Skeletal Physiology Dysregulation in Obesity: The Role of Growth Hormone

NCT# 01724489

Version Date: November 7th, 2019

Statistical Power and Statistical Analysis

Statistical power*: For Aim 1, primary endpoint,* with an n=75 and 50 completers, we will have >80% chance of detecting a difference in the rate of BMD AP lumbar spine increase of 0.024g/cm^2/y at a 2-sided p=0.05 level. This compares favorably with the rate of BMD AP spine increase of 0.036 g/cm^2/y in a study by Baum *et al.*[54] administering GH for 18 months to hypopituitary subjects with GH deficiency. This was calculated with software developed by co-investigator [redacted] for sample size for random slopes models, available in Matlab [redacted] The variance covariance matrix of the random effects, Ψ, and residual variance, σ , were estimated from longitudinal data from Baum *et al.*[54] that had originally been analyzed by [redacted]. The variance of the intercept was 0.024, slope 0.000211, covariance 0.00199 and residual variance, σ , 0.00092 and an exponential drop-out rate of 0.36/yr, which is greater than we expect. The program uses the fact that the variance covariance matrix for the fixed effects of the model are: $\mathbf{V}(B) = \{\mathbf{X}'[\sigma^2] + \mathbf{Z}\Psi\mathbf{Z}']^{-1}\mathbf{X}\}f^{-1}$. One incorporates drop outs by replacing $\mathbf{X}'[\sigma^2] + \mathbf{Z}\Psi\mathbf{Z}']^{-1}\mathbf{X}$ by a weighted sum for different follow-up times with weights equal to the follow-up probability -- a special case of the general model described in Pinheiro and Bates.[121] *For Aim 2, primary endpoint,* with 50 completers we can detect 0.81 x SD of the difference in bone marrow fat content over a year; *For secondary endpoints:* for trabecular thickness (TbTh) and cortical thickness (CoTh), based on our preliminary data in obese men receiving placebo for 6 months in which the standard deviation (SD) of the change in TbTh was 0.0047mm and of CoTh was 0.0015mm, we will be able to detect an annual increase of 0.008mm and 0.012mm, respectively; for FEA we will have a power of 80% at a two-sided p=0.05 significance level, if the true rate of difference in estimated failure load between the treatments is 316 units, based on an assumption that the SD is 570 units, as derived from Boutroy *et al.*[89] With 50 completers, we will have an 80% power at a 2-sided p=0.05 level to see a difference in P1NP of 38.6 ng/ml, total abdominal fat of 50.8 cm^2, lean mass of 1.2 kg, vitamin D of 0.025 ng/ml, hsCRP of 0.025 mg/l, and apoB of 18.6 mg/dl per year. These were calculated from our preliminary data administering GH to obese subjects for 6 months. The *power for the mediation analyses* (see below) depends on the effect size of the treatment on BMD and on the correlation of the secondary endpoint on BMD. We derived a formula using the delta method. With at least 40 subjects we will have an 80% chance of achieving significance at a two-sided p=0.05 level if the effect size is 2 units and the correlation is 0.5. *For the 6-month withdrawal period*, we will have adequate power to test whether the BMD remains significantly different from placebo at 24 months. We will have 80% power to detect a difference of 0.04 g/cm^2 assuming that there are at least 43 patients still in the study at 18 months (we expect 50).

Data analysis plan: We will pool data from men and women using a random slopes model; we will consider a model where each patient has a random slope and intercept. The mean intercept and slope may be different for men and women and in each treatment group. *The primary analysis* is a pooled analysis of treatment effect across gender with weights equaling the frequency of men and women in the sample. A *secondary analysis* will test for the significance of the treatment-gender interaction on the rate of change of BMD and all other endpoints. For all analyses, an intent-to-treat analysis will be performed. We plan to apply the guidelines developed by the National Research Council regarding missing data.[122] Our primary analysis, the random slopes model, assumes that data is missing at random. We also will conduct two sensitivity analyses. The first assumes that the missing data is a function of the unobserved random effect[123] and the second tests the assumptions in this model under other missingness mechanisms.[124] We will also determine which of the secondary endpoints appears to *mediate* the treatment effect on an endpoint, such as BMD, microarchitectural parameters, and bone marrow fat by examining the effect of treatment on the potential mediator and of the potential mediator on the endpoint; examples of this will include whether Pref-1 mediates the effects of GH on bone marrow fat or strength, and whether bone marrow fat, body composition variables, inflammatory markers, lipids or vitamin D mediate the effects of GH on bone strength. The effect on one endpoint by a potential mediator is the product of the two effects. The Sobel test will test whether this effect is zero.[125] As there are no known moderators of GH treatment in this population and the population has been purposely selected for osteopenia, we will not control for covariates other than gender in our primary analysis. We will perform exploratory analyses to answer the following: 1. Among an osteopenic population are there covariates

(such as family history, smoking, alcohol, exercise, free testosterone or estradiol) that predict problems such as anomalies in bone microarchitecture and predicted strength? 2. Are these factors moderators of the treatment effect of GH? Both of these analyses will use stepwise variable selection using a random effects model as above. 3. Is the effect of GH on BMD mediated by IGF1? We will perform a mediation analysis using a random effects model with final GH dose and a final GH dose time interaction in the treated group.

Timeline: Screening, enrollment, and study visits will occur from year 1 through year 5. Data management will occur in years 1-5, and analyses and manuscript preparation will be performed in years 4-5.