NOVARTIS

Clinical Development

ZOL446 zoledronic acid, Aclasta®

CZOL446H2337 / NCT00799266

A multicenter, randomized, double-blind, placebo-controlled efficacy and safety trial of intravenous zoledronic acid twice yearly compared to placebo in osteoporotic children treated with glucocorticoids

RAP Module 3 – Detailed Statistical Methodology

Author: [redacted], [redacted] Statistician
[redacted], Novartis Trial Statistician

Document type: RAP Documentation

Document status: V1.3

Release date: 21AUG2017

Statistical methods planned in the protocol and determination of sample size

Statistical and analytical plans

This document describes the statistical methodology for analyses of efficacy and safety in Study CZOL446H2337. The purpose of the study is to determine the bone strengthening effect of zoledronic acid compared to placebo in osteoporotic children treated with glucocorticoids for Duchenne muscular dystrophy or chronic inflammatory conditions including rheumatic conditions and inflammatory bowel disease. To ensure adequate intake of calcium and vitamin D, all children enrolled in this study will be provided with conservative intervention (diet with or without supplements), regardless of their treatment assignment.

The primary objective of the study is to demonstrate that 0.05 mg/kg (max 5 mg per dose) zoledronic acid administered every 6 months is superior to placebo for the change in lumbar spine (LS) bone mineral density (BMD) Z-score at Month 12 relative to baseline.

Categorical data will be presented as frequencies and percentages. For continuous data, mean, standard deviation, median, minimum, and maximum will be presented.

General considerations

Baseline definition

For both efficacy and safety analyses, unless specified otherwise, baseline is defined as the last non-missing measurement obtained on or prior to administration of the first infusion of the study drug. All assessments obtained after the administration of first infusion of the study drug are considered as post-baseline unless otherwise specified.

Study day

The general rules for calculating the study day are as follows:

- If the event date occurs before the first infusion date, then the study day is defined as the date of the event minus the first infusion date
- If the event date occurs after the first infusion date (inclusive), then the study day is defined as the date of the event minus the first infusion date plus 1.

Visit schedule

Summaries by visit will be based on the visits collected on the CRF and no visit window will be applied. Only scheduled visits will be included in the summary, unless specified otherwise.

Subjects and treatments

The following analysis sets will be defined:

- Intention-to-treat (ITT) population

- Modified Intention-to-treat (MITT) population
- Per-protocol (PP) population
- Safety population

Intention-to-treat (ITT) population

The ITT population will consist of all randomized patients. Following the ITT principle, patients will be analyzed according to the treatment they are assigned at randomization. Disposition, protocol deviation, demographics, baseline disease characteristics, medical history, and disease-related background information will be based on ITT population.

Modified Intention-to-treat (MITT) population

The MITT population will consist of all randomized patients who have both baseline and at least one post-baseline LS-BMD Z-score. All efficacy analyses will be based on MITT population including primary efficacy analysis.

Safety population

The Safety population will consist of all patients who have been exposed to at least one infusion of study drug. Patients will be analyzed according to the treatment actually taken. All safety analyses will be based on Safety population.

Per-protocol (PP) population

The PP population will include all patients in the MITT population who do not have any protocol deviations that could confound the interpretation of analyses conducted on the MITT population. Prior to unblinding, the Clinical Trial Team and Trial Statistician will identify the major protocol deviations that will exclude patients from the PP population. Primary efficacy endpoint will also be analyzed in the PP population.

Important protocol deviations for exclusion from analysis sets are specified in Appendix section “PD and non-PD criteria leading to exclusion from analysis sets” and others will be identified by the clinical team before the final database lock.

Pooling of centers

Centers will be pooled into four regions based on the similarity of clinical practice for randomization and analysis purpose. The regions are as follows:

- North America region
  - United States
  - Canada
- EMEA America region
  - Belgium
  - Finland
  - Germany
  - Hungary
  - Italy
  - Poland
  - Romania

  - Russia
  - South Africa
  - Sweden
  - Turkey
  - United Kingdom
- Latin America region
  - Brazil
- APAC region
  - Australia
  - New Zealand

Patient disposition

The number and percentage of patients in the following categories will be tabulated and listed by treatment group within the ITT population:

- Completed study
- Discontinued study

Discontinuations will be further broken down by reason.

In addition, recruitment will be summarized by country and then center.

A complete listing of patient disposition will be provided for all ITT patients.

Patients who have signed informed consent but are not randomized are considered screening failures. Demographic information and reasons for discontinuing will be provided in a listing for screen failure patients.

Protocol deviations

Protocol deviations will be summarized in a table and will be listed for all ITT patients by treatment group. The number of patients with any protocol deviation will also be summarized.

Patients excluded from analyses

The number and percentage of patients in each analysis population (ITT, MITT, PP, Safety) will be summarized by treatment group. Listing of these analysis populations will be provided as well as a listing of patients that belong to the ITT population but are excluded from the efficacy analysis.

Demographic and other baseline characteristics

Descriptive statistics for the following demographic variables and baseline characteristics will be summarized by treatment group using the ITT and MITT populations:

- Age at randomization (years)
- Age group at randomization (> = 5 years to <= 8 years; >=9 years to <=17 years)
- Sex

- Race
- Ethnicity
- [redacted]
- Weight (kg)
- [redacted]
- Tanner evaluation (breast development, genital stage, pubic hair, and whether menarche has occurred)
- Lumbar spine BMD Z-score
- Lumbar spine BMD Z-score category
  - <= -2.0
  - > -2.0 to <= -1.0
  - > -1.0 to <= -0.5
- Lumbar spine Bone Mineral Content (BMC) (g)
- Total body BMC (g)
- Serum 25-hydroxy vitamin D (ng/dL)
- Serum calcium level (mg/dL)
- Serum NTX (nmol/L)
- Serum TRAP-5b (U/L)
- Serum BSAP (U/L)
- Serum P1NP (ng/mL)
- Vertebral morphometry
- Non-vertebral fractures
- Metacarpal cortical width
- [redacted]
- Pain score

In addition, the baseline comparability between treatment groups will be evaluated for demographic and other baseline characteristics variables. Categorical variables will be evaluated using Fisher's exact test, and the continuous variables will be evaluated using a Wilcoxon rank-sum test.

Note that these tests of comparability are performed for descriptive purpose only, and will not be considered to define any formal basis for determining factors which should be included in statistical analysis models. However, when these tests yield significant results they can be used as extra information in interpreting any treatment-by-factor interactions that are observed in the sensitivity analyses performed on the primary and secondary efficacy variables.

All demographic and baseline characteristic data will be listed as presented above.

Disease-related background information

Disease-related background and historical information will be summarized by the following type and sub-type of disease as collected on the CRF.

- Rheumatic conditions:
  - Systemic onset Juvenile Idiopathic Arthritis (JIA)
  - Polyarticular rheumatoid factor positive JIA
  - Polyarticular rheumatoid factor negative JIA
  - Psoriatic arthritis JIA
  - Enthesitis-related arthritis JIA
  - Oligoarticular arthritis JIA
  - Unclassified JIA
  - Systemic lupus erythematosus (SLE)
  - Juvenile dermatomyositis
  - Scleroderma (generalized or localized)
  - Overlap syndromes (including mixed connective tissue disease)
  - Sjogren's syndrome
  - Giant cell (temporal) arteritis
  - Takayasu's arteritis
  - Polyarteritis nodosa
  - Wegener's granulomatosis
  - Churg-Strauss syndrome
  - Microscopic polyangiitis
  - Essential cryoglobulinemic vasculitis
  - Cutaneous leukocytoclastic angiitis
  - Behcet's disease
  - Other vasculitis
- Inflammatory bowel disease:
  - Crohn's disease
  - Ulcerative colitis
- Duchenne muscular dystrophy
- Other

Frequency counts will be provided by treatment group as well as the total in a table. A listing will also be generated.

In addition, the following disease types and sub-types will be derived based on the preferred term from the medical history/current medical conditions panel.

- Chronic rheumatic conditions:
    - Juvenile Idiopathic Arthritis (JIA),Systemic onset; JIA
    - Polyarticular rheumatoid factor positive; JIA
    - Polyarticular rheumatoid factor negative; JIA
    - Psoriatic arthritis; JIA
    - Enthesitis-related arthritis; JIA
    - Oligoarticular arthritis; JIA Unclassified
    - Systemic lupus erythematosus (SLE)
    - Juvenile dermatomyositis
    - Scleroderma (generalized or localized)
    - Overlap syndromes (including mixed connective tissue disease)
    - Sjogren's syndrome
    - Systemic vasculitis
    - Giant cell (temporal) arteritis
    - Takayasu's arteritis
    - Polyarteritis nodosa
    - Wegener's granulomatosis
    - Churg-Strauss syndrome
    - Microscopic polyangiitis
    - Essential cryoglobulinemic vasculitis
    - Cutaneous leukocytoclastic angiitis
    - Behcet's disease
    - Other vasculitis
- Inflammatory bowel disease:
    - Crohn's disease
    - Ulcerative colitis
- Duchenne muscular dystrophy
- Nephrotic syndrome
- Other

Frequency counts will be provided for the derived disease types and sub-types by treatment group as well as the total in a table.

Relevant medical history and current medical conditions

Relevant medical history/current medical conditions will be summarized by primary system organ class, preferred term and treatment group with frequencies and percentages displayed. A patient with multiple occurrences of a medical condition for a preferred term or system organ class is counted only once in each specific category. A corresponding listing will also be provided.

Extent of exposure

A summary table will be provided for the exposure to study medication, including number of patients with any infusion, duration for 1st infusion, duration for 2nd infusion, volume for 1st infusion, and volume for 2nd infusion.

Prior/concomitant medications and significant non-drug therapies

The number and percentage of patients receiving prior and concomitant medications and significant non-drug therapies will be summarized by treatment group and preferred term according to the WHO Drug Reference List employing the Anatomical Therapeutic Chemical (ATC) classification system. Prior medications are defined as those ending prior to the first infusion of study drug; medications ending on or after the first infusion of study drug are concomitant medications. Prior medications will be presented separately from concomitant medications. The summary table will first be sorted by ATC class alphabetically and then by preferred term in decreasing order of frequency with respect to usage in the zoledronic acid group.

Furthermore, summarizations by ATC class will be provided with regard to a group of specified concomitant medications/significant non-drug therapies, which include osteoporosis-related medications, nutritional supplements, and glucocorticoids. The summary table will be sorted by ATC class in decreasing order of frequency in the zoledronic acid group.

Missing date convention

If the medication start and end dates are recorded in the CRFs, medications that start on or prior to the first infusion in study and end on or after the first infusion in the study including ongoing at end of the study were included in both prior to and after start of study drug administration in the study, respectively.

Partial medication start dates will be imputed as follows:

- Only day is missing: the partial start date will be imputed using 15MONYYYY if the month and year indicate that the medication started prior to the first infusion of study medication or 01MONYYYY if the month and year indicate that the medication started on or after the first infusion of study medication.
- Both day and month are missing: the partial start date will be imputed using 01JULYYYY if the year indicates that the medication started prior to the first infusion of study or 01JANYYYY if the year indicates that the medication started on or after the first infusion of study medication.

Partial medication end dates will be imputed as follows:

- Only day is missing: the partial end date will be imputed using the last date of the month.
- Both day and month are missing: the partial end date will be imputed as December 31 if the end year is prior to the year of first infusion of study drug or one day after the

last study infusion date if end year is equal to or after the year of first infusion of study drug.

Efficacy evaluation

Analyses of all efficacy endpoints will be performed on the MITT population. Confirmation of the result for the primary efficacy endpoint will be conducted using the PP population without imputation for missing data. All efficacy measurements will be presented in listings.

Analysis for the primary objective

The primary objective is to demonstrate that zoledronic acid administered every 6 months at 0.05 mg/kg (max 5 mg) is superior to placebo for the change in lumbar spine (LS) bone mineral density (BMD) Z-score at Month 12 relative to baseline.

Variable

The primary efficacy variable will be the change in LS-BMD Z-score at Month 12 relative to baseline, in patients treated with twice yearly intravenous (i.v.) zoledronic acid compared to patients treated with placebo. It is derived as (LS-BMD Z-score at Month 12 – LS-BMD Z-score at baseline). The LS-BMD Z-score will be assessed at Screening, Month 6 and Month 12.

Positive changes from baseline indicate an improvement in condition. Treatment difference will be calculated as <zoledronic acid minus placebo>.

Statistical model, hypothesis, and method of analysis

Let μ_Z and μ_P be the population mean of the change in LS-BMD Z-score at Month 12 relative to baseline for patients treated with zoledronic acid and patients receiving placebo, respectively. The null hypothesis of no difference between treatment groups (H_0: $\mu_Z = \mu_P$) will be tested against the alternative hypothesis that there is difference between treatment groups (H_A: $\mu_Z \neq \mu_P$). The test will be conducted at a nominal significance level of α=0.05.

The primary efficacy analysis will be conducted in the MITT population based on an analysis of covariance (ANCOVA) model with treatment, pooled centers, underlying condition treated with glucocorticoids and baseline LS-BMD Z-score as explanatory variables and pooled centers as random effect. The t-test statistic, $|d|/SE$, will be computed and compared with $t_{df,\ 0.975}$ to determine significance, where d and SE are the estimated treatment difference and the estimated standard error of d from the ANCOVA model, respectively, df is the associated degrees of freedom computed from the error term in the model, and $t_{df,\ 0.975}$ is the 97.5% quantile of the t distribution with df degrees of freedom.

The test and confidence intervals based on this test statistic are robust to the assumptions of normality and of equal variances when within-center treatment group sample size are equal, which is generally the case that treatment allocation is nearly equal within each center because of the manner in which the randomization is carried out. Nevertheless, the underlying assumption of normality and variance homogeneity for the primary analyses will be reviewed and tested by Shapiro's test and Levene's test, respectively, prior to unblinding the data. If the underlying assumption is violated, a transformation or non-parametric approach, such as van

Elteren or Cochran-Mantel-Haenszel test adjusted for center, will be applied to the data as appropriate for the hypothesis evaluated.

From the model, the treatment difference between zoledronic acid group vs. placebo group will be estimated and presented together with two-sided 95% confidence interval (CI). P-value will be rounded to the 4th decimal place. In the same table, the Least Squares means and standard errors for change from baseline in lumbar spine Z-score at Month 12 will be presented for each treatment group. Descriptive statistics will also be shown for baseline and Month 12 visits.

A plot will be generated for LS BMD Z-score over time.

Handling of missing values/censoring/discontinuations

In the MITT population, last observation carried forward (LOCF) will be used to impute a value for LS-BMD Z-score at Month 12 for any patient who does not have a Month 12 assessment; however, if a patient prematurely discontinues from the study during the double-blind phase without providing any LS-BMD measurements at Month 6 or Month 12, then a zero percent change in LS-BMD will be assumed at Month 12.

Given that a pediatric population is being studied and patients are expected to experience bone growth during the study, an assumption of zero change over 12 months and LOCF from Month 6 measurements would be considered conservative.

Supportive analyses

Confirmation of the result for the primary efficacy variable will be performed using the per-protocol population without imputation for missing data using the ANCOVA model applied for the primary efficacy analysis.

As another supportive analysis to the primary analysis, the homogeneity of treatment effects across centers will be evaluated by adding the treatment-by-center interaction term into the ANCOVA model employed in the primary efficacy analysis. If $p < 0.1000$ for the treatment-by-center interaction test, which indicates heterogeneous treatment effects across centers, then an additional table will be presented showing the ANCOVA model-based treatment comparisons for primary efficacy variable by each center. Furthermore, interactions between treatment and underlying condition, treatment and baseline tanner stage in pubic hair, treatment and baseline tanner stage in breast development within female population, and treatment and baseline tanner stage in genital stage within male population will be evaluated and presented in a similar way.

In addition, a responder analysis will be done by comparing the proportion of patients who increase their LS-BMD Z-score at Month 12 compared to baseline in the zoledronic acid group relative to the placebo group using the MITT population. Between-treatment differences will be evaluated using a logistic regression model with treatment, pooled centers, underlying condition treated with glucocorticoids and baseline LS-BMD Z-score as explanatory variables.

Analysis for secondary objectives

The analyses of secondary efficacy variables will be performed on the MITT population without imputation for missing values unless specified otherwise. Descriptive statistics will be presented by treatment group for each continuous variable and the number and percentage of patients will be presented by treatment group for each categorical variable, unless otherwise specified. All hypothesis tests will be evaluated at a 0.05 level of significance, and the p-values will be rounded to the 4th decimal place. Adjustment for multiple comparisons will not be made for any of the secondary efficacy variables since none of these variables is intended to be used for any promotional claims.

The secondary efficacy variables and their analysis methods are described below.

- **Change in LS-BMD Z-score at Month 6 relative to baseline, zoledronic acid group vs. placebo**

The change in LS-BMD Z-score at Month 6 relative to baseline between zoledronic acid and placebo will be evaluated using an ANCOVA model with treatment, pooled centers, underlying condition treated with glucocorticoids and baseline LS-BMD Z-score as explanatory variables.

From the model, the treatment difference between zoledronic acid group vs. placebo group will be estimated and presented together with two-sided 95% CI and p-value. In the same table, the LS means and standard errors for change from baseline in LS-BMD Z-score at Month 6 will be presented for each treatment group. Descriptive statistics will also be shown for baseline and Month 6 visits.

- **Change in LS and total body BMC at Months 6 and 12 relative to baseline, zoledronic acid group vs. placebo**

The change in BMC at the lumbar spine and total body at Month 6 and Month 12 relative to baseline between zoledronic acid and placebo will be evaluated using an ANCOVA model with treatment, pooled centers, underlying condition treated with glucocorticoids and baseline BMC as explanatory variables.

From the model, the treatment difference between zoledronic acid group vs. placebo group will be estimated and presented together with two-sided 95% CI and p-value. In the same table, the LS means and standard errors for change from baseline in lumbar spine or total body BMC at Month 6 and Month 12 will be presented for each treatment group. Descriptive statistics will also be shown for baseline, Month 6 and Month 12 visits.

- **Relative change in biochemical markers of bone turnover (Serum NTX, TRAP-5b, BSAP and P1NP) at Months 6 and 12 relative to baseline, zoledronic acid group vs. placebo**

The analysis of changes from baseline to Month 6 and Month 12 in the biochemical markers, serum NTX, TRAP-5b, BSAP, and P1NP will be performed as follows: a transformation of the $\log_e$ ratio of post-baseline value vs. baseline value at each visit will be used to normalize the distribution of the biochemical marker parameters. An ANCOVA model with treatment group, pooled centers, underlying condition treated with glucocorticoids, and $\log_e$ (baseline

value) as explanatory variables will be used on the transformed data at each post-baseline time point.

From the model, the treatment difference between zoledronic acid group vs. placebo group in log scale will be estimated and presented together with two-sided 95% CI and p-value. Estimated treatment difference and two-sided 95% CI will then be transformed back to the original scale and displayed on the table. In the same table, the LS means and standard errors in both log scale and original scale for change from baseline for the biomarkers at Month 6 and Month 12 will be presented for each treatment group. Descriptive statistics will also be shown for baseline, Month 6 and Month 12 visits in both original scale and log scale. A plot will be generated for each biochemical marker over time respectively.

- **Incidence of new vertebral fractures at Month 12, zoledronic acid group vs. placebo**

The number and percentage of patients with new vertebral fractures at Month 12 will be presented by treatment group. Between-treatment differences will be evaluated using Fisher's exact test.

- **Change in vertebral morphometry at Month 12 relative to baseline, zoledronic acid group vs. placebo**

The change in vertebral morphometry at Month 12 relative to baseline will be evaluated using an ANCOVA model with treatment, pooled centers, underlying condition treated with glucocorticoids and baseline vertebral morphometry as explanatory variables.

From the model, the treatment difference between zoledronic acid group vs. placebo group will be estimated and presented together with two-sided 95% CI and p-value. In the same table, the LS means and standard errors for change from baseline in vertebral morphometry at Month 12 will be presented for each treatment group. Descriptive statistics will also be shown for baseline and Month 12 visits.

- **Change in 2nd metacarpal cortical width at Month 12 relative to baseline, zoledronic acid group vs. placebo**

The change in 2nd metacarpal cortical width at Month 12 relative to baseline will be evaluated using an ANCOVA model with treatment, pooled centers, underlying condition treated with glucocorticoids and baseline metacarpal cortical width as explanatory variables.

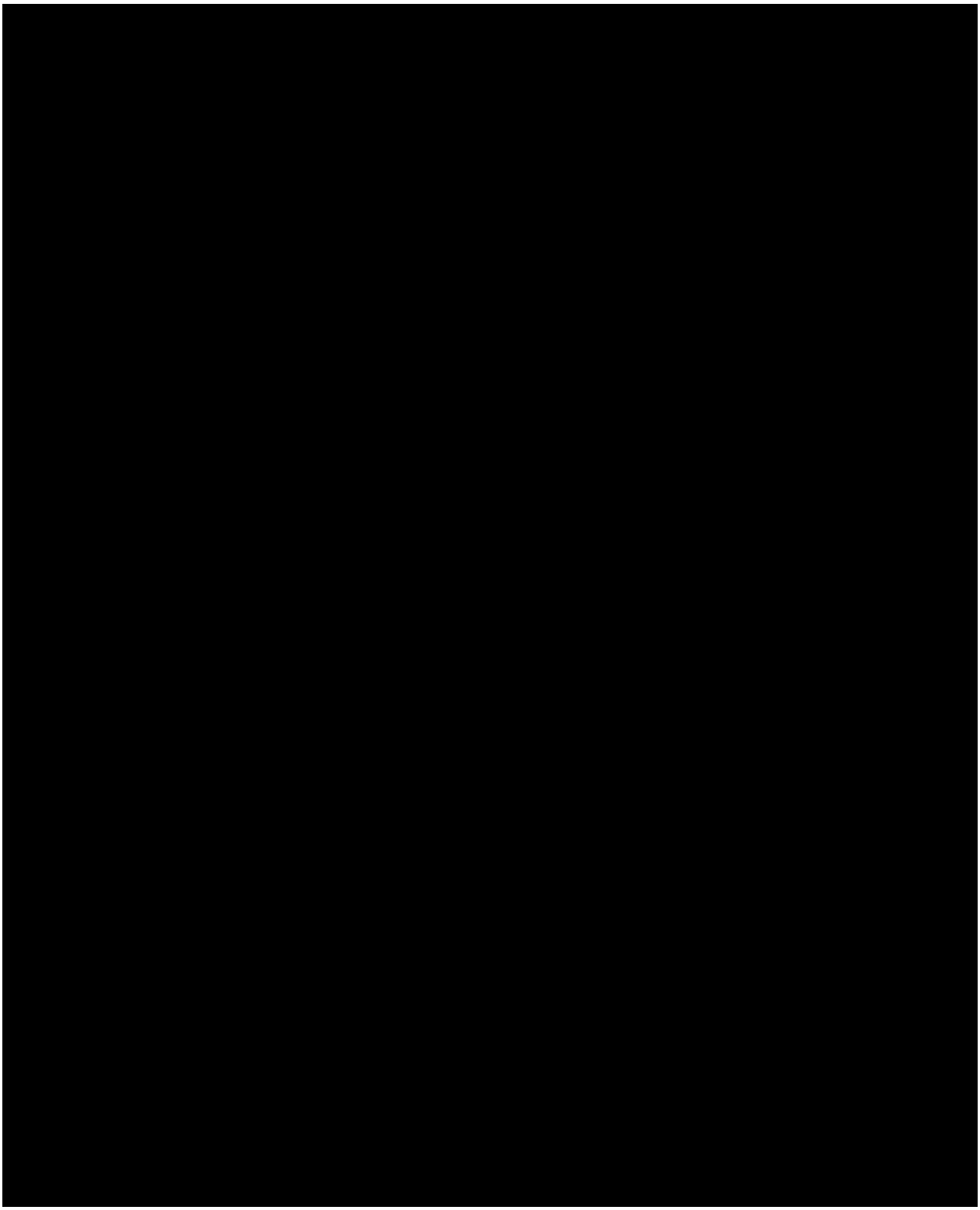
From the model, the treatment difference between zoledronic acid group vs. placebo group will be estimated and presented together with two-sided 95% CI and p-value. In the same table, the LS means and standard errors for change from baseline in metacarpal cortical width at Month 12 will be presented for each treatment group. Descriptive statistics will also be shown for baseline and Month 12 visits.

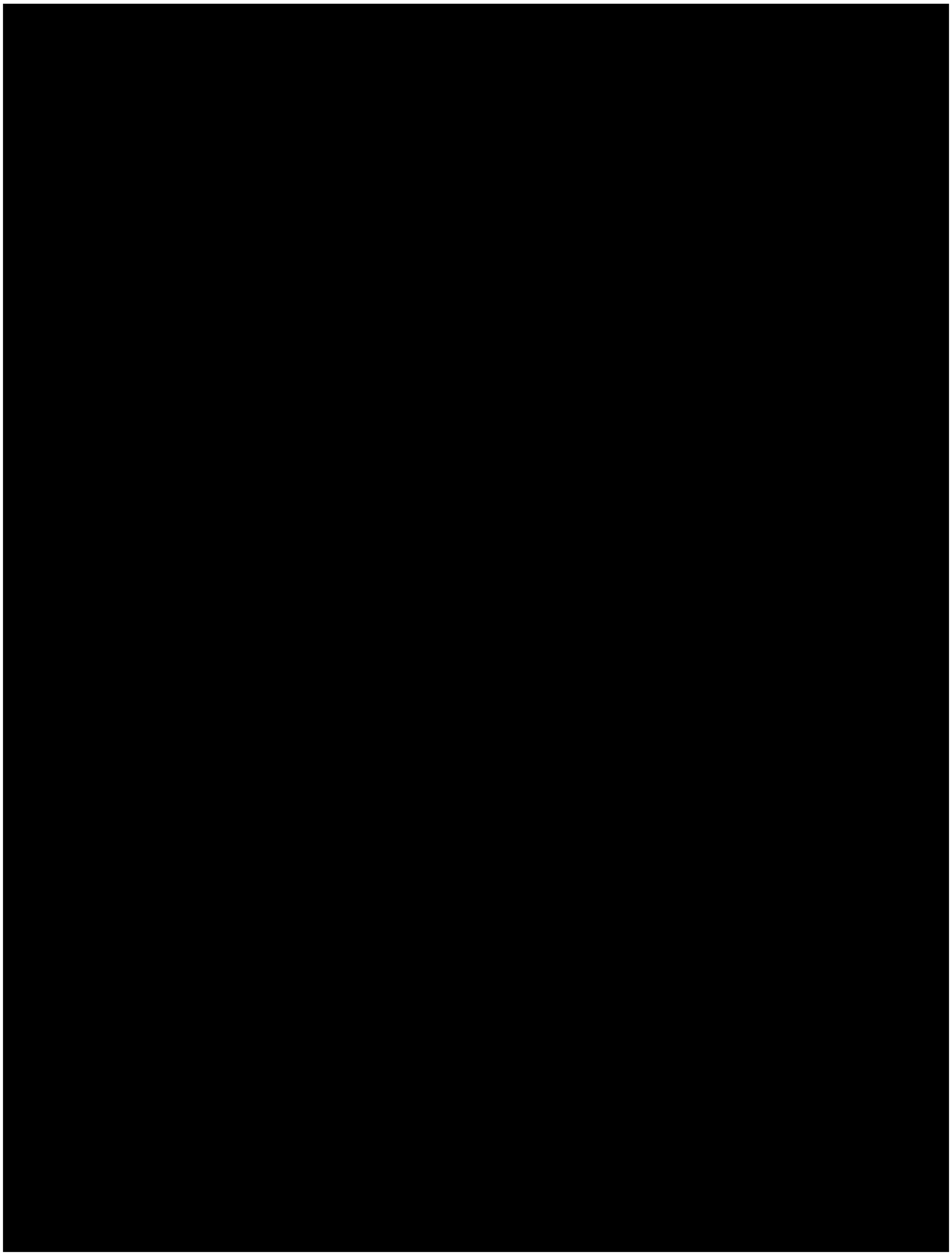
- **Reduction in pain from baseline, zoledronic acid group vs. placebo**

Pain is assessed using Faces Pain Scale-Revised (FPS-R) at randomization, Month 3, 6, 9 and 12, which ranges from 0 (No Pain) to 10 (Very Much Pain). Details of FPS-R can be found in Appendix section "Faces Pain Scale – Revised" in the protocol.

The reduction in pain from baseline by visit will be evaluated based on whether or not patients have a decrease in their FPS-R from baseline. If pain remains the same or worsens from

baseline a patient will be classified as ‘0’ and if the pain scale decreased then the patient will be classified as ‘1’. Whether or not a patent experiences a decrease in pain will be evaluated using a logistic regression model with treatment, pooled centers, underlying condition treated with glucocorticoids and baseline pain score as explanatory variables.





Pharmacokinetic (PK) evaluations

Descriptive summary statistics will be provided by treatment for urine concentration of zoledronic acid at Month 12, including mean (arithmetic and geometric), standard deviation (arithmetic and geometric), coefficient of variation (CV) (arithmetic and geometric), median, first quartile (Q1), third quartile (Q3), minimum and maximum. Furthermore, frequency (n, %) of zoledronic acid concentrations below the Lower Limit of Quantification (LLOQ) will be displayed, which is defined as 10 ng/mL. Concentrations below LLOQ will be treated as zero in summary statistics. A geometric mean will not be reported if the dataset includes zero values. The same summary statistics will also be provided for PK parameter Aet1-t2 at Month 12, which is calculated as urine volume (t1-t2) x urine concentration. A listing of concentrations of zoledronic acid at Month 12 will also be provided.

Safety evaluation

All safety analyses will be performed on the Safety Population. The assessment of safety will include adverse events, laboratory tests (hematology, blood chemistry and urine values), vital signs, and tanner staging score. Patient listings will be provided for all safety data.

Adverse events

An adverse event (AE) is any untoward medical occurrence in a subject or clinical investigation subject after providing written informed consent for participation in the study until the end of study visit. Treatment-emergent adverse events will be defined as adverse events that were absent prior to the first infusion and occurred after the first infusion during the study or that were present prior to the first infusion and occurred at increased severity after the first infusion during the study. Summarization will be based on treatment-emergent adverse events. Adverse events will be coded using the Medical dictionary for regulatory activities (MedDRA). The number and percentage of patients who report adverse events will

be summarized by treatment according to the primary system organ class and preferred term. If a patient reports more than one adverse event within the same primary system organ class (preferred term), the patient will be counted only once with the greatest severity at the system organ class (preferred term) level. Serious adverse events and selected adverse events will be presented by treatment group code. Listing of these events will be provided.

Adverse event tables will be sorted alphabetically by system organ class and then according to decreasing frequency of preferred term in the zoledronic acid group. If only system organ classes or preferred terms are presented, these will be sorted according to the decreasing frequency in the zoledronic acid group.

The following sets of adverse events will be summarized:

- All adverse events (by primary system organ class and preferred term)
- Adverse events suspected to be related to study drug (by primary system organ class and preferred term)
- Adverse events that occurred after first infusion and before second infusion (by primary system organ class and preferred term)
- Most frequent adverse events (at least 5% in any group) (by primary system organ class and preferred term)
- All adverse events (by primary system organ class, preferred term and maximum severity)
- All adverse events (by underlying disease group, primary system organ class and preferred term)
- Adverse events that result in death (by primary system organ class and preferred term)
- Serious adverse events (by primary system organ class and preferred term)
- Adverse events causing permanent discontinuation of study drug (by primary system organ class and preferred term)
- Adverse events associated with a change in renal function (by preferred term). Time to first adverse event associated with a change in renal function will be presented separately.
- [redacted]
- Hypocalaemia adverse events (by preferred term). Time to first hypocalaemia occurrence based on AE and central lab calcium will be presented separately.

Incidence of risk as defined in the case retrieval strategy, during the study period, will be summarized by treatment group, and MedDRA levels. Relative risk of zoledronic acid group

vs. placebo will be presented along with corresponding 95% confidence intervals using the Mantel-Haenszel test.

Missing date convention

Partial AE start dates: If an AE start date is completely missing, then no imputation is implemented. AEs with completely missing start dates will be treated as treatment emergent. If the year of an AE start date is missing but the month is available, then the date will be imputed as the start date of the first study drug infusion + 1.

If the year is not missing and the month is missing, then date will be imputed as follows:

- If the AE year is less than the year the study drug infusion was administered, then July 1st of the year will be used.
- If the AE year is equal to the year the study drug infusion was administered, then the start date will be imputed as the start date of the first study drug infusion + 1.
- If the AE year is greater than the year the study drug infusion was administered, then the Jan 1st of the year will be used.

If the year is available and the month is non-missing, then the following rules will be applied relative to the month that study drug treatment was given:

- If the month in which the AE started is less than the month that the study drug infusion was administered, then:
    - If the AE year is less than or equal to the year the study drug infusion was administered, then the 15th day of the month will be used.
    - If the AE year is greater than the year the study drug infusion was administered, then the maximum of the first day of the month and the treatment start date + 1 will be used.
- If the month in which the AE started is equal to the month that the study drug infusion was administered, then:
    - If the AE year is less than the year the study drug infusion was administered, then the 15th day of the month will be used.
    - If the AE year is greater than or equal to the year the study drug infusion was administered, then the maximum of the first day of the month and the treatment start date + 1 will be used.
- If the month in which the AE started is greater than the month that the study drug infusion was administered, then:
    - If the AE year is less than the year the study drug infusion was administered, then the 15th day of the month will be used.
    - If the AE year is greater than or equal to the year the study drug infusion was administered, then the maximum of the first day of the month and the treatment start date + 1 will be used.

In case that the AE end date was completed, the AE start date imputed as above was after the AE end date, the AE start date was re-imputed using the AE end date.

Partial AE end dates: There is no STL convention to impute partial AE end date. Partial AE end date was imputed using (a) the last date of the month if only the day of AE end date is missing, (b) December 31 if only the year of AE end date is present. If a patient dies and the imputed end date is after the date of death, the date of death was imputed as the AE end date. When a patient dies, the end date for ongoing AEs was imputed using the date of death.

Adverse events associated with change in renal function

Adverse events associated with change in renal function are defined based on predefined search list which describe a clinically significant change in renal function. A search of the clinical trial database will occur based on this pre-specified list of MedDRA preferred terms. These adverse events will be summarized by preferred term. In addition, the event rate and 25^{th} percentile (95% CI) of time to first adverse event associated with change in renal function will be summarized. The hazard ratio and 95% CI based on a Cox proportional hazards model with treatment as a factor will be presented. Between-treatment differences in time to first adverse event associated with change in renal function will be evaluated using the log-rank test. If a patient experiences an adverse event associated with change in renal function, the event time is defined as the earliest of all events if there are multiple. Otherwise, a patient is considered censored at the last contact date. Kaplan-Meier estimates will be plotted over time if at least 1% of overall population has experienced the event.

Patient listings will be provided for all adverse events associated with clinically significant change in renal function.

Hypocalcaemia adverse events and laboratory events

Hypocalcaemia adverse events will be searched based on preferred term and summarized. Patients with reported laboratory measurements associated with a signal of hypocalcaemia will be summarized by visit. Lab hypocalcaemia (low serum calcium levels) is present if the value of serum calcium from central lab is less than the lower limit of normal range provided by the central lab.

The number and percentage of patients who met these criteria will be summarized by treatment group overall.

The event rate and 25^{th} percentile (95% CI) of time to first hypocalaemia occurrence based on AE and central lab calcium will be summarized. The hazard ratio and 95% CI based on a Cox proportional hazards model with treatment as a factor will be presented. Between-treatment differences in time to first hypocalaemia occurrence based on AEs and central lab will be evaluated using the log-rank test. If a patient experiences a hypocalaemia and low calcium value based on the central lab, the event time is defined as the earlier of these events. Otherwise, a patient is considered censored at the last contact date. Kaplan-Meier estimates will be plotted over time if at least 1% of overall population has experienced the event.

Patient listings will be provided for all adverse events associated with hypocalcaemia.

Laboratory parameters

The summary of laboratory evaluations will be presented with respect to three groups of laboratory tests (hematology, serum chemistry, and urinalysis).

- Hematology parameters: hemoglobin, hematocrit, red blood cell count, white blood cell count with differential counts, and platelet count.

- Serum chemistry parameters: non-fasting glucose, creatinine, serum urea, uric acid, total protein, SGOT (AST), SGPT (ALT), alkaline phosphatase, sodium, potassium, chloride, phosphorus, magnesium, albumin, calcium; and Glomerular Filtration Rate (GFR).

- Urinalysis parameters: amorphous crystals, urine bilirubin, urine blood, calcium oxalate crystals, urine color, urine glucose, urine ketone, urine leukocytes, urine nitrite, pH, urine protein dipstick, specific gravity, urine urobilinogen, and WBC Casts (Urine) /LPF.

GFR will be calculated using the Schwartz equation at all visits where serum creatinine will be measured (Visits 1, 3, 5, 6 and 8) by the Central Laboratory.

$$\text{GFR (mL/min/1.73 m}^2\text{)} = k[\text{Height (m)}]/\text{Serum Creatinine (mg/dL)}, \text{ where } k=0.41$$

Descriptive summary statistics for baseline, each post-baseline visit, and change from baseline to each visit will be presented by laboratory test group and treatment group. Note that if either the baseline or the post-baseline value is missing, then the calculation of change from baseline will also be missing.

Also, shift tables for hematology and serum chemistry will be provided in order to compare a patient's baseline laboratory evaluation relative to each study visit. For the shift tables, the normal laboratory values will be used to evaluate whether a particular laboratory test value is normal, low, or high for each visit value relative to whether or not the baseline value is normal, low, or high. These summaries will be presented by laboratory test group and treatment group.

In addition, a summary table will be presented for the incidence rates of newly occurring clinically notable laboratory abnormalities, where patients will be counted if they have normal or missing baseline value and with at least one post-baseline assessment satisfying the clinically notable criteria. Incidence rates of newly occurring hypocalcaemia (based on central lab) by visit will be summarized similarly for patients who have normal or missing baseline level of serum calcium and clinically notable abnormal serum calcium level at each scheduled post-baseline visit. The definition of clinically notable laboratory abnormalities can be found in Appendix section "Clinically notable laboratory values and vital signs" in the protocol.

Furthermore, a summary table will be provided for the number and percentage of patients with liver enzyme abnormalities by each study visit and with at least one post-baseline assessment satisfying the liver enzyme abnormalities.

Listings of all laboratory parameter measurements will be displayed. In addition, listings of hematology and biochemistry assessments will be provided for patients with newly occurring clinically notable laboratory abnormalities. In these listings, all laboratory values will be

displayed for any patient who has at least one newly occurring clinically notable laboratory abnormality. The abnormal value will be flagged, e.g., L/H for a value being below/above normal range, and the lower and upper limit of normal values will be presented as well. Central lab calcium values will also be listed for patients with at least one post-baseline laboratory signal of hypocalcaemia and the below normal serum calcium results will be flagged as L.

For each numerical laboratory parameter, the number of decimal places to be used for laboratory values should be the same as that for the normal range. For continuous variables recorded as <*lower limit*>, these will be imputed as being half of the lower limit.

All laboratory parameters will be expressed using SI units.

Vital signs, [redacted] and weight

Descriptive summary statistics for the baseline, each study visit, and change from baseline to each study visit will be presented for sitting systolic blood pressure, sitting diastolic blood pressure, pulse rate, and body weight. [redacted]. The number and percentage of patients who have vital sign values that meet the criteria for being clinically notable after the first infusion of study medication will be presented by treatment group. In addition to the listing of all vital sign data, a patient listing will also be provided for patients with newly occurring clinically notable vital signs. The definition of clinically notable vital sign abnormalities can be found in Appendix section "Clinically notable laboratory values and vital signs" in the protocol.

Tanner staging scale

Tanner staging scale data will be summarized at baseline (Visit 2), Month 6 (Visit 5) and End of study (Visit 8). The number and percentage of patients in each Tanner staging scale category will be included in the summary. The following tanner staging scales will be summarized:

- Female
    - breast development
    - pubic hair
    - menarche occurrence
- Male
    - genital stage
    - pubic hair

A listing of tanner staging scale by treatment and gender will also be provided.

Determination of sample size

The sample size is calculated to show superiority of zoledronic acid relative to placebo at Month 12.

The null hypothesis of no difference between treatment groups in the mean change in lumbar spine BMD Z-score at Month 12 relative to baseline. In testing this hypothesis, it is assumed

that the standard deviations for the two treatment groups is 0.93 under H_0 and H_A. Based on a primary endpoint of the change in lumbar spine BMD Z-score at 12 months relative to baseline assuming a 0.63 increase in lumbar spine BMD Z-score over 12 months with zoledronic acid and no change in the Z-score in the placebo group and a common standard deviation of 0.93, 82 patients are required to show a significant increase in lumbar spine BMD Z-score at the 0.05 level with 85% power. Adjusting for a dropout rate of 10%, approximately 46 patients per treatment group are required.

The amount of information from controlled trials in a pediatric GIO population is limited. However, the 0.63 increase in lumbar spine BMD Z-score is based on the increase observed with pamidronate-treated children on chronic glucocorticoid therapy presented in (Acott et al 2005).

The common standard deviation of 0.93 was estimated from the pamidronate-treated children and untreated control evaluated in (Acott et al 2005). This standard deviation is also consistent with the estimate of the standard deviation obtained from the zoledronic acid group of pediatric osteogenesis imperfecta patients in Study CZOL446H2202.

References

Acott DA, Lang BA, Wong JA, et al (2005) Pamidronate treatment of pediatric fracture patients on chronic steroid therapy. Pediatric Nephrology, vol 20 : 368 - 373

Analysis of covariance (ANCOVA) model

The following is an example of the SAS code used to perform these analyses.

```
ods output lsmeans=lsmeans
           diffs=lsmeandiffcl;
proc mixed data=xxx;
  class treat center gluco ;
random center;
  model CHGVAR = treat center gluco BASVAR;
  lsmeans treat / pdiff=control ('1') alpha=.05 cl;
quit;
```

Adjusted change from baseline

Mean (SE): lsmeans.LSMean (lsmeans.STDERR)

Diff, 95% CI: lsmeandiffcl.DIFFERENCE (lsmeandiffcl.LowerCL, lsmeandiffcl.UpperCL)

p-value (pairwise): lsmeans.ProbtDiff

Logistic regression model

The following is an example of the SAS code used to perform these analyses.

```
ods output parameterestimates=EST oddsratios=ODDS;
proc logistic data=xxx descending;
format _all_;
  class treat center gluco/param=ref ref=first ;
  model RESP = treat center gluco BASVAR;
run;
```

Odds ratio: ODDS.oddsratioest where EFFECT contains 'treat'

95% CI for odds ratio: ODDS.LOWERCL and ODDS.UPPERCL where EFFECT contains 'treat'

P-value: EST.PROBCHISQ where VARIABLE='treat'

Appendix

PD and non-PD criteria leading to exclusion from analysis sets

Analysis Set	PD ID (that causes subjeects to be excluded)	Non-PD criteria (that causes subjeects to be excluded)
ITT		Not randomized
MITT		Not in ITT; No baseline BMD Z-score; No post-baseline Z-score
SAF		No study drug taken
PPS	All inclusion and exclusion criteria; S01; S02; S03; M01	Not in MITT