Statistical Analysis Plan

Protocol Title:	Prospective, Multicenter, Single-arm Study to Evaluate Efficacy, Safety, and Pharmacokinetics of Denosumab in Children With Osteogenesis Imperfecta
Short Protocol Title:	Prospective, Multicenter, Single-arm Study to Evaluate Efficacy, Safety, and Pharmacokinetics of Denosumab in Children With Osteogenesis Imperfecta
Protocol Number:	20130173
NCT Number:	NCT02352753
Authors:	[redacted]
Sponsor:	AMGEN, Inc. One Amgen Center Drive Thousand Oaks, CA 91320, USA Tel: +1 (805) 447-1000
SAP Date:	**Document Version** — **Date** **Amendment 4 (v5.0)** — **27 Apr 2022**

Version Number	**Date (DDMMMYYYY)**	**Summary of Changes, including rationale for changes**
Original (v1.0)	23 Oct 2015	NA. The first version.
Amendment 1 (v2.0)	13 Feb 2020	This statistical analysis plan (SAP) was updated based on the protocol amendment 4 for Denosumab Study 20130173 dated 24 March 2016 and the German country specific substudy protocol supplement dated 06 August 2015 to perform primary analysis of Denosumab in lumbar spine bone mineral density at 12 months on a 6-month dosing regimen.
Amendment 2 (v3.0)	13 Apr 2020	The following changes are made in version 2 of the SAP. Section 2.2- Secondary objectives related to incidence fractures were updated. Section 4.1- Secondary efficacy endpoints related to incidence fractures were updated. Section 7- The DXA analysis set and non-vertebral fracture analysis set were updated.
Amendment 3 (v4.0)	12 Nov 2021	This statistical analysis plan (SAP) is updated based on the protocol amendment 7 for Denosumab study 20130173 dated 22 October 2021 and the German country specific substudy protocol supplement dated 28 October 2021 to perform final analysis.
Superseding amendment 3 (V4.1)	18 Feb 2022	This statistical analysis plan (SAP) is updated to remove by age cohort summary except for hypercalcemia

		adverse event and to remove the pooled adverse events summary for the combined dosing regimen. Also its updated to provide the listings for 3-month dosing regimen german subjects
Amendment 4 (V5.0)	**27 Apr 2022**	**The following changes are made in version 5 of the SAP.** **Section 6.4.8- Added the definition for German substudy safety analysis set.** **Section 5.1- The definition for treatment emergent adverse event for 6-month dosing regimen is updated to include all the adverse events.** **Section 5.7.4 and 9.5.2.2- Provided the details for source data of non-vertebral fractures and long bone fractures to include the maximum fractures. Appendix C- Updated the upper bound for renal laboratory in visit window.**

Table of Contents

List of Table

List of Appendices

List of Abbreviations

Abbreviation	Explanation
AE	Adverse Event
BMC	Bone mineral content
BMD	Bone mineral density
BMDS	Biomedical data stewardship standard
BMI	Body mass index
BSAP	Bone-specific alkaline phosphatase
BTM	Bone turnover marker
CHAQ	Childhood Health Assessment Questionnaire
CHQ-PF	Child Health Questionnaire - Parent Form
CT	Computerized tomography
DMC	Data monitoring committee
DoB	Date of birth
DPN	Deoxypridinoline
DXA	Dual-energy X-ray absorptiomet
DPD	Deoxypyridinoline
EC	Ethics committee
eCRF	Electronic case report form
IC	Informed consent
IP	Investigational product
IVRS	Interactive voice response system
MRI	Magnetic resonance imaging
OI	Osteogenesis imperfecta
NTX	Cross-linked N-telopeptides of type 1collagen
NA	Not Applicable
PK	Pharmacokinetic
PKDM	Pharmacokinetic and drug metabolism
Q6M	**6-month dosing regimen**
Q3M	**3-month dosing regimen**
SAP	Statistical analysis plan
sCTX	Serum type I collagen C-telopeptide
SQ	Semiquantitative
SAS	Statistical Analysis System

WBFPRS	Wong-Baker FACES Pain Rating Scale

1. Introduction

The purpose of this Statistical Analysis Plan (SAP) is to provide details of the statistical analyses that have been outlined within the protocol amendment 7 for Denosumab study 20130173 dated 22 October 2021 and the German country specific sub study protocol supplement dated 28 October 2021. The scope of this plan includes the final analysis and will be executed by the Biostatistics department unless otherwise specified.

2. Objectives, Endpoints and Hypotheses

2.1 Objectives and Endpoints

Objectives	Endpoints
Primary	
To evaluate the effect of denosumab in lumbar spine bone mineral density (BMD) Z-score at 12 months, as assessed by dual-energy X-ray absorptiometry (DXA), in children 2 to 17 years of age (at the time of screening) on a 3-month dosing regimen with Osteogenesis Imperfecta	Change from baseline in lumbar spine BMD Z-score, as assessed by DXA, at 12 months in subjects receiving the 3-month dosing regimen
Secondary To evaluate denosumab in children 2 to 17 years of age (at the time of screening) with OI on a 3-Month Dosing Regimen with respect to:	
Change in lumbar spine BMD Z-score, as assessed by DXA, at 6 months	Change from baseline in lumbar spine BMD Z-score, as assessed by DXA, at 6 months in subjects receiving the 3-month dosing regimen
Change in proximal femur BMD Z-score, as assessed by DXA, at 6 and 12 months (in subjects 5 years of age and older)	Change from baseline in proximal femur BMD Z-score, as assessed by DXA, at 6 and 12 months (in subjects 5 years of age and older) in subjects receiving the 3-month dosing regimen
Incidence of X-ray confirmed long bone and new and worsening vertebral	Incidence of X-ray confirmed long bone and new and worsening vertebral

fractures from last 12 months on 6-month dosing regimen to 12 months on 3-month dosing regimen	fractures from last 12 months on 6-month dosing regimen to 12 months on 3-month dosing regimen
Incidence of X-ray confirmed new and worsening vertebral fractures from last 12 months on 6-month dosing regimen to 12 months on 3-month dosing regimen	Incidence of X-ray confirmed new and worsening vertebral fractures from last 12 months on 6-month dosing regimen to 12 months on 3-month dosing regimen
Incidence of X-ray confirmed new vertebral fractures from last 12 months on 6-month dosing regimen to 12 months on 3-month dosing regimen	Incidence of X-ray confirmed new vertebral fractures from last 12 months on 6-month dosing regimen to 12 months on 3-month dosing regimen
Incidence of X-ray confirmed improving vertebral fractures from baseline of 3-Month Dosing Regimen to 12 months on 3-Month Dosing Regimen	Incidence of X-ray confirmed improving vertebral fractures from baseline of 3-Month Dosing Regimen to 12 months on 3-Month Dosing Regimen
Incidence of vertebral and non-vertebral fractures from last 12 months on 6-month dosing regimen to 12 months on 3-month dosing regimen (in subjects 5 years of age or older)	Incidence of vertebral and non-vertebral fractures from last 12 months on 6-month dosing regimen to 12 months on 3-month dosing regimen (in subjects 5 years of age or older)
Change in Child Health Questionnaire-Parent Form-50 (CHQ-PF-50) Physical Summary score at 12 months	Change from baseline in CHQ-PF-50 Physical Summary score at 12 months in subjects receiving the 3-month dosing regimen
Change in CHQ-PF-50 Psychological Summary score at 12 months	Change from baseline in CHQ-PF-50 Psychological Summary score at 12 months in subjects receiving the 3-month dosing regimen
Change in Childhood Health Assessment Questionnaire (CHAQ) Disability Index score at 12 months	Change from baseline in CHAQ Disability Index score at 12 months in subjects receiving the 3-month dosing regimen

Change in Wong-Baker Faces Pain Rating Scale (WBFPRS) at 12 months	Change from baseline in WBFPRS at 12 months in subjects receiving the 3-month dosing regimen
Change in growth velocity (determined by calculating age-adjusted Z-scores for height, weight, and body mass index [BMI]) at 12 months	Change from baseline in growth velocity (determined by calculating age-adjusted Z-scores for height, weight and BMI) at 12 months in subjects receiving the 3-month dosing regimen
Serum denosumab concentration and bone turnover markers (BTM) collected at 1(pre-dose),10, 30, and 60 days after the first dose of 3-month dosing regimen and every 3 months(pre-dose) in all subjects	Serum concentration of denosumab and serum BTM on days 1(pre-dose), 10, 30, 60 after the first dose and every 3 months(pre-dose), in subjects on 3-month dosing regimen

Exploratory	
[redacted]	

Safety To evaluate denosumab in children 2 to 17 years of age (at the time of screening) with OI on a 3-Month Dosing Regimen or 6-Month Dosing Regimen with respect to:	
Adverse events and serious adverse events	Subject incidence of adverse events and serious adverse events
Laboratory parameters	Change from baseline in laboratory values
Vital signs	Change from baseline in vital signs
Antidenosumab antibodies	Subject incidence of antidenosumab antibodies
Metaphyseal index	Subject incidence of metaphyseal index Z-score above age-appropriate normal range
Molar eruption and mandibular shaping	Subject incidence of abnormal molar eruption and Subject incidence of abnormal mandibular shaping
Hypercalcemia	Subject incidence of hypercalcemia

2.2 Hypotheses and/or Estimations

The hypothesis of the study is that the change from baseline in lumbar spine BMD Z-score following 12 months of denosumab treatment with 3-month dosing regimen in children 2 to 17 years of age (at the time of screening) with OI will be greater than that from historical control (ie, placebo or untreated subjects) (Letocha et al, 2005; Rauch et al, 2009; Ward et al, 2011) which was estimated to be 0.01 (95% CI: -0.17, 0.18).

3. Study Overview

3.1 Study Design

This is a prospective, single-arm, multicenter, phase 3 study to demonstrate the superiority of denosumab administration compared with historical controls as measured by change from baseline in lumbar spine BMD Z-score at 12 months and to evaluate the safety, efficacy, and pharmacokinetics in children with OI. The effect of denosumab administered every 3 months will be evaluated based on comparison in BMD Z-score change from baseline of 3-month dosing regimen to 12 months with historical controls derived from published randomized, controlled trials (Letocha et al, 2005; Rauch et al, 2009; Ward et al, 2011).

6-month dosing regimen

Approximately 150 subjects between age 2 to 17 will be enrolled, including at least 30 subjects younger than 7 years old. All subjects will receive denosumab 1 mg/kg (up to a maximum of 60 mg) subcutaneously every 6 months for 36 months, and all subjects will receive appropriate calcium and vitamin D (study design and treatment schema are provided at end of protocol synopsis section). Subjects meeting eligibility criteria will be enrolled in the study to receive treatment with denosumab. Although enrollment in the study will not be stratified, it will be gated as follows:

Older children (11 to 17 years) will be enrolled first because they will be the most skeletally mature and are more verbal and cognizant to describe adverse events.

The first 5 subjects enrolled in the age 11 to 17 year cohort (sentinel cohort) will undergo intensive mineral homeostasis monitoring for the first 14 days after administration of investigational product, to permit adaptation of subsequent mineral homeostasis assessments for all study subjects, as appropriate. Once these 5 subjects have completed 14 days of mineral homeostasis monitoring, enrollment will be open to the remainder of the age 11 to 17 year cohort.

When the first 5 subjects in the age 11 to 17 cohort have been treated for 12 months, enrollment will be open to the next age cohort (7 to 10 years).

Enrollment will be open to children 2 to 6 years old after the first 5 subjects between 7 to 10 years old have been treated for 6 months.

3-month dosing regimen

Sixty previously enrolled subjects will transition to a 3-month dosing regimen, among them approximately 20 subjects younger than 7 years of age (at time of initial enrollment). All subjects will receive denosumab 1 mg/kg (up to a maximum of 60 mg) subcutaneously every 3 months for a minimum of 12 months, and all subjects will receive age appropriate calcium and vitamin D.

All subjects who remain on study (have not previously completed month 36/EOS under the 6-month dosing regimen) will be eligible to transition to the 3-month dosing regimen unless they ended treatment due to an adverse event. Subjects may be consented and transition to the 3-month dosing regimen up to and including the date they attend for a month 36 visit under the 6-month dosing regimen. Subjects who previously ended treatment but remain on study will be eligible to transition to the 3-month dosing regimen unless they ended treatment due to an adverse event. Subjects will have the option to either consent to the 3-month dosing regimen or to end study participation.

Based on the life-threatening events of hypercalcemia observed, Amgen evaluated that benefit/risk profile of denosumab is not favorable. Per DMC recommendation and Amgen decision, an urgent safety measure was issued to stop IP with the EOS visit 24 weeks after the last dose of IP.

3.2 Sample Size

3.2.1 6-month dosing regimen

A study sample size of 150 subjects will be enrolled to collect adequate efficacy and safety information for the ad hoc analysis at 12 months. Assuming that 20% of subjects will not be evaluable at month 12 for the ad hoc efficacy endpoint due to dropout, the analysis set for the ad hoc efficacy endpoint will include at least 120 subjects.

3.2.2 3-month dosing regimen

Sixty previously enrolled subjects were previously transitioned to a 3-Month Dosing Regimen. Approximately 60 previously enrolled subjects were to transition to a 3-Month Dosing Regimen.

This sample size will provide more than 95% power to demonstrate that the difference in the change from baseline lumbar spine BMD Z-score at month 12 between denosumab and historical control is significantly greater than zero using 1-sample 2-sided t-test with a significance level of 0.05 based on the following assumptions:

- Mean change in lumbar spine BMD Z-score at 12 months for denosumab is 1.31 (SD = 0.91), which represents the lower bound of the 95% CI for the effect of a bisphosphonate (zoledronic acid) (EMA assessment report for Zometa Procedure no EMEA/H/C/000336/II/0031).
- Weighted mean (SE) change in lumbar spine BMD Z-score at 12 months for historical control is 0.01(0.09) (95% CI: -0.17, 0.18), estimated using meta-analysis based on data from the placebo or control of 3 studies in children with OI (Ward et al, 2011; Rauch et al, 2009; Letocha et al, 2005).

Assuming the sample size will provide at least 10% of the subjects evaluable for the primary efficacy endpoint in most subgroups of interest, the effect of denosumab on lumbar spine BMD Z-score also may be summarized descriptively in these subgroups.

4. Covariates and Subgroups

4.1 Planned Covariates

The following covariates are assumed to have prognostic value with respect to the lumbar spine BMD Z-score as assessed by DXA.

- Age (years) at the time of transition to 3-month dosing regimen
- Baseline lumbar spine BMD Z-score of 3-month dosing regimen

The covariate's association with the primary endpoint will be investigated.

4.2 Subgroups

The following subgroups will be used to analyze the change from baseline of 3-month dosing regimen in lumbar spine BMD Z-score as assessed by DXA at 12 month. Selected subgroups may be used to analyze other endpoints.

- Age cohorts at the time of screening (2 to 6, 7 to 10**,** 11 to 17 years)
- Gender (females or males)
- Baseline OI severity (Level 1, Level 2 or Level 3)
    - **Severity level 1**: blue sclera, normal height (3-97percentile) at screening, and does not meet the required criteria for greater severity (see below).

- **Severity level 2**: blue or grey sclera, >2 fractures/year averaged over lifetime, height <50 percentile for age at screening and does not meet the required criteria for greater severity (see below).
- **Severity level 3**: fractures in utero or at birth, saber shins and/or significant bowing, typical for type III per investigator, dentinogenesis imperfecta, >2 fractures/year averaged over lifetime, and height <25 percentile for age at screening.

If there are < 10 subjects within a subgroup, this subgroups analysis will not be provided.

5. Definitions

5.1 Basic Definitions

Investigational Product

Denosumab (1 mg/kg, up to a maximum of 60 mg).

Interactive Voice Response System (IVRS)

The system used to screen, screen failure and enroll subjects, provide investigational product (IP) box number, and volume administered, as well as to manage the cohorts and denosumab drug supply at the site and to track subject study termination data.

Exposure-adjusted subject incidence rate (EAIR)

Exposure-adjusted subject incidence rate (EAIR) for fractures during last 12 months on 6- month dosing regimen will be calculated as the number of subjects with a fracture divided by the total time at risk for the fracture. Time at risk will be calculated as sum of the total time to first fracture or total exposure of last 12 months on Q6M if no event across all subject.

Exposure-adjusted subject incidence rate (EAIR) for fractures during12 months on 3-month dosing regimen will be calculated as the number of subjects with a fracture divided by the total time at risk for the fracture. Time at risk will be calculated as the sum of total time to first fracture or total exposure of 12 month or the end of study whichever earlier if no event across all subjects.

Exposure-adjusted subject incidence rate (EAIR) will be displayed per 100 subjects-years.

Exposure-adjusted event rate (EAER)

Exposure-adjusted event rate (EAER) for fractures during last 12 months on 6- month dosing regimen will be calculated as the number of fractures divided by the total time at risk for the fracture. Time at risk will be calculated as sum of subjects years of exposure on last 12 months on Q6M.

Exposure-adjusted event rate (EAER) for fractures during 12 months on 3-month dosing regimen will be calculated as the number of fractures divided by the total time at risk for the fracture. Time at risk will be calculated as sum of total time from first dose of Q3M to month 12 or to the end of study whichever earlier across all subjects.

Exposure-adjusted Event rate (EAER) will be displayed per 100 subjects-years.

Treatment-emergent adverse event (TEAE) for 6-month dosing regimen

- **Subject who did not transition to 3-months dosing regimen and discontinued the study early; Treatment-emergent adverse event are events categorized as Adverse Events (AEs) starting on or after first dose of investigational product as determined by "Did event start before first dose of investigational product" equal to "No" or missing on the Events eCRF up to end of study.**
- **Subject who transition to 3-months dosing regimen Treatment-emergent adverse event are events categorized as Adverse Events (AEs) starting on or after first dose of investigational product as determined by "Did event start before first dose of investigational product" equal to "No" or missing on the Events eCRF up to the day before first dose of 3-month dosing regimen .**

Treatment-emergent adverse event (TEAE) for 3-month dosing regimen

Treatment-emergent adverse event are events categorized as Adverse Events (AEs) starting on or after first dose of investigational product on 3-month dosing regimen and up to the end of study.

5.2 Study Points of Reference

Baseline for 3-month dosing regimen

Baseline for 3-month dosing regimen will be Day 1 of 3-month dosing regimen.

Baseline for 6-month dosing regimen

Baseline for 6-month dosing regimen will be the closest recorded measurement before the administration of the first dose of 6-month dosing regimen.

Baseline for Vertebral Fractures

Baseline for vertebral fractures will be the x-rays obtained on day 1 of 3-month dosing regimen.

Baseline DXA Assessments

Baseline DXA will be performed at day 1 visit (in duplicate) of the 3-month dosing regimen. Because the baseline DXA will be assessed in duplicate, the baseline BMD/BMC value for analysis purpose will be the average of the BMD/BMC values from all baseline DXA scans for the same skeletal site.

If the baseline (before first dose of 3-month dosing regimen) DXA assessment is missing, only then the DXA scans obtained after the first 3-month dosing regimen but no more than 60 days from the first 3-month dosing regimen may be considered as baseline DXAs.

Baseline for fractures occurred during last 12 months on 6-month dosing regimen.

Day x to day 1 of Q3M dosing regimen where “x” is the day of the assessment closest to the previous 12 months in 6-month dosing.

Study Day 1 for 3-month dosing regimen.

3-month dosing regimen day 1 is defined as the day that the initial dose of IP for 3-month dosing regimen is administered to the subject.

Study Day 1 for 6-month dosing regimen.

6-month dosing regimen day 1 is defined as the day that the initial dose of IP for 6-month dosing regimen is administered to the subject.

Study Day

The number of days from Study Day 1 of respective dosing regimen, inclusive:

$$\text{Study Day} = (\text{Date of Interest} - \text{Date of Study Day 1}) + 1$$

End of Treatment (EOT)

End of Treatment (EOT) defined as “the last assessment for the protocol-specified treatment phase of the study for an individual subject.”

EOS for subject who did not transition to 3-month dosing regimen

The date recorded on the End of Study eCRF and no record of 3-month dosing regimen in Investigational Product Administration eCRF with volume > 0.

EOS for Subject who did transition to 3-month dosing regimen:

The date recorded on the End of Study eCRF and has record of 3-month dosing regimen in Investigational Product Administration eCRF with volume > 0

As of 30 September 2021, due to the life-threatening risk of hypercalcemia, subjects on study who had not completed their EOS visit were immediately discontinued from IP, and will be followed for safety for 24 weeks following the last dose of IP per the updated Schedule of Assessments. For subjects who have been on alternative therapy for more than 6 months or whose last dose of IP was greater than 24 weeks, no additional 24-week safety follow-up is required. These subjects can follow the previous protocol (protocol amendment 6) and complete the EOS CRF. Subjects whose last dose was between 12 and 24 weeks ago can complete the EOS assessments at week 24 per the updated Schedule of Assessments.

Visit Windows

For analysis purpose the analysis visit windows defined in Appendix C will be used to assign evaluations to the most appropriate nominal visit.

Ad-hoc analysis

Analysis based on all subjects who completed 12 months on 6-month dosing regimen based on as-is snapshot data referred as ad hoc analysis.

5.3 Study Dates

Informed Consent (IC) Dates for 6-month dosing regimen

The date on which the study informed consent forms, and subject assent form if applicable, is signed.

Screening Date

The screening date is the informed consent (IC) for 6-month dosing regimen / subject assent date.

Enrollment Date

The date on which the enrolment call is made using IVRS.

First Dose Date For 3-month dosing regimen

The date of administration of first dose of IP for 3-month dosing regimen will be day 1 of 3-month dosing regimen (ie, the first date recorded on the Investigational Product Administration eCRF with volume > 0).

Last Dose Date

The date of administration of last dose of IP (ie, the last date recorded on the Investigational Product Administration eCRF with volume > 0).

End of Study Date

The date recorded on the End of Study eCRF.

5.4 Study Time Intervals

Screening Period

The time between the date of informed consent and the enrolment.

Study Period

The time period between the first dose date and the end of study date, inclusive.

Last 12 months on 6- month dosing regimen for fracture endpoints is the 12 months prior to first dose date of 3-month dosing regimen.

12 months on 3-month dosing regimen for fracture endpoints starts at first dose date of 3 months dosing regimen and ends at months 12 or end of study date for 3-month dosing regimen, whichever comes first.

5.5 Subject Disposition

Enrolled

Individuals are considered enrolled when an enrollment call is made using IVRS. Enrolled individuals are referred to as ‘subjects’.

Exposed to IP

Subjects are defined as exposed to IP if the total amount of IP volume received during the study is greater than zero.

5.6 Arithmetic Calculations

Age at Screening (Years)

Age as collected on screening period based on screen date, and to allocate subjects to the appropriate cohort (11 to 17, 7 to 10, or 2 to 6 years) at enrollment.

Age at Screening (months)

Age in months is derived as collected age on screening period in years multiplied by number of months in year.

Age at Visit

Age at visit in years is derived as –

((visit date - Date of Birth)+1) / 365.25

If Date of birth is incomplete then use below formula,

Collected age at screening period in years + (visit date - screening date) / 365.25

Age at visit in months is derived as –

((visit date - Date of Birth)+1) / 30.4375

If Date of birth is incomplete then use below formula,

Collected age at screening period in months + (visit date - screening date) / 30.4375

Age at the time of transition to 3-month dosing regimen in years is derived as-

((Day 1 visit date of 3-month dosing regimen - Date of Birth)+1) / 365.25

If Date of birth is incomplete then use below formula,

Collected age at screening period in years +(Day 1 visit date of 3-month dosing regimen - screening date) / 365.25

Age at the time of transition to 3-month dosing regimen in months is derived as-

((Day 1 visit date of 3-month dosing regimen - Date of Birth)+1) / 30.4375

If Date of birth is incomplete then use below formula,

Collected age at screening period in months + (Day 1 visit date of 3-month dosing regimen - screening date) / 30.4375

5.7 Study Endpoints

5.7.1 Fracture-related Trauma Severity Definitions

Low Trauma Severity

Assessed by the investigator and collected on the Clinical Fracture Summary eCRF for each clinical fracture event and includes:

- Fall from standing height or less than 20 inches

- Minimal or Moderate trauma other than a fall from standing height or less than 20 inches
- Unknown/don't Know

High Trauma Severity

Assessed by the investigator and collected on the Clinical Fracture Summary

eCRF for each clinical fracture event and includes

- Severe trauma other than a fall

Pathologic Fracture

Assessed by the investigator and collected on the Clinical Fracture Summary eCRF for each clinical fracture event and includes:

- Pathologic fractures

5.7.2 DXA Assessments

Only GE Lunar or Hologic bone densitometers will be allowed for this study. The same DXA machine must be used for all DXA assessments for a particular subject. The left side should be used for proximal femur, unless prohibited (eg, hip implant). If another side must be used or is inadvertently used during screening, then the same side must be used consistently throughout the study.

Lumbar spine scans should include L1 through L4. Individual vertebral levels may be excluded due to artifact. A vertebral level excluded from one visit will be excluded from all visits with the total lumbar spine BMD calculated based on the evaluable vertebral levels. Baseline DXAs will be performed as duplicate scans; DXA assessments at all other visits will be performed as single scans. After analysis of the scans by the central imaging vendor, the study site may be asked to re-acquire a scan, because of poor positioning or other technical reasons.

5.7.3 Vertebral Fractures

Lateral radiographs of the thoracic and lumbar spine will be obtained at times to determine clinical or morphometric vertebral fractures.

The lateral spine radiographs will be assessed for prevalent, new, worsening and improving vertebral fractures using the Genant Semiquantitative (SQ) scoring method:

- Grade 0: normal (approximately < 20% reduction in anterior, middle, and/or posterior height)

- Grade 1: mild fracture (approximately 20% - 25% reduction in anterior, middle, and/or posterior height)
- Grade 2: moderate fracture (approximately 25% - 40% reduction in anterior, middle, and/or posterior height)
- Grade 3: severe fracture (> 40% reduction in anterior, middle, and/or posterior height)

All vertebral fractures data will be verified by Bioclinica for analysis purpose.

Prevalent Vertebral Fracture

A subject has prevalent vertebral fracture if any vertebra from T4 to L4 has a grade of ≥ 1 at baseline. A subject does not have prevalent vertebral fracture when all 13 grades from T4 to L4 are 0 on the first evaluable spinal radiograph collected during the study. Otherwise, the subject will have an unknown status for prevalent vertebral fracture.

Any vertebra(e) with a missing grade at baseline can be assumed having a grade of 0 if the subsequent x-ray shows a grade of 0 for the same vertebra(e).

New Vertebral Fractures

A new vertebral fracture is identified when there is ≥ 1 grade increase from the baseline grade of 0 in any vertebra from T4 to L4, excluding any symptomatic new vertebral fracture associated with high trauma severity or a pathologic fracture Because the vertebral fractures can improve in the OI population, the SQ score of the fractured vertebra should remain ≥ 1 at the last assessment on or before the timepoint of interest. So for example, if a vertebra is assessed as having 0, 2 SQ scores at baseline and 12 months, respectively, then the vertebra will be identified as having a new vertebral fracture at the 12-month assessment because the SQ score remained > 0 at 12 months.

A subject will be identified as having multiple new vertebral fractures when more than one thoracic or lumbar vertebra is identified as having a new vertebra fracture at the timepoint of interest, and the total number of new vertebral fractures will be the sum of new vertebra fractures at the timepoint of interest.

Worsening Vertebral Fractures

A worsening vertebral fracture is identified when there is ≥ 1 grade increase from the baseline grade of ≥ 1 in any vertebra from T4 to L4, excluding any symptomatic worsening vertebral fracture associated with high trauma severity or a pathologic

fracture. For example, if a vertebra is assessed as having SQ scores of 2, 3 at baseline and 12 months respectively the vertebra will be assessed as having a worsening vertebral fracture at the 12-month assessments.

A subject will be identified as having multiple worsening vertebral fractures when more than one thoracic or lumbar vertebra is identified as having a worsening vertebra fracture at the timepoint of interest, and the total number of new vertebral fractures will be the sum of worsening vertebral fractures at the timepoint of interest.

New or Worsening Incident Vertebral Fracture

A new and worsening vertebral fracture is identified when there is ≥ 1 grade increase from the baseline grade in any vertebra from T4 to L4, excluding any symptomatic vertebral fracture associated with high trauma severity or a pathologic fracture.

A subject will be identified as having multiple new or worsening vertebral fractures when more than one thoracic or lumbar vertebra is identified as having a new or worsening vertebra fracture at the timepoint of interest, and the total number of new or worsening vertebral fractures will be the sum of new or worsening vertebra fractures at the timepoint of interest.

Improving Vertebral Fractures

At 12 months (an improvement in a prevalent vertebral fracture is identified when the SQ grade decreases by at least 1 relative to baseline SQ score, ie the last SQ score assessment in the period of interest is smaller than its baseline SQ score.

At 12 months a subject is identified as having an improving vertebral fracture if at least one thoracic or lumbar vertebra is assessed as an improving vertebral fracture at the visit of interest, and the vertebra SQ score remains < baseline SQ at the last assessment on or before the timepoint of interest. For example, if a vertebra is assessed as having 2, 1 SQ scores at baseline and 12 months respectively, then the vertebra will be identified as having an improving vertebral fracture at the 12 month assessments.

Clinical Vertebral Fractures

Clinical vertebral fracture is any new or worsening vertebral fracture assessed at either a scheduled or unscheduled study visit, and associated with any signs and/or symptoms indicative of a fracture on the Clinical Fracture Summary eCRF, excluding

any fracture associated with high trauma severity or a pathologic fracture. Clinical vertebral fracture is also called symptomatic vertebral fracture.

5.7.4 Nonvertebral Fracture

Nonvertebral fracture **(other than long bone fractures),** reported on the Clinical Fracture Summary eCRF, is defined for post-treatment period as any nonvertebral (excluding cervical vertebrae, thoracic vertebrae and lumbar vertebrae) fracture confirmed by the central imaging vendor based on a copy of radiographs or other diagnostic images such as computerized tomography (CT) or magnetic resonance imaging (MRI), and/or documented in a copy of the radiology report, surgical report, or discharge summary, excluding fractures associated with high trauma severity or pathologic fractures.

Long bone fractures are subset of nonvertebral fractures including the femur, humerus, radius, ulna, tibia and fibula bone.

Long Bone fractures reported on the Clinical Fracture Summary eCRF, is defined for post-treatment period as any long bone fracture confirmed by imaging excluding fractures associated with high trauma severity or pathologic fractures.

See Appendix D for the list of nonvertebral fracture locations.

Historical Fracture

All historical fractures are defined based on the fracture locations listed on the Subject Fracture History eCRF. Nonvertebral fractures should be substantiated by radiographs where available or with a medical report in the event of unavailable radiographs.

Subject fracture history eCRF page will be used for the classification of nonvertebral fractures for last 12 months on 6- month dosing regimen excluding spine fracture, ankle fracture, facial bone fracture, skull fractures and for the classification of long bone fractures exclude forearm fracture, lower leg (Not knee or ankle) fracture, upper arm fracture and upper leg (Not hip) fracture.

Any Historical Fracture

Any fracture recorded on the Subject Fracture History eCRF regardless of trauma severity.

Historical Nonvertebral Fracture

Any nonvertebral fracture recorded on the Subject Fracture History eCRF and not associated with known high trauma severity or pathologic fractures.

5.7.5 Growth Velocity Endpoints

Height-for-age Z-score is defined as the difference between the subject's height and the median height for the population with the same age and gender, divided by the population standard deviation:

$$Height - for - age\ \ Z - score = \frac{(measured\ value\ -\ median\ of\ reference\ population)}{standard\ deviation\ of\ the\ reference\ population}$$

Growth velocity is given by the change from baseline in height-for-age Z-scores at 12 months. During normal growth, the change in growth velocity score should equal 0. Growth acceleration is indicated by a positive change, and growth deceleration is indicated by a negative change.

The definitions of growth velocity based on weight and BMI are analogously calculated; weight-for-age and BMI-for-age Z-scores will be calculated for each subject.

To programmatically calculate the Z-scores, the National Center for Health Statistics percentiles growth charts, based on the 2000 Center for Disease Control and Prevention (CDC) (http://www.cdc.gov/growthcharts/cdc_charts.htm), and the CDC Anthropometric Software Package 3.0 Z scores will be used. See details in Appendix E.

5.7.6 PRO Questionnaires

CHQ-PF-50

The CHQ-PF50 is a 50-item questionnaire to be completed by the parents or guardian of children between 5 and 18 years of age. The 50 questions measure 14 domains which are summarized as the physical and psychological summary scores. Each domain is scored from 0 to 100, with higher score indicating better physical and psychosocial health.

CHAQ Disability Index score

The disability domain (questions 1-54) of the CHAQ is used to measure the subject's assessment of physical functioning or the parent's assessment of the child's physical functioning. The disability index comprises 8 categories (dressing and grooming, arising, eating, walking, hygiene, reaching, gripping, and activities). Scoring ranges from 1 to 5; 1 is "without any difficulty," 2 is "with some difficulty," 3 is "with much

difficulty," and 4 is "unable to do." An answer of "not applicable" is scored as a 5, but is not counted. If a child requires assistance from another person or uses an aid or other device for any of the 8 categories, the minimum score for that category will be recorded as a 3. The CHAQ questions will then be converted to a 0 to 3 scale score.

Missing values will not be imputed. If any of the domain questions are missing, but the aids/device indicator is non-missing, the domain can still be computed. However, if all domain questions are missing (or are scored as "not applicable") and the aids/device indicators are missing, then the domain score is considered missing. The CHAQ score is not computed when the subject/parent provides answers in fewer than 6 domains (normally coded as "UNK" in the database).

For more information on the contribution of the 54 questions to calculate the CHAQ, see Appendix B.

WBFPRS

The WBFPRS is validated for use in children 3 years of age or older. At each timepoint, subjects are asked to choose the face that best describes their own pain. See details in Appendix B.

5.7.7 Metaphyseal Index Z-score of the Distal Femur

To assess a subject's eligibility at screening, AP radiographs of both knees (unless prohibited by the presence of hardware such as implants) is to be used to calculate the metaphyseal index Z-score of each knee; the knee for assessment during the study at 6 and 12 months should be the one with the higher Z-score at day 1 of 6-month dosing regimen, unless prohibited by presence of hardware, in which case an AP radiograph of the contralateral knee may be obtained. Subject incidences of metaphyseal index Z-score above age-appropriate normal range (ie, metaphyseal index Z-score > +2) will be provided at each timepoint. The metaphyseal index (MI) will be calculated by the central imaging vendor as the ratio of the femoral width (W) over the distal femoral growth plate width (GPW), and the Z-score for each subject, relative to the subject's age as:

$$\text{MI Z-score} = (\text{subject value} - \text{mean})/\text{SD},$$

where mean and standard deviation (SD) are the corresponding values based on Ward *et al* database, (Ward, 2005), for the subject's age group at the time of the assessment.

5.7.8 Mandibular Shaping

Lateral cephalogram will be performed to enable assessment of mandibular shaping at Day 1 of 6-month dosing regimen, Day 1 of 3-month dosing regimen and EOS. The lateral cephalogram is a profile X-ray of the skull and soft tissues and is used to assess the relation of the teeth in the jaws, the relation of the jaws to the skull, and the relation of the soft tissues to the teeth and jaws.

The following anatomical angles will be measured to evaluate the correct proportions of the mandible and its position relative to the skull/maxilla:

- Gonial angle
- SNA angle
- SNB angle
- ANB Angle = SNA - SNB.

5.7.9 Molar Eruption

Radiographic assessment of molars will be performed at all visits for the 3-month dosing regimen and 6-month dosing regimen to assess the risk for unerupted molars. In addition, each subject will undergo a visual inspection under natural light for the presence of the first and second molars. The subject should be referred to a dentist to perform radiographic assessment of the unerupted molar(s) if:

- A subject is age 7 to 12 years and appears to have an unerupted upper or lower first molar (ie, not all 4 first molars are visible/detectable).
- A subject is age 13 years or older and appears to have an unerupted upper or lower (first or) second molar (ie, not all 4 first molars and all 4 second molars are visible/detectable).

6. Analysis Sets

6.1 Full Analysis Set

The full analysis set (FAS) is defined according to intent-to-treat analysis to include all subjects enrolled into the study.

6.2 Safety Analysis Set

6.2.1.1 6-Month Dosing Regimen Safety Analysis Set

The 6-month dosing regimen safety analysis set includes all subjects in the FAS who received ≥ 1 dose of IP.

6.2.1.2 3-Month Dosing Regimen Safety Analysis Set

The 3-month dosing regimen safety analysis set includes all subjects in the FAS who received ≥ 1 dose of 3-month dosing regimen**.**

6.3 PK Analysis set

The PK analysis set includes all subjects in the 3-month dosing regimen safety analysis set who have ≥ 1 serum denosumab reported result on 3-month dosing regimen.

6.4 Study-specific Analysis Sets

6.4.1 DXA Analysis Set

The DXA analysis set includes all subjects in the FAS with baseline and ≥ 1 postbaseline DXA assessment on 3-month dosing regimen for the endpoint of interest (lumbar spine and proximal femur) as provided by the central imaging vendor**.**

6.4.2 Vertebral Fracture Analysis set

The vertebral fracture analysis set includes all subjects in the FAS who have a readable non-missing baseline and ≥ 1 non-missing postbaseline X-ray vertebral evaluation on 3-month dosing regimen as provided by the central imaging vendor. This analysis set will be used to analyze incidence of vertebral fracture endpoints.

6.4.3 PRO Analysis set

For each PRO (CHQ-PF-50, CHAQ disability index score and WBFPRS), the PRO analysis set includes all subjects in the FAS with baseline and ≥ 1 postbaseline valid PRO response on 3-month dosing regimen for the appropriate PRO. Moreover, the CHQ-PF-50 analysis set will only include subjects 5 years of age and older at the time of transition to 3-month dosing regimen; the questionnaire has not been validated in younger children (4 years of age or younger), and so this subgroup will not be included in the CHQ-PF-50 analysis set.

6.4.4 Growth Velocity Analysis set

The analysis set includes all subjects in the FAS who have evaluable data (age in total months, and weight, height, and BMI) at baseline and ≥ 1 post baseline assessment on 3-month dosing regimen for each growth velocity endpoint (weight-for-age, height-for-age, and BMI-for-age Z-scores).

6.4.5 Metaphyseal Analysis set

This analysis set includes all subjects in the 3-month dosing safety analysis set with open growth plates (and no hardware preventing accurate calculation of metaphyseal index) at baseline and X-ray of the knee at baseline and post-baseline.

6.4.6 BTM Analysis set

The BTM analysis set includes all subjects in the 3-month dosing regimen safety analysis set who have baseline and ≥ 1 postbaseline assessments for the endpoint of interest on 3-month dosing regimen.

6.4.7 DXA Per Protocol Analysis Set

The DXA per protocol analysis set includes all subjects in the FAS with baseline and month 12 DXA assessment on 3-month dosing regimen for lumbar spine as provided by the central imaging vendor.

6.4.8 German Substudy Safety Analysis Set

[redacted]

6.4.9 PK Substudy Analysis Set

The subset includes all subjects in the 6-Month Dosing Regimen safety analysis set who enroll in the PK/BTM substudy and have ≥ 1 serum denosumab reported result.

6.4.10 BTM Substudy Analysis Set

The subset includes all subjects in the safety analysis set who enroll in the PK/BTM substudy and have baseline and ≥ 1 postbaseline assessments for the endpoint of interest on 6-Month Dosing Regimen**.**

7. Planned Analyses

Descriptive statistics will be provided for demographics and subject characteristics, efficacy, PROs, and safety data. Descriptive statistics of continuous measurements will include means, medians, standard deviations, and ranges, while categorical data will be summarized using counts and percentages.

7.1 Interim Analysis and Early Stopping Guidelines

Not applicable for the study.

7.2 Primary Analysis

The primary analysis (or final analysis since there is only 1 milestone analysis for this study) will occur after the last subject completes the 24-week safety follow up visit following the last dose of IP.

7.3 Final Analysis

The final analysis is primary analysis for this study.

8. Data Screening and Acceptance

8.1 General Principles

The objective of the data screening is to assess the quantity, quality, and statistical characteristics of the data relative to the requirements of the planned analyses.

8.2 Data Handling and Electronic Transfer of Data

The Amgen Global Study Operations-Data Management (GSO-DM) department will provide all data to be used in the planned analyses except for the imaging and laboratory data. This study will use the RAVE database. The central imaging vendor will provide DXA scans, and X-ray (lateral thoracic, knees, lateral cephalogram, and panoramic radiograms or radiographic assessment) data to Amgen cumulatively.

The central laboratory vendor will provide all lab-related data, except for the samples analyzed at local laboratories. All laboratory values will be transferred cumulatively from the central laboratory.

Data screening will be performed periodically during the conduct of the study. The objective of the data screening is to assess the quantity, quality, and statistical characteristics of the data relative to the requirements of the planned analyses. Any questionable values or situations will be reported to Clinical Data Management (CDM) for review or confirmation.

As part of the data acceptance procedure, all datasets, planned tables, listings, and graphs will be generated and reviewed to identify any additional data issues. Any critical issues identified must be resolved with CDM before final acceptance of the data.

Cumulative data from the central vendors and Amgen will be transferred via a secure electronic transfer directly to the Independent Biostatistical Group supporting the DMC on a monthly basis initially. The frequency may change later on the study.

8.3 Handling of Missing and Incomplete Data

Subjects may have missing specific data points for a variety of causes. The general procedures outlined below describe what will be done when data point is missing or incomplete.

8.3.1 Missing Data

Missing baseline or post-baseline endpoints will not be imputed, unless otherwise specified.

8.3.2 Incomplete Dates

NA

8.3.3 Other Incomplete Dates

The following data will be imputed using the following algorithm:

- Adverse Events
- Concomitant Medications

Imputation Rules for Partial or Missing Start Dates

Start Date		Stop Date						
		Complete: yyyymmdd		Partial: yyyymm		Partial: yyyy		Missing
		< 1st dose	≥ 1st dose	< 1st dose yyyymm	≥ 1st dose yyyymm	< 1st dose yyyy	≥ 1st dose yyyy	
Partial: yyyymm	= 1st dose yyyymm	2	1	n/a	1	n/a	1	1
	≠ 1st dose yyyymm		2	2	2	2	2	2
Partial: yyyy	= 1st dose yyyy	3	1	3	1	n/a	1	1
	≠ 1st dose yyyy		3		3	3	3	3
Missing		4	1	4	1	4	1	1

1=Impute the date of first dose; 2=Impute the first of the month; 3=Impute January 1 of the year; 4=Impute January 1 of the stop year
For subjects who were never treated (first dose date is missing), partial start dates will be set to the first day of the partial month or first day of year if month is also missing.
Note:

1. The above imputation rule will apply to the subject respective dosing regimen during which the event has started.

2. If the imputation leads to an overlap of the imputation rules with respect to Q3M or Q6M, imputation rules will be applied in the following order.

- If a subject's transition happened during a month from Q6M to Q3M and AE start date is partial, check the end date of AE if it is before the Q3M administration start date, then the AE start date imputation should be as per the Q6M administration start date.
- If the AE end date is missing, the AE start date imputation should be as per the Q3M administration start date.

Imputation Rules for Partial or Missing Stop Dates

Initial imputation

- If the month and year are present, impute the last day of that month.
- If only the year is present, impute December 31 of that year.
- If the stop date is entirely missing, assume the event or medication is ongoing.

If the imputed stop date is before the start date, set stop date to missing.
If the imputed stop date is after the death date, impute as death date.

8.4 Detection of Bias

Given the open-label nature of the study, changes to the study conduct and statistical analyses should be kept to a minimum to avoid undermining the study credibility. Any change to the protocol that is data driven should be avoided, except for safety related events and changes related to inclusion/exclusion criteria to speed up enrollment, if needed.

8.5 Outliers

Scatter plots will be examined to identify outliers in continuous variables. Frequencies of the categorical data will be examined to identify questionable values. Outliers resulting from data entry error will be corrected in the database. Outliers that are not due to data entry error will be included in the analysis. The validity of any questionable values or outliers will be confirmed. No valid measurement will be excluded from descriptive or inferential analyses. However, sensitivity analyses may be conducted to evaluate the influence of extreme values in the data. These analyses will be documented in the study report.

8.6 Distributional Characteristics

The assumption of normality underlying the primary analysis is considered quite robust with this sample size, nonetheless the normality of the denosumab sample will be tested.

8.7 Validation of Statistical Analyses

Programs will be developed and maintained and output will be verified in accordance with current risk-based quality control procedures.

Tables, figures and listings will be produced with validated standard macro programs where standard macros can produce the specified outputs.

The production environment for statistical analyses consists of Amgen-supported versions of statistical analysis software, for example the safety analysis set System version 9.2 or later.

9. Statistical Methods of Analysis

9.1 General Considerations

Descriptive statistics will be provided for demographics and subject characteristics, efficacy, patient reported outcomes and safety data. Descriptive statistics of continuous measurements will include mean, standard deviation, minimum, 25th percentile, median (50th percentile), 75th percentile, maximum and number of nonmissing observations. Nominal and ordinal categorical variables will be summarized using counts and percentages.

The primary endpoint will be compared with historical control (Ward et al, 2011; Rauch et al, 2009; Letocha et al, 2005) using a one-sample t-test, with a significance level of 0.05. This is the only pre-planned comparison in this single-arm study.

9.2 Subject Accountability

The disposition of all subjects who enrolled in this study and who transitioned to 3-month dosing regimen will be tabulated . Disposition, for successfully completing IP administration, and completing the study will be included. The disposition of subjects will also include the number of subjects who withdrew from the IP and their reasons for withdrawal and the number of subjects who withdrew from study and their reasons for withdrawal. The number of subjects enrolled within each region and country will be summarized. The number of transitioned subjects participating in each of the sub studies will be summarized.

9.3 Important Protocol Deviations

Important Protocol Deviations (IPDs) categories are defined by the study team before the first subject's initial visit and updated during the IPD reviews throughout the study prior to database lock. These definitions of IPD categories, subcategory codes, and descriptions will be used during the course of the study. Eligibility deviations are defined in the protocol. Summary table and listing of the deviations from eligibility criteria will be generated for the subjects who received 3-month dosing regimen.

9.4 Demographic and Baseline Characteristics

Subject demographic and baseline disease characteristics will be summarized based on FAS.

Below characteristics are collected at the time of initial screening :

- Race
- Ethnicity
- Gender
- Type and details of Osteogenesis Imperfecta

Below characteristics are based on baseline of 6-month dosing regimen

- Body composition (height [cm], weight [kg], and BMI [kg/m^2])
- Growth Velocity (height-for-age, weight-for-age, and BMI-for-age z-scores)
- Armspan
- Vital signs (pulse, respiration rate, temperature)
- Selected laboratory assessments
- [redacted]
- DXA assessments of the lumbar spine and proximal femur (these DXAs will be performed as duplicate scans at baseline only)
- Fracture history including any fractures, vertebral fractures, and nonvertebral fractures

- Prevalent vertebral fracture at pre-treatment (based on screening spinal radiograph)
- Number of prevalent vertebral fractures (0, 1, ≥ 2; based on screening spinal radiograph)

Subject demographic and baseline disease characteristics will be summarized based on 3-Month Dosing Regimen Safety Analysis Set .

- Age (in years) at the time of transition to 3-month dosing regimen
- Age (in months) at the time of transition to 3-month dosing regimen
- Age cohorts (2 to 6, 7 to 10, and 11 to 17) at the time of screening
- Body composition (height [cm], weight [kg], and BMI [kg/m^2])
- Growth Velocity (height-for-age, weight-for-age, and BMI-for-age z-scores)
- Armspan
- Vital signs (pulse, respiration rate, temperature, systolic and diastolic blood pressure)
- Selected laboratory assessments
- [redacted]
- DXA assessments of the lumbar spine and proximal femur (these DXAs will be performed as duplicate scans at baseline only)
- Fracture history including any fractures, vertebral fractures, and nonvertebral fractures from last 12 months on 6-month dosing regimen
- Number of Prevalent vertebral fracture (based on Day 1 of 3-month dosing regimen spinal radiograph)

9.5 Efficacy Analyses

Most analyses will be descriptive given the lack of a comparator in this study as well as lack of historical comparator data for many endpoints. In addition, the magnitude of any natural change in most efficacy endpoints is unknown. However, for the primary endpoint of change from baseline in lumbar spine BMD Z-score at 12 months in subjects on 3-month dosing regimen, comparison with a weighted estimate of historical

controls will be performed at the 5% significance level based on a 2-sided test (see Section 9.5.1.1). The weighted mean change in lumbar spine BMD Z-score at 12 months from historical controls is 0.01 (95% CI: -0.17, 0.18). This suggests that the natural progression over 12 months with respect to the primary efficacy endpoint is minimal, and any bias in estimating the treatment effect, if exist, will likely to be trivial.

All descriptive statistics will be provided by visit based on the observed values, with no imputation for missing values. Analysis visit windows are defined in Appendix C for each endpoint.

Categorical endpoints will be summarized using the number and percentage in each category, and continuous endpoints will be summarized using descriptive statistics including mean, standard deviation, minimum, 25th percentile, median (50th percentile), 75th percentile, maximum, and number of nonmissing observations (n).

Table A Primary Efficacy Endpoint Summary Table

Endpoint	Primary Summary and Analysis Method (Specify Analysis Set if FAS is Not Used)	Sensitivity Analysis
Change from baseline in lumbar spine BMD Z-score, as assessed by DXA, at 12 months in subjects receiving the 3-month dosing regimen	The change from baseline of 3-month dosing regimen in BMD Z-score at 12 months in subjects on 3-month dosing regimen will be analyzed based on the DXA analysis set using repeated measures analysis with visit (6 and 12 months), baseline age, and baseline BMD Z-score as fixed effects.	• The primary analyses will be repeated substituting baseline age with baseline height to the fixed effects in the model. • The DXA per protocol analysis set will be used in this analysis.

Table B Secondary Efficacy Endpoint Summary Table

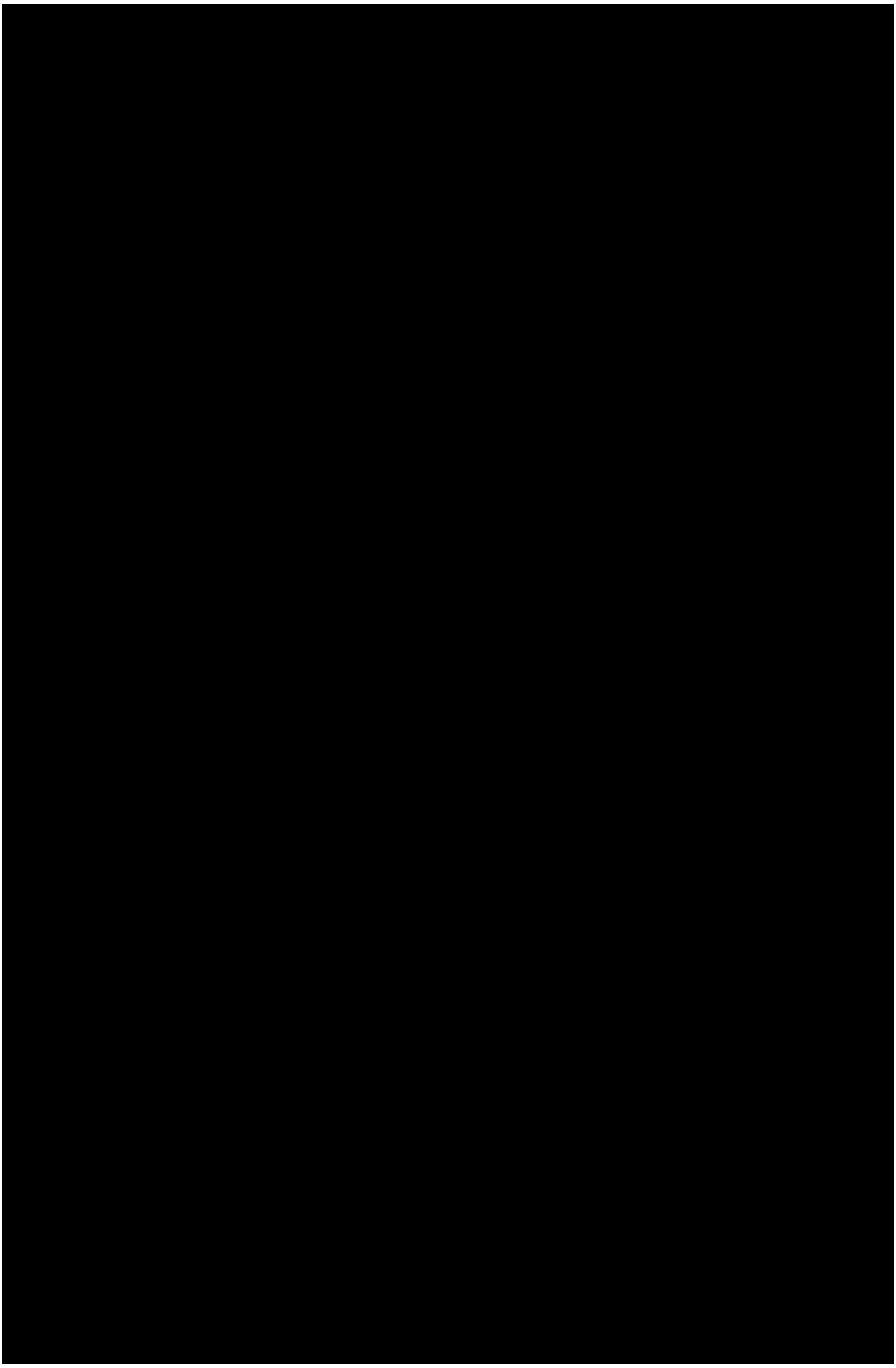
Endpoint	Primary Summary and Analysis Method (Specify Analysis Set if FAS is Not Used)	Sensitivity Analysis
Change from baseline in lumbar spine BMD Z-score, as assessed by DXA, at 6 months in subjects receiving the 3-month dosing regimen	The change from baseline of 3-month dosing regimen in lumbar spine BMD Z-score in subjects on 3-month dosing regimen will be analyzed based on the DXA analysis set using repeated measures analysis with visit (6 and 12 months), baseline age, and baseline BMD Z-score as fixed effects. Visit will be treated as a categorical variable.	
Change from baseline in proximal femur BMD Z-score, as assessed by DXA, at 6 and 12 months (in subjects 5 years of age and older) in subjects receiving the 3-month dosing regimen	The change from baseline of 3-month dosing regimen in proximal femur BMD Z-score in subjects on 3-month dosing regimen will be analyzed based on the DXA analysis set using repeated measures analysis with visit (6 and 12 months), baseline age, and baseline BMD Z-score as fixed effects. Visit will be treated as a categorical variable.	
Incidence of X-ray confirmed long bone and new and worsening vertebral fractures from last 12 months on 6- month dosing regimen to 12	Exposure adjusted subject incidence rate and exposure adjusted event rate of long bone and new and worsening vertebral fractures will be summarized based on the 3-month dosing regimen safety analysis set.	

Endpoint	Primary Summary and Analysis Method (Specify Analysis Set if FAS is Not Used)	Sensitivity Analysis
months on 3-month dosing regimen		
Incidence of X ray confirmed new and worsening vertebral fractures from last 12 months on 6-month dosing regimen to 12 months on 3-month dosing regimen	Exposure adjusted subject incidence rate and exposure adjusted event rate of new and worsening vertebral fractures will be summarized based on the vertebral fracture analysis set	
Incidence of X ray confirmed new vertebral fractures from last 12 months on 6-month dosing regimen to 12 months on 3-month dosing regimen	Exposure adjusted subject incidence rate and exposure adjusted event rate of new vertebral fractures will be summarized based on the vertebral fracture analysis set.	
Subject incidence of improving vertebral fractures from baseline of 3-Month Dosing Regimen to 12 months on 3-month dosing regimen	The proportion of subjects with improved vertebral fracture from baseline of 3-month dosing regimen will be summarized at each timepoint based on the vertebral fracture analysis set.	
Incidence of vertebral and nonvertebral fractures from last 12 months on 6-month	Exposure adjusted subject incidence rate and exposure adjusted event rate of vertebral and nonvertebral fractures will be	

Endpoint	Primary Summary and Analysis Method (Specify Analysis Set if FAS is Not Used)	Sensitivity Analysis
dosing regimen to 12 months on 3-month dosing regimen (in subjects 5 years of age or older)	summarized based on the 3-month dosing regimen safety analysis set.	
Change from baseline in CHQ-PF-50 Physical Summary score at 12 months in subjects receiving the 3-month dosing regimen	Patient-reported outcomes (CHQ-PF-50,) and their changes from 3-month dosing regimen will be summarized at month 12 based on their respective PRO analysis set.	
Change from baseline in CHQ-PF-50 Psychological Summary score at 12 months in subjects receiving the 3-month dosing regimen	Patient-reported outcomes (CHQ-PF-50) and their changes from 3-month dosing regimen will be summarized at month 12 based on their respective PRO analysis set.	
Change from baseline in CHAQ Disability Index score at 12 months in subjects receiving the 3-month dosing regimen	Change from baseline in CHAQ Disability Index score at 12 months in subjects receiving the 3-month dosing regimen will be summarized at month 12 based on their respective PRO analysis	
Change from baseline in WBFPRS at 12 months in subjects receiving the 3-month dosing regimen	Patient-reported outcomes (WBFPRS) and their changes from 3-month dosing regimen will be summarized at month 12	

Endpoint	Primary Summary and Analysis Method (Specify Analysis Set if FAS is Not Used)	Sensitivity Analysis
	based on their respective PRO analysis set.	
Change from baseline in growth velocity (determined by calculating age-adjusted Z scores for height, weight and BMI) at 12 months in subjects receiving the 3-month dosing regimen	Change from baseline of 3-month dosing regimen in age-adjusted growth velocity endpoints will be summarized at 12 months based on the growth velocity analysis set.	
Serum concentration of denosumab on days 1(pre-dose), 10, 30, 60 and every 3 months (pre-dose) in subjects on 3-month dosing regimen	Serum denosumab concentrations will be summarized by visit based on PK analysis set	
Serum BTM on days 1(pre-dose), 10, 30, 60, and every 3 months (pre-dose), in subjects on 3-month dosing regimen	Serum BTM will be summarized by visit based on BTM analysis set	

Table C Exploratory Efficacy Endpoint Summary Table



[redacted]

9.5.1 Analyses of Primary Efficacy Endpoint(s)

9.5.1.1 Primary Analysis

The change from baseline of 3-month dosing regimen in lumbar spine BMD Z-score at 12 months in subjects on 3-month dosing regimen will be analyzed based on the DXA analysis set using repeated measures analysis with visit (6 and 12 months), baseline age, and baseline BMD Z-score as fixed effects. Visit will be treated as a categorical variable. Missing baseline and post-baseline BMD Z-scores will not be imputed; only subjects with baseline and at least one lumbar spine DXA assessment at 6-month or 12-month visit will be included in the analysis. The estimated 12-month least squares (LS) mean will be presented with a 95% confidence interval (CI). See Appendix E (Code Fragment).

The LS mean will be compared to weighted estimated mean (SE) of 0.01 (0.09) from historical controls (Ward et al, 2011; Rauch et al, 2009; Letocha et al, 2005) using one sample t test with the historical mean used as the null hypothesis with a significance level of 0.05. See Appendix E for details.

All analyses will be based on the site-specific DXA analysis set.

9.5.1.2 Per Protocol Analysis

The above analyses will be repeated for the primary endpoint based on the DXA Per Protocol analysis set.

9.5.1.3 Subgroup Analysis

Summary statistics will be provided by the subgroups defined in Section 4.2.

9.5.1.4 Sensitivity Analysis

The primary analyses above will be repeated substituting baseline age with baseline height to the fixed effects in the model. The DXA Per Protocol analysis set will be used in this analysis.

9.5.2 Analyses of Secondary Efficacy Endpoint(s)

9.5.2.1 Other Changes From Baseline in BMD Z-score by DXA

Similarly as above, for each skeletal site (lumbar spine, total hip and femoral neck), the change from baseline of 3-month dosing regimen in BMD Z-score in subjects on 3-month dosing regimen will be analyzed using repeated measure analysis with visit (6, 12 months), baseline age, and baseline BMD Z-score as fixed effects. Visit will be treated as categorical variable.LS mean (95% CI) will be presented for total hip and femoral neck, and for each visit excluding the 12-month visit for lumbar spine.

All analyses will be based on the skeletal site specific DXA analysis set. Missing baseline and post-baseline BMD Z-scores will not be imputed; only subjects with baseline and at least one skeleton-specific DXA assessment visit will be included in the analysis.

9.5.2.2 New and Worsening Vertebral Fractures, and Vertebral and Nonvertebral Fractures

Fracture analysis will use the data as mentioned below :

1. **Vertebral Fractures** : central imaging vendor **i.e fractures confirmed by vendor**
2. **Nonvertebral Fractures other than long bone fracture** : **fractures reported on the Clinical Fracture Summary eCRF which is confirmed by the central imaging vendor**
3. **Long bone Fractures** : **fractures reported on the Clinical Fracture Summary eCRF i.e fractures confirmed by imaging**

Summary of the incidence of fractures for each endpoint (new and worsening vertebral fractures, vertebral and nonvertebral fractures , long bone and vertebral fracture) will be provided for 12 month.Exposure-adjusted incidence rate (EAIR) and exposure-adjusted event rate (EAER) for fractures occurred during last 12 months on 6- month dosing regimen and 12 months on 3-month dosing regimen will be summarized during the 12 months interval.

For details on the analysis code see Appendix E.

9.5.2.3 Incidence of Improving Vertebral Fractures

The proportion of subjects with improved vertebral fracture from baseline of 3-month dosing regimen will be summarized at month 12 based on the vertebral fracture analysis set on 3-month dosing regimen.

9.5.2.4 PRO Measurements

Patient-reported outcomes (CHQ-PF-50, CHAQ disability index score and WBFPRS) and their changes from 3-month dosing regimen will be summarized at month 12 based on their respective PRO analysis set. There is no imputation for missing baseline or post baseline data.

WBFPRS, CHQ-PF-50 physical and psychological summary scores, and CHAQ disability score will be calculated as detailed in the Appendix B.

9.5.2.5 Growth Velocity

Descriptive statistics of height-for-age, weight-for-age and BMI-for-age Z-scores, and the respective changes from baseline at 12 months will be summarized for 3-month dosing regimen. There is no imputation for missing baseline or postbaseline data.

9.5.2.6 Denosumab Serum Concentration

Descriptive statistics of denosumab serum concentration by visit (days 1 (pre-dose), 10, 30 ,60 and every 3 months (pre-dose) in subjects on 3-month dosing regimen) will be provided and may be evaluated by cohort based on the PK analysis set.

9.5.3 Analyses of Exploratory Efficacy Endpoint(s)

[redacted]

9.6 Safety Analyses

9.6.1 Analyses of Primary Safety Endpoint(s)

Table D Safety Endpoint Summary Table

All safety endpoints will be summarized for all subjects on 6-Month Dosing Regimen and subjects who received ≥ 1 dose of 3-Month Dosing Regimen separately.

Endpoint	Primary Summary and Analysis Method (Specify Analysis Set if Safety Analysis Set is Not Used)	Sensitivity Analysis
Subject incidence of adverse events and serious adverse events	All TEAEs will be summarized for all subjects on 6-Month Dosing Regimen using 6-Month Dosing Regimen Safety Analysis Set and subjects who received ≥ 1 dose of 3-Month Dosing Regimen separately using 3-Month Dosing Regimen Safety Analysis Set.	
Change from baseline in laboratory values	Descriptive statistics of the actual values and changes from baseline in laboratory values will be presented by visit.	
Change from baseline in vital signs	Descriptive statistics of the actual values and changes from baseline in vital signs (pulse, respiration rate, temperature) will be presented by visit. In addition, systolic and diastolic blood pressure will be summarized for 3-Month Dosing Regimen.	
Subject incidence of antidenosumab antibodies	The percentages of subjects who tested positive (binding or neutralizing) for antidenosumab antibodies will be descriptively summarized.	

Subject incidence of metaphyseal index Z-score above age-appropriate normal range	Subject incidence of the metaphyseal index Z-score above age-appropriate normal range safety events will be tabulated.	
Subject incidence of abnormal molar eruption	Subject incidence of the abnormal molar eruption safety events **based on radiological findings** will be tabulated.	
Subject incidence of abnormal mandibular shaping	**Descriptive summary statistics for mandibular shaping parameters will be tabulated.**	
Subject incidence of hypercalcemia	Subject incidence of hypercalcemia will be summarized per baseline age cohort and overall.	

9.6.2 Adverse Events

The Medical Dictionary for Regulatory Activities (MedDRA) version 24.0 or later will be used to code all events categorized as adverse events to a system organ class and a preferred term. All TEAEs will be summarized for all subjects on 6-Month Dosing Regimen using 6-Month Dosing Regimen Safety Analysis Set and subjects who received ≥ 1 dose of 3-Month Dosing Regimen separately using 3-Month Dosing Regimen Safety Analysis Set.

The subject incidence of adverse events will be summarized for all treatment-emergent adverse events with CTCAE grades, serious adverse events, adverse events leading to discontinuation of investigational product, fatal adverse events ,treatment-related adverse events with CTCAE grades, treatment-related serious adverse events, treatment-related fatal adverse events and treatment-related adverse events leading to withdrawal of IP.

Subject incidence of all treatment-emergent adverse event related to investigator product (IP) (including all treatment-emergent AEs, serious AEs and AEs leading to withdrawal of investigational product) will also be summarized by SOC and PT in descending order of frequency.

Summaries of treatment-emergent and serious AEs occurring in at least 5% of the subjects by preferred term will be provided in descending order of frequency in the overall group.

The exposure-adjusted subject incidence will be summarized for all subjects on 6-Month Dosing Regimen using 6-Month Dosing Regimen Safety Analysis Set and subjects who received ≥ 1 dose of 3-Month Dosing Regimen separately using 3-Month Dosing Regimen Safety Analysis Set.

9.6.2.1 Adverse Events of Special Interest

Subject incidence of following adverse events of interest will be summarized based on 6-Month Dosing Regimen using 6-Month Dosing Regimen Safety Analysis Set and subjects who received ≥ 1 dose of 3-Month Dosing Regimen separately using 3-Month Dosing Regimen Safety Analysis Set.

This tentative list has been based on the events of interest (EOI) identified for denosumab in the osteoporosis indication, and may be revised and changed later on to better reflect the OI population.

Event of Special Interest	Search Strategy	Narrow or Broad Search?
Hypocalcemia	Hypocalcemia AMQ	Narrow
Hypercalcemia	Hypercalcemia AMQ	Narrow
Positively adjudicated osteonecrosis of the jaw (ONJ)	Positively adjudicated cases in the database	NA
Hypersensitivity	Hypersensitivity SMQ	Narrow
Bacterial Cellulitis (skin infection)	Bacterial cellulitis AMQ with serious criteria of “hospitalization”	Narrow
Typical OI Femur Fractures (TOIFF)	Atypical femoral fractures AMQ	Narrow

In addition, Subject incidence for each adverse event of interest (EOI) will be presented by PT except for Positively adjudicated ONJ. Subject incidence of hypercalcemia will be summarized per baseline age cohort and overall. Listing of Positively adjudicated ONJ and typical OI femur fractures (TOIFF**)** will be presented.

9.6.2.2 Other Safety Endpoints

The subject incidence of the following safety events will be tabulated for all subjects on 6-Month Dosing Regimen using 6-Month Dosing Regimen Safety Analysis Set and subjects who received ≥ 1 dose of 3-Month Dosing Regimen separately using 3-Month Dosing Regimen Safety Analysis Set.:

- Metaphyseal index Z-score above age-appropriate normal range (ie, metaphyseal index Z-score > +2)
- Severe or symptomatic hypocalcemia (ie. serious hypocalcemia)
- Abnormal molar eruption of the first or second molars
- Abnormal mandibular shaping

9.6.3 Laboratory Test Results

Actual values and changes from baseline in each parameter will be descriptively summarized at each visit. For serum calcium, phosphorus and alkaline phosphatase, summary of the percent change from baseline will be provided as well. Shifts in laboratory parameters between baseline and the most extreme post-baseline values will be assessed based on the Common Terminology Criteria for Adverse Events v4.0 (CTCAE). See **Appendix A**. Graphical representations of serum calcium, BSAP, sCTX, Urine calcium may also be presented.

All laboratory analyses will be summarized for all subjects on 6-Month Dosing Regimen using 6-Month Dosing Regimen Safety Analysis Set and subjects who received ≥ 1 dose of 3-Month Dosing Regimen separately using 3-Month Dosing Regimen Safety Analysis Set.

Listings of actual values, change from baseline and percent change from baseline for serum calcium, phosphorus and alkaline phosphatase will be provided as well for German substudy **safety** analysis set.

Visit windows will be used for these summaries as described in Appendix C.

9.6.3.1 Renal Clinical Laboratory Test Results (German substudy only)

For 24-hour urine calcium and 24-hour urine phosphate the mean of all values during the 24-hour period will be calculated and used for analysis.

Listing will be presented for urine calcium, including spot urine calcium-to-creatinine ratio and, if applicable, 24-hour urine calcium at each visit. Change from baseline and

percent change from baseline will also be listed for German substudy **safety** analysis set.

Nephrocalcinosis grade and shifts from baseline, if applicable, will be listed by visit based on the German substudy **safety** analysis set**.**

Visit windows will be used for these summaries as described in Appendix C.

9.6.4 Vital Signs

Descriptive statistics of the actual values and changes from baseline in vital signs (pulse, respiration rate, temperature) will be presented by visit for all subjects on 6-Month Dosing Regimen using 6-Month Dosing Regimen Safety Analysis Set and subjects who received ≥ 1 dose of 3-Month Dosing Regimen separately using 3-Month Dosing Regimen Safety Analysis Set. In addition, systolic and diastolic blood pressure will be summarized during 3-Month Dosing Regimen.

9.6.5 Antidenosumab Antibodies

Subjects receiving at least 1 dose of denosumab will be tested for antidenosumab antibodies. The percentages of subjects who tested positive (binding or neutralizing) for antidenosumab antibodies will be descriptively summarized for all subjects on 6-Month Dosing Regimen using 6-Month Dosing Regimen Safety Analysis Set and subjects who received ≥ 1 dose of 3-Month Dosing Regimen separately using 3-Month Dosing Regimen Safety Analysis Set.

9.6.6 Exposure to Investigational Product

Descriptive statistics will be produced to describe the exposure to investigational product using 3-Month Dosing Regimen Safety Analysis Set..

9.6.7 Exposure to Non-investigational Product

Not Applicable for this study

9.6.8 Exposure to Other Protocol-required Therapy

Not Applicable for this study

9.6.9 Exposure to Concomitant Medication

Not Applicable for this study

9.7 Other Analyses

Not applicable for this study

9.7.1 Pharmacokinetic/Pharmacodynamic Analysis

The subgroups described in Section 4.2 will be used to descriptively summarize the serum denosumab concentrations by scheduled visit. Graphs of summary statistics by scheduled visit will also be produced.

Serum denosumab concentration and bone turnover marker data collected from the PK substudy may be used in combination with data from other clinical studies with denosumab in a population PK or pharmacokinetic and pharmacodynamic (PKPD) analyses. The results of the analysis, if conducted, and details regarding objectives, data handling, and methodology will be provided in a separate population PK or PKPD analysis plan.

9.7.2 Analyses of Clinical Outcome Assessments

Not applicable for this study

9.7.3 Analyses of Health Economic Endpoints

Not applicable for this study

9.7.4 Analyses of Biomarker Endpoints

Not applicable for this study

10. Changes From Protocol-specified Analyses

Removed the German substudy analysis set from section 6.4.8 as this definition was applicable to efficacy endpoints considering baseline and post baseline information for endpoint of interest. However to analyze the safety endpoints related German substudy the new definition is added as below in section 6.4.8,

German substudy safery analysis set - "The subset includes all subjects in the 3-month dosing regimen safety analysis set who enroll in the German substudy."

Instead of Subject incidence of abnormal mandibular shaping, we will provide descriptive summary statistics for mandibular shaping parameters.

To provide the maximum incidence of fractures the Section 9.5.2.2 updated as follows,

Fracture analysis will use the data as mentioned below :

- **Vertebral Fractures : central imaging vendor i.e fractures confirmed by vendor**
- **Nonvertebral Fractures other than long bone fractures: fractures reported on the Clinical Fracture Summary eCRF which is confirmed by the central imaging vendor**

- **Long bone Fractures : fractures reported on the Clinical Fracture Summary eCRF i.e fractures confirmed by imaging**

11. Literature Citations / References

Centers for Disease Control and Prevention, National Center for Health Statistics. CDC growth charts: United States. May 2000 http://www.cdc.gov/growthcharts/2000 Centers for Disease Control and Prevention Growth Charts for the United States:Methods and Development:http: //www.cdc.gov/growthcharts/2000growthchart-us.pdf

Letocha AD, Cintas HL, Troendle JF, et al. Controlled trial of pamidronate in children with types III and IV osteogenesis imperfecta confirms vertebral gains but not short-term functional improvement. J Bone Miner Res. 2005;20:977-986.

Rauch F, Munns CF, Land C, Cheung M, Glorieux FH Risedronate in the Treatment of Mild Pediatric Osteogenesis Imperfecta: A Randomized Placebo-Controlled Study. J Bone Miner Res. 2009;24:1282-1289.

Ward LM, Rauch F, Whyte MP, et al. Alendronate for the treatment of pediatric osteogenesis imperfecta: a randomized placebo-controlled study. J Clin Endocrinol Metab. 2011;96:355-364.

Ward K, Cowell CT and Little DG. Quantification of metaphyseal modeling in children treated with bisphosphonates. Bone 2005; 36: 999-1002

12. Prioritization of Analyses

Not applicable for this study

13. Data Not Covered by This Plan

Analysis for denosumab concentrations and additional pharmacokinetic analyses will not be described or analyzed under this plan. Pharmacokinetic and Drug metabolism (PKDM) group will provide a separate report for those analyses. The PKDM group will perform summary statistics on the serum denosumab PK samples collected in this study.

14. Appendices

Appendix A. Reference Values/Toxicity Grades

The Common Terminology Criteria for Adverse Events (CTCAE) are available at the following link:

https://evs.nci.nih.gov/ftp1/CTCAE/CTCAE_4.03/CTCAE_4.03_2010-06-14_QuickReference_5x7.pdf

Appendix B.Health Economic Forms/Instruments

B.1 Wong-Baker FACES Pain Rating Scale (WBFPRS)

At each assessment (baseline, 12 months), subjects are asked to choose a face that best describes their own pain; the corresponding number (0, 2, 4, 6, 8, or 10) is recorded.

This rating scale is recommended for subjects age 3 years and older, therefore only subjects at the age of 3 or older at baseline will be asked to respond.

There is no imputation for missing values.

Change from baseline will be calculates for those subjects with baseline and post-baseline assessments at the timepoint of interest (12 months).

B.2 Child Health Questionnaire – Parent Form 50

The scales and corresponding items for the CHQ-PF50 are listed below in the order in which they appear in the health survey. Specific instructions for their scoring are given in the Child Health Questionnaire Scoring and Interpretation Manual ©, 2013 HealthActCHQ Inc Boston MA. A summary of those instructions is provided below.

Scale/Item Name (and abbreviation)	Questions	No. of Items	Questionnaire Section
Global Health (GH)	1.1	1	Your Child's Global Health (1)
Physical Functioning (PF)	2.1.a - 2.1.f	6	Your Child's Physical Activities (2)

Roles/Social Limitations-Emotional/Behavioral (REB)	3.1.a - 3.1c	3	Your Child's Everyday Activities (3.1)
Roles/Social Limitations-Physical (RP)	3.2.a - 3.2.b	2	Your Child's Everyday Activities (3.2)
Bodily Pain/Discomfort (BP)	4.1, 4.2	2	Pain (4)
Behavior (BE)	5.1.a - 5.1.e,	6	Behavior (5.1)
Global Behavior Item (GBE)	5.2	1	Behavior (5.2)
Mental Health (MH)	6.1.a - 6.1.e	5	Well-Being (6)
Self Esteem (SE)	7.1.a - 7.1.f	6	Self Esteem (7)
General Health Perception (GH)	8.1.a - 8.1.e	5	Your Child's Health (8.1)
Change in health (CH)	8.2	1	Your Child's Health (8.2)
Parental Impact – Emotional (PE)	9.1.a – 9.1.c	3	You and Your Family (9.1)
Parental Impact – Time (PT)	9.2.a – 9.2.c	3	You and Your Family (9.2)
Family Activity (FA)	9.3a – 9.3 f	6	You and Your Family (9.3)
Family Cohesion (FC)	9.4	1	You and Your Family (9.4)

Items should be recoded and recalibrated:

- Recoding is necessary to ensure that all items are positively scored so that a higher score indicates better health. Scores vary from 5 to 1.
- Recalibrating is performed for two items (Question 1.1 and 5.2) with an "excellent to poor" response as given below.

Table Recalibration

Response choices	Precoded Value	Final Value
Excellent	1	5.0
Very good	2	4.4
Good	3	3.4
Fair	4	2.2
Poor	5	1.0

- Recalibrating is performed for one item (Question 4.1) with a “none to very severe” response as given below.

Response choices	Precoded Value	Final Value
None	1	6
Very mild	2	5
Mild	3	4
Moderate	4	3
Severe	5	2
Very severe	6	1

- Recalibrating is performed for one item (Question 4.2) with a “none of the time to every/almost every day” response as given below.

Response choices	Precoded Value	Final Value
None of the time	1	6
Once or twice	2	5
A few times	3	4
Fairly often	4	3
Very often	5	2
Every/almost every day	6	1

- Recalibrating is performed for one item (Question 6.1.e) with an “all of the time to none of the time” response as given below.

Response choices	Precoded Value	Final Value
All of the time	1	5
Most of the time	2	4
Some of the time	3	3
A little of the time	4	2
None of the time	5	1

- Recalibrating is performed for six items (Question 7.1.a - 7.1.f) with a “very satisfied to very dissatisfied” response as given below.

Response choices	Precoded Value	Final Value
Very satisfied	1	5
Somewhat satisfied	2	4
Neither satisfied nor dissatisfied	3	3
Somewhat dissatisfied	4	2
Very dissatisfied	5	1

- Recalibrating is performed for two items (Question 8.1.b, 8.1.d) with a “definitely true to definitely false” response as given below.

Response choices	Precoded Value	Final Value
Definitely true	1	5
Mostly true	2	4
Don't know	3	3
Mostly false	4	2
Definitely false	5	1

- Recalibrating is performed for a item (Question 8.2) with a “Much better now than 1 year ago” to “Much worse now than 1 year ago” response as given below.

Response choices	Precoded Value	Final Value
Much better now than 1 year ago	1	5
Somewhat better now than 1 year ago	2	4
About the same now as 1 year ago	3	3
Somewhat worse now than 1 year ago	4	2
Much worse now than 1 year ago	5	1

- Recalibrating is performed for three items (Question 9.1.a - 9.1.c) with a “none at all to a lot” response as given below.

Response choices	Precoded Value	Final Value
None at all	1	5
A little bit	2	4
Some	3	3
Quite a bit	4	2
A lot	5	1

For subjects that complete half or more of the items in each scale, the mean of the scale will be calculated based on the available responses by adding all non-missing values and divided by the number of non-missing items; otherwise scale is set to

missing. Mean is provided with 1 decimal place. After calculating the mean for each subject, it will need to be standardized using the formula as: [(Mean -1)/N]*100. The transformed score should be rounded to 0 decimal place.

For each of the 10 scales in Table below, from the manual, each subject's non-missing mean is then standardized by using the Mean and SD of the reference population. For example, the mean for REB is standardized as:

REB_Z = (REB - 90.4013015) / 19.5067502

CHQ-PF50 Scale and Factor Score Coefficients

CHQ-PF50 Scale	No. of Completed Items*	N	Mean	SD	Factor Score Coefficients	
					PhS	PsS
GH	3	4	66.6958379	19.3564297	0. 29460	-0.05547
PF	3	3	90.8525408	16.3826344	0. 37138	-0. 09243
REB	2	3	90.4013015	19.5067502	-0.01178	0. 21155
RP	1	3	91.4951246	18.9079749	0. 34493	-0.06973
BP	1	5	78.6833515	20.7355708	0. 27883	-0. 05514
BE	3	4	72.3086051	17.1447913	-0.12675	0. 27911
MH	3	4	77.2595806	13.6861999	-0.08263	0. 25335
SE	3	4	79.2555314	17.8308361	-0.09480	0. 24792
PE	2	4	73.9788476	21.406013	0. 06063	0. 19823
PT	2	4	83.8816188	20.2901603	0. 09113	0. 16944

*The raw scores can be created by computing the algebraic mean of items when the respondents who completed at least the number of items listed in the table; otherwise both the raw scores and transformed scores should be set to missing.

Values are given with 2 decimal places.

The aggregate CHQ-PF50 physical summary score (PhS) consists of multiplying each CHQ-PF50 scale z-score by its respective physical factor score coefficient (PhS) and

summing the ten products. Result is given with 6 decimal places and is called PhS_raw. The formula of calculating aggregate PhS (PhS_raw) is:

PhS_raw = (GH_Z * 0.29460) + (PF_Z * 0.37138) + (REB_Z * -0.01178) + (RP_Z * 0.34493) + (BP_Z * 0.27883) + (BE_Z * -0.12675) + (MH_Z * -0.08263) + (SE_Z * -0.09480) + (PE_Z * 0.06063) + (PT_Z * 0.09113)

And finally the transformed physical (PhS) is calculates as:

PhS = 50 + (PhS_raw * 10)

PhS should be rounded to 1 decimal place.

Similarly, the aggregate CHQ-PF50 psychological summary score (PsS) consists of multiplying each CHQ-PF50 scale z-score by its respective psychological factor score coefficient (PsS) and summing the ten products. Result is given with 6 decimal places and is called PsS_raw. The formula of calculating aggregate PsS (PsS_raw) is:

PsS_raw = (GH_Z * -0.05547) + (PF_Z * -0.09243) + (REB_Z * 0.21155) + (RP_Z * -0.06973) + (BP_Z * -0.05514) + (BE_Z * 0.27911) + (MH_Z * 0.25335) + (SE_Z * 0.24792) + PE_Z * 0.19823) + (PT_Z * 0.16944)

And finally the transformed physical (PsS) is calculates as:

PsS = 50 + (PsS_raw * 10)

PsS should be rounded to 1 decimal place.

For each subject, the summary scale scores (PhS and PsS) are set to missing if the subject is missing any one of the ten CHQ-PF50. To minimize the number of summary scores missing and for a further comparison with original PhS and PsS, each of the ten scale scores is calculated if half or more of the items are completed.

B.3 Childhood Health Assessment Questionnaire (CHAQ)

The disability domain (questions 1-54) of the CHAQ was used to measure the subject's assessment of physical functioning or the parent's assessment of the child's physical functioning. The disability index comprises eight categories (dressing and grooming, arising, eating, walking, hygiene, reaching, gripping, and activities). Scoring ranged from 1 to 5; 1 was "Without ANY difficulty," 2 was "With SOME difficulty," 3 was "With MUCH difficulty," and 4 was "UNABLE to do". An answer of "not applicable" was scored as a 5, but was not counted. If a child required assistance from another person or used an aid or other device for any of the eight categories, the minimum score for that

category was to be recorded as a 3. The CHAQ questions would then need to be converted to a 0 to 3 scaled score.

Afterwards, the CHAQ score using the aids (and/or devices) indicator is computed by taking the maximum score of the questions in each domain (range: 0, 3) and whether or not aids/devices are used (range: 0, 1):

A = max(dressing & grooming domain questions, 2*aids indicator) +

+ max(rising domain questions, 2*aids indicator) +

+ max(eating domain questions, 2*aids indicator) +

+max(walking domain questions, 2*aids indicator) +

+ max(hygiene domain questions, 2*aids indicator) +

+ max(reach domain questions, 2*aids indicator) +

+ max(grip domain questions, 2*aids indicator) +

+ max(usual activities domain questions, 2*aids indicator).

CHAQ= A/(total number of sub-domains with at least 6 non-missing)

The CHAQ score is given with 2 decimal places. No missing values are imputed. If any of the domain questions are missing, but the aids/device indicator is non-missing, the domain can still be computed. However, if a domain is missing all domain questions (or is scored as “not applicable”) and is missing aids/device indicators, then the domain score is considered missing. The CHAQ score is not computed when the subject/parent provides answers in fewer than six domains (normally coded as “UNK” in the database).

The following table shows the contribution of the 54 questions used to calculate the CHAQ:

CHAQ Domains and Scores

CHAQ Domains	**Domian Components: At least 6 domains must have scores to compute the CHAQ.**		**CHAQ Component Score Definition:**
	Domain Questions	**Aids/Devices Questions**	

Dressing/Grooming	1, 2, 3, 4	16, 20	First set the Domain Component to missing if the original question=5 for Scale Range Questions. Then, convert values of each question to a scale of 0-3 (Without ANY difficulty=0, With SOME difficulty=1, With MUCH difficulty=2, UNABLE to do=3). If Aids/Devices=Yes for Yes/No questions, then set the minimum CHAQ Domain to be equal to 2. Specific terms are mapped to current component for question 19, 53 or 54. Set max of the Domain Questions and Aid/Devices Questions as the component score. Note: Questions 19, 53 and 54 use format to define which category.
Arising	5, 6	18, 21	
Eating	7, 8, 9	17, 22	
Walking	10, 11	12, 13, 14, 15, 23	
Hygiene	24, 25, 26, 27, 28	43, 44, 46, 48, 49	
Reach	29, 30, 31, 32	47, 50	
Grip	33, 34, 35, 36, 37	45, 51	
Activity	38, 39, 40, 41, 42	52	

Appendix C.Analytical Windows

Per protocol, all tests and procedures scheduled on Study Day 10, 30 and 60 will be performed ± 3 days of the scheduled day and ± 7 days for visits at Months 3, 6, 9,12 and EOS. To allow for variations in scheduling, the following sets of visit windows will be used to assign evaluations to a most appropriate nominal visit for analysis. Furthermore, there will be no gaps between visit windows in order to include as many data points as possible for summarization.

Regardless of the width of the visit window, if more than 1 visit falls within the defined window, the result from the visit closest to the target day will be used. If 2 evaluations are of the same distance from the target day, the result from the later visit will be used. Results not included for analysis will be included in listings.

Study day 1 is defined as the day of initial dose of IP administration during the respective dosing regimen.

Visit Windows of 3-month dosing regimen

Table C.1.1 Lateral Thoracic and Lumbar Spine X-ray,CHQ-PF50, CHAQ Disability Score, Non-vertebral Fractures, WBFPS and Armspan

Nominal Visit	Target Day	Window Definition (Study Day)
Q3M Baseline	1	Study Day 1
Q3M Month 12	366	Study Day 2 to 458
Q3M Month 18	549	After Study Day of 458

Table C.1.2 DXA, AP Knees

Nominal Visit	Target Day	Window Definition (Study Day)
Q3M Baseline	1	Study Day 1
Q3M Month 6	183	Study Day 2 to 275
Q3M Month 12	366	Study Day 276 to 458
Q3M Month 18	549	After Study Day of 458

Table C.1.3 Tanner stage, Physical Examination, Oral Visual Inspection, Dental Xray (molars), Haematology, Height, Weight, Pregnancy Test, Urine calcium, Denosumab Administration and Dispensation of calcium and vitamin D

Nominal Visit	Target Day	Window Definition (Study Day)
Q3M Baseline	1	Study Day 1
Q3M Month 3	92	Study Day 2 to 138
Q3M Month 6	183	Study Day 139 to 229
Q3M Month 9	275	Study Day 230 to 321
Q3M Month 12	366	Study Day 322 to 412
Q3M Month 15	458	Study Day 413 to 504
Q3M Month 18[a]	549	After Study Day of 504

[a]This visit is not for Denosumab Administration and Dispensation of calcium and vitamin

Table C.1.4 Vital Signs, Bone Turnover Marker

Nominal Visit	Target Day	Window Definition (Study Day)
Q3M Baseline	1	Study day 1
Q3M Day 10	10	Study Day 2 to 20
Q3M Day 30	30	Study Day 21 to 45
Q3M Day 60	60	Study Day 46 to 76
Q3M Month 3	92	Study Day 77 to 138
Q3M Month 6	183	Study Day 139 to 229
Q3M Month 9	275	Study Day 230 to 321
Q3M Month 12	366	Study Day 322 to 412
Q3M Month 15	458	Study Day 413 to 504
Q3M Month 18	549	After Study Day of 504

Table C.1.5 Serum Chemistry

Nominal Visit	**Target Day**	**Window Definition (Study Day)**
Q3M Baseline	1	Study day 1
Q3M Day 10	10	Study Day 2 to 20
Q3M Day 30	30	Study Day 21 to 61
Q3M Month 3	92	Study Day 62 to 97
Q3M Month 3 Day 10	102	Study Day 98 to 112
Q3M Month 3 Day 30	122	Study Day 113 to 153
Q3M Month 6	183	Study Day 154 to 188
Q3M Month 6 Day 10	193	Study Day 189 to 203

Nominal Visit	**Target Day**	**Window Definition (Study Day)**
Q3M Month 6 Day 30	213	Study Day 204 to 244
Q3M Month 9	275	Study Day 245 to 321
Q3M Month 12	366	Study Day 322 to 412
Q3M Month 15	458	Study Day 413 to 504
Q3M Month 18	549	After Study Day of 504

Table C.1. 6 Cephalogram and Panoramic Radiograph

Nominal Visit	Target Day	Window Definition (Study Day)
Q3M Baseline	1	Last evaluation on or before Study Day 1
Q3M Month 12	366	After Study Day of 2

Table C.1.7 Renal laboratory value

Nominal Visit	Target Day	Window Definition (Study Day)
Q3M Baseline	1	Study Day 1
Q3M Week 6	42	Study Day 2 to 49
Q3M Week 8	56	Study Day 50 to 63
Q3M Week 10	70	Study Day 64 to 77
Q3M Week 12	**84**	**Study Day 78 to 109**
Q3M Month 3 Week 6	**134**	**Study Day 110 to 141**
Q3M Month 3 Week 8	**148**	**Study Day 142 to 155**
Q3M Month 3 Week 10	**162**	**Study Day 156 to 169**
Q3M Month 3 Week 12	**176**	**Study Day 170 to 201**
Q3M Month 6 Week 6	**225**	**Study Day 202 to 232**
Q3M Month 6 Week 8	**239**	**Study Day 233 to 246**
Q3M Month 6 Week 10	**253**	**Study Day 247 to 260**
Q3M Month 6 Week 12	**267**	**Study Day 261 to 292**
Q3M Month 9 Week 6	**317**	**Study Day 293 to 324**
Q3M Month 9 Week 8	**331**	**Study Day 325 to 338**
Q3M Month 9 Week 10	**345**	**Study Day 339 to 352**
Q3M Month 9 Week 12	**359**	**Study Day 353 to 384**
Q3M Month 12 Week 6	**408**	**Study Day 385 to 415**
Q3M Month 12 Week 8	**422**	**Study Day 416 to 429**
Q3M Month 12 Week 10	**436**	**Study Day 430 to 443**
Q3M Month 12 Week 12	**450**	**Study Day 444 to 500**
Q3M Month 18	549	**After Study Day of 500**

Visit Windows of 6-month dosing regimen

The below visit windows are applicable only for the assessments collected during 6-month dosing regimen. For subjects transitioned to 3-month dosing regimen the upper bound (UB) of last visit on 6-month dosing regimen will be assigned as below. There will not be subsequent visits on 6-month dosing regimen for the transitioned subjects.

UB for Month xx = minimum(UB of month xx defined in table C.2.x, (first dose date of 3-month dosing regimen - first dose date of 6-month dosing regimen + 1)).

Table C.2.1 Lateral Thoracic and Lumbar Spine X-ray,CHQ-PF50, CHAQ Disability Score, Non-vertebral Fractures ,WBFPS, Oral Visual Inspection and Dental X-ray , Cephalogram and Panoramic Radiograph

Nominal Visit	Target Day	Window Definition (Study Day)
Baseline	1	Last evaluation on or before Study Day 1
Month 12	365	Study Day 2 to 548
Month 24	731	Study Day 549 to 913
Month 36	1096	Study Day 914 to 1278

Table C.2.2 DXA

Nominal Visit	Target Day	Window Definition (Study Day)
Baseline [a]	1	Last evaluation on or before Study Day 60
Month 6	183	Study Day 61 to 275
Month 12	366	Study Day 276 to 458
Month 18	549	Study Day 459 to 641
Month 24	732	Study Day 642 to 915
Month 36	1098	Study Day 916 to 1281

[a] only applicable when baseline (before first dose) is missing.

Table C.2.3 Tanner stage and X-ray AP Knees

Nominal Visit	Target Day	Window Definition (Study Day)
Baseline	-38	Last evaluation on or before Study Day 1
Month 6	183	Study Day 2 to 275
Month 12	366	Study Day 276 to 458
Month 18	549	Study Day 459 to 641
Month 24	732	Study Day 642 to 824
Month 30	915	Study day 825 to 1007
Month 36	1098	Study Day 1008 to 1190

Table C.2.4 Bone Turnover Marker

Nominal Visit	Target Day	Window Definition (Study Day)
Baseline	1	Study day 1 to 5
Study Day 10	10	Study day 6 to 20
Month 6	183	Study Day 21 to 275
Month 12	366	Study Day 276 to 458
Month 18	549	Study Day 459 to 641
Month 24	732	Study Day 642 to 824
Month 30	915	Study day 825 to 1007
Month 36	1098	Study Day 1008 to 1190

Table C.2.5 Vital Signs, Bone Turnover Marker Substudy and Serum Chemistry

Nominal Visit	Target Day	Window Definition (Study Day)
Screening	-38	Last evaluation on or before Study Day 1
Baseline	1	Study day 1 to 5
Study Day 10	10	Study day 6 to 20
Study Day 30	30	Study day 21 to 61
Month 3	92	Study day 62 to 138
Month 6	183	Study Day 139 to 275
Month 12	366	Study Day 276 to 458
Month 18	549	Study Day 459 to 641
Month 24	732	Study Day 642 to 824
Month 30	915	Study day 825 to 1007
Month 36	1098	Study Day 1008 to 1190

Table C.2.6 Physical Examination and Haematology

Nominal Visit	Target Day	Window Definition (Study Day)
Screening	-38	Last evaluation on or before Study Day-1
Baseline	1	Study day 1 to 5
Month 3	91	Study day 6 to 137
Month 6	183	Study Day 138 to 275
Month 12	366	Study Day 276 to 458
Month 18	549	Study Day 459 to 641
Month 24	732	Study Day 642 to 824
Month 30	915	Study day 825 to 1007
Month 36	1098	Study Day 1008 to 1190

Table C.2.7 Weight

Nominal Visit	Target Day	Window Definition (Study Day)
Screening	-38	Last evaluation on or before Study Day-1
Baseline	1	Study Day 1 to 5
Month 6	183	Study Day 6 to 275
Month 12	366	Study Day 276 to 458
Month 18	549	Study Day 459 to 641
Month 24	732	Study Day 642 to 824
Month 30	915	Study Day 825 to1007
Month 36	1098	Study Day 1008 to 1190

Table C.2.8 Pregnancy Test, Denosumab Administration

Nominal Visit	Target Day	Window Definition (Study Day)
Screening[a]	-38	Last evaluation on or before Study Day-1
Baseline	1	On study day 1.
Month 6	183	Study Day 2 to 275
Month 12	366	Study Day 276 to 458
Month 18	549	Study Day 459 to 641
Month 24	732	Study Day 642 to 824
Month 30	915	Study Day 825 to 1007

[a] Only for Pregnancy Test

Table C.2.9 Height and Armspan

Nominal Visit	Target Day	Window Definition (Study Day)
Screening	-38	Last evaluation on or before Study Day 1
Month 12	366	Study Day 2 to 549
Month 24	732	Study Day 550 to 915
Month 36	1098	Study Day 916 to 1190

Table C.2.10 Renal laboratory value

Nominal Visit	**Target Day**	**Window Definition (Study Day)**
Baseline	**-38**	**Last evaluation on or before Study Day 1**
Week 6	**42**	**Study Day 2 to 49**
Week 8	**56**	**Study Day 50 to 63**
Week 10	**70**	**Study Day 64 to 77**
Week 12	**84**	**Study Day 78 to 155**
Month 6 Week 6	**225**	**Study 156 to 232**
Month 6 Week 8	**239**	**Study 233 to 246**
Month 6 Week 10	**253**	**Study 247 to 260**
Month 6 Week 12	**267**	**Study 261 to 338**
Month 12 Week 6	**408**	**Study 339 to 415**
Month 12 Week 8	**422**	**Study 416 to 429**
Month 12 Week 10	**436**	**Study 430 to 443**
Month 12 Week 12	**450**	**Study 444 to 521**
Month 18 Week 6	**591**	**Study 522 to 598**
Month 18 Week 8	**605**	**Study 599 to 612**
Month 18 Week 10	**619**	**Study 613 to 626**

Month 18 Week 12	**633**	**Study 627 to 704**
Month 24 Week 6	**774**	**Study 705 to 781**
Month 24 Week 8	**788**	**Study 782 to 795**
Month 24 Week 10	**802**	**Study 796 to 809**
Month 24 Week 12	**816**	**Study 810 to 887**
Month 30 Week 6	**957**	**Study 888 to 964**
Month 30 Week 8	**971**	**Study 965 to 978**
Month 30 Week 10	**985**	**Study 979 to 992**
Month 30 Week 12	**999**	**Study 993 to 1070**

Appendix D.Bone Codes for Non-vertebral Fractures

Table D1 Bone Codes for Non-vertebral Fractures[a]

Code	Name	Code	Name
100	Skull	480	Acetabulum
120	Facial	502	Femur Neck
122	Mandible	503	Femur Intertrochanter
210	Cervical Vertebrae	504	Femur Subtrochanter
240	Sacrum	512	Femur Midshaft
250	Coccyx	513	Femur Distal
310	Ribs	520	Patella
320	Sternum	530	Fibula
401	Clavicle	531	Fibula Proximal
402	Scapula	532	Fibula Shaft
410	Humerus	533	Fibula Distal
411	Humerus Proximal	540	Tibia
412	Humerus Shaft	541	Tibia Proximal
413	Humerus Distal	542	Tibia Shaft
420	Radius	543	Tibia Distal
421	Radius Proximal	550	Metatarsus
422	Radius Shaft	560	Tarsus
423	Radius Distal	570	Toe Phalanges
430	Ulna	620	Ilium
431	Ulna Proximal	630	Ischium
432	Ulna Shaft	640	Pubis
433	Ulna Distal	888	Other (Specify)
450	Carpus		
460	Metacarpus		
470	Finger Phalanges		

[a]From BioClinica Charter for Independent Imaging Assessment

Table D2. Mapping Nonvertebral Fracture Locations[a]

BioClinica Location	eCRF Locations
Carpus	FRACTURED CARPAL
Cervical Vertebrae	CERVICAL VERTEBRAL
Clavicle	CLAVICLE **FRACTURE**
Coccyx	**FRACTURED** COCCYX
Facial	FACIAL BONES **FRACTURE**
Femur Distal	FEMUR DISTAL **FRACTURE**
Femur Intertrochanter	FEMUR INTERTROCHANTER **FRACTURE**
Femur Midshaft	FEMUR SHAFT **FRACTURE**
Femur Neck	FEMORAL NECK **FRACTURE**
Femur Subtrochanter	FEMUR FRACTURE SUBTROCHANTERIC
Fibula	FIBULA **FRACTURE**
Fibula Distal	
Fibula Proximal	
Fibula Shaft	
Finger Phalanges	FRACTURED FINGER
Humerus	HUMERUS **FRACTURE**
Humerus Distal	
Humerus Proximal	
Humerus Shaft	
Ilium	ILIUM **FRACTURE**
Ischium	FRACTURED ISCHIUM
Mandible	FRACTURED MANDIBLE
Metacarpus	FRACTURED METACARPAL
Metatarsus	FRACTURED METATARSAL
Other (Specify)	-
Patella	PATELLA FRACTURE
Pubis	PUBIS FRACTURE
Radius	RADIUS FRACTURE
Radius Distal	
Radius Proximal	
Radius Shaft	
Ribs	RIB FRACTURE
Sacrum	FRACTURED SACRUM
Scapula	SCAPULA FRACTURE
Skull	SKULL FRACTURE
Sternum	STERNAL FRACTURE
Tarsus	FOOT FRACTURE

BioClinica Location	**eCRF Locations**
Tibia	TIBIA FRACTURE
Tibia Distal	
Tibia Proximal	
Tibia Shaft	
Toe Phalanges	FRACTURED TOE
Acetabulum	ACETABULUM
Ulna	ULNA FRACTURE
Ulna Distal	
Ulna Proximal	
Ulna Shaft	
Wrist	

[a]From BioClinica Charter for Independent Imaging Assessment and eCRF Clinical Fracture

Appendix E.Code Fragment

[redacted]