Revised Statistical Analysis Plan

1. Data Analysis and Statistical Methods

a. Interim Monitoring. The study will be monitored by the Data Safety Monitoring Board (DSMB), an independent group that will periodically review the results in order to assess safety. No interim efficacy assessment is planned.

b. Statistical Analyses

i. Descriptive Analyses. We will calculate descriptive measures for each of the variables. Means, medians, standard deviations, range, minima, maxima, and interquartile ranges will be computed for all continuous variables. Frequencies will be calculated for categorical variables.

ii. Baseline Comparability. We will use summary statistics and graphical techniques, such as boxplots, to compare the baseline characteristics of treatment groups. Pre-treatment values will be compared between the placebo- and testosterone-groups using Analysis of Variance (ANOVA), Kruskal-Wallace tests or Chi-square tests, as appropriate, to determine if the groups are balanced with respect to baseline characteristics.

iii. Compliance Analysis. Subject's compliance with treatment will be assessed by the number of gel packages and tablets used, expressed as a percent of the total number of gel packages or tablets that should have been used during treatment. Percent compliance will be averaged across subjects within each treatment group to obtain group means.

iv. Testing Specific Hypotheses. All analyses will use an intent-to-treat approach. The primary outcome variable is the rate of atherosclerosis progression. We will calculate per-subject rate of change in the right distal CCA far wall IMT and in total CAC score by EBCT. The rate of change in the right distal CCA far wall IMT will be computed for each subject by fitting a regression line of IMT on years since baseline. The estimated slope of the regression line will be used as that subject's IMT rate of change. For CAC, the 18 month and 3 year rates of change, calculated as (post-treatment score – baseline score)/years of treatment will be treated as separate dependent variables in the analysis of variance. Analyses of trial end points will be performed on the intent-to-treat sample, which will be defined as all randomized subjects who had a baseline and at least 1 follow-up carotid IMT measurement taken after randomization. Statistical assumptions of the planned analyses will be verified. Alternative analyses will be performed if the assumptions of the planned analyses are not justified. Key assumptions that will be verified are: 1) normality of error terms in the model; 2) homogeneity of variances between treatment groups; and, 3) deviations from linearity in the regression of carotid IMT measurements against time in-study.

The effects of on-trial end point variables other than treatment modality on the per-subject rate of change in the right distal CCA far wall IMT will be assessed. The parameters to be tested will include laboratory data (lipids, total testosterone, estradiol and SHBG levels, etc.), clinical variables (blood pressure, body mass), and lifestyle variables (smoking). Baseline and on-trial values of these data will be collected in order to address questions of clinical significance within each treatment group such as: 1) which baseline characteristics have differential effectiveness on the rate of change in IMT or CAC score; and, 2) what is the role of changes in these variables on the rate of change in IMT and CAC score.

The primary analytic strategy will assess end-of-treatment differences in outcomes as a function of randomization, controlling for stratification factors (age range and site). This method has the advantage that it can consider all observations of each outcome simultaneously while accounting for the violation of strict independence characterized by participant-level associations between repeated measures of the outcomes. Control for stratification factors is consistent with the experimental design.

The precise form of the regression model will be as follows: subjects will be treated as random effects (i.e. be assigned random intercepts) nested within study site, which will also be treated as random effects.

Treatment randomization, age group and time, and treatment-by-time interactions will be treated as fixed effects in the conventional way.
To diagnose the form of the model of within-subject correlation we will use a correlogram and other exploratory techniques; though we anticipate that outcomes with more than 3 measurements (including carotid IMT) may evince an autocorrelative structure (i.e. those measurements taken more closely together in time will be more strongly associated with each other than those further apart), this may not necessarily be the case. If no convincing pattern is observed, we will estimate an unstructured correlation matrix. The treatment effect will be formally estimated using a treatment-by-time interaction term and associated 95% confidence interval; hypothesis tests will be evaluated allowing a maximal type-I error probability of 0.05. The mixed effects model will provide consistent estimates of all effects provide data are missing at random; though we anticipate few missing records, we will evaluate evidence for violation of this assumption using exploratory techniques and a model of the missing data process.

A secondary analysis will consider whether treatment effects are meaningfully altered by control for baseline diabetes status, baseline testosterone level, and other covariates as suggested by the exploratory comparisons described above.