham and Women's Hospital	Shalender Bhasin, MD	Shehzad Basaria, MD	Richard Eder Erica Appleman Kathleen Ann Halley
14	Northwestern University	Mark Molitch, MD	Daniel Toft, MD Amisha Wallia, MD	Diane Larsen Elaine Massaro, MS, RN Daphne Adelman, BSN, MBA
15	University of Alabama at Birmingham	Cora E. Lewis, MD, MSPH	James Shikany, DrPH Peter Kolettis, MD	Phillip Johnson Margaret N. Pike, RN Isabelle J. Joffrion, RN
16	University of California at Los Angeles	Ronald Swerdloff, MD	Christina Wang, MD	Xiaodan Han, RN Jamila Ashai
17	University of California at San Diego	Elizabeth Barrett-Connor, MD	Karen Herbst, MD Heather Hofflich, DO J. Kellogg Parsons, MD	Noralinda Kamantigue, BSN, RN Mary Lou Carrion-Peterson, BSN, RN Gabriela Reno Lauren Claravall Jean Smith, RN
18	University of Florida	Marco Pahor, MD	Susan Nayfield, MD, M.Sc Stephen D. Anton, Ph.D. Todd Manini, Ph.D Philip Dahm, MD Michael Marsiske, Ph.D Bhanuprasad Sandesara, MD	Melissa Lewis Mieniecia L. Black, MPH Jeffrey Knaggs William Marena Jane Ching-ju Lu, MPH
19	University of Minnesota	Kristine E. Ensrud, MD, MPH	Susan J. Diem, MD, MPH Howard Fink, MD, MPH Christopher Warlick, MD	Sandra Potter, CCRC Luanne Welch, RN Pamela Van Covering, MA Kristi Lee Jacobson Lisa Miller
20	University of Pittsburgh	Jane Cauley, DrPH	Mara J. Horwitz, MD Susan L. Greenspan, MD Thomas M. Jaffe, MD	Linda Prebehalla, RN Janet T. Bonk, RN, MPH Jennifer L. Rush, MPH
21	University of Washington	Alvin M. Matsumoto, MD	N/A	Janet Gilchriest, RN Kathy Winter Magdalena Wojtowicz, RN
22	Yale University	Thomas M. Gill, MD	Natalie deRekeneire, MD Susan Kashaf, MD Lee Katz, MD Hamid Mojibian, MD	Joanne McGloin Karen Wu Dismayra Martinez, MA Denise Shepard, BSN, MBA

Data and Safety Monitoring Board

Member	Affiliation
Michael O. Thorner, MD	Department of Medicine, University of Virginia
Eric Klein, MD	Cleveland Clinic
Wayne A. Meikle, MD	University of Utah
Wayne J. Hellstrom, MD	Department of Urology, Tulane University
J. Philip Miller, PhD (Chair)	Washington University
Richard Chappell, PhD	Department of Biostatistics, University of Wisconsin
George A. Kuchel, MD	Division of Geriatrics, University of Connecticut
Manuel D. Cerqueira, MD	Cleveland Clinic Nuclear Medicine
E. Magnus Ohman, MD	Duke University Medical Center

BRIGHAM RESEARCH ASSAY CORE

Steroid Assay Procedures

Measurement of Total Testosterone by a Centers for Disease Control (CDC) - Certified Liquid Chromatography Tandem Mass Spectrometry (LC-MS/MS) Assay

Serum total testosterone concentration was measured using a previously published LC-MS/MS assay, which has been certified by the Centers for Disease Control's HoST Program.[1-4] Testosterone in human serum is extracted by solid phase extraction (SPE), separated by high performance liquid chromatography (HPLC), and measured by mass spectrometry (MS) in electrospray ionization (ESI^+) source. Deuterated testosterone (16,16, 17-D_3 testosterone) was used as an internal standard for the calibration of assay.

The lower limit of quantitation was 1 ng/dL, and linear range 1 to 1000 ng/dL. The recovery was calculated by adding known amounts of testosterone to charcoal stripped serum and analyzing by LC-MS/MS was 102 +/- 2%. The correlation between the amount added and the amount measured by LC-MS/MS was 0.998. The cross-reactivity of DHEA, DHEAS, DHT, androstenedione and estradiol in the testosterone assay is negligible at ten times the circulating concentrations of these hormones. There is no measurable carryover effect on the blank after running the highest standard (2000 ng/dL). The assay has been validated for human serum and human heparin plasma. These two matrices exert no significant interference. The interassay coefficient of variation is 7.9% at 48.6 ng/dL, 7.7% at 241 ng/dL, 4.4% at 532 ng/dL, and 3.3% at

1016 ng/dL respectively. The interassay CV was <5% for the NIST female standard 971 and <2% for NIST male standard 971. As part of the Centers for Disease Control's (CDC) Testosterone Assay Harmonization Initiative, quality control samples provided by the CDC were run every three months; the coefficient of variation in quality control samples with testosterone concentrations in 100 to 1000 ng/dL range has consistently been less than 6.2%.

The reference range (280-873 ng/dL) was derived from a reference sample of healthy men, 19-40 years of age, participating in the Framingham Heart Study.[1-4] The assay was cross-calibrated against the CDC standards in September 2013, and the reference ranges were then adjusted using Deming's regression.

SHBG

Serum SHBG concentration was measured using a two-site immunofluorometric assay (Beckman Instruments). The inter-assay CVs were 8.3%, 7.9%, and 10.9%, and intra-assay CVs 7.3%, 7.1% and 8.7%, respectively, in the low, medium, and high pools. The analytical sensitivity of the assays was 0.5 nmol/L. The assay was been validated for human serum. The reference range for SHBG in men of the Framingham Heart Study is 15-72 nmol/L.

Free Testosterone

Free testosterone was measured by equilibrium dialysis.[2-3] 200 uL serum was mixed with 3H-testosterone (10,000 cpm) and then dialyzed for 24 hours at 37C against a buffer mimicking the composition of protein-free plasma and containing gelatin using a dialysis membrane with a cut off of 10,000 daltons. The tracer was purified by celite chromatography prior to equilibrium dialysis to remove any 3H-water. Free testosterone was calculated as a product of total testosterone measured using LC-MS/MS and percent free testosterone determined using equilibrium dialysis. Reference range for free testosterone in healthy young men, 21-50 years of age, is 58-294 pg/ml.

5-alpha-dihydrotestosterone (DHT)

Serum DHT concentration was measured using a derivatized LC-MS/MS assay. The serum samples were extracted using solid phase extraction (SPE), eluted by high performance liquid chromatography (HPLC), and determinated by mass spectrometry (MS) in Electrospray Ionization (ESI^+) source. Deuterated DHT was utilized as an internal standard for the calibration of assay. The recovery was $\geq$ 81%, linear range 1 to 100 ng/dL, and the lower limit of quantitation was 1 ng/dL. The inter-assay CV in quality control samples with concentrations of 5.2 ng/dL, 22.0 ng/dL, and 44.1 ng/dL were 6.1%, 6.5%, and 8.6%, respectively. The reference range, generated in a convenience sample of 100 healthy young men, 21-50 years of age, who were participating in a randomized trial, was 22 to 110 ng/dL.

Estradiol

Estradiol was measured by LC-MS/MS after derivatization with dansyl chloride.[6-8] 20µL estrone-d4 and estradiol-d3 was added to 200µL serum, extracted with 1.2mL methyl t-butyl ether (5), derivatized using dansyl chloride (3.7mmol/L) in sodium carbonate (10mmol/L, pH10.5) at 60ºC for 10-min, diluted in 50µL acetonitrile and water (1:1 volume ratio of acetonitrile and water), and injected in a C1 cartridge. Two-dimensional chromatographic separation is used. First-dimension separation was performed using a mobile phase consisting of methanol with 10mmol/L formic acid and 10mmol aqueous formic acid. The effluent is directed to second dimension separation on a Germini Phenyl C6 column 100×2.0 mm with 3µm particles; the mobile phase consisted of acetonitrile with 10mmol/L formic acid and 10mmol/L aqueous formic acid.[6] Effluent of the first column is directed to waste during >90% of the analysis time; therefore, unreacted dansyl chloride and most derivatization byproducts are not transferred to the analytical column. Effluent from the analytical column is directed onto an API 4000 triple-quadrupole mass spectrometer (Applied Biosystems/MDS Sciex) with turbo ion spray ion source.

Two MS/MS transitions were monitored for each compound: for estrone, 504>171 and 504>156; for estrone-d4, 508>171 and 508>156; for estradiol, 506>171 and 506>156; for estradiol-d5, 509>171 and 509>156. The peak area ratio of the two transitions was used to assess possible interferences in each sample.[6] The linear range of the assay is 1 to 500 pg/mL. The lower limit of quantitation for both hormones is 2 pg/mL. Inter-assay CVs for estrone are 4.5%, 7.7%, and 6.9% at concentrations of 8, 77, and 209 pg/mL, respectively, and for estradiol 6.9%, 7.0%, and 4.8% at concentrations of 8, 77,

and 206 pg/ml, respectively. The cross-reactivity of testosterone, DHEA, DHEAS, DHT, androstenedione in the assay is negligible at ten times the circulating concentrations of these hormones. The reference range for estradiol in community dwelling men, 19-40 years of age, was 15-50 pg/mL.[7]

References

1. Bhasin S, Pencina M, Jasuja GK, Travison TG, Coviello AD, Orwoll E, Wang PY, Nielson C, Wu F, Tajar A, Labrie F, and the EMAS group, Vesper H, Zhang, A, Ulloor J, Singh R, D'Agostino R, Vasan RS. Reference ranges for testosterone in nonobese men generated using liquid chromatography tandem mass spectrometry in a community-based sample of healthy non-obese young men in the Framingham Heart Study and applied to three geographically distinct cohorts. J Clin Endocrinol Metab 2011;96:2430-9
2. Ohlsson C, Lunetta KL, Stolk L, Perry JRB, Koster A, Petersen A-K, Eriksson J, Lehtimäki T, Huhtaniemi IT, Hammond GL, Maggio M, Coviello AD, Ferrucci L, Heier M, Hofman A, Holliday KL, Jansson J-O, Kähönen M, Karasik D, Karlsson MK, Kiel DP, Liu Y, Ljunggren O, Lorentzon M, Lyytikäinen L-P, Meitinger T, Mellström D, Melzer D, Miljkovic I, Nauck M, Nilsson M, Penninx B, Pye SR, Vasan R, Reincke M, Rivadeneira F, Tajar A, Teumer A, Uitterlinden AG, Ulloor J, Viikari J, Völker U, Völzke H, Wichmann HE, Zhuang WV, Ziv E, Wu FCW, Raitakari O, Eriksson A, Bidlingmaier M, Harris TB, Murray A, de Jong F, Murabito JM, Bhasin S, Vandenput L, Haring R. Genetic determinants of serum

testosterone concentrations in men. PLOS Genet 2011(10):e1002313. Epub 2011 Oct 6.

3. Bhasin S, Travison TG, Storer TW, Lakshman K, Kaushik M, Mazer NA, Ngyuen AH, Davda MN, Jara H, Aakil A, Anderson S, Knapp PE, Hanka S, Mohammed N, Daou P, Miciek R, Ulloor J, Zhang A, Brooks B, Orwoll K, Hede-Brierley L, Eder R, Elmi A, Bhasin G, Collins L, Singh R, Basaria S. Effect of testosterone supplementation with and without a dual 5α-reductase inhibitor on fat-free mass in adult men with suppressed testosterone production: a randomized trial. JAMA 2012;307:931-9.

4. Vesper HW, Bhasin S, Wang C, Tai SS, Dodge LA, Singh RJ, Nelson J, Ohorodnik S, Clarke NJ, Salameh WA, Parker CR Jr, Razdan R, Monsell EA, Myers GL. Interlaboratory comparison study of serum total testoserone measurements performed by mass spectrometry methods. Steroids. 2009;74:498-503.

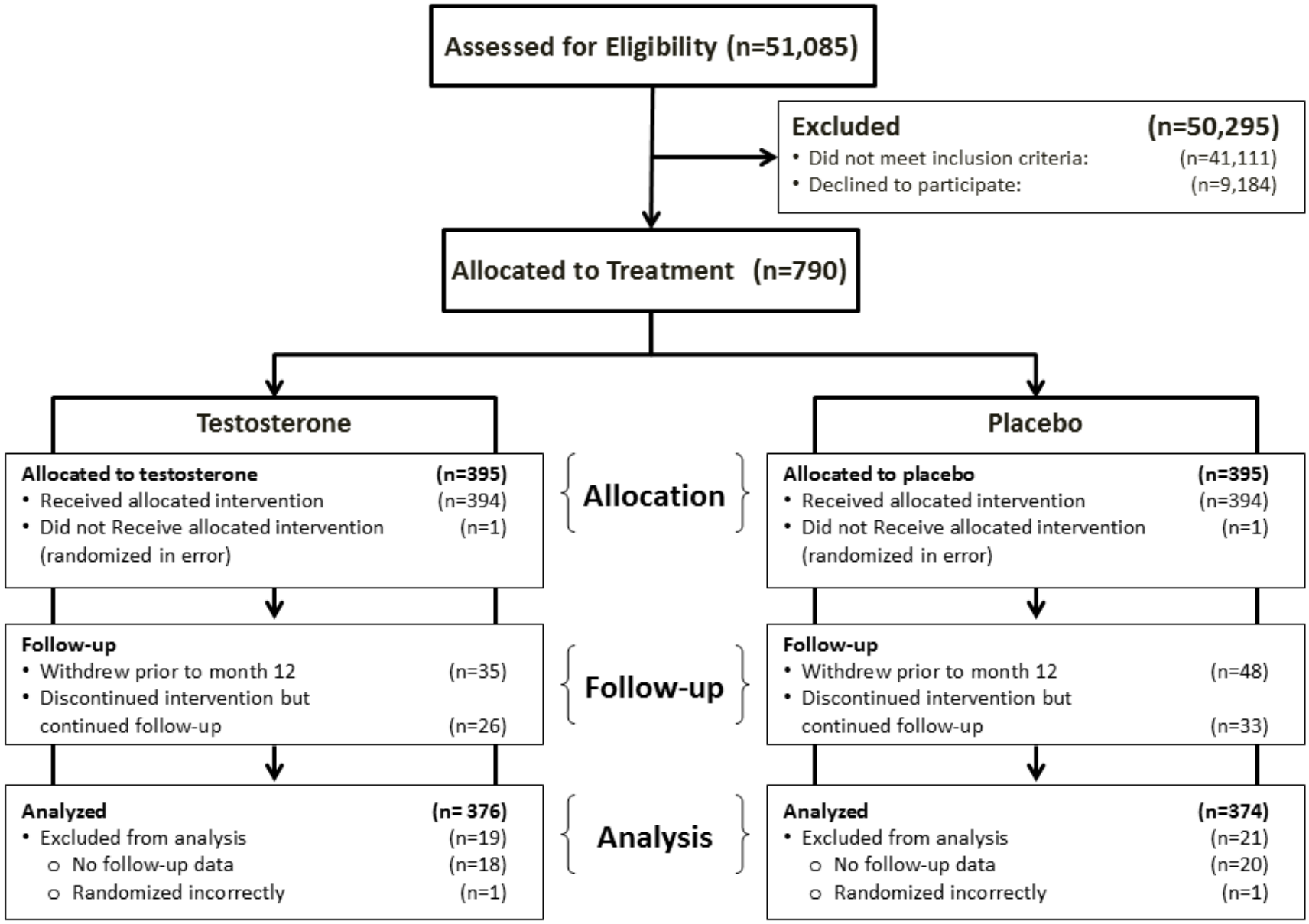
5. Bhasin S, Jasjua GK, Pencina M, D'Agostino R, Sr., Coviello AD, Vasan RS, Travison TG. Sex-hormone binding globulin, but not testosterone, is associated prospectively and independently with incident metabolic syndrome in men: The Framingham Heart Study. Diabetes Care 2011;34:2464-70.

6. Kushnir MM, Rockwood AL, Bergquist J, et al. High-sensitivity tandem mass spectrometry assay for serum estrone and estradiol. Am J Clin Pathol. 2008;129:530-9.

7. Jasuja GK, Travison TG, Davda M, Zhang A, Kushnir MM, Rockwood AL, Meikle W, Pencina MJ, Coviello A, Rose AJ, D'Agostino R, Vasan RS, Bhasin S. Estrone levels are associated with diabetes risk in men of the Framigham Heart Study. Diabetes Care 2013 in press

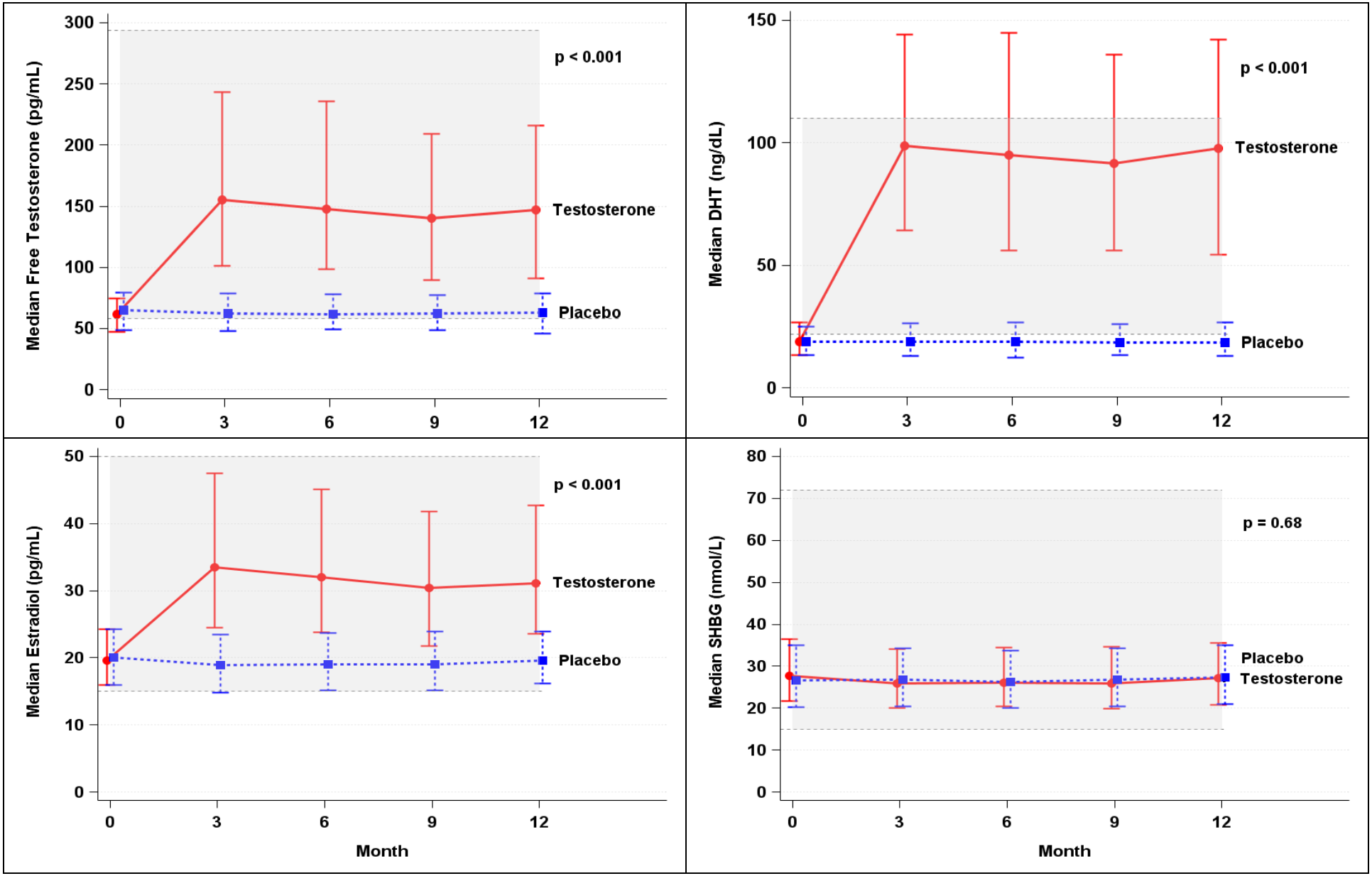
8. Jasuja GK, Travison TG, Davda M, Murabito JM, Basaria S, Zhang A, Kushnir MM, Rockwood AL, Meikle W, Pencina MJ, Coviello A, Rose AJ, D'Agostino R, Vasan RS, Bhasin S. Age Trends in estradiol and estrone levels measured using liquid chromatography tandem mass spectrometry in community-dwelling men of the Framingham Heart Study. J Gerontol A Biol Sci Med Sci 2013;68(6):733-40.

Figure S1. Screening and Enrollment



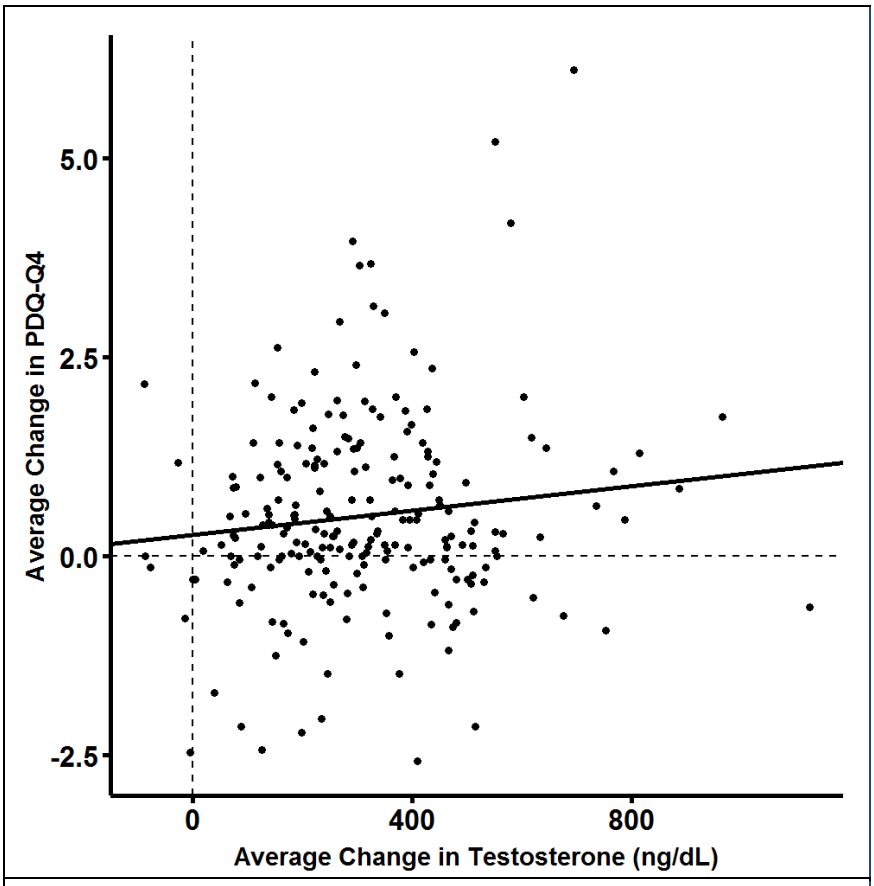
Screening and retention of participants

Figure S2. Serum Concentrations of Free Testosterone, Estradiol, Dihydrotestosterone (DHT) and Sex Hormone Binding Globulin (SHBG)

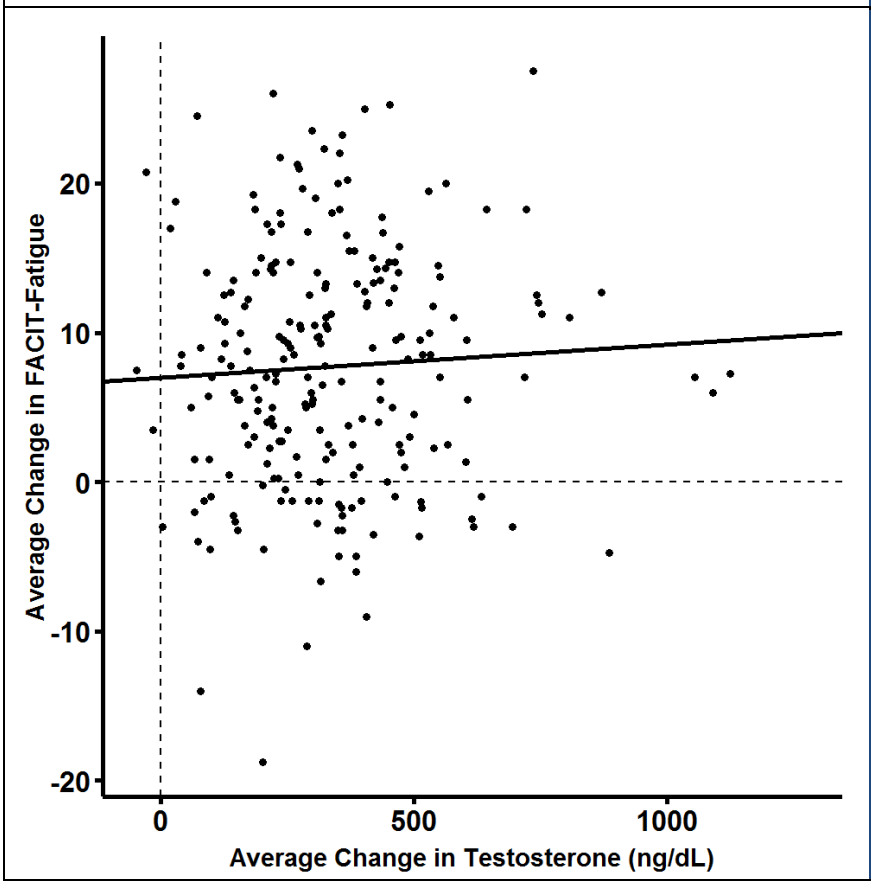


Median (interquartile ranges) serum concentrations of free testosterone, estradiol, dihydrotestosterone and sex hormone binding globulin in response to testosterone or placebo treatment from baseline to 12 months in men participating in The Testosterone Trials. To convert the values for free testosterone, estradiol and dihydrotestosterone to nmol/L, multiply by 3.47, 3.67 and 0.0344, respectively.

Figure S3. Association between changes in testosterone and changes in Psychosexual Daily Questionnaire and in FACIT-Fatigue Scale

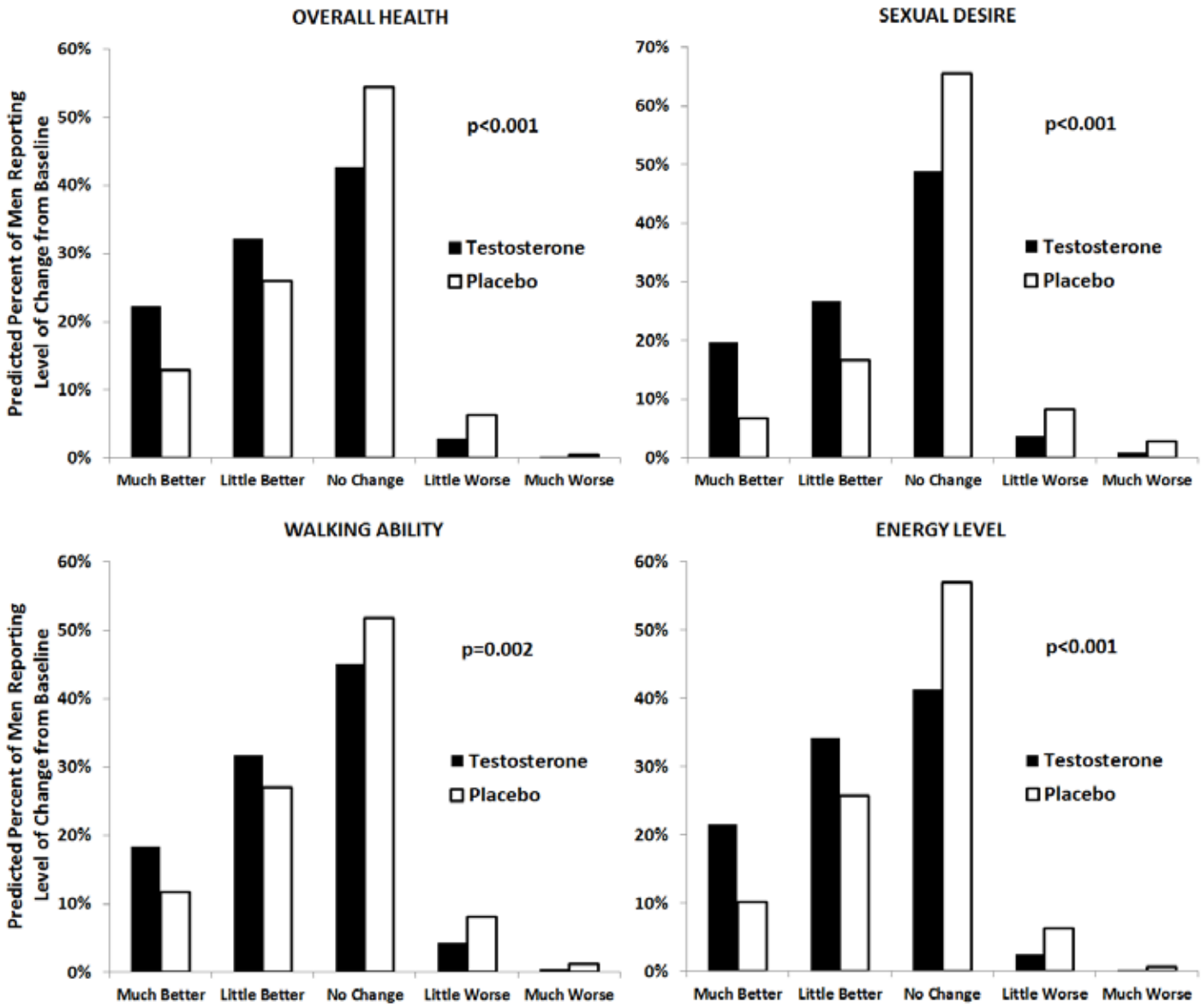


The top panel shows the unadjusted association between the average change in total testosterone and the average change in the Question 4 of the Psychosexual Daily Questionnaire (PDQ-Q4) during months 3-12 in men in the testosterone arm of the Sexual Function Trial. Instrumental variable methods indicated a significant effect of the change in total testosterone on the change in PDQ-Q4 ($p<0.001$).



The bottom panel shows the unadjusted association between the average change in total testosterone and the average change in the FACIT-Fatigue scale during months 3-12 in men in the testosterone arm of the Vitality Trial. Instrumental variable methods indicated a significant effect of the change in total testosterone on the change in the FACIT-Fatigue scale ($p=0.02$).

Figure S4. Patient Global Impression of Change



Responses to patient global impression of change questions. At months 3, 6, 9 and 12, participants were asked, “Since the beginning of the trial, has your (overall health, or sexual desire, or walking ability, or energy) been “much better”, “a little better”, not changed”, “a little worse” or “much worse”. The predicted proportion of participants responding in each category by treatment arm was estimated by multinomial regression adjusted for balancing factors. Presented predicted proportions were averaged over time.

Table S1. Characteristics at Baseline of Men Enrolled in Testosterone Trials

Characteristics	Treatment Group	
	Placebo	Testosterone
N	395	395
Demographics		
Age (yr)	72.3 ± 5.8	72.1 ± 5.7
Race		
Caucasian (%)	351 (88.9%)	349 (88.4%)
African-American (%)	20 (5.1%)	21 (5.3%)
Other (%)	24 (6.1%)	25 (6.3%)
Ethnicity		
Hispanic (%)	10 (2.5%)	18 (4.6%)
Non-Hispanic (%)	385 (97.5%)	376 (95.2%)
College graduate (%)	198 (50.1%)	214 (54.2%)
Married or living with partner (%)	304 (77.0%)	290 (73.4%)
Concomitant conditions		
BMI (kg/m^2)	31.0 ± 3.6	31.0 ± 3.5
BMI >30 (%)	246 (62.3%)	251 (63.5%)
Alcohol Use (no. drinks/week)	3.4 ± 5.0	3.0 ± 4.3
Smoking[1]		
Current smoker (%)	34 (8.6%)	30 (7.6%)
Ever smoker (%)	268 (67.9%)	256 (64.8%)
Diabetes (%)	144 (36.5%)	148(37.5%)
Hypertension (%)	280 (70.9%)	286 (72.4%)
History of myocardial infarction (%)	63 (16.0%)	53 (13.4%)
History of stroke (%)	17 (4.3%)	16 (4.1%)
Sleep apnea	76 (19.2%)	78 (19.8%)
International Prostate Symptom Score	9.6 ± 5.3	9.0 ± 5.2
Risk(%) of all prostate cancer	17.6 ± 6.0	17.3 ± 6.0
Risk(%) of high-grade prostate cancer	3.0 ± 1.7	2.9 ± 1.7
Medication Use		
Alpha blocking agents (%)	47(12.0%)	56 (14.0%)
5-alpha reductase inhibitors (%)	16 (4.1%)	15 (3.8%)
Phosphodiesterase inhibitors (%)	36 (9.1%)	30 (7.6%)
Sex Hormones		
Testosterone (ng/dL)	236 ± 67	232 ± 63
Free testosterone (pg/mL)	65.0 ± 23.4	62.0 ± 21.4
Dihydrotestosterone (ng/dL)	20.9 ±13.0	21.3 ±11.6
Estradiol (pg/mL)	20.4 ± 6.4	20.3 ± 6.7
Sex hormone binding globulin (nM)	29.5 ± 14.7	31.4 ± 15.2
Performance		
Mini-mental State Examination	28.4 ± 1.7	28.4 ± 1.7
Gait speed (m/s)	1.1 ± 0.2	1.1 ± 0.2
Derogatis Interview for Sexual Function - Sexual Desire	14.0 ± 7.8	13.9 ± 7.7
FACIT-Fatigue	36.8 ± 8.8	37.0 ± 8.6

1 Includes cigarettes or cigars

Table S2. Characteristics of Men Who Completed 12 Months and Those of Men Who Did Not

Characteristics			
		Completed	**Did Not Complete**
N		705	85
Demographics			
Age (yr)		72.0 ± 5.7	73.2 ± 6.0
Race			
Caucasian (%)		626 (88.8%)	74 (87.1%)
African-American (%)		35 (5.0%)	6 (7.1%)
Other (%)		44 (6.2%)	5 (5.9%)
Ethnicity	Hispanic (%)	27 (3.8%)	1 (1.2%)
	Non-Hispanic (%)	677 (96.0%)	84 (98.8%)
College graduate (%)[1]		358 (50.8%)	54 (63.5%)
Married or living with partner (%)		527 (74.7%)	67 (78.8%)
Concomitant conditions			
BMI (kg/m^2)		31.0 ± 3.5	31.0 ± 3.7
BMI >30 (%)		444 (63.0%)	53 (62.4%)
Alcohol Use (no. drinks/week)		3.2 ± 4.8	3.1 ± 4.0
Smoking	Current smoker (%)[2]	62 (8.8%)	2 (2.4%)
	Ever smoker (%)	465 (66.0%)	59 (69.4%)
Diabetes (%)		264 (37.4%)	28(32.9%)
Hypertension (%)[3]		495 (70.2%)	71 (83.5%)
History of myocardial infarction (%)		107 (15.2%)	9 (10.6%)
History of stroke (%)		29 (4.1%)	4 (4.7%)
Sleep apnea		143 (20.3%)	11 (12.9%)
International Prostate Symptom Score[1]		9.1 ± 5.1	10.6 ± 6.0
Risk(%) of all prostate cancer		17.6 ± 6.0	16.8 ± 5.7
Risk(%) of high-grade prostate cancer		2.9 ± 1.7	2.9 ± 1.6
Medication Use			
Alpha blocking agents (%)		71(10.1%)	7 (8.2%)
5-alpha reductase inhibitors (%)		27 (3.8%)	4 (4.7%)
Phosphodiesterase inhibitors (%)		62 (8.8%)	4 (4.7%)
Sex Hormones			
Testosterone (ng/dL)		233 ± 65	245 ± 64
Free testosterone (pg/mL)		63.0 ± 22.2	67.0 ± 24.5
Dihydrotestosterone (ng/dL)		21.1 ±12.6	20.7 ±10.0
Estradiol (pg/mL)		20.3 ± 6.4	21.1 ± 7.3
Sex hormone binding globulin (nM)		30.2 ± 14.9	32.8 ± 15.7
Performance			
Mini-mental State Examination		93.1 ± 5.7	92.2 ± 6.5
Gait speed (m/s)		1.1 ± 0.2	1.0 ± 0.3
Derogatis Interview for Sexual Function - Sexual Desire		14.0 ± 7.9	12.8 ± 7.5
FACIT-Fatigue		37.0 ± 8.7	36.5 ± 8.5

[1]$p = 0.03$, [2]$p = 0.04$, [3]$p = 0.01$

Table S3. Sensitivity Analyses	Testosterone (vs. Placebo)		
SEXUAL FUNCTION TRIAL	**Effect**	**Standard error**	**P-value**
Primary Analysis	0.5788	0.1041	<0.0001
Completers	0.5146	0.1081	<0.0001
Non-completers	0.5258	0.4834	0.2813
Patter n Mixture Effect (Completers and Non-completers)	0.5158	0.1094	<0.0001
Shared Random Effects Model	0.6123	0.1772	0.0006
SEXUAL FUNCTION – ALL PARTICIPANTS			
Primary Analysis	0.6234	0.08826	<0.0001
Completers	0.6315	0.09261	<0.0001
Non-completers	0.4747	0.3573	0.1869
Patter n Mixture Effect (Completers and Non-completers)	0.6154	0.0908	<0.0001
Shared Random Effects Model	0.6201	0.1369	<0.0001
PHYSICAL FUNCTION TRIAL	**Log Odds Ratio**	**Standard error**	**P-value**
Primary Analysis	0.3531	0.2756	0.2008
Completers	0.3635	0.2882	0.2080
Non-completers	-0.5563	1.3893	0.6916
Patter n Mixture Effect (Completers and Non-completers)	0.2815	0.2903	0.3322
Shared Random Effects Model	0.4135	0.4545	0.3633
PHYSICAL FUNCTION – ALL PARTICIPANTS			
Primary Analysis	0.5665	0.1923	0.0033
Completers	0.5977	0.2010	0.0030
Non-completers	-0.3451	0.8266	0.6775
Patter n Mixture Effect (Completers and Non-completers)	0.5042	0.1988	0.0112
Shared Random Effects Model	0.7806	0.3317	0.0187
VITALITY TRIAL	**Log Odds Ratio**	**Standard error**	**P-value**
Primary Analysis	0.2090	0.2035	0.3049
Completers	0.1644	0.2145	0.4440
Non-completers	0.6624	0.8550	0.4422
Patter n Mixture Effect (Completers and Non-completers)	0.2215	0.2137	0.3000
Shared Random Effects Model	0.3513	0.3080	0.2544
VITALITY – ALL PARTICIPANTS			
Primary Analysis	0.2040	0.1657	0.2185
Completers	0.1663	0.1750	0.3421
Non-completers	0.4264	0.7845	0.5883
Patter n Mixture Effect (Completers and Non-completers)	0.1967	0.1797	0.2737
Shared Random Effects Model	0.3435	0.2642	0.1938

Table S4. Adverse Events During the First (Treatment) Year, Second (Observation) Year, and Total of Both Years

Event	Year 1		Year 2		Total	
	Placebo	Testosterone	Placebo	Testosterone	Placebo	Testosterone
N	394	394	360	347	394	394
Prostate						
PSA Increase ≥1.0 ng/mL	8	23	5	6	11[1]	28[1]
Prostate biopsy	0	1	2	2	2	3
Prostate cancer	0	1[2]	1[3]	2[2]	1	3
IPSS >19[4]	26	27	N/A	N/A	26	27
Hemoglobin ≥17.5 g/dL	0	7	N/A	N/A	0	7
Cardiovascular[5]						
Myocardial infarction						
Definite MI	0	2	8	1	8	3
Probable MI	1	0	0	0	1	0
Possible MI	1	0	0	0	1	0
Cannot Determine	0	0	0	1	0	1
Stroke						
Definite	5	4	0	0	5	4
Probable	0	1	0	1	0	2
Cardiovascular death	1	0	1	0	2	0
Total MI, stroke, CV death	7	7	9	2	16[1]	9
Hospitalization for arrhythmia	4	7	3	4	7	11
Peripheral artery disease	2	2	1	2	3	4
Unstable angina	4	0	2	1	6	1
Cardiac intervention	6	6	6	5	11[1]	11
Carotid artery disease[6]	4	1	1	1	5	2
Decompensated heart failure	6	2	2	0	8	2
Venous thromboembolism	3	2	2	2	5	4
Aneurysm	0	1	0	1	0	2
Serious Adverse Events						
Death	7	3	4	4	11	7
Hospitalization	79	68	47	59	109[1]	114[1]
Other[7]	6	6	3	6	9	12

[1] Participants who had more than one event in a category or who had events in both year 1 and year 2 are counted only once in the total.
[2] Diagnosed following increase in PSA ≥1.0 ng/mL above baseline
[3] Diagnosed following detection of a prostate nodule
[4] The International Prostate Symptom Score (IPSS) is an 8 question screening tool used to identify symptoms of benign prostatic hyperplasia (BPH)
[5] Cardiovascular adverse events were collected by a specific questionnaire administered at each visit during treatment and for one year after treatment. Cardiovascular adverse events were also identified from the Adverse Event log and SAE Report Forms. The questionnaire, adverse event log and serious adverse event report form are included in the Supplemental materials. Cardiovascular events were adjudicated. Myocardial infarction, stroke and death were adjudicated by two adjudicators, and all other events above by one adjudicator.
[6] Carotid artery disease that resulted in revascularization
[7] "Other" consists of those Serious Adverse Events labeled as Congenital Anomaly, Disability, Important Medical Event or Life Threatening

Table S5. Incidence of Adverse Events >5% according to MedDRA System Organ Class (Year 1)

	Testosterone (n = 395)			Placebo (n = 395)		
Body System	**n**	**%**	**Events**	**n**	**%**	**Events**
Cardiac disorders	45	11.4 %	82	45	11.4 %	88
Infections and infestations	152	38.5 %	221	145	36.7 %	220
Eye disorders	23	5.8 %	27	27	6.8 %	36
Gastrointestinal disorders	65	16.5 %	99	71	18.0 %	116
General disorders and administration site conditions	74	18.7 %	115	75	19.0 %	99
Injury, poisoning and procedural complications	87	22.0 %	143	89	22.5 %	160
Investigations	45	11.4 %	59	37	9.4 %	54
Metabolism and nutrition disorders	24	6.1 %	31	29	7.3 %	34
Musculoskeletal and connective tissue disorders	145	36.7 %	272	155	39.2 %	293
Neoplasms benign, malignant and unspecified (including cysts and polyps)	28	7.1 %	35	22	5.6 %	27
Nervous system disorders	75	19.0 %	101	78	19.7 %	107
Psychiatric disorders	26	6.6 %	41	27	6.8 %	37
Renal and urinary disorders	43	10.9 %	59	42	10.6 %	53
Reproductive system and breast disorders	49	12.4 %[1.2]	63	19	4.8 %	22
Respiratory, thoracic and mediastinal disorders	74	18.7 %	110	67	17.0 %	106
Skin and subcutaneous tissue disorders	88	22.3 %	156	52	13.2 %	75
Surgical and medical procedures	36	9.1%	55	31	7.8 %	43
Vascular disorders	33	8.4 %	42	34	8.6 %	43

[1]p<0.001 compared to placebo
[2]Imbalances between the treatment arms in Reproductive System and Breast Disorders were primarily in breast tenderness and nipple pain.
[3]Imbalances between the treatment arms in Skin and Subcutaneous Tissue Disorders were primarily in acne, abnormal hair growth, erythema, and pruritus.

Table S6. Incidence of Adverse Events > 5% according to MedDRA System Preferred Term (Year 1)

	Testosterone (n = 395)			Placebo (n = 395)		
Preferred Term	**n**	**%**	**Events**	**n**	**%**	**Events**
Oedema peripheral	23	5.8 %	28	16	4.1 %	19
Upper respiratory tract infection	59	14.9 %	66	64	16.2 %	83
Fall	33	8.4 %	42	36	9.1 %	45
Arthralgia	41	10.4 %	55	43	10.9 %	58
Back pain	31	7.8 %	33	23	5.8 %	27
Musculoskeletal pain	19	4.8 %	22	20	5.1 %	21
Pain in extremity	26	6.6 %	28	36	9.1 %	43
Hair growth abnormal	20	5.1 %	22	2	0.5 %	3
Rash	27	6.8 %	39	17	4.3 %	25