THE BONE TRIAL OF THE
THE TESTOSTERONE TRIAL

Sponsors	**National Institute on Aging (NIA) and Abbott Laboratories**
NIH Grant Number	**RO1AG037679-01A1**
University of Pennsylvania Protocol Number	**808676 (T Trial) and 814217 (Bone Trial)**
Principal Investigator	**Peter J. Snyder, MD**
Study Drug Provider	**Abbott Laboratories**
IND Number	**104707**
Clinical Trials.gov Number	**NCT00799617**

Bone Trial Original Protocol	July 29, 2011	Version Number: 1.0
Administrative Update	September 9, 2011	Version Number: 1.1
Administrative Update	January 30, 2012	Version Number: 1.2
Administrative Update	**February 3, 2012**	**Version Number 1.3**

Table of Contents

Study Summary

Title	**The Bone Trial of The Testosterone Trial**
Protocol Number	808676 (University of Pennsylvania, Data Coordinating Center)
Study Design	Randomized, placebo-controlled, double-blind
Study Duration	Five years
Study Centers	Nine clinical sites participating in the T Trial
Objectives	To test the hypothesis that testosterone treatment for one year will increase volumetric trabecular bone mineral density (vBMD) of the lumbar spine as measured by quantitative computed tomography (QCT) compared with placebo treatment
Number of Subjects	200
Diagnosis & Main Inclusion Criteria	Participants in this study will be men enrolled in the T Trial, a randomized, placebo-controlled, double-blind study of seven coordinated trials in men >65 years of age, using AndroGel or placebo gel for one year. Men in the trial have unequivocally low testosterone concentration (average of 2 morning testosterone values, <275ng/dL), and symptoms and objective manifestations of mobility disability, low libido, or low vitality.
Study Product, Dose, Route, Regimen	AndroGel ® 1%, testosterone in an alcohol-water gel, administered transdermally in doses from 5 to 10 grams per day, adjusted as necessary to maintain the serum testosterone concentration within the range of normal for young men.
Duration of Administration	AndroGel or placebo will be administered for one year.

A. Introduction

This document is a protocol for a human research study. This study is to be conducted according to US and international standards of Good Clinical Practice (FDA Title 21 part 312 and International Conference on Harmonization guidelines), applicable government regulations and Institutional research policies and procedures.

The Bone Trial is being conducted as a study within The Testosterone Trial (TTrial). The TTrial is a randomized, placebo-controlled, double-blind study of seven coordinated trials in men greater the 65 years of age using testosterone gel or placebo gel for one year. The participants in this trial will be men presently enrolled in the TTrial who agree to also participate in the Bone Trial. As a result, the parameters of the TTrial protocol will be referenced throughout this description of the Bone Trial.

1. Background

As men get older, they experience many conditions, often together, that eventually result in the inability to perform many activities of daily living, an increased propensity to fall, and decreased independence. These conditions include mobility disability and low vitality. Elderly men also experience increased anemia, metabolic syndrome, decreased sexual function, memory impairment, and osteoporosis. These conditions likely have multiple causes, but one cause that could contribute to all of them is a low serum testosterone concentration. When young hypogonadal men are treated with testosterone, they experience improvements in sexual function, muscle mass and strength, bone mineral density, sense of well-being, and anemia. However, the benefits of testosterone therapy in older men with age-related decline in testosterone concentration are not known and are the subject of this investigation.

B. Specific Aims

1. Study Aims

The overall specific aim of the Bone Trial is to test the hypothesis that testosterone treatment for one year, compared with placebo treatment, of 200 men ≥65 years who have serum testosterone concentrations <275 ng/dL at Screening Visit 1 and < 300 ng/dL at Screening Visit 2, and an average serum testosterone concentration of < 275 ng/dL, and are participating in The Testosterone Trial, will improve bone quality.

The primary specific aim of the Bone Trial is to test the hypothesis that testosterone treatment for one year of elderly men with serum testosterone concentrations <275 ng/dL at 8AM will increase volumetric trabecular bone mineral density (vBMD) of the lumbar spine as measured by quantitative computed tomography (QCT) compared with placebo treatment. vBMD is the focus of the study because it is more responsive to testosterone than areal BMD (aBMD); because it separates the trabecular compartment from cortical shell; and because it avoids the artifacts of osteophytes and aortic calcification so common in elderly men.

The secondary specific aims are to test the hypotheses that testosterone treatment of these men will:

- Increase vBMD of the hip, as determined by QCT, because of the clinical importance of hip fractures and because this measure has been shown to predict hip fracture.
- Increase bone strength and strength-to-density ratio of the spine and hip, as determined by finite element analysis of the QCT data, because these parameters predict fractures better than bone density.
- Increase areal BMD of the lumbar spine and proximal hip, as measured by DXA. This assessment is proposed, despite its limitations, because it is the current standard clinical method of evaluating bone and because it does predict fracture risk.

Exploratory aims are to test the hypotheses that testosterone treatment will increase cortical bone in the spine and hip by QCT and reduce the incidence of radiographically confirmed clinical fractures. Another aim is to apply the fracture data from all 800 men in The Testosterone Trial for a sample size estimate for a larger, longer trial in which the effect of testosterone on the incidence of clinical fractures would be the primary end point.

2. Significance

The Bone Trial within the Testosterone Trial will be the first study of testosterone treatment in elderly hypogonadal men to employ QCT and thereby avoid the artifacts of DXA, but more importantly, the first to assess bone strength. Demonstration that testosterone treatment improves bone strength in elderly men would provide convincing scientific rationale for a larger, longer trial to determine if testosterone treatment reduces clinical fractures as well. The Bone Trial will also provide data on clinical fractures that will inform sample size estimation for a larger, longer trial.

3. Innovation

The Bone Trial is the first trial designed to adequately test the hypothesis that testosterone treatment of elderly hypogonadal men will improve their bone quality, because it would be the first to enroll elderly men who are unequivocally hypogonadal; the first to assess bone quality by a method most likely to detect an effect of testosterone in this population; and the first to assess bone strength in this population.

C. Study Design

The Bone Trial is a randomized, placebo-controlled trial to test the hypothesis that testosterone treatment for one year of men ≥65 years who have unequivocally low serum testosterone concentrations will improve their bone quality. Two hundred (200) men will be recruited for The Bone Trial from among 800 men who have been recruited, screened, and enrolled in The Testosterone Trial.

The Testosterone Trial has the following major features:

- Recruit men who are ≥65 years old and have unequivocally low early morning serum testosterone concentrations (average <275 ng/mL) on two days.
- Enroll 800 subjects in one or more of the three trials (Physical Function, Sexual Function, or Vitality) for which they qualify and consent to participate.

- Allocate men to treatment with testosterone or placebo gel double-blindly for one year; adjust the dose to maintain the serum testosterone within the reference range for young men while maintaining blinding.
- Exclude men who have diseases that testosterone could exacerbate and monitor those who do enroll for the development of these diseases.

The Bone Trial proposes the following *additional* features:

- Recruit 200 men from among 800 who have qualified for The Testosterone Trial.
- Exclude men who have conditions or are taking medications known to affect bone.
- Assess bone quality of these 200 men by QCT of the spine and hip, DXA of the spine and hip, and clinical fractures.

1. Trial Subjects

A total of 200 subjects will be recruited for The Bone Trial at nine of the 12 Testosterone Trial sites. Subjects who are screened and qualify for The Testosterone Trial will be asked if they wish to participate in the Bone Trial at the time of the baseline visit before using T or placebo gel. Those who do will undergo further screening for The Bone Trial.

2. Inclusion Criteria

Upon assessment of eligibility for the T Trial, interested subjects will be evaluated for eligibility for the Bone Trial. There will be no specific inclusion criteria for the Bone Trial. Participants will be required to sign an informed consent form specifically for the Bone Trial.

3. Exclusion Criteria

T Trial participants who have the following conditions or take the following medications will not be eligible to participate in the Bone Trial:

- Bone mineral density by DXA at the lumbar spine, total hip or femoral neck lower than -3.0.
- Elevated serum calcium (>10.5 mg/dL) at Screening Visit 1
- Medications that that could influence bone, eg anticonvulsants, glucocorticoids (prednisone >20 mg/d >2 wk/year), bisphosphonates (alendronate, risedronate, ibandronate), denosumab, and teriparatide. Calcium and OTC vitamin D supplements will be allowed.
- Any procedure or condition wherein lumbar vertebrae 1-4 are not available for analysis (e.g. lumbar laminectomy or fusion or metal in the lumbar area)

4. Study Medication

Testosterone (AndroGel III ®) or placebo (identical in appearance to AndroGel) will be applied to the abdomen, shoulders or upper arms once a day at the same time to dry, intact skin. Subjects will be instructed to wash their hands after application and to let the gel dry before dressing. It is important not to have contact with women or children while the gel is wet. They will also be

asked not to bathe or get this area wet for five hours after application. Subjects will be taught how to apply the gel and they will be provided with written instructions and precautions. This information will be reviewed at each contact and visit.

The initial dose of AndroGel is 5.0 g once a day as described in the T Trial protocol. During the course of T Trial participation, the dose will be adjusted (higher or lower) to achieve a T level between 400 – 800 ng/dl.

Participation in the Bone Trial does not alter the dose or application procedures of the T or placebo gel for study subjects.

5. Participating Study Sites

The following nine (9) T Trial participating clinical sites will enroll men in the Bone Trial:

- Boston University, Boston, MA
- University of California at Los Angeles, Los Angeles, CA
- University of California at San Diego, San Diego, CA
- University of Alabama, Birmingham, AL
- University of Minnesota, Minneapolis, MN
- University of Pittsburgh, Pittsburgh, PA
- Yale University, New Haven, CT
- Baylor University, Houston Texas
- Northwestern University, Chicago, Illinois

The University of Pennsylvania, Philadelphia, PA, will serve as the Data Coordinating Center (DCC) for the Bone Trial.

6. Reading Centers

The DXA reading center will be the SF Coordinating Center DXA Quality Assurance Group in the Department of Epidemiology and Biostatistics at the University of California at San Francisco.

The QCT reading center will be ON Diagnostics LLC, in Berkeley, CA.

Both reading centers will provide quality control for the respective procedures.

D. Study Procedures

1. Study Visits

Participants in the Bone Trial will follow the established T Trial visit schedule. Two (2) additional visits for testing are required during the one year gel use phase: the first at the time of the baseline visit and the second visit at the end of the treatment phase at the Month 12 visit.

- QCT of the spine and hip will be performed at the baseline visit before the initiation of gel treatment and again after 12 months of treatment.
- DXA scan of the spine and hip will also be performed at the baseline visit before the initiation of gel treatment and again after 12 months of treatment.

It is an important objective of the study to collect complete data on all subjects who enroll in the Bone Trial and sites will be asked to identify and approach men who are likely to be reliable subjects. However, a subject who enrolls in the Bone Trial may withdraw from the study without impacting his participation in the T Trial.

2. Quantitative Computerized Tomography (QCT)

a. Rationale

QCT will be the primary tool for assessing efficacy, because it can distinguish trabecular bone, which testosterone appears to affect primarily, from cortical bone. In addition, it is not artefactually influenced by osteophytes and aortic calcification, which are common in elderly men and confound the interpretation of DXA results. Finally, QCT data can be used for estimation of bone strength by Finite Element Analysis (FEA).

QCT of the spine will be the primary end point, because testosterone appears to have its greatest effect at this site. QCT of the hip will also be performed as a secondary end point, because of the clinical importance of hip fractures and because QCT of the hip predicts hip fractures.

b. Technique

Subjects will be instructed to wear metal free clothing and to remove any metal objects or jewelry from their body. They will be instructed to lie fully clothed on the scan table.

The QCT scans, covering the lumbar spine and pelvic region, will be acquired for each patient at 120 kvp. Subjects will be in the supine position in the scanner, arms above head, with an external mineral calibration phantom placed in the table cushion beneath. Reconstructions at 1mm thickness and spacing will be formed using a large field of view (FOV) and a standard (soft tissue type) kernal. The QCT exam will be completed in approximately 15 minutes. The same acquisition parameters will be used at all sites. Scans will be identified by Participant ID# and initials only and sent via DICOM format directly to O.N. Diagnostics for BCT analysis. Quantitative measures of the QCT scans will be transmitted to the DCC in batches of 10.

All patient CT scans and radiographs will be reviewed by a staff radiologist at the clinical facility. All incidental findings will be made available to the patient or patient's primary care physician, as appropriate.

c. Quality Control

Several steps will be taken to ensure consistency in data collection at the 7 imaging facilities. Each facility will undergo training to familiarize personnel with the study protocol. Site selection was influenced by the facility's ability to provide reasonable assurance that the same CT scanner will be available for use with all patients at both time points. CT phantoms will rotate between sites every two months to be scanned using the same acquisition settings as for subjects. An external calibration phantom containing chambers of known bone mineral density concentration (Mindways, Model 3) will be used to convert each pixel value in the QCT scan into units of bone mineral density (mg/cm^3). A torso phantom (Mindways, Model 3) will be used to

correct for field nonuniformities due to beam hardening. These phantom scans will provide a means for cross-calibrating scanners to minimize machine differences across sites. If a scanner is changed during the study, additional phantom scans will be acquired before and after the change. O.N. Diagnostics will review the first five scans acquired at each facility, and every tenth scan thereafter, to identify and rectify any deviations from the protocol or loss in image quality.

d. Bone Strength by Finite Element Analysis

Several measures will be used to assess overall bone strength.

- Finite element models of the L1 vertebra will be constructed from digitized CT scans (1, 2). Each vertebral image (less posterior elements) will be segmented from the scan, rotated into a standard coordinate system, and resampled to 1 x 1 x 1 mm^3 voxels. The CT number of each voxel will be converted into a bone mineral density value using the included external mineral phantom. The finite element mesh will be created by converting each voxel into an 8-noded brick element. The mean BMD across all voxels in the FE mesh will be used as the integral measure of volumetric density (vBMD).
- For each voxel, the axial elastic modulus (E_z) is calculated using the empirical correlation between elastic modulus and BMD for human vertebral trabecular bone (3). Elastic anisotropy of the bone is accounted for by assuming fixed ratios of the various elastic constants with respect to E_z(1).
- Material failure of the bone will be modeled by assigning an elastic-perfectly-plastic von Mises failure criterion. A thin layer of polymethylmethacrylate (PMMA) will be virtually placed at the ends of each vertebral model to create planoparallel surfaces on which uniform compressive displacement boundary conditions are applied.
- The compressive strength of the vertebra will then be computed as the total reaction force generated at 2% strain (applied displacement divided by bone height) in a non-linear stress analysis.
- The *peripheral compartment*s are defined as all cortical bone plus any trabecular bone within 2 mm of the periosteal surface, and the *trabecular compartment* as all remaining trabecular bone. To assess the relative contribution to vertebral strength due only to the peripheral compartment, a fixed value of density is assigned to the trabecular compartment. Similarly, a fixed value of density will be assigned to the peripheral compartment to compute “trabecular strength.” To assess the relative contribution to vertebral strength due only to differences in external bone geometry, models will be created with a single uniform material property assigned throughout the entire vertebral body. In addition, vBMD of the cortical and trabecular bone combined (“integral” density), and the peripheral and trabecular compartments will be measured.
    - Finite element models will be constructed to estimate femoral strength **(Figure 2)**(4-6). The left proximal femur will be segmented from the calibrated CT images and resampled into 1.5 mm-sided voxels. Finite element mesh will be constructed by converting each voxel into an 8-noded brick element. Boundary conditions will be applied to simulate an unprotected fall to the side of the hip, the diaphysis angled at 15° with respect to the

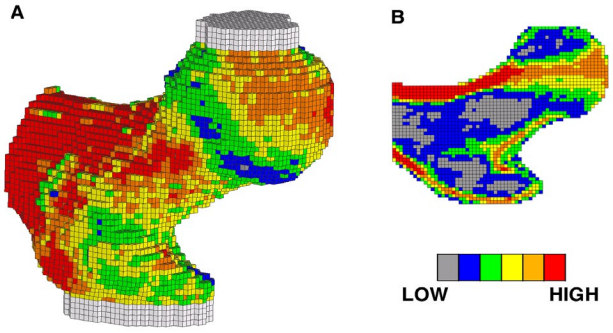

Figure 2: Typical finite element model of the femur, showing 3D (A) and 2D sectional (B) views. The color-coding shows the spatial variation of material strength assigned to the individual finite elements.

ground with 15° of internal rotation. To ensure consistent orientation, each bone will be registered to a reference bone in the fall orientation. The cortical and trabecular bone regions in these models will be distinguished from each other on the basis of apparent density (cut point of 1.0 g/cm^3) and element-specific isotropic material properties will be derived from the calibrated volumetric BMD values using empirical relations, with different relations being used for the cortical and trabecular bone (7-9) From a nonlinear stress analyses femoral strength will calculated from the resulting force-deformation curve as the force at 4% deformation of the femoral head with respect to the greater trochanter.

Geometric strength, and strength and density of the peripheral and trabecular compartments, will be measured as described above for the vertebra with the exception that the peripheral compartment will consist of all cortical bone plus any trabecular bone within 3 mm of the periosteal surface.

3. *Dual Energy X-Ray Absorptiometry (DXA)*

a. Rationale

DXA of the spine and hip will be secondary end points, in spite of the limitations of DXA in elderly men, because DXA is the current standard method of assessing bone in clinical practice, because it does predict clinical fractures, and because previous studies of the effect of testosterone on bone suggested that testosterone treatment of elderly men should improve aBMD of the spine in men who have sufficiently low pretreatment serum testosterone concentrations.

b. Technique

Subjects will be asked to lie as still as possible on a padded table during the scans of the spine and hip. The detector is scanned over the area, generating images on a computer monitor. During the scan of the spine, the legs are supported to flatten the pelvis and lumbar spine. To scan the hip, the technologist will place the foot so that the hip rotates inward. The scan will take approximately 20 to 30 minutes to complete. Scans will be identified by Participant ID# and initials only. Analysis will be performed locally. Analyzed scans will be sent directly to UCSF for quality control and incorporation into the DXA dataset. Routine DXA scans will be transmitted to the DCC once a month. Baseline scans with a lumbar spine, total hip or femoral neck T score <3.0 and 12-month scans that show a <8% decrease in BMD will be sent immediately.

Subjects will be given the report of their baseline scans by the clinical site. Subjects will not be given their 12-month scans unless the BMD decreases by 8%, confirmed by the DXA reading center.

c. Quality Control

All study sites will use a Hologic scanner to minimize site differences in the measurement of change in aBMD. All measures will be obtained and analyzed using Hologic specifications and a study-specific protocol. Quality control will be performed at the UCSF Coordinating Center, DXA QA Group (Dr. Ann Schwartz, PI) using procedures similar to those in the SOF, FIT, WHI, Health ABC, and PaTH studies. Densitometry operators will be certified in the study-specific protocol, and performance will be monitored by central review of selected scans. A complete scan database will be maintained at the CC. Sites will report possible cases of low BMD at baseline or the 12-month visit (BMD T-score < -3.0 at the total hip, femoral neck, or lumbar spine) to UCSF for confirmation. Sites will also report possible cases of excessive bone loss (change from baseline to 12-months >8%) to the DXA QA center for review. If confirmed, UCSF will notify the clinic coordinator who will notify the participant of low BMD or excessive bone loss.

Densitometer performance will be monitored with longitudinal spine phantom scans. Measures of bone loss will be derived from analyses of paired scans using the 'compare' feature, so that baseline and follow-up scans are analyzed with the same software version. However, to the greatest extent possible, changes in scanner software will be avoided during the study.

4. Calcium and Vitamin D Treatment

Many older men have inadequate intake of calcium and vitamin D. Therefore, participation in the Bone Trial includes providing calcium and OTC vitamin D supplements to all men in the Bone Trial. It is likely that this supplementation will improve bone density in the treatment as well as the placebo group. As a result, the sample size estimation has been calculated based on a study that employed similar supplementation. Subjects will be reminded to take calcium and vitamin D at each study visit.

a. Calcium

Subjects will be given a supply of calcium supplements and instructed to take 1200 mg of elemental calcium a day, which is the amount recommended by the NIH and the National Osteoporosis Foundation.

b. Vitamin D

Subjects will be given a supply of OTC vitamin D and instructed to take 800 units of vitamin D a day.

5. Clinical Fractures

Although the study will not likely have enough subjects followed for a sufficiently long time to have adequate power to detect a difference between treatment groups in the incidence of fractures, clinical fractures in all 800 T Trial subjects will be tracked and evaluated for trends.

The same procedures for fracture ascertainment and adjudication will be followed as in the MrOS study, which are:

- Information on fracture history will be obtained prior to randomization
- After randomization, each trial site will inquire specifically about fractures at each contact
- If informed of a fracture, the trial site will inquire about the circumstances (e.g. fall, motor vehicle accident, etc.), anatomic site of fracture, and if and where x-rays were taken; obtain the radiology report and other clinical documentation, such as operative and orthopedic notes, etc.
- Clinical sites will send reports and other clinical documentation films directly to the UCSF Reading Center, where Dr. Douglas Bauer will compile the data. Fractures associated with major trauma, e.g., motor vehicle accident, may be excluded from analysis.

E. Adverse Events

1. General

All aspects of recording and reporting adverse events described in the T Trial protocol apply to the Bone Trial. At each Bone Trial visit, adverse event data will be collected and reported in the same manner as in the T Trial with the identical oversight and safeguards.

2. Additional Risks

The additional risks associated with the Bone Trial participation are related to the additional radiographic tests.

a. Quantitative CT of Bone

Radiation exposure from CT scans introduces a potential risk to study participants. The effective dose to each subject is expected to be approximately 1.8 mSv during the standard lumbar CT exam protocol and 2.4 mSv during the pelvic CT exam protocol. Assuming an annual natural background radiation dose of 2.4 mSv, the total per patient dose for this study (~4.2 mSv) has the same relative risk as about 1¾ years of natural background radiation. The total radiation exposure for the study will be less than typical exam doses (Table 1) and well below the maximal acceptable dose for normal subjects. The subjects will be counseled about the amount of radiation they will receive as a result of participation in the study as part of Informed Consent procedures.

The QCT test of the spine and hip will expose the subject to a dose of 4.2 mSV of additional radiation.

Table 1: Typical CT Exam Doses

Exam	Dose (mSv)
Colonography *	2.4-7.8
Body CTA *	6.1
Routine Liver *	8.0
Coronary CTA *	14.5
Proposed QCT Spine & Hip	4.2

* Dose estimates made using CT-Expo v2.0 (copyright Georg Stamm and Hans Dieter Nagel, 2001-2011) and protocols recommended by Knollmann and Coakley, "Multislice CT Principles and Protocols", Elsevier, 2006.

b. DXA Scan

DXA of the spine and hip will expose subjects to a very low dose of radiation, 2.2 microSv for the lumbar spine scan and 5.1 microSv for the hip scan, which is less than that of a standard chest x-ray.

Time Burden

The tests proposed in the Bone Trial will not likely be burdensome by themselves, but because the men who will participate in the Bone Trial will also be participating, and have a large number of tests in the overall Testosterone Trial, the total subject burden could be considerable. The site staff will be instructed to be cognizant that some subjects could have difficulty participating in The Bone Trial tests, as well as the overall Testosterone Trial, and to consider spreading the testing over more than one day.

3. Protection Against Risk

Care is taken during QCT and DXA imaging tests to use the lowest radiation dose possible while producing the best images for evaluation. National and international radiology protection councils continually review and update the technique standards used by radiology professionals.

State-of-the-art x-ray systems have tightly controlled x-ray beams with significant filtration and dose control methods to minimize stray or scatter radiation. This ensures that those parts of a participant's body not being imaged receive minimal radiation exposure.

The clinical sites participating in the Bone Trial have been chosen for their use of state-of-the-art scanning equipment and techniques and for their imaging experience in many similar trials.

F. Statistical Considerations

1. Treatment allocation and balance

The standard approach to treatment allocation is randomization, balancing on prognostic factors by stratification, but this approach is not practical for The Testosterone Trial is actually several trials in one, which would greatly complicate balancing by stratified randomization, since

there would be more than a thousand strata. Stratifying using fewer variables would risk imbalance with regard to one or more key prognostic factors. The minimization has been selected as a more suitable alternative in this setting. Minimization is a covariate-adaptive procedure that can be employed even when there are a large number of factors to be balanced (10, 11), because balance is achieved within strata defined by the sum of categories for the balancing factors rather than the product, as with stratified randomization. Minimization is based on real-time calculations of treatment assignment, balanced with respect to each important prognostic variable. The subject is assigned to the treatment that produces the best balance overall, considering all relevant factors, with a specified high probability. Because the treatment assignments generated by this method are no more predictable than they would be with stratified randomization, the assignments can be considered essentially random.

Minimization will be used with a random component, for treatment allocation, assigning subjects to the optimal balancing treatment with probability 80%. Factors for balancing will include a dichotomous variable for each of the primary efficacy trials in which a subject may participate, study site, baseline testosterone concentration, age, and current use of an antidepressant and PDE-5 inhibitor; balancing factors will be accounted for as covariates in all analyses. Using this approach, simulation studies have performed simulations that verify that it provides the desired balance with regard to the prognostic variables planned for use.

2. Treatment blinding

Several methods are used to maintain blinding. Treatment assignment will be known only to the data coordinating center and research pharmacy. The testosterone and placebo preparations will look, smell, and feel the same. Maintenance of blinding when the testosterone dose is adjusted is achieved by adjusting the dose in a participant using placebo gel. Clinical sites will not have access to the testosterone-dependent laboratory tests, e.g. hemoglobin or PSA.

3. Analytical methods and sample size estimations

Each of the individual efficacy trials in The Testosterone Trial, such as The Bone Trial, is considered a separate trial, so the results will be analyzed separately. Data in The Bone Trial will be analyzed as continuous variables. All analyses will incorporate accounting for balancing covariates.

a. Analytical methods, sample size and power estimations for The Bone Trial

The **primary end point** for The Bone Trial will be change from baseline to 12 months in trabecular volumetric bone mineral density (vBMD) in the spine as measured by quantitative computed tomography (QCT). The analytic approach will be comparison of treatment groups using ANCOVA, accounting for balancing factors, the same as used for balancing in treatment allocation. Data from a prior study in hypogonadal men showed an increase in mean trabecular vBMD of 14% after 18 months of testosterone treatment, with a SEM of 3% (12). The testosterone levels of men in that study were similar to those in the proposed study. It is conservatively assumed that improvement will be linear and will be 9% by 12 months, with the same SEM. To detect such an effect with 90% power will require 86 subjects per arm or 172

subjects for both arms. An additional 15% accrual will compensate for dropout and noncompliance, leading to a total sample size of 200.

Secondary end points will include trabecular vBMD of the hip by QCT and bone strength and strength-to-density ratio by finite element analysis in the spine and hip by QCT. The power to detect differences between the AndroGel and the placebo groups in these end points is based on data from a study of ibandronate versus placebo treatment of postmenopausal women for one year (5). Assuming a sample size of 172, this study has 90% power to detect a similar difference between treatment groups in hip trabecular vBMD, 98% power in hip trabecular strength, >99% power in hip strength/density ratio, and 90% in spine trabecular strength.

Areal bone mineral density of the lumbar spine and hip by DXA will also be secondary end points. Data from two prior trials (13, 14) suggest a threefold improvement in aBMD of the spine compared with a placebo arm, given a pretreatment serum testosterone <275 ng/dL. Given a sample size of 172, there should be 90% power to detect a doubling (a more conservative estimate) between the two groups.

Exploratory end points will be changes in cortical bone in the spine and hip by QCT and clinical fractures. Serum markers of bone turnover tests will be conducted at a later date if additional funding becomes available. Clinical fractures will be assessed in all 800 subjects in The Testosterone Trial. Even if testosterone treatment should be associated with a reduction in clinical fractures, the mechanism could be an improvement in muscle strength and reduced tendency to fall. For this reason, and because falls is an endpoint in all 800 men, finding fewer falls in the testosterone group than the placebo group might be predictive of a lower fracture rate. Regression analyses will be performed to assess whether certain baseline characteristics, including baseline vBMD and aBMD values and age, are associated with greater change in trabecular vBMD of the spine at 12 months.

Another exploratory objective is development of a **sample size estimate for a longer, larger trial** in which the effect of testosterone on clinical fracture incidence in elderly men is the primary end point. The design of such a trial requires data on the expected fracture rate in untreated men and an estimate of the potential treatment effect in order to develop sample size requirements. The entire 800 men in The Testosterone Trial treated for one year and followed for two years will provide data on fracture incidence. The T Trial results will be compared to those from the MrOS study, in which clinical fractures occurred at a rate of 19/1000 person-years (Cauley, personal communication). If clinical fracture incidence in The Testosterone Trial is similar to that in MrOS, weighted according to expected age distribution, 25-35 fractures may be observed.

The Bone Trial will also provide an estimate of the effect of testosterone treatment on fracture incidence. Although the number of fractures in the two treatment groups will be too small to draw firm conclusions, any trend will be of interest. The association between bone strength by finite element analysis (FEA) and clinical fracture data will be determined, because of the findings from MrOS of a strong association between bone strength by FEA and incidence of hip fracture. If the T Trial fractures data confirm those of MrOS, and also predict that the

testosterone group would have fewer fractures than the placebo group, these results would provide both a rationale and a quantitative basis for conducting a larger, longer study with the incidence of clinical fractures the primary outcome.

4. Managing Missing Values

The primary analysis will include only men who have measurements at both baseline and 12 months; however, a variety of sensitivity analyses will be performed to assess the potential impact of missing 12-month outcomes. Baseline characteristics of men who do and who do not have a 12-month measurement will be compared to help determine any possible biases attributable to the pattern of missing values.

5. Intent-to-Treat

Primary analyses will follow the "intent-to-treat" (ITT) principle; individuals will be analyzed according to their assigned treatment group whether or not they continue treatment. Every attempt will be made to follow and evaluate all enrolled subjects.

6. Multiple analyses

Because a large number of analyses are likely to be conducted – e.g., individual domains contributing to the instrument scores, alternative statistical approaches, and interest in particular subsets will lead to multiple analyses –the likelihood of multiplicity-induced false positive outcomes will be explored. The approach to multiple comparisons will be based on controlling the false discovery rate according to the methods of Benjamini and Hochberg.

G. Study Management

The Data Coordinating Center (DCC) at the University of Pennsylvania will develop a secure electronic data management system for the collection, storage, and management of the Bone Trial data. Lab data will be transferred from Quest using the secure server. The DCC will follow all of the Standard Operating Procedures (SOP) that have been established in the T Trial to achieve the same level of data integrity and security.

1. Data Management System Components

The Data Coordinating Center (DCC) at the University of Pennsylvania will develop an electronic data management system module for the collection, storage, and management of Bone Trial data, as consistent with the main Testosterone Trial Protocol.

Manual of Procedures (MOP) – A MOP will describe the technique for conducting QCT and DXA scans on participants. This manual will provide detailed information for testers in the preparation, acquisition, transfer, and data quality assessment of images acquired in the Bone Trial. The MOP will also provide instruction in case report form completion, use of the electronic data management system. Review of this manual will be part of the initial training conducted by the reading center.

Training and certification procedures – The DXA and QCT Reading Center Directors will conduct training in these respective procedures. Each DXA and QCT technologist involved in the Trial should have a complete understanding of the protocol and experience in acquiring images and operating the scanner. To ensure quality control, each site should designate specific technicians to perform the scans. The Trial will provide training in study-specific procedures, but operators are assumed to be proficient in image acquisition.

DXA training will be supervised by Dr. Ann Schwartz, the Director of the DXA Reading Center, who will prepare a manual of procedures for DXA. She will conduct training of DXA technicians using a standard PowerPoint presentation and teleconference system. Using the standard interactive presentation and teleconference, each site can get specific training and follow-up as needed. Further, quality assurance will be performed on scan (see below), and necessary feedback given to the specific technologist promptly after the scan is received at the Reading Center. This will allow for rapid identification of improper scan techniques and has been shown to dramatically decrease inadequate data sets.

QCT training will be conducted by Dr. David Kopperdahl of ON Diagnostics. He will prepare a manual of procedures for QCT and visit each site at least once. During these visits he will include in the training a supervisor as well as the technicians who will perform the scans, so if the technicians who are trained leave, someone who knows the protocol will be able to help train a new technician.

The DCC will establish a data structure and transfer process for each reading center and test these procedures at the start of the trial. The DCC will monitor quality control of data received from the QCT and DXA reading centers. A data validation plan, rule set specifications, and programming logic to implement data validation rules will be applied to the data.

2. Routine reports

The DCC will develop a set of standard enrollment, tracking, quality review, and safety monitoring reports. Adverse event reports, DSMB reports and reports for statistical analysis will be developed and produced on an appropriate schedule.

3. Data Security

The data management system will be the same system as the main Testosterone Trial Study and is designed to prevent unauthorized access to trial data and to prevent data loss due to equipment failure or catastrophic events. The procedures to do so encompass user account management, user privilege assignment, data loss prevention (database backup), computer systems validation, performance monitoring, and DMS change management. User access will be controlled by assignment of confidential usernames, passwords and role assignment. The system will meet the applicable Federal regulatory requirements and those described in the E6 Good Clinical Practice Guidelines to ensure the confidentiality of trial subjects.

a. Maintaining Anonymity of Submitted Medical Records

Clinical site personnel will de-identify all medical records before sending them to the DCC by obliterating any Protected Health Information (PHI). Upon receipt, DCC personnel will review the records to ensure that no PHI is visible.

b. Confidentiality

Information about study subjects will be kept confidential and managed according to the requirements of the Health Insurance Portability and Accountability Act (HIPAA) of 1996.

H. Reference List

1. Crawford RP, Cann CE, Keaveny TM 2003 Finite element models predict in vitro vertebral body compressive strength better than quantitative computed tomography. Bone 33:744-750

2. Melton LJ, 3rd, Riggs BL, Keaveny TM, Achenbach SJ, Hoffmann PF, Camp JJ, Rouleau PA, Bouxsein ML, Amin S, Atkinson EJ, Robb RA, Khosla S 2007 Structural determinants of vertebral fracture risk. J Bone Miner Res 22:1885-1892

3. Kopperdahl DL, Morgan EF, Keaveny TM 2002 Quantitative computed tomography estimates of the mechanical properties of human vertebral trabecular bone. J Orthop Res 20:801-805

4. Keaveny TM, Hoffmann PF, Singh M, Palermo L, Bilezikian JP, Greenspan SL, Black DM 2008 Femoral bone strength and its relation to cortical and trabecular changes after treatment with PTH, alendronate, and their combination as assessed by finite element analysis of quantitative CT scans. J Bone Miner Res 23:1974-1982

5. Lewiecki EM, Keaveny TM, Kopperdahl DL, Genant HK, Engelke K, Fuerst T, Kivitz A, Davies RY, Fitzpatrick LA 2009 Once-monthly oral ibandronate improves biomechanical determinants of bone strength in women with postmenopausal osteoporosis. J Clin Endocrinol Metab 94:171-180

6. Orwoll ES, Marshall LM, Nielson CM, Cummings SR, Lapidus J, Cauley JA, Ensrud K, Lane N, Hoffmann PR, Kopperdahl DL, Keaveny TM 2009 Finite element analysis of the proximal femur and hip fracture risk in older men. J Bone Miner Res 24:475-483

7. Bayraktar HH, Morgan EF, Niebur GL, Morris GE, Wong EK, Keaveny TM 2004 Comparison of the elastic and yield properties of human femoral trabecular and cortical bone tissue. J Biomech 37:27-35

8. Morgan EF, Bayraktar HH, Keaveny TM 2003 Trabecular bone modulus-density relationships depend on anatomic site. J Biomech 36:897-904

9. Morgan EF, Keaveny TM 2001 Dependence of yield strain of human trabecular bone on anatomic site. J Biomech 34:569-577

10. Pocock SJ, Simon R 1975 Sequential treatment assignment with balancing for prognostic factors in the controlled clinical trial. Biometrics 31:103-115

11. Taves DR 1974 Minimization: a new method of assigning patients to treatment and control groups. Clin Pharmacol Ther 15:443-453

12. Katznelson L, Finkelstein JS, Schoenfeld DA, Rosenthal DI, Anderson EJ, Klibanski A 1996 Increase in bone density and lean body mass during testosterone administration in men with acquired hypogonadism. JClinEndocrinolMetab 81:4358-4365

13. Amory JK, Watts NB, Easley KA, Sutton PR, Anawalt BD, Matsumoto AM, Bremner WJ, Tenover JL 2004 Exogenous testosterone or testosterone with finasteride increases bone mineral density in older men with low serum testosterone. J Clin Endocrinol Metab 89:503-510

14. Snyder PJ, Peachey H, Hannoush P, Berlin JA, Loh L, Holmes JH, Dlewati A, Staley J, Santanna J, Kapoor SC, Attie MF, Haddad JG, Jr., Strom BL 1999 Effect of testosterone treatment on bone mineral density in men over 65 years of age. J Clin Endocrinol Metab 84:1966-1972.