STATISTICAL ANALYSIS PLAN

NCT #NCT03512262

A Randomized, Double-blind, Placebo-controlled, Phase 3 Multicenter Study to Evaluate the Safety and Efficacy of Abaloparatide-SC for the Treatment of Men with Osteoporosis

Protocol BA058-05-019

Protocol Version and Date:	Amendment 6 (30 March 2021)
Name of Test Drug:	Abaloparatide-SC
Phase:	Phase III
Methodology:	International, Multicenter, Randomized, Double-Blind, Placebo-Controlled
Sponsor:	Radius Health, Inc. 22 Boston Wharf Road, 7th floor Boston, MA 02210 Tel: 617-551-4000 Fax: 617-551-4701
Analysis Plan Date:	25 August 2021
Analysis Plan Version:	Final 1.0

Protocol Title: A Randomized, Double-blind, Placebo-controlled, Phase 3 Multicenter Study to Evaluate the Safety and Efficacy of Abaloparatide-SC for the Treatment of Men with Osteoporosis

Protocol Number: BA058-05-019, Amendment 6 (30 March 2021)

Study Drug: Abaloparatide-SC or Placebo

Version (Document Date): Final 1.0 (25 August 2021)

Prepared By:

[redacted] Date

[redacted]

Radius Health, Inc.

REVIEWED AND APPROVED BY:

[redacted] Date

[redacted]

Radius Health, Inc.

[redacted] Date

[redacted]

Radius Health, Inc.

[redacted] Date

[redacted]

Radius Health, Inc.

TABLE OF CONTENTS

Tables Included in the Text

Figures Included in the Text

LIST OF ABBREVIATIONS

Abbreviation	Term
°C	Degree Celsius
°F	Degree Fahrenheit
µg	Microgram
µmol	Micromole
ADA	Anti-Drug Antibody
AE	Adverse event
ALT	Alanine aminotransferase
AST	Aspartate aminotransferase
BMD	Bone mineral density
BMI	Body mass index
bpm	Beats per minute
cm	Centimeter
CSR	Clinical study report
s-CTX	Serum carboxy-terminal cross-linking telopeptide of type I collagen
DXA	Dual energy X-ray absorptiometry
ECG	Electrocardiogram
eCRF	Electronic case report form
FDA	Food and Drug Administration
g	Gram
HLT	High level term
ITT	Intent-to-treat
IRT	Interactive response technology
IU	International unit
IV	Intravenous
kg	Kilogram
L	Liter
LDH	Lactate dehydrogenase
LOCF	Last observation carried forward
MedDRA	Medical dictionary for regulatory activities
µL	Microliter
mg	Milligram
mL	Milliliter
mmHg	Millimeter of mercury
msec	Millisecond
NAb	Neutralizing antibody
ng	Nanogram
PD	Protocol deviation
pg	Picogram

Abbreviation	Term
PI	Principal Investigator
s-PINP	Serum procollagen type I N propeptide
PK	Pharmacokinetic
PT	Preferred term
PTH	Parathyroid hormone
PTHrP	Parathyroid hormone related peptide
PTT	Partial thromboplastin time
QT	Total depolarization and repolarization time
QTc	Total depolarization and repolarization time corrected with heart rate
SAE	Serious adverse event
SAP	Statistical Analysis Plan
SC	Subcutaneous
SD	Standard deviation
SHBG	Sex-hormone binding globulin
SMQ	Standardized MedDRA Query
SOC	System organ class
TEAEs	Treatment emergent adverse events
TSH	Thyroid stimulating hormone
ULN	Upper limit of normal
WBC	White blood cells
WHO	World Health Organization

1 INFORMATION FROM THE STUDY PROTOCOL

1.1 Introduction and Objectives

1.1.1 Introduction

Abaloparatide (marketed as TYMLOS® in the United States) is a novel, synthetic, 34 amino acid peptide designed to be a potent and selective activator of the PTH/PTH-related protein (PTHrP) type 1 receptor ($PTHR_1$) signaling pathway with 41% homology to PTH[1–34] and 76% homology to human PTHrP[1–34]. It was developed with the expectation that it would be effective at increasing bone mineral density (BMD) and reducing fracture risk in individuals with osteoporosis, but with a limited effect on bone resorption and a reduced risk of hypercalcemia.

Abaloparatide has been approved in the United States for the treatment of women with postmenopausal osteoporosis at high risk for fracture defined as a history of osteoporotic fracture, multiple risk factors for fracture, or subjects who have failed or are intolerant to other available osteoporosis therapy. While other agents from both anabolic and antiresorptive classes of anti-osteoporosis therapies have been approved for the treatment of osteoporosis in men, the safety and efficacy of abaloparatide has yet to be demonstrated in osteoporotic men. However, given the existence of an enhanced therapeutic response to abaloparatide-SC in postmenopausal women at high risk of fracture, as well as the substantive morbidity and mortality burden associated with osteoporotic fractures, confirmation of a similarly robust effect of abaloparatide-SC in men is a clinical priority. The present study seeks to determine if abaloparatide-SC is effective in increasing bone mass in men with osteoporosis.

1.1.2 Study Objectives

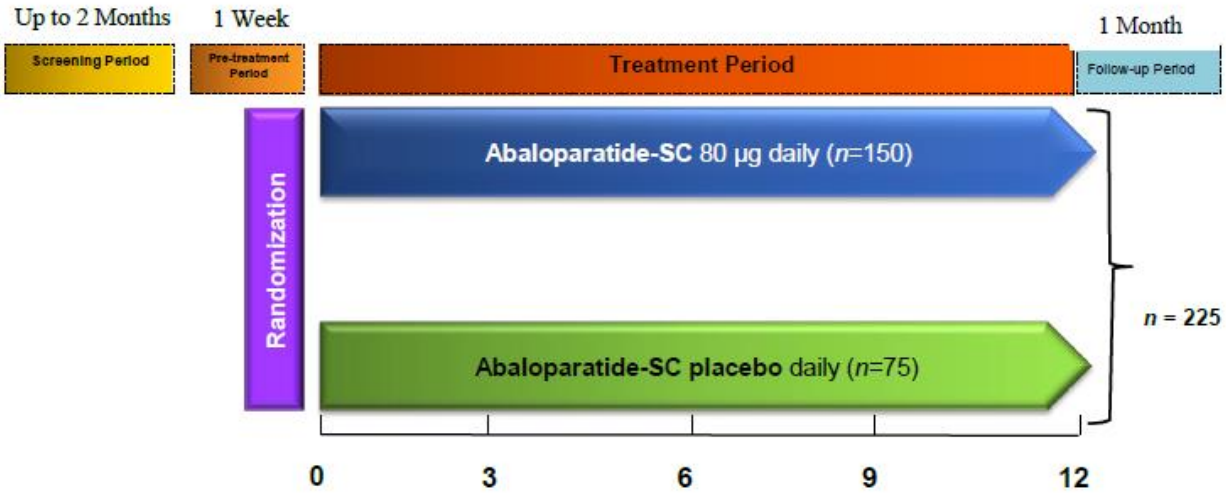
The primary objective of this prospective controlled study is to evaluate the efficacy and the safety of abaloparatide-SC 80 µg per day compared to placebo in men with osteoporosis.

1.2 Study Design

1.2.1 Synopsis of Study Design

This is a randomized, double-blind, placebo-controlled, Phase 3, multicenter study enrolling approximately 225 subjects at up to 40 study centers designed to evaluate the efficacy and safety of abaloparatide-SC 80 µg per day for the treatment of osteoporosis in men at high risk for fracture. The study will consist of a Screening Period (up to 2 months), a Pretreatment Period (1 week), a Treatment Period (12 months), and a final study visit one month after the last dose of study medication (Follow-up visit). Eligible subjects will be randomized, on a 2:1 basis, to receive either abaloparatide-SC 80 µg per day or placebo injections for 12 months to determine the effect of abaloparatide on lumbar spine BMD and other bone health-related endpoints. All subjects will receive calcium 500-1000 mg/day and vitamin D 400-800 IU/day, or a dose determined by the Investigator and agreed by the Sponsor Medical Monitor, according to the subject's need. Both treatment groups of subjects will undergo protocol-specified procedures (Schedule of Events, Table 1), including BMD and fracture assessment. The study design is presented in Figure 1.

Figure 1: Study BA058-05-019 Study Design



1.2.2 Randomization Methodology

It is planned to enroll approximately 225 subjects at up to 40 study sites in 3 to 5 countries to receive either abaloparatide-SC 80 µg or placebo in a 2:1 ratio on Day 1 using permuted block randomization.

Study medication kits will be shipped to each study site. Once a subject has been deemed eligible to be randomized into the study on Day 1 (Visit 3), the study site will use the Interactive Response Technology (IRT) system to assign a medication kit to the subject. Throughout the study, medication kits will be assigned to the subject based on information provided by the IRT.

1.2.3 Blinding

Abaloparatide-SC 80 µg and placebo study medications will be prepared in a blinded fashion and will be supplied to the study sites in identical packaging. Information regarding treatment assignment will reside solely within the IRT, as part of the study blinding.

The central imaging laboratory responsible for measuring BMD will be blinded to treatment assignment throughout the study. Radius will not receive post-baseline BMD measurements until database lock.

The Pharmacokinetic (PK) samples and the blood samples for serum markers of bone metabolism (s-PINP and s-CTX) will be analyzed during study. The results of the PK and bone turnover markers data will only be provided to Radius after database lock.

In order to follow subjects who are antibody positive at their completion visit (Visit 9, or Visit 8 for early terminations), the appropriate procedures for maintaining the integrity of the study blind are documented in the Blind Management Plan for Immunogenicity Testing and the Immunogenicity Testing Operations Manual.

In the event of a medical emergency where the identity of study medication is necessary to appropriately treat the subject, the Investigator may unblind the subject's treatment through the IRT. Unblinding at the end of the study will occur after the data cleaning has occurred, the SAP has been finalized, and the database is locked.

1.2.4 Study Procedures

The schedule of assessments, as outlined in the study protocol, is provided in Table 1 and Table 2.

Table 1: Schedule of Assessments and Procedures

	Visit	1	2	3	4	5	6	7	8	9
Procedure	***Study Day/Month:***	**Screening**	**Pretreatment**	**Day 1**	**Month 1**	**Month 3**	**Month 6**	**Month 9**	**Month 12/Early Termination**	**Month 13 Follow-up**
Visit Window (Days)		**N/A**	**± 4**	**± 4**	**± 4**	**± 4**	**± 4**	**± 4**	**± 4**	**± 7**
Informed consent		X								
Verification of entry criteria		X	X							
Physical examination[1]		X								
Review of medical history[2]		X								
Symptom directed physical examination			X	X	X	X	X	X	X	X
Vital signs[3]		X	X	X	X	X	X	X	X	X
Weight measurement		X		X	X		X		X	
Height measurement[4]		X	X	X	X	X	X	X	X	X
Electrocardiogram[5]		X		X					X	X
Urinalysis (dipstick)[6]		X			X		X		X	
Chemistry blood collection		X			X		X		X	
Hematology blood collection		X			X		X		X	
Coagulation (Prothrombin Time and PTT) blood collection		X							X	
PTH (1–84)		X							X	
25-hydroxyvitamin D level		X							X	
1,25-dihydroxy vitamin D level		X							X	
Total testosterone and SHBG		X							X	
Estradiol		X								
Thyroid stimulating hormone[7]		X								
PK samples (see Table 2)							X	X	X	
Study medication assignment via IRT				X						

Table 1: Schedule of Assessments and Procedures (Continued)

	Visit	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**
Procedure	***Study Day/Month:***	**Screening**	**Pretreatment**	**Day 1**	**Month 1**	**Month 3**	**Month 6**	**Month 9**	**Month 12/Early Termination**	**Month 13 Follow-up**
Visit Window (Days)		**N/A**	**± 4**	**± 4**	**± 4**	**± 4**	**± 4**	**± 4**	**± 4**	**± 7**
Injection training for subjects			X							
Calcium and vitamin D supplements			Daily Administration							
Study medication administration				Daily administration						
Serum markers of bone metabolism (s-PINP and s-CTX)				X	X	X	X		X	
Serum calcium and albumin[8]				X	X	X	X	X		
24-hour urine collection (for calcium:creatinine and creatinine clearance)[9]				X		X				
Radiologic (lumbar and thoracic vertebrae) assessments		X								
Symptom driven spinal radiologic assessment			At any time							
Clinical assessment of new fractures[10]			X	X	X	X	X	X	X	
BMD of lumbar spine, total hip and femoral neck by DXA[11]		X				X	X		X	
BMD of wrist by DXA[12]				X		X	X		X	
vBMD of hip by QCT[13]				X					X	
Sample for immunogenicity testing[14]				X	X	X	X	X	X	X
Investigator assessment of local tolerance (dermal reactions assessment)				X	X	X	X	X		

	Visit	1	2	3	4	5	6	7	8	9
Procedure	***Study Day/Month:***	**Screening**	**Pretreatment**	**Day 1**	**Month 1**	**Month 3**	**Month 6**	**Month 9**	**Month 12/Early Termination**	**Month 13 Follow-up**
Visit Window (Days)		**N/A**	**± 4**	**± 4**	**± 4**	**± 4**	**± 4**	**± 4**	**± 4**	**± 7**
Subject assessment of local tolerance[15]				X				X		
Subject diary review[16]				X	X	X	X	X	X	
Document AEs and concomitant[17] medications		At any time, question subjects at every visit								
Drug supply/resupply/accountability				X	X	X	X	X	X	

1.A complete physical examination includes a review of the following systems: head/neck/thyroid, eyes/ears/nose/throat, respiratory, cardiovascular, lymph nodes, abdomen, skin, musculoskeletal, and neurological. Breast, anorectal, and genital examinations will be performed when medically indicated. After screening, any clinically significant abnormal findings in physical examinations should be reported as AEs.
2. Including alcohol and tobacco use assessment.
3. Blood pressure, heart rate, body temperature, and respiration rate are to be recorded pre-dose at each study visit. Only blood pressure, heart rate and respiration rate are to be recorded one-hour post-dose at each study visit during the Treatment Period. All blood pressure assessments need to be orthostatic.
4. Height is to be measured at each visit in the standing position using a medical stadiometer.
5. ECGs are to be performed pre-dose and one-hour post-dose during the Treatment Period.
6 All routine urinalysis will be performed on a sample freshly voided during the visit and sent to a central lab for microscopy if test is positive for micro-organisms via dipstick.
7. Reflex testing of T3 and free T4 will be performed by the central laboratory for any out-of-range TSH result.
8. Serum calcium and albumin will be measured pre-dose from the standard chemistry panel on Month 1 and Month 6, and from a separate blood draw at Day 1, Month 3 and Month 9 and four hours post-dose at Day 1, Month 1, Month 3, Month 6 and Month 9.
9. A 24-hour urine collection will be collected at Day 1 and Month 3, and will be used for urinary calcium and urinary creatinine measurements. Subjects will discard the first void and begin a 24-hour urine collection on the day prior to the clinic visit.
10. If the subject reports that a fracture has occurred, remind the subject to bring X-rays and any medical reports of the fracture to the next clinic visit. Documentation must be obtained on all new fractures that occur during the study. This documentation should be maintained in the source documents.
11. Each DXA for a given subject should be performed on the same machine, and if available, preferably by the same technician. For screening purposes, DXA scans of the lumbar spine, total hip, and femoral neck taken up to 35 days prior to the beginning of the Screening Period may be used to determine study eligibility.
12. The first wrist DXA will be performed prior to study drug administration on Day 1 and should be performed on the same machine, and if available, preferably by the same technician.
13. QCT performed in a subset of subjects at selected centers.
14. Samples will be drawn prior to treatment on Day 1, Months 1, 3, 6, 9,12 and the Follow-up Visit scheduled at 1 month after the last dose of study medication.

15. The subject will maintain a diary of their assessment of local tolerance beginning on Day 2 and continuing through Day 7 in Month 1, and beginning on Day 2 and continuing through Day 7 in Month 9.
16. The subject medication diary will be reviewed by study personnel at each study visit to ensure subject compliance.
17. AEs and SAEs will be recorded on the case report forms starting from the time of subject entry into the Screening Period (Visit 1) of the study (signed informed consent) until 30 days after the last dose of study medication. All AEs will be followed until resolution or stabilization. Any SAEs that occur at any time after completion of the study, which are considered by the Investigator to be related to study treatment, must be reported to the Sponsor or its designee.

AE = adverse event; BMD = bone mineral density; DXA = dual energy X-ray absorptiometry; IRT = Interactive response technology; PTH = Parathyroid hormone; PTT = Partial thromboplastin time; QCT = quantitative CT; s-CTX = serum carboxyl-terminal cross-linking telopeptide of type I collagen; s-PINP = serum procollagen type I N-terminal propeptide; SAE = serious adverse event; SHBG = Sex hormone binding globulin; vBMD = volumetric bone mineral density

Table 2: Schedule for Pharmacokinetic Sample Collections

Study Visit	**Predose**	**Postdose**				
	Within 10 minutes before injection	20 min ± 10 min	45 min ± 15 min	1.5 hours ± 0.5 hours	2.5 hours ± 0.5 hours	4 hours ± 0.5 hours
Month 6			X		X	
Month 9		X				X
Month 12	X			X		

Note: This schedule only applies to those subjects who consent to have PK samples collected during the study.
min = minute; PK = pharmacokinetics

1.2.5 Efficacy and Safety Parameters

Efficacy Parameters

The primary efficacy endpoint is:

- Percent change from baseline in lumbar spine bone mineral density (BMD) at 12 months.

The key secondary efficacy endpoints are:

- Percent change from baseline in total hip BMD at 12 months
- Percent change from baseline in femoral neck BMD at 12 months

Additional secondary efficacy endpoints are:

- Percent change from baseline in:
  - Lumbar spine BMD at 3 and 6 months
  - Total hip BMD at 3 and 6 months
  - Femoral neck BMD at 3 and 6 months
  - Ultra-distal radius BMD at 3, 6 and 12 months
  - Distal one-third radius BMD at 3, 6 and 12 months
- Log ratio of post-baseline over baseline in:
  - Serum procollagen type I N-terminal propeptide (s-PINP) at 1, 3, 6 and 12 months
  - Serum carboxy terminal cross linking telopeptide of type I collagen (s-CTX) at 1, 3, 6 and 12 months.
- Proportion of subjects experiencing BMD gains from baseline of >0%, >3%, and >6% at the lumbar spine, femoral neck, and total hip at 3, 6, and 12 months
- Incidence of new clinical fractures at 12 months
- Proportion of subjects converting from the categories of Osteoporosis to Osteopenia and Normal at end of treatment, where:
  - Osteoporosis is defined as lumbar spine or total hip BMD T-score ≤ -2.5,
  - Osteopenia is defined as:
    - lumbar spine > - 2.5 and total hip BMD T-score > -2.5 and < -1.0, or
    - lumbar spine > - 2.5 and < - 1.0 and total hip BMD T-score > -2.5,
  - Normal is defined as lumbar spine and total hip BMD T-score ≥ -1.0.
- The percent change in total hip and femoral neck volumetric BMD as measured by quantitative CT (QCT) from baseline to 12 months. QCT will be performed in a subset of subjects at a selected number of study sites.

For those subjects who consent to PK sample collection, the PK endpoints are:

- The plasma concentration of abaloparatide based on sparse PK sampling at the following visits:
  - Month 6: post-dose PK samples collected 45 minutes (±15 minutes) and 2.5 hours (±0.5 hour) after abaloparatide-SC injection
  - Month 9: post-dose PK samples collected 20 minutes (±10 minutes) and 4 hours (±0.5 hour) after abaloparatide-SC injection

- Month 12: a pre-dose sample and a post-dose sample collected 1.5 hours (±0.5 hour) after abaloparatide-SC injection

Safety Parameters

Safety evaluations will be based on adverse events (AEs), vital signs, electrocardiograms (ECGs), laboratory evaluations (serum chemistry, hematology, urinalyses, and coagulation), investigator assessment of local tolerance, subject assessment of local tolerance, and anti-drug antibody assessments.

In addition, the presence of hypercalcemia will be assessed. Hypercalcemia is defined as a serum calcium (albumin-corrected) value ≥ 0.3 mg/dL (or ≥ 0.08 mmol/L) above the upper limit of normal. The difference in the percentage of subjects with one or more incidents of hypercalcemia over the 12-month treatment period will be assessed.

The presence of hypercalciuria will also be assessed. Subjects with a urine calcium/creatinine ratio > 0.4 mg/mg (or > 1.131 mmol/mmol) and a urine calcium/creatinine ratio > 0.3 mg/mg (or > 0.848 mmol/mmol) will be tabulated by treatment group.

2 ANALYSIS POPULATIONS

2.1 Population Definitions

Four subject populations will be considered in the statistical analyses of this study.

Intention-to-treat (ITT) population

The Intention-to-Treat (ITT) population is defined as all subjects randomized into this study.

Safety population

The Safety population is defined as all subjects who received at least one dose of study medication.

Antibody Population

The antibody population will include those abaloparatide-SC treated subjects in the Safety population who have at least one post-baseline anti-abaloparatide antibody assessment.

Per-Protocol (PP) population

The Per-Protocol (PP) population will include all ITT subjects who did not have any significant protocol deviations (PDs) (described in Section 2.2).

The PP population will be a supportive population for the efficacy analyses.

2.2 Protocol Deviation

A protocol deviation (PD) is any change, divergence, or departure from the study design or procedures defined in the protocol. The PD categories related to study conduct are described in the Management of Protocol Deviations/Waivers section of the Medical Management Plan. PDs are recorded by the contract research organization (ICON) in their Clinical Trial Management System (CTMS). Subjects with at least one protocol deviation will be presented in a data listing, with indicators for key and non-key PDs (as defined in Appendix 1 of the Medical Management Plan, where a key PD is a deviation requiring reporting to the Functional Lead, Medical Monitor and or Sponsor). A summary table of the key PDs identified during the study will be presented.

Significant PDs are PDs that have the potential to impact the efficacy of the treatment on the subject.

The algorithm to determine significant PDs is in Appendix A of this SAP. Subjects with at least one significant PD will be excluded from the PP population. A Per-Protocol Memo with the algorithm in Appendix A together with the programming specification ("Per Protocol Population Data Set Development") will be developed and finalized prior to database lock to detail how each of the criteria in the algorithm can be programmed in SAS. The list of subjects with significant protocol deviations identified after applying the algorithm to the unblinded SDTM data after database lock will be reviewed and approved internally at Radius in a separate memo. A summary table of reasons for exclusion from the PP population will be presented. A data listing for subjects excluded from the PP population together with the exclusion reason(s) will be presented.

3 GENERAL STATISTICAL METHODS

3.1 Sample Size Planned and Specified in the Protocol

A study sample size of 225 subjects (150 in the abaloparatide-SC group and 75 in the placebo group) will provide at least 99% power to detect a mean difference of 6.5% in the percentage change from baseline in lumbar spine BMD at 12 months between the abaloparatide-SC and the placebo groups at a two-sided alpha level of 0.01, assuming a standard deviation (SD) of 6.0% and a drop-out rate of 10%. For the key secondary endpoints, a total sample size of 225 subjects will have 96% power to detect a 2.2% mean difference (SD=3.5%) of percent change in total hip BMD at 12 months, and 80% power to detect a 2.0% mean difference (SD=4.1%) of percent change in femoral neck BMD at 12 months.

3.2 General Methods

Summary tables will be presented by treatment group (abaloparatide-SC and placebo) and overall, when appropriate. For categorical data, summary tabulations of the number and percentage of subjects within each category of the parameter will be presented. When tabulating categorical data, "missing" will be included as a category and the number of subjects with missing data will be presented. For continuous data, the number of subjects, mean, median, standard deviation (SD), minimum, interquartile range (Q1 and Q3), and maximum will be presented.

All data listings that contain an evaluation date will contain a relative study day. Screening, pre-treatment, and on-treatment study days will be numbered relative to the day of the first dose of study medication which is designated as Day 1. Study days prior to the first dose of study drug will be calculated as: [date of assessment – date of first dose of study drug]. Study days on or after the first dose of study drug will be calculated as: [date of assessment – date of first dose of study drug + 1].

Where applicable, confidence intervals about a parameter estimate will be presented using the same number of decimal places as the parameter estimate and p-values will be rounded to 4 decimal places prior to assessment of statistical significance.

If the date of last dose is missing, then the earlier of the following 2 dates will be used as the last dose date for analysis and summary purposes: the final visit/early termination visit date or the last study drug dispense date + 30 days.

For subjects in the abaloparatide-SC group, the serum samples for immunogenicity testing will be analyzed and reported separately. All the planned efficacy and safety analyses related to immunogenicity results and antibody status described in this SAP will be reported separately.

This study is being conducted during the COVID-19 pandemic. The COVID-19 outbreak emerged in December 2019 quickly became a global pandemic as declared by the World Health Organization in early March 2020. At that time, a variety of mitigations to assure safety of participants and address operational issues were implemented by Radius. To investigate the effect of the COVID-19 pandemic on study conduct, additional analyses comparing the subjects who completed or withdrew from the study before the onset of the COVID-19 pandemic, i.e, before 01 March 2020, with the subjects who were ongoing or randomized during the COVID-19 pandemic, i.e, on or after 01 March 2020, will be conducted where applicable.

3.3 Computing Environment

All descriptive and statistical analyses will be performed using SAS statistical software Version 9.4, or later version. Medical history and AEs will be coded using the Medical Dictionary for Regulatory Activities (MedDRA), Version 24.0. Concomitant medications will be coded using the World Health Organization Drug Dictionary (WHO DRUG DDE+HD/2017-DEC-B3).

3.4 Baseline Definitions

The baseline value is defined as the last non-missing value obtained prior to the first dose of study medication, including the screening value, if necessary. If a randomized subject is not treated, the baseline value is defined as the last non-missing assessment.

3.5 Data Pooling

Data from all study sites and countries will be pooled together for all analyses, unless otherwise specified.

3.6 Adjustments for Covariates

Adjustments for covariates will be employed in the Analysis of Covariance (ANCOVA) and Mixed-Effect Model for Repeated Measures (MMRM) analyses. Any adjustments will be pre-specified in the description of the statistical analysis (Section 4).

3.7 Fixed-Sequence Tests of the Primary and Key Secondary Efficacy Endpoints for Multiple Comparisons

The primary and key secondary efficacy endpoints will be analyzed using a fixed-sequence testing procedure to control the family-wise error rate for any joint distribution of hypothesis test statistics. To claim statistical significance at the 2-sided level of 1%, the following 3 fixed-sequence tests will be performed in sequential order. At any step of the sequential testing, if the treatment difference is not statistically significant at the 1% level then all the subsequent comparisons following the fixed sequence cannot be claimed statistically significant.

1. Percent change from baseline in BMD at the lumbar spine at 12 months
2. Percent change from baseline in BMD at the total hip at 12 months
3. Percent change from baseline in BMD at the femoral neck at 12 months

P-values for treatment comparisons of all other efficacy endpoints will be generated to support the study findings without any adjustments for multiplicity.

3.8 Subgroups

The following subgroups will be created for possible use in selected data analyses. Section 4.2 will describe the specific analyses to be performed for each subgroup.

- Age (<65, 65 - <75, ≥75)
- Race (White, Black or African-American, Asian, and Other Races [i.e., American Indian or Alaska Native, Native Hawaiian or Other Pacific Islander, Other, and Multiple (more than 1 race category chosen)])
- BMI (< 25, 25 - < 30, ≥ 30)

- Region (North America including USA and Canada, Europe including Poland and Italy)
- Prior clinical fracture (Yes = Subjects with a record in medical history with "fracture" in the preferred term, No)
- Prevalent vertebral fracture (Yes = Subjects with any vertebral fracture with a score> 0, No)
- Prior fracture (Yes =Subjects with a prevalent vertebral fracture and / or a prior clinical fracture, No)
- Smoking status (current, former, never)
- Lumbar spine BMD T-score at baseline (≤ -2.5, > -2.5)
- Total hip BMD T-score at baseline (≤ -2.5, > -2.5)
- Femoral neck BMD T-score at baseline (≤ -2.5, > -2.5)
- s-PINP at baseline (tertile values)
- Total testosterone level (<9.36 nmol/L, ≥9.36 nmol/L)
- Estradiol level (≤73.4 pmol/L, >73.4 pmol/L)

3.9 Anti-Abaloparatide Antibody Status

At Day 1, Month 1, Month 3, Month 6, Month 9, Month 12/EOT, and Month 13/EOS, for subjects in the abaloparatide-SC group, the anti-abaloparatide antibody (ADA) assessment in the serum samples for immunogenicity testing will be categorized by two outcomes:

- Negative for the ADA (ADA-)
- Positive for the ADA (ADA+)

For those subjects with a positive ADA (i.e, ADA is present), the serum samples will be further assessed for the presence of neutralizing anti-abaloparatide antibodies (NAb), and the results will be categorized by two outcomes:

- Negative for the NAb (NAb-)
- Positive for the NAb (NAb+)

Subjects' ADA status and presence of NAb by Month 12 (ADA-, ADA+/NAb+, ADA+/NAb-, All ADA+) will be defined as follows:

- Subjects will be classified as ADA- if they were ADA- in all post-baseline immunogenicity assessments during the 12 months treatment period.
- Subjects will be classified as ADA+ and NAb+ if they were ADA+ at any post-baseline immunogenicity assessment and NAb+ at any subsequent tests during the 12 months treatment period.
- Subjects will be classified as ADA+ and NAb- if they were ADA+ at any post-baseline immunogenicity assessment and NAb- at all subsequent tests during the 12 months treatment period.

3.10 Withdrawals, Dropouts, Loss to Follow-up

Subjects who are withdrawn or discontinue from the study will not be replaced.

3.11 Missing, Unused, and Spurious Data

Statistical methods for addressing missing data for the analyses of the primary and secondary efficacy endpoints are described in Section 4.3.

Rules for imputation of missing dates for onset of AEs and start of concomitant medications are included in Appendix D of this SAP. Unless otherwise specified, missing dates for onset of AEs and start of concomitant medications will be handled according to those imputation rules.

In the event that incorrect data are discovered after database lock, the impact of such incorrect data will be assessed and a determination will be made as to whether the data need to be corrected in the database (and subsequently in the analysis). In cases where it is decided that the data need to be corrected, the decision will be documented, and the documentation will be saved in the Trial Master File.

All available data will be included in data listings.

3.12 Visit Windows

In general, the nominal visit will be used for the safety and efficacy data summaries and analyses.

During the COVID-19 pandemic, if a subject's DXA and/or QCT assessment cannot be performed during the protocol specified visit windows, the site will contact the study Medical Monitor and the subject will be allowed to make up the visits within an allowed time period. The allowable time period for collection of DXA and QCT assessments are shown in Table 3, where the ideal visit date is defined as the date that represents the ideal number of days from the date of first dose of study drug to the nominal visit.

Table 3: Allowable Time Period for Collection of DXA and QCT Assessments.

Visit	Ideal Visit Day	BMD or vBMD Assessment
Month 3	90	Up to 30 days after ideal visit day
Month 6	180	Up to 90 days after ideal visit day
Month 12	360	Up to 30 days after ideal visit day

Due to the expanded time periods for collection of efficacy assessments, analysis visit windows are defined below for the analyses of BMD, vBMD, and serum markers of bone metabolism (PINP and CTX) data to determine the assessment that will be used for a given study visit.

Table 4: Analysis Visit Windows for BMD, vBMD, s-PINP and s-CTX.

Visit / Month	Allowable days from first dose of study drug		
	BMD lumbar spine, total hip and femoral neck, BMD Wrist	s-PINP, s-CTX	vBMD
Visit 3 / Baseline	≤ 1	≤ 1	≤ 1
Visit 4 / Month 1		2—60	
Visit 5 / Month 3	2—120	61—120	
Visit 6 / Month 6	121—270	121—270	
Visit 7 / Month 9			
Visit 8 / Month 12	≥ 271	≥ 271	≥2

It is possible that more than 1 assessment for a parameter may be obtained within the same analysis visit window. In such circumstances, the assessment with the date closest to the ideal visit day will be used. In the event that 2 assessments are equidistant to the ideal visit day, the later of the assessments will be used.

The analysis visit window convention will not be applied to the data listings. The data listings will display all data as collected with the corresponding nominal visit.

3.13 Interim Efficacy Analyses

No interim analysis for efficacy is planned for this study.

4 STUDY ANALYSES

4.1 Subject Disposition

Subject disposition will be tabulated by treatment group including the number screened, the number randomized, the number treated, the number in each analysis population, the number who completed the study, the number who discontinued from the study treatment together with the primary reason for discontinuation, and the number who discontinued from the study together with the primary reason for discontinuation as recorded in the eCRF. The number of subjects excluded from the PP population together with the exclusion reasons will also be summarized.

A Kaplan-Meier plot of the time from randomization to withdrawal from study will be presented by treatment. Subjects who complete the study will be censored at the date of their study completion.

The total number of subjects randomized will be tabulated by region (North America, Europe), country, study site, and treatment group.

Subject disposition data will be carefully reviewed with respect to the impact of the COVID-19 pandemic on study conduct. Significant protocol deviations and study/treatment discontinuation related to the COVID-19 pandemic will be identified during the study. A summary of subject disposition by subject's randomization date relative to COVID-19 pandemic onset (subjects randomized prior to 01 March 2020 vs subjects randomized on or after 01 March 2020) will be presented. A summary of subject disposition based on subject's study status relative to COVID-19 pandemic onset (subjects who completed or withdrew from the study before 01 March 2020 vs subjects who completed or withdrew from the study on or after 01 March 2020) will also be presented. The number (%) of subjects with assessments impacted by COVID-19 pandemic as recorded on the COVID-19 Impact Log eCRF page will be summarized by visit and the following categories:

- Altered Protocol Assessment/Procedure
- Discontinuation
- IP – Missed/Modified
- Missing Data
- Missed Visit
- Out of Window Visit
- Remote /Virtual Visit
- Assessment Not Done
- Other

A listing of subjects with assessments impacted by COVID-19 pandemic as indicated on the COVID-19 Impact Log eCRF page will be provided. The listing will include subject ID, age, race, site, visit, CRF form used, impact category, and comment.

A by-subject data listing of study completion information, including the primary reasons for premature study withdrawal as recorded in the eCRF, will be presented. A by-subject data listing of subject excluded from PP population and the reasons for exclusion will also be presented.

A summary of screening failures and listings of inclusion/exclusion criteria responses for screening failure subjects will also be presented.

4.2 Demographics and Baseline Characteristics

Demographic and baseline characteristics will be summarized and presented by treatment group and overall. Age will be calculated as of the date the subject provided signed informed consent for this study. Age, age group (<65, 65-<75, ≥75), height, weight, body mass index (BMI), and BMI group (BMI < 25, 25 ≤ BMI < 30, BMI ≥ 30) will be summarized using descriptive statistics (number of subjects, mean, SD, median, interquartile range, minimum and maximum). The number of subjects in each ethnicity and race category will also be presented. Subjects who report more than one race category will be counted under the race category "Multiple". The usage of alcohol, tobacco, and other substances at baseline will be summarized by treatment group.

The number and percentage of subjects with the following fracture history will be summarized:

- Prior clinical fracture: Subjects with a record in medical history with "fracture" in the preferred term.
- Prevalent vertebral fracture: Subjects with any vertebral fracture with a score > 0.
- Prior fracture: Subjects with a prevalent vertebral fracture and / or a prior clinical fracture
- Prior clinical vertebral fracture: Subjects with a record in medical history with "vertebral fracture" or "spinal fracture" in the preferred term
- Prior vertebral fracture: Subjects with a prior clinical vertebral fracture or a prevalent vertebral fracture.

Baseline values for BMD measurements and biomarker parameters including lumbar spine BMD T-score and T-score group (≤ -2.5, > -2.5), total hip BMD T-score and T-score group (≤ -2.5, > -2.5), femoral neck BMD T-score and T-score group (≤ -2.5, > -2.5), s-PINP and s-PINP group (tertile values), s-CTX, total testosterone level and total testosterone level group (<9.36 nmol/L, ≥9.36 nmol/L), and estradiol level and estradiol level group (<73.4 pmol/L, ≥73.4 pmol/L) will also be summarized for each treatment group and overall.

BMD measurements via Quantitative Computed Tomography (QCT) will be collected in a subset of subjects at a selected number of centers. The summary of demographics and baseline characteristics will be repeated for this subset of subjects.

Medical history will be presented by MedDRA system organ class (SOC), preferred term (PT), and treatment group, summarizing the proportion of subjects in each treatment group who have a condition noted.

Results from the baseline physical examination will be summarized by body system as recorded in the eCRF and treatment group.

All analyses described in this section will be performed on the ITT population, and repeated for the Safety and PP populations.

All analyses described in this section will also be repeated for the Antibody population. The presentation of these analyses will be based on subjects' ADA status and presence of neutralizing

antibody (NAb) during the study with columns for ADA-, ADA+/NAb+, ADA+/NAb-, All ADA+, and Overall.

Additionally, demographics data and baseline characteristics will be tabulated by all subgroups listed in Section 3.8 using the ITT population. To address the effect of the COVID-19 pandemic, summaries of demographic data and baseline characteristics by subject's randomization date relative to COVID-19 pandemic onset (subjects randomized prior to 01 March 2020 vs subjects randomized on or after 01 March 2020) will be presented. Summaries of demographic data and baseline characteristics based on subject's study status relative to COVID-19 pandemic onset (subjects who completed or withdrew from the study before 01 March 2020 vs subjects who completed or withdrew from the study on or after 01 March 2020) will also be presented.

Demographics including gender, race, ethnicity, age, and country for screen failure subjects will also be summarized descriptively.

All subject data collected for demographics and baseline characteristics will be presented by treatment, study site, and subject number in data listings.

4.3 Efficacy Analyses

The ITT population will be the primary population for all efficacy analyses. The PP population will be used for supportive analyses. In efficacy analyses, subjects will be analyzed according to the treatment to which they were randomized.

The primary and BMD-related secondary efficacy endpoints will be derived from the Bioclinica-corrected BMD measurements.

BMD scans may be deemed unevaluable at the first reading and a repeat BMD scan may be performed. The re-scan BMD assessment will be recorded as an unscheduled visit for the corresponding nominal visit. To determine the assessment that will be used for a given study visit by the analysis window convention, all BMD assessments (including unscheduled BMD assessments) will be considered. In the efficacy analyses by nominal visit, unscheduled BMD assessments will be used if the BMD assessment at the nominal visit is not available but an unscheduled BMD assessment for the same nominal visit is available.

Analyses of the BMD-related efficacy endpoints and the serum markers of bone metabolism (s-PINP and s-CTX) will be conducted using the analysis visit windows defined in Table 4 of Section 3.12 as the primary analysis. Analyses of the BMD-related efficacy endpoints and the serum markers of bone metabolism (s-PINP and s-CTX) using the nominal visits will also be presented as supportive analyses.

All confidence intervals, statistical tests, and p-values will be reported as nominal 2-sided. The primary and key secondary efficacy endpoints will be assessed at the significance level of 0.01. All other efficacy endpoints will be assessed at a significance level of 0.05.

4.3.1 Primary Efficacy Endpoint

The primary efficacy endpoint is the percent change from baseline in lumbar spine BMD at 12 months. The mean of percent changes from baseline in lumbar spine BMD to 12 months will be estimated for each treatment group, and the abaloparatide-SC will be compared to placebo using

difference in group means. The primary comparison of abaloparatide-SC with placebo will be based on the ITT population.

The missing data from this study will be reviewed carefully to evaluate any missing data patterns. As this study is not designed to systematically collect data from subjects who discontinue treatment early, the "retrieved dropouts" approach for handling missing data is not applicable to this study. To appropriately address the possibility of data missing not at random, the primary analysis will be based on a pattern mixture model (PMM) analysis with multiple imputation using the wash-out imputation method. This method uses sequential regression and wash-out imputation methodology to impute missing values after a subject's discontinuation from the study. The missing primary endpoint values for subjects in the abaloparatide-SC group will be imputed with observed baseline and data (including imputed values) from the placebo group; no intermediate values from the abaloparatide-SC group will be used in the imputation for the abaloparatide-SC group. For subjects in the placebo group, intermediate observed values from the completers in the placebo group will be used while imputing missing values during the 12-month treatment period. The PROC MI methodology for imputation of monotone missing data patterns will be used to impute the outcome variables at consecutive visits in a sequential (chain) manner. Each of these imputed datasets will be analyzed using analysis of covariance (ANCOVA) model with treatment and DXA instrument manufacturer as fixed effects, and the baseline lumbar spine BMD as a covariate. Results from all imputed datasets will be combined using PROC MIANALYZE for overall statistical inference. The statistical test will be 2-sided comparing abaloparatide-SC to placebo; Both 99% and 95% confidence intervals (CIs) will be presented together with the estimated p-values. The detailed imputation and analysis steps are described below:

1. Intermittent missing data in the placebo treatment group will be imputed using MCMC methods, assuming MAR. SAS PROC MI will be utilized for this step using the MCMC impute=monotone option. A total of 100 datasets will be created. These datasets will be utilized in Step #2.
2. The remaining missing values in the placebo group with a monotone missing data pattern will be imputed in this step. Missing data will be imputed assuming data are MAR. Only data (both observed and imputed) from subjects in the placebo group will be utilized in this step. SAS PROC MI will be used to impute missing values utilizing the monotone reg option. This will be performed for the 100 datasets. After this step, the 100 datasets will be fully imputed for the placebo treatment group.
3. The missing values at Month 12 in the abaloparatide-SC group will be imputed in this step. Control-based PMM imputation will be performed. With this imputation model, the missing values in the abaloparatide-SC group will not be constructed from the observed data in the abaloparatide-SC group but rather from the observed baseline and data (including imputed values) from the placebo group at Month 12. The MNAR statement in SAS PROC MI will be used to impute missing values. This will be performed for the 100 datasets.
4. A total of 100 fully imputed datasets will be created for both treatment groups. Since multiple imputation is a stochastic method, slight differences in output can be expected for different initial states of the random number generator. The seed numbers will be identified in the SAS programs to allow for reproducibility.

5. After the missing data imputation is completed using the above steps, percentage change values will be calculated in each of the imputed datasets at each visit.
6. These 100 datasets will be analyzed using ANCOVA models with fixed effect of treatment group and DXA instrument manufacturer, and baseline lumbar spine BMD as a covariate for the primary efficacy endpoint.
7. Treatment effects (difference in LS means between treatments) from these 100 analyses will then be combined using Rubin's Method via SAS PROC MIANALYZE procedure for the primary endpoint.

As a sensitivity analysis, the primary efficacy endpoint will also be analyzed using a Mixed Model for Repeated Measures (MMRM) analysis. The MMRM model will include the fixed effects of treatment, DXA instrument manufacturer, visit, and treatment-by-visit interaction, with the baseline lumbar spine BMD as covariate. Imputation for missing data is not necessary for the MMRM analysis because the model uses a restricted maximum likelihood (REML) based on the repeated-measures approach. For the MMRM model, an unstructured variance-covariance matrix shared between the two treatment groups will be used to model the within-subject errors over the visits. The Kenward-Roger approximation will be used to estimate denominator degrees of freedom and to adjust standard errors (Kenward, 1997). Analyses will be implemented using the SAS PROC MIXED procedure. The treatment comparison will be obtained by testing the contrast (difference in least squares mean) between the two treatment groups at Month 12 (Visit 8). The statistical test will be 2-sided comparing abaloparatide-SC to placebo; 99% CIs will be presented together with the estimated p-values.

As supportive analysis, the primary efficacy endpoint will also be analyzed using an ANCOVA model using the last observation carried forward (LOCF) method. The ANCOVA model will include treatment and DXA instrument manufacturer as fixed effects, and the baseline lumbar spine BMD as covariate. To test for statistically significant differences between the treatment groups, the treatment comparison will be derived by testing the contrast (difference in least squares mean) between the two treatment groups at the Month 12 (Visit 8) visit. Missing data imputation based on the LOCF method will be performed as follows:

- If baseline BMD is missing, no imputation for post-baseline value will be done and this subject's data will be excluded from the analysis.
- If baseline BMD is observed, missing post-baseline BMD will be imputed by the last observed value prior to the missing assessment. That is, if a value is missing at a post-baseline time point, the post-baseline BMD will be imputed using the last value observed at any time (could be a value obtained at an unscheduled visit) between baseline (inclusive) and the specified post-baseline time point.

The primary analysis using the PMM with the wash-out imputation method, and the other analyses using the MMRM model and the ANCOVA model using LOCF method, will be repeated using the PP population to support the findings from the ITT population.

For the primary efficacy endpoint using PMM analysis with wash-out imputation, subgroup analyses for each of the subgroups defined in Section 3.8 will be performed. The treatment groups will be compared within each subgroup. A forest plot of the estimated least square (LS) mean differences between the treatment groups (and corresponding 99% CIs) will be presented

by subgroup. Additional exploratory analyses examining subgroup effect and treatment by subgroup interaction may be performed, if warranted.

The efficacy analysis using MMRM model will also be repeated for the Antibody population, in which the subjects who are ADA+ at any post-baseline time point will be compared with subjects who are ADA- at all post-baseline time points. The MMRM model will include the fixed effects of ADA status, DXA instrument manufacturer, visit, and ADA status-by-visit interaction, with the baseline lumbar spine BMD as covariate.

To investigate the effect of the COVID-19 pandemic, the following sensitivity analyses will also be performed:

a) PMM analysis with wash-out imputation for subjects who completed or withdrew from the study before 01 March 2020 and for subjects who completed or withdrew from the study on or after 01 March 2020. The analysis visit windows described in Table 4 will be used.

b) PMM analysis with wash-out imputation using the nominal visit.

c) PMM analysis with wash-out imputation excluding BMD measurements collected outside the allowable time period for data collection. In this analysis, BMD assessments are allocated to a Visit based on the inclusion rules shown in Table 5 below re: Sensitivity Analysis c).

d) PMM analysis with wash-out imputation excluding BMD measurements which are more than 30 days after the ideal visit day. In this analysis, BMD assessments are allocated to a Visit based on the inclusion rules shown in Table 5 below re: Sensitivity Analysis d).

Table 5: Visit Windows for DXA Assessments for Sensitivity Analyses.

Visit	Ideal Visit Day	Sensitivity Analysis c): Allowable days from first dose of study drug	Sensitivity Analysis d): Allowable days from first dose of study drug
Month 3	90	2 —120 days	2 —120 days
Month 6	180	121—270 days	121—210 days
Month 12	360	271—390 days	271—390 days

It is possible that more than 1 assessment for a parameter may be obtained within the same visit window. In such circumstances, the assessment with the date closest to the ideal visit day will be used. In the event that 2 assessments are equidistant to the ideal visit day, the later of the assessments will be used.

4.3.2 Secondary Efficacy Endpoints

The key secondary efficacy endpoints are:

- Percent change from baseline in total hip BMD at 12 months
- Percent change from baseline in femoral neck BMD at 12 months.

Additional secondary efficacy endpoints are:

- Percent change from baseline in
  - Lumbar spine BMD at 3 and 6 months

  - Total hip BMD at 3 and 6 months
  - Femoral neck BMD at 3 and 6 months
  - Ultra-distal radius BMD at 3, 6 and 12 months
  - Distal one-third radius BMD at 3, 6 and 12 months.
- Log ratio of post-baseline over baseline in
  - s-PINP at 1, 3, 6 and 12 months
  - s-CTX at 1, 3, 6 and 12 months.
- Proportion of subjects experiencing BMD gains from baseline of >0%, >3%, and >6% at the lumbar spine, femoral neck, and total hip at 3, 6, and 12 months
- Incidence of new clinical fractures at 12 months
- Proportion of subjects converting from the categories of Osteoporosis to Osteopenia and Normal at end of treatment, where:
  - Osteoporosis is defined as lumbar spine or total hip BMD T-score ≤ -2.5,
  - Osteopenia is defined as:
    - lumbar spine > - 2.5 and total hip BMD T-score > -2.5 and < -1.0, or
    - lumbar spine > - 2.5 and < - 1.0 and total hip BMD T-score > -2.5,
  - Normal is defined as lumbar spine and total hip BMD T-score ≥ -1.0.
- The percent change in total hip and femoral neck volumetric BMD as measured by quantitative CT (QCT) from baseline to 12 months. QCT will be performed in a subset of subjects at a selected number of study sites.

Analyses of Key Secondary Efficacy Endpoints:

The primary population for the analysis of the key secondary efficacy endpoints will be the ITT population. The analyses will be performed using PMM analysis with wash-out imputation similar to the analysis model for the primary efficacy endpoint, with the appropriate baseline covariate.

Analyses of the primary and key secondary efficacy endpoints will follow a fixed-sequence testing approach to control the overall Type 1 error rate at the 2-sided significance level of 1%, as described in Section 3.7.

Sensitivity analyses using a MMRM model, sensitivity analyses for addressing the effect of the COVID-19 pandemic, and supportive analyses using LOCF method, as with the primary efficacy endpoint, will also be conducted. For the estimated treatment effects of the key secondary endpoints, 99% CIs will be presented together with the estimated p-values.

The analyses of each of the key secondary efficacy endpoints using the PMM method with wash-out imputation, the MMRM model, and the ANCOVA model using LOCF method will be repeated using the PP population to support the findings from the ITT population.

Subgroup analyses using PMM analysis with wash-out imputation will be performed on the key secondary efficacy endpoints using the ITT population for each of the subgroups defined in Section 3.8. The analyses will use analysis models as described above using the subset of ITT subjects belonging to the specific subgroup. A forest plot of the estimated least square (LS) mean

differences between the treatment groups (and corresponding 99% CIs) will be presented by subgroup for each of the key secondary efficacy endpoints.

The efficacy analysis using MMRM model for the key secondary endpoints will also be repeated for the Antibody population, in which the subjects who are ADA+ at any post-baseline time point will be compared with subjects who are ADA- at all post-baseline time points.

Analyses of Additional Secondary Efficacy Endpoints:

All analyses performed for the primary efficacy endpoint for the ITT and PP populations will also be performed for all additional secondary efficacy endpoints involving the percent change from baseline in BMD parameters. The primary analyses will use the PMM analysis with wash-out imputation similar to the analysis model for the primary efficacy endpoint, with the appropriate baseline covariate. Sensitivity analyses, subgroup analyses, and other supportive analyses will be performed as described in Section 4.3.1; however, a two-sided alpha level of 0.05 will be used. For these analyses, 95% CIs for the estimated treatment effects will be presented along with the estimated p-values.

While no forest plot will be generated for the additional endpoints, LS mean (±SE) percent changes from baseline for lumbar spine, total hip, and femoral neck will be plotted by treatment group over time using the ITT population.

Percent (%) changes from baseline in BMD will be summarized using descriptive statistics by visit (Month 3, 6, and 12) for each treatment group. LS Mean (±SE) percent changes from baseline BMD will be plotted by treatment group over time using the ITT population.

The primary population for the analysis of the secondary efficacy endpoints involving the bone formation marker, s-PINP, and the bone resorption marker, s-CTX, will be the ITT population. These parameters will be analyzed based on the ratio of the post-baseline value relative to the baseline value at each visit (using analysis visit windows described in Table 4). The transformation of the $\log_e$ ratio of post-baseline versus baseline value (derived by dividing the post-baseline value by the baseline value and then applying the natural log transformation) will be used to normalize the distributions of the s-PINP and s-CTX parameters. The analysis comparing abaloparatide-SC with placebo will use PMM analysis with wash-out imputation as described in Section 4.3.1. Because their distribution violates normality assumptions, the bone marker data will be transformed using natural log transformation before imputation; and the observed and imputed data will be back-transformed after imputation. The relative treatment effect, defined as the exponential of the least squares mean difference in $\log_e$ ratio between the two treatment groups at each visit will be presented along with the estimated p-values and 95% confidence intervals. Sensitivity analyses using a MMRM model, as with the primary efficacy endpoint, will also be conducted. Any value of s-PINP or s-CTX that is provided as "< *numeric value*" or "> *numeric value*" will be imputed as the *numeric value* in all summary tables and analyses; but will be shown with the actual result in the listings.

Geometric mean values, geometric mean values (±SE) relative to baseline, mean (±SD) change from baseline values, and mean (±SD) percent change from baseline values at each visit for s-PINP and s-CTX will be tabulated using descriptive statistics. Geometric mean (±SE) of the ratio relative to baseline as well as mean (±SE) percent changes from baseline for s-PINP and s-CTX will be plotted by treatment group over time using the ITT population. Plots will also be

presented for the median (± interquartile range) for the raw values of s-PINP and s-CTX over time by treatment group.

A BMD responder analysis will be performed for exploratory purposes. This analysis will be based on the ITT population. Three categories of BMD response will be evaluated: % BMD increase >0%, >3%, and >6%. The by-visit responses (using analysis visit windows described in Table 4) will be assessed at the lumbar spine, total hip, and femoral neck. For the "all sites" analysis, subjects who have the pre-specified percentage BMD increase at all three anatomical sites at the same visit will be considered responders. A subject must have non-missing BMD results at all three sites at the same visit to be included in the "all sites" analysis. No imputation of missing data will be implemented. For the "all sites" analysis, the BMD response status at end of study will also be summarized. In this analysis, the "all sites" response status will be based on the subject's last BMD assessment during the study. If the last BMD assessment is missing the result for one or two anatomical site(s), the subject will be classified as a non-responder at End of Study. The denominator for the percentage of BMD responders at End of Study will be the number of subjects in the ITT population who had both a baseline and a post-baseline BMD.

BMD responder analyses will also be performed at each of the three anatomical sites separately, using the same three categories of BMD response used in the "all sites" analysis. A subject must have non-missing BMD results at a given visit to be included in the by-visit analysis. No imputation of missing BMD data will be implemented.

The primary time point for all BMD responder analyses is Month 12. The chi-square test will be used to explore the difference in the number (%) of responders between the two treatment groups at each visit (Months 3, 6 and 12) for each category of response. If the number of responders is less than five in any treatment group, the Fisher's exact test will be used. No multiplicity adjustment to the p-values for the BMD responder analysis will be implemented.

Analysis of the time to new clinical fracture during the study will be performed for exploratory purposes. Duration in days from the date of randomization to the first incidence of clinical fracture will be derived. The crude incidence rates of new clinical fracture, Kaplan-Meier estimates of the incidence rate at Month 12, hazard ratio (95% CI) based on a Cox proportional hazard model, Kaplan-Meier curves, and p-values from the log-rank test for the comparison of the treatment groups will be generated using the ITT and PP populations for this time-to-event variable. If a subject does not experience any clinical fracture over the 12 months of treatment plus the 30-day follow-up, this subject will be censored at the last known day in the study up to the follow-up visit.

The number and percentage of subjects for whom disease status has changed from osteoporosis at baseline to osteopenia or to normal (Refer to Section 4.3.2 for the definitions of osteoporosis, osteopenia, and normal), or no change at end of treatment will be tabulated; and will be analyzed using the Cochran-Mantel-Haenszel **(**CMH) test stratified by baseline disease status for the ITT population. No imputation of missing BMD data will be implemented.

The percent change from baseline in total hip and femoral neck volumetric BMD (vBMD) at Month 12, as measured by quantitative computed tomography (QCT), will be analyzed using an ANCOVA model with fixed effects of treatment and DXA instrument manufacturer, and with the relevant vBMD baseline value as covariate for the ITT population. No imputation of missing vBMD data will be implemented.

4.4 Population PK Analyses

PK data will be analyzed and reported separately. Details of the PK data summaries and analysis methods will be provided in a separate analysis plan.

4.5 Safety Analyses

The primary population for all safety analyses will be the Safety population. In safety analyses, treatment classification will be based on the actual treatment received. Unless otherwise specified, no formal statistical hypothesis testing will be performed for safety endpoints, and the presentations will be based on descriptive summaries.

The impact of the COVID-19 pandemic and the implementation of a variety of mitigations for study conduct will be reviewed and evaluated carefully for the safety assessment. The pandemic-related protocol deviations including missed or delayed visits, missed safety-related tests, use of alternative assessment procedures for visit execution, use of alternative data collection or alternative modalities, use of rescue medications, or non-adherence to the planned course of treatment will be identified. Additional data summaries and analyses addressing the impact of the pandemic may be performed if necessary.

4.5.1 Study Drug Exposure

All doses of study medication are to be self-administered or administered by an individual trained in giving the injection. Study personnel may administer the injection on days of clinic visits. The overall duration of study drug exposure will be calculated using the data recorded on the eCRF as follows:

- Overall duration of study drug exposure (days**)** = (date of last dose – date of first dose) + 1.

The total number of doses delivered and study drug administration compliance will be calculated in two ways as follows:

- **Based on missed doses in subject diary:**
  - Number of doses delivered = Overall duration of study drug exposure – Total number of missed doses per subject diary
  - Percent compliance (%) to study drug administration = (Number of doses delivered / Overall duration of study drug exposure) x 100%
- **Based on measurement of drug remaining in pen:**
  - Number of doses delivered = [(45.5 mm × Number of used pens) - Sum of measurements of drug remaining in all used pens)] /1.08 mm. Here the 45.5mm is the initial measurement of unused study drug in an unused injection pen and 1.08mm is the distance the plunger travels for each 80 ug dose from the pen.
    - Note: If the measurement of drug remaining in any used pen is missing, then 100% compliance will be assumed for the treatment period for which the pen with missing measurement of drug remaining was dispensed. The number of doses taken from the pen with missing measurement of drug remaining is a value less than or equal to 30, which will be based on the number of doses taken from the pen(s) with non-missing measurement of drug remaining dispensed at the same visit as the pen with missing

measurement of drug remaining. The number of doses delivered would be the minimum of the following:

  - Total doses taken from pens with non-missing measurements of drug remaining (calculated as: [45.5 mm × Number of pens with non-missing measurements of drug remaining - Sum of measurements of drug remaining from all pens with non-missing measurements of drug remaining] /1.08 mm) + 30 × number of pens with missing measurement of drug remaining
  - Duration of the treatment period for which the pen with missing measurement of drug remaining was dispensed.

- Percent compliance (%) to study drug administration = (Number of doses delivered / Overall duration of study drug exposure) x 100%

The duration of study drug exposure, duration category in months (≤ 1 month, 1< to ≤ 3 months, 3 to ≤ 6 months, 6 to ≤ 9 months, and > 9 months), number of doses delivered, and percent compliance to study drug administration will be summarized by treatment group for the Safety, ITT, and PP populations.

To evaluate the relationship between study drug exposure and the antibody status, the duration of study drug exposure and percent compliance to study drug administration will also be summarized by antibody status (ADA-, ADA+/Nab+, ADA+/Nab-, All ADA+, and Overall) using the Antibody Population.

If a subject is shown to be <80% compliant with study drug administration using both calculations of study drug administration compliance, this subject will be considered as having a significant protocol deviation and thus excluded from the Per-Protocol Population.

Subjects' drug exposure and compliance will be carefully reviewed against the onset of the COVID-19 pandemic. Each subject's study drug exposure and compliance prior to 01 March 2020 and on or after 01 March 2020 will be calculated based on information collected from the subject diary:

- For subjects with last dose date prior to 01 March 2020 and subjects with first dose date on or after 01 March 2020, their duration of study drug exposure (days), number of doses delivered, and percent compliance to study drug administration relative to 01 March 2020 will be calculated as above based on missed doses in subject diary.
- For subjects who started taking the study drug prior to 01 March 2020 and continued on or after 01 March 2020, their duration of study drug exposure and compliance before 01 March 2020 and on or after 01 March 2020 will be calculated as follows:
  - Duration of study drug exposure relative to 01 March 2020:
    - Duration before 01 March 2020 = 29 February 2020 – date of first dose + 1
    - Duration on or after 01 March 2020 = Overall duration of study drug exposure – duration before 01 March 2020; where overall duration of study drug exposure = last dose date – first dose date + 1
  - The number of doses delivered relative to 01 March 2020:

- Number of doses delivered before 01 March 2020 = A – {B + ([C x D] / E)}, and will be rounded to the nearest integer; where:
  - A = Duration of study drug exposure before 01 March 2020
  - B = Total number of missed doses from pens dispensed before 01 March 2020 and with scheduled return dates before or on 01 March 2020
  - C = Number of missed doses from pens dispensed during the last drug dispensing visit prior to 01 March 2020 but with scheduled return date after 01 March 2020
  - D = 01 March 2020 – last drug dispensing date prior to 01 March 2020
  - E = (number of dispensed pens during the last drug dispensing visit prior to 01 March 2020 – 1) x 30.
- Number of doses delivered on or after 01 March 2020 = Number of doses delivered during whole study – number of doses delivered before 01 March 2020.

– The percent compliance (%) to study drug administration relative to 01 March 2020:
- Percent compliance before 01 March 2020 = (Number of doses delivered before 01 March 2020 / Duration of study drug exposure before 01 March 2020) x 100%
- Percent compliance on or after 01 March 2020 = (Number of doses delivered on or after 01 March 2020 / Duration of study drug exposure on or after 01 March 2020) x 100%.

The study drug exposure and percent compliance to study drug administration before 01 March 2020 and on or after 01 March 2020 will be summarized by treatment group respectively.

The study drug exposure and compliance will also be summarized based on subject study status during the pandemic (completed or withdrew from the study before 01 March 2020 vs ongoing on or after 01 March 2020) and by region (North America and Europe).

All study drug administration and accountability data will be listed by treatment, study site, and subject number.

4.5.2 Nutritional Supplements Exposure

Subject exposure to Vitamin D and calcium supplements will be tabulated separately in terms of total amount of IU (for Vitamin D) and mg received (for calcium), respectively, (equal to total dispensed minus total returned) and duration of exposure (date of last dose – date of first dose + 1) by treatment group using the Safety population.

Summary statistics for the average Vitamin D and calcium supplements exposure per day (= ([total dispensed minus total returned] / Duration of exposure) will also be presented separately by treatment group. Similarly, summary statistics will be provided for the average calcium supplements exposure per day.

4.5.3 Adverse Events

All adverse events (AEs) will be coded by System Organ Class (SOC) and Preferred Term (PT) using the MedDRA (Version 24.0) coding system. Analyses of AEs will be performed for those events that are considered to be Treatment-Emergent Adverse Events (TEAEs). A TEAE is defined as any AE that was absent (i.e., had not occurred) or had resolved prior to the start of

study drug, and which occurred on or after the date of the first dose of study drug and within 30 days of the last dose of study drug; or any AE that started prior to the first dose of study drug, was ongoing after study treatment started, and increased in severity after the start of study drug and within 30 days of the last dose of study drug. AEs with missing onset dates will be handled according to the Conventions for Imputing Missing Dates in Appendix D.

In any tabulation of adverse events, a subject contributes only once to the count for a given SOC or PT. For summaries by severity, a subject with multiple occurrences of an AE will be represented under the most severe occurrence. For summaries by relationship to study drug, a subject with multiple occurrences of an AE will be represented under the most related occurrence.

An overall summary of the adverse events will be provided, in which the number of subjects with any adverse events, number of subjects with any TEAEs, number of subjects with any treatment-related TEAEs, number of subjects with any severe TEAEs, number of subjects with any treatment-related severe TEAEs, number of subjects with any serious TEAEs, number of subjects with any treatment-related serious TEAEs, number of subjects with any TEAES leading to study drug withdrawal, number of subjects with any TEAEs leading to study drug interruption, number of subjects with any TEAEs leading to study discontinuation, and number of subjects with any TEAEs leading to death will be summarized by treatment groups.

The following summaries will also be presented:

- TEAEs by SOC and PT
- TEAEs by SOC, PT, and region (North America and Europe)
- TEAEs by PT. Note: TEAEs will be sorted in descending order of total number of subjects with the preferred term in the abaloparatide-SC treatment group
- TEAEs by SOC, PT, and maximum severity
- Severe TEAEs by SOC and PT
- Treatment-related severe TEAEs by SOC and PT
- TEAEs by SOC, PT, and relationship to study drug
- Treatment-related TEAEs by SOC and PT
- TEAEs leading to study drug interruption by SOC and PT
- TEAEs leading to study drug withdrawal by SOC and PT
- TEAEs leading to study discontinuation by SOC and PT
- TEAEs leading to death by SOC and PT
- Serious TEAEs by SOC and PT
- Treatment-related serious TEAEs by SOC and PT
- Most common (>=5% in any treatment group) TEAEs by SOC and PT
- Non-TEAEs by SOC and PT

SOCs will be sorted in alphabetical order. Within an SOC, AEs will be sorted in descending order of the total number of subjects with the preferred term in the abaloparatide group.

To explore the relationship between TEAEs and the duration of study drug exposure, the incidence of TEAEs by time of onset (≤1 month, 1< to ≤3 months, 3< to ≤6 months, 6< to ≤9 months, and >9 months) will be summarized by treatment group, SOC and PT. Subjects with multiple occurrences of an AE within a time period will be counted once in each time period that an onset of the AE occurs.

To evaluate the relationship between TEAEs and the antibody status, the incidence of TEAEs will be summarized by antibody status (ADA-, ADA+/Nab+, ADA+/Nab-, All ADA+, and Overall), SOC, and PT using the Antibody Population. A summary of TEAEs by PT and antibody status will also be presented, in which TEAEs will be sorted in descending order of total number of subjects with the preferred term in the overall abaloparatide-SC treatment group.

During the COVID-19 pandemic, AEs may be assessed through phone/video by the investigator conducting the study visit with subjects remotely. To explore the effect of the COVID-19 pandemic on reported AEs, summaries of TEAEs by SOC and PT, TEAEs leading to treatment discontinuation, and serious TEAEs will also be presented:

- by subject's study status relative to COVID-19 pandemic onset (subjects who completed or withdrew from the study before 01 March 2020, vs subjects who completed or withdrew from the study on or after 01 March 2020)
- by AE onset dates relative to COVID-19 pandemic onset (AE occurs before 01 March 2020, vs AE occurs on or after 01 March 2020)
- by region (North America and Europe)

The number of subjects with any TEAEs, the number of subjects with any SAEs, and the number of subjects with TEAEs leading to study drug withdrawal will also be summarized by treatment group, SOC, PT and age group (< 65, 65 - < 75, and ≥ 75 years). Similar tables will be provided by treatment group, SOC, PT and BMI group (BMI < 25, 25 ≤ BMI < 30, and BMI ≥ 30).

Subject listings will be provided for all AEs by SOC and PT. Listings will also be produced for deaths, SAEs, AEs leading to study drug withdrawal and study discontinuation, and severe AEs. A listing for subjects who are anti-PTH or anti-PTHrP positive and had adverse events will also be presented.

A data listing for COVID-19 related AEs (AE PTs related to COVID-19 using MedDRA Version 24.0) will be prepared by SOC and PT. A listing of subjects who experience an Unanticipated Adverse Device Effect (UADE) will be provided from the Pharmacovigilance database.

Adverse Events of Special Interest

An adverse event of special interest (AESI) is an AE that is designated to be of special medical or scientific interest to the Sponsor. There are four major categories of AESIs as follows:

- Injection site reaction AEs -- Skin reaction AEs at the injection site
- Skin AESIs -- AEs of eschar, ulcer, or non-healing wounds at the injection site
- Hypersensitivity AESIs -- TEAEs in the Hypersensitivity Standardized MedDRA Query (SMQ narrow)
- Other AESIs:
  - Orthostatic hypotension

- Tachycardia
- Nausea
- Hypercalcemia
- Hypercalciuria
- Hypophosphataemia

The MedDRA PTs, High Level Terms (HLT), and SMQs used to identify the AESIs are stored in the study folder in the Biometrics SAS Environment.

An overall summary of the treatment emergent AESIs will be provided, in which the number of subjects with any adverse events, number of subjects with any TEAEs, number of subjects with any treatment-related TEAEs, number of subjects with any severe TEAEs, number of subjects with any treatment-related severe TEAEs, number of subjects with any serious TEAEs, number of subjects with any treatment-related serious TEAEs, number of subjects with any TEAES leading to study drug withdrawal, number of subjects with any TEAEs leading to study drug interruption, number of subjects with any TEAEs leading to study discontinuation, and number of subjects with any TEAEs leading to death will be summarized by treatment group.

The treatment emergent AESIs will be summarized by PT and SOC for each of the AESIs. Summary of the onset time and duration of the treatment emergent AESIs will be provided by treatment group using descriptive statistics (number of subjects, mean, SD, median, interquartile range, minimum and maximum). The number of subjects in each of the onset time groups (1 day, 2-3 days, 4-7 days, 8-15 days, 16-29 days, 1 to <2 months, 2 to <6 months, 6 to <10 months, and >=10 months) and duration groups (1 day, 2-3 days, 4-7 days, 8-15 days, 16-29 days, and >= 30 days) will also be summarized by treatment group.

To evaluate the relationship between AESIs and the antibody status, the incidence of treatment emergent AESIs will also be summarized by antibody status (ADA-, ADA+/Nab+, ADA+/Nab-, All ADA+, and Overall), SOC, and PT for the abaloparatide SC group.

An overall summary of the treatment emergent AESIs, as described above, will also be presented by treatment group (and overall) and the following subgroups:

- Age group (< 65, 65 - < 75, and ≥ 75 years).
- BMI group (BMI < 25, 25 ≤ BMI < 30, and BMI ≥ 30)
- Region (North America and Europe).

Subject listings will be provided for each of the AESIs.

4.5.4 Anti-Abaloparatide Antibody

For subjects in the abaloparatide-SC group, the anti-abaloparatide antibody assessment in the serum samples for immunogenicity testing will be categorized by two outcomes:

- Negative for anti-abaloparatide antibodies
- Positive for anti-abaloparatide antibodies

The number (%) of subjects in the above categories of abaloparatide antibodies will be summarized at Day 1 (Visit 3), Month 1 (Visit 4), Month 3 (Visit 5), Month 6 (Visit 6), Month 9 (Visit 7), Month 12 (Visit 8) and the Follow-up Visit scheduled approximately 1 month after the last dose of study medication (Month 13/Visit 9) using the Antibody population.

For those subjects with anti-abaloparatide antibody present (i.e, are ADA positive), the titer level, the presence of neutralizing anti-abaloparatide antibodies (negative or positive), and the assessment outcomes for cross-reactivity to PTH (negative or positive) and PTHrP antibodies (negative or positive) together with the neutralizing antibody status in subjects that are positive for PTH (negative or positive for the presence of PTH-neutralizing antibodies) or PTHrP (negative or positive for the presence of PTHrP-neutralizing antibodies) will be summarized by study visit. Figures for the incidence of antibody and neutralizing antibody for the anti-abaloparatide, anti-PTH, and anti-PTHrP will be presented by nominal visit.

All antibody information collected for the abaloparatide-SC group will be presented in a data listing.

The results of antibody analyses will be presented in a separate report.

4.5.5 Local Tolerance

Local tolerance will be assessed by the investigator and by the subject separately.

Investigator Assessment of Local Tolerance

For each local skin reaction symptom, descriptive statistics (number of subjects, mean, SD, median, interquartile range, minimum and maximum) will be used to summarize the responses on the 4-point scale (0=None, 1=Mild, 2=Moderate,3=Severe) by treatment, visit and timepoint. In addition, for each symptom, the number (%) of subjects in each response category (none, mild, moderate, and severe) will be provided by treatment per visit and timepoint.

Investigator assessment of local tolerance was not scheduled for Month 12 in the protocol. Investigator assessments recorded on Month 12/Early Termination (Month 12/ET) CRFs will be summarized under the "End of Treatment" visit.

The maximum severity of each symptom of local skin reaction per subject will be summarized with the number (%) of subjects in each response category (none, mild, moderate, and severe) by treatment and visit. The maximum severity of each symptom of local skin reaction per subject will also be summarized with the number (%) of subjects in each response category by treatment and timepoint (and at post-dose timepoint).

The maximum severity of symptom scores will also be summarized descriptively (number of subjects, mean, SD, median, interquartile range, minimum and maximum) for each response on the 4-point scale (0=None, 1=Mild, 2=Moderate, 3=Severe) by treatment and visit, and by treatment and timepoint, respectively.

During the COVID-19 pandemic, the investigator's assessment of the symptoms of local skin reactions may be assessed through alternative methods other than direct on-site assessment. To explore the effect of the COVID-19 pandemic, summaries of the investigator's assessment of local tolerance will also be presented by:

- region (North America and Europe)
- subject's study status relative to COVID-19 pandemic onset (subjects who completed or withdrew from the study before 01 March 2020, vs subjects who completed or withdrew from the study on or after 01 March 2020)

- excluding local tolerance assessments made via alternative methods other than direct on-site assessment

Subject Assessment of Local Tolerance

Subject assessment of local tolerance using a 4-point scale will be done within 5 minutes and at 1 hour following administration beginning on Day 2 and continuing through Day 7 of Month 1 and Month 9, as recorded in the subject diary.

For each local skin reaction symptom, descriptive statistics (number of subjects, mean, SD, median, interquartile range, minimum and maximum) will be used to summarize the responses on the 4-point scale (0=None, 1=Mild, 2=Moderate,3=Severe) collected during Day 2 to Day 7 at Month 1 and at Month 9:

- by treatment, visit, day, and timepoint.
- maximum severity score by treatment, visit and day
- maximum severity score by treatment, visit and timepoint
- maximum severity score by treatment and visit
- maximum severity score by treatment and day across visits
- maximum severity score by treatment and timepoints

In addition, for each symptom, the number (%) of subjects in each response category (none, mild, moderate, and severe) will be provided by treatment per visit, day, and timepoint in a similar manner as described above for the summary of responses by the 4-point scale. The maximum severity of symptom will be similarly summarized for each response category (none, mild, moderate, and severe):

- by treatment, visit, and day
- by treatment, visit, and timepoint
- by treatment and visit
- by treatment and days across visits
- by treatment and timepoint across visits (and at post dose timepoint).

All injection site local tolerability assessment scores will be presented in a listing.

4.5.6 Laboratory Data

Clinical laboratory values will be provided in conventional units and in standard international (SI) units by the central laboratory.

Summaries of laboratory data using descriptive statistics by treatment group and study visit will be presented based on the SI units, including absolute results and changes from baseline. This includes serum chemistry, hematology, coagulation, hormones (including total testosterone, estradiol, SHBG, and PTH), 25-hydroxyvitamin D, 1,25-dihydroxy vitamin D, and urinalysis (for quantitative urinalysis assays). In the event of repeat assessments during a visit time point, the last non-missing value per study day/time will be used in the summary statistics. In the event of duplicate assessments at a visit time point, the worst value per study day/time will be used in the summary statistics. Results from unscheduled visits will not be included in the summaries.

Shift analyses of laboratory data from baseline to the worst post-baseline value and from baseline to Month 12/ET (Visit 8) will be performed by treatment group. Results for serum chemistry,

hematology, coagulation, and urinalysis will be presented by category (above normal limit, within normal limit, below normal limit). For shift tables from baseline to worst post-baseline value, and for baseline to Month 12, percentages will only be calculated based on the number of subjects with valid data for baseline and post-baseline; subjects who are missing either assessment will not be included in the percentage calculation (numerator or denominator). Results from unscheduled visits will be included in the shift analyses, as applicable.

To explore the effect of the COVID-19 pandemic, summaries of all laboratory parameters will also be presented for the following:

- By region (North America and Europe)
- By subject's study status relative to COVID-19 pandemic onset (subjects who completed or withdrew from the study before 01 March 2020, vs subjects who completed or withdrew from the study on or after 01 March 2020)

Any laboratory parameter value that is provided as "< *numeric value*" or "> *numeric value*" will be imputed as the *numeric value* in all summary tables and analyses; but will be shown with the actual result in the listings.

All laboratory data, including repeated values and results from unscheduled visits, will be presented in data listings, with indication of higher or lower than the associated normal range of each laboratory test, where applicable.

Potentially Clinically Significant Laboratory Values

Criteria for determining values of laboratory tests that are considered Potentially Clinically Significant (PCS) are in Appendix B of this SAP. The number (%) of subjects with at least 1 post-baseline PCS values who did not have PCS values at baseline will be presented for each laboratory parameter. Results from unscheduled visits will be included in identification of subjects with post-baseline PCS values, as applicable. For the calculation of the percentages of subjects, the denominator will be based on the number of subjects who had no PCS values at baseline and had at least one post-baseline assessment for the specific laboratory parameter of interest. The numerator will be based on the number of subjects from the denominator who had at least one PCS value post-baseline for the specific laboratory parameter of interest. A listing of subjects with any PCS values will be provided.

Hypercalcemia

A hypercalcemic event (hypercalcemia) is defined as an elevated serum calcium (albumin-corrected) value ≥ 0.3 mg/dL (≥ 0.08 mmol/L) above the upper limit of the normal range (ULN).

Three cutoff levels of serum calcium (albumin-corrected):

a) ≥ ULN
b) ≥ ULN+ 0.3 mg/dL (or + 0.08 mmol/L)
c) ≥ ULN + 1.0 mg/dL (or + 0.25 mmol/L)

The following summaries and analyses will be provided by treatment group for each of the cutoff levels above:

- Number (%) of subjects with at least one occurrence of a serum calcium (albumin-corrected) value above the cutoff level at each visit, timepoint, and overall (entire treatment period):

  - The corresponding 95% CI based on Wilson's Score method will be provided per treatment group.
  - For the number of subjects with at least one occurrence of a serum calcium (albumin-corrected) value above the cutoff level during the entire treatment period:
    - The treatment groups will be compared using the chi-square test.
    - The risk reduction between the treatment groups (calculated as Abaloparatide-SC − Placebo) will be provided together with the 95% CI based on Newcombe's method.
    - The relative risk reduction between the treatment groups (calculated as (Abaloparatide-SC − Placebo)/Placebo) will be provided together with the 95% CI based on Wald's method.
- Number of occurrences of a serum calcium (albumin-corrected) value above the cutoff level per subject during the entire treatment period
- The ratio of the total number of occurrences of a serum calcium (albumin-corrected) value above the cutoff level to the total number of serum calcium (albumin-corrected) blood samples per subject during the entire treatment period.

Mean (±SE) serum calcium (albumin-adjusted) values will be plotted over study visit by treatment group. All serum calcium (albumin-corrected) results will be included in the plot, which will indicate the results at pre-dose and 4-hours post-dose at each visit.

To evaluate the relationship between age and hypercalcemia, the number (%) of subjects with at least one occurrence of hypercalcemia at each visit and overall (entire treatment period) will be summarized by treatment group and age group (< 65, 65 - < 75, and ≥ 75 years).

Hypercalciuria

The number (%) of subjects with at least one occurrence of hypercalciuria (urine calcium/creatinine ratio > 0.4 mg/mg (> 1.131 mmol/mmol)) at Baseline (Day 1), Month 3, and overall will be summarized by treatment group. A 95% CI for the percentage based on the Wilson's Score method will be calculated for each treatment group at each visit and overall. This analysis will be repeated for subjects with one or more occurrences of urine calcium/creatinine ratio > 0.3 mg/mg (> 0.848 mmol/mmol).

To evaluate the relationship between age and hypercalciuria, the number (%) of subjects with at least one occurrence of hypercalciuria at each visit and overall will be summarized by treatment group and age group (< 65, 65 - < 75, and ≥ 75 years).

Creatinine and Calcium Excretion and Creatinine Clearance

The 24-hour urine for calcium:creatinine and creatinine clearance are collected at Day 1 and Month 3. The urine sample values will be adjusted for the time period to derive the urine result that is standardized for a 24-hour period. The 24-hour urine results will be summarized descriptively by treatment group and visit. If the end of the collection time of the 24-hour urine is on the same day as the first dose date or within 24 hours from the time of first dose, then the 24-hour urine sample will be considered as a Day 1 sample and will be treated as the baseline sample. Otherwise, the 24-hour urine data will be considered as a post-Day 1 (i.e, post-baseline) sample.

The parameters to be summarized include:

- Urine calcium:creatinine ratio
- 24-hour urine calcium excretion (calcium concentration x 24-hour urine volume)
- 24-hour urine creatinine excretion (creatinine concentration x 24-hour urine volume)
- Creatinine clearance (CCr) calculated in two ways:
  a) Based on 24-hour urine collection
  b) Based on serum creatinine, calculated using the Cockcroft-Gault estimate, CCr:

$$\text{CCr (mL/min)} = [(140 - \text{age}) * \text{weight} * 1.23] \div \text{SCr}$$

  Where:
  - Age in years
  - Weight in kilograms
  - SCr = Serum creatinine in µmol/L.

The following categories of renal function will be defined:

	Cockcroft-Gault Estimated Serum Creatinine Clearance
Normal	≥ 90 mL/min (≥ 1.5 mL/s)
Mild	≥ 60 mL/min to < 90 mL/min (≥ 1.0 mL/s to < 1.5 mL/s)
Moderate	≥ 30 mL/min to < 60 mL/min (≥ 0.5 mL/s to < 1.0 mL/s)
Severe	≥ 15 mL/min to < 30 mL/min (≥ 0.25 mL/s to < 0.5 mL/s)

The number (%) of subjects in each renal function category at each visit and overall (across all visits) will be presented by treatment group. A similar summary table will be presented by treatment group and age group (< 65, 65 - < 75, and ≥ 75 years).

4.5.7 Vital Signs and Physical Examination

The actual value, the change from baseline to each visit, the change from baseline to the worst post-baseline value during the treatment period, and the change from baseline to Month 13 (Follow-up visit) will be summarized for vital signs using descriptive statistics by treatment group. If a subject has repeat assessments on the vital signs at a particular scheduled visit/time point, the last non-missing value will be used for the by-visit summary. Unscheduled assessments will be included in the summaries for the worst post-baseline value or PCS, they will not be included in the descriptive by-visit summary tables.

The number (%) of subjects experiencing post-dose orthostatic hypotension will be summarized by treatment group, visit and time point, and across all visits. Orthostatic hypotension will be defined as a decrease in systolic blood pressure (SBP) of ≥ 20 mmHg from supine to standing or in diastolic blood pressure (DBP) of ≥ 10 mmHg supine to standing. Results from unscheduled visits will be included in the summaries.

The number (%) of subjects experiencing heart rate increase from pre-dose to post-dose and the number of occurrences per subject will be summarized with various threshold values (ranging from >5 bpm to >40 bpm, with threshold values increasing by 5 BPM). The median, minimum and maximum heart rate increase (in BPM) will be provided per threshold value. The number of occurrences per subject will be summarized using the median, minimum and maximum values.

All heart rate data, including repeat and unscheduled assessments, will be included in determining the increase from pre-dose in this summary.

Criteria for determining values of vital sign measurements that are considered PCS are in Appendix C of this SAP. The number (%) of subjects who have at least 1 PCS vital sign values for each vital sign parameter of interest after the first dose of study drug, but had no PCS values at baseline, will be presented by treatment. Results from unscheduled visits will be included in identification of subjects with post-baseline PCS values, as applicable. For the calculation of the percentages of subjects, the denominator will be based on the number of subjects who had no PCS values at baseline and had at least one post-baseline assessment for the specific vital sign parameter of interest.

A listing of subjects with any PCS vital sign values will be provided.

To explore the effect of the COVID-19 pandemic, the above planned summaries will also be presented for the following:

- By region (North America and Europe)
- By subject's study status relative to COVID-19 pandemic onset (subjects who completed or withdrew from the study before 01 March 2020, vs subjects who completed or withdrew from the study on or after 01 March 2020)

To evaluate the relationship between age and orthostatic hypotension, the number (%) of subjects with at least one occurrence of post-dose orthostatic hypotension at each visit, timepoint, and overall (across all visits) will be summarized by treatment group and age group (< 65, 65 - < 75, and ≥ 75 years).

To evaluate the relationship between age and heart rate increase, the number (%) of subjects with at least one occurrence of heart rate increase from pre-dose to post-dose will be summarized with various threshold values (ranging from >5 bpm to >40 bpm, with threshold values increasing by 5 bpm) by treatment group and age group (< 65, 65 - < 75, and ≥ 75 years). Descriptive statistics for heart rate increase per threshold value and the number of occurrences of heart rate increase per subject per threshold value will be provided.

Physical examination is performed only at screening visit and will be summarized as described in Section 4.2.

4.5.8 Electrocardiogram

Electrocardiogram (ECG) parameters collected at study visits are heart rate, PR interval, QRS duration, QT, and QTcF (QT corrected by Fridericia's formula). The actual value, change from baseline value, changes from baseline to worst value during the treatment period, and change from baseline to Month 13 (follow-up visit) will be summarized for quantitative ECG results by treatment group, study visit, and time point (pre-dose and 1-hour post-dose). The overall ECG assessments will be summarized categorically (normal, abnormal without clinical significance, and abnormal with clinical significance) as a shift table by treatment group, study visit, and timepoint. With the exception of change from baseline to the worst post-baseline value during the treatment period, results from unscheduled visits will not be included in the summaries.

Summaries will also be provided for the change from predose values for ECG heart rate, where the predose value is the last value obtained prior to dosing at each visit where ECG is assessed. If a subject has multiple ECG results at a particular visit/time point, the repeat measurement reason will be checked to determine if the first or the last non-missing value will be used for the summary. If the reason for the repeat assessment is abnormal value, then the first value will be used. If the reason is misplacement of the limb lead electrodes, the last value will be used. Results from unscheduled visits will not be included in these descriptive summary tables.

The number (%) of subjects experiencing heart rate increase from pre-dose to post-dose and the number of occurrences per subject will be summarized with various threshold values (ranging from >5 bpm to >40 bpm, with threshold values increasing by 5 bpm). The median, minimum, and maximum heart rate increase (in bpm) will be provided per threshold value. The number of occurrences per subject will be summarized using the median, minimum, and maximum values. All ECG heart rate data, including repeat and unscheduled assessments, will be included in determining the increase from pre-dose in this summary. To evaluate the relationship between age and heart rate increase, the above summary of heart rate increase with various threshold values will also be repeated by treatment and age group (< 65, 65 - < 75, and ≥ 75 years).

The number (%) of subjects with the following clinically important QTcF values will be presented by treatment group, visit, and timepoint. Results from repeat/unscheduled visits will be included in shift analyses, as applicable:

- Post-baseline QTcF > 450 msec
- post-baseline QTcF > 480 msec
- post-baseline QTcF > 500 msec
- change from baseline > 30 msec in QTcF
- change from baseline > 60 msec in QTcF
- post-baseline QTcF > 500 msec and change from baseline >60 msec in QTcF

Shift tables from baseline to post-baseline QTcF values using the following categories will be presented for each post-baseline visit and time point. Results from repeat/unscheduled visits will be included in shift analyses, as applicable

- ≤ 450 msec and > 450 msec
- ≤ 480 msec and > 480 msec
- ≤ 500 msec and > 500 msec.

All ECG data will be provided by subject in a data listing. Listings will also be provided for those subjects who have any QTcF values greater than 450 msec, 480 msec, and 500 msec, and those with QTcF change from baseline values >30 msec and >60 msec.

During the COVID-19 pandemic, ECG may be conducted at a subject's home (home visit) by authorized personnel and then be processed and assessed by a central reader. To explore the effect of the COVID-19 pandemic, summaries for ECG parameters will also be presented for the following:

- By region (North America and Europe)
- By subject's study status relative to COVID-19 pandemic onset (subjects who completed or withdrew from the study before 01 March 2020, vs subjects who completed or withdrew from the study on or after 01 March 2020)

4.5.9 Concomitant Medications

Medication history and concomitant medications will be coded using the WHO Drug Dictionary. Medications will be tabulated by anatomic therapeutic class (ATC), preferred term (PT), and treatment group. Medications and concomitant medications will be summarized using the Safety population. All recorded medications will be presented in a by-subject listing.

Summary of concomitant medications will include medications taken from the date of the first dose of study medication (Day 1) until 30 days after the last dose of study medication. Medications that did not end prior to the first dose of study medication will be included in the summary.

Prior medications (those with start date prior to the first dose of study medication) will also be summarized by ATC, PT, and treatment group.

If both start date and stop date are missing, the medication will be assumed to occur both prior and concomitantly.

To explore the effect of the COVID-19 pandemic, summaries for the prior medications and concomitant medications will also be presented for the following:

- By region (North America and Europe)
- By subject's study status relative to COVID-19 pandemic onset (subjects who completed or withdrew from the study before 01 March 2020, vs subjects who completed or withdrew from the study on or after 01 March 2020).
- By medication start dates relative to COVID-19 pandemic onset (medications start before 01 March 2020, vs medications start on or after 01 March 2020). For the calculation of percentages, the denominator includes all subjects in the safety population for the relevant treatment group.

5 CHANGES FROM THE PROTOCOL-SPECIFIED ANALYSES

The following changes from the protocol-specified analyses are included in this SAP:

- In Section 1.2.5, a word "Serum" is added to the additional secondary efficacy endpoint "carboxy-terminal cross-linking telopeptide of type I collagen (s-CTX) at 1, 3, 6, and 12 months"
- In Section 1.2.5, the words "12 months" are replaced by "end of treatment" in the additional secondary efficacy endpoint "Proportion of subjects converting from the categories of osteoporosis to osteopenia or to normal at 12 months" to clarify the timepoint at which the endpoint is assessed.
- Section 2.2 defines the Antibody Population
- Section 3.12 defines analysis visit windows for the analyses of BMD, vBMD, and serum markers of bone metabolism (s-PINP and s-CTX) data to determine the assessment that will be used for a given study visit.
- In Section 4.3.1, the primary analysis will use a pattern mixture model (PMM) analysis with multiple imputation using the wash-out imputation method, instead of the protocol-specified ANCOVA model using LOCF technique to impute missing values. The LOCF method has been shown to lead to substantial bias in estimating treatment effects and is no longer recommended by regulatory authorities. While the PMM analysis with wash-out imputation is an appropriate approach to address the possibility of data missing not at random. The PMM analysis with wash-out imputation method was recommended by the FDA in their comment on the primary analysis of this study: *Advice-Information Request IND 73176, dated 05 March 2021*.
- In Section 4.3.1, age is removed from the MMRM and ANCOVA models as a covariate. From previous clinical studies, there is little evidence to suggest that age is associated with the BMD response in subjects who have taken abaloparatide-SC.
- In Section 4.5.3, definitions and summary of AESIs are added. AESIs were not defined in the protocol.
- Additional analyses to evaluate the impact of the COVID-19 pandemic on study conduct were added for safety and efficacy endpoints where applicable.
- The results of antibody analyses will be presented in a separate report.

All other statements in this SAP are enhancements to the statistical considerations described in the protocol.

Appendix A. Algorithm to Identify Significant Protocol Deviations that Exclude Subjects from the Per-Protocol (PP) Population

A protocol deviation is any change, divergence, or departure from the study design or procedures defined in the protocol. Significant PDs are a subset of PDs that have the potential to impact the effect of the treatment on the subject.

The below algorithm outlines the criteria for identifying significant protocol deviations. Subjects with at least one of these significant protocol deviations will be excluded from the PP population.

A. Deviations of subject eligibility (Inclusion/Exclusion criteria) that will result in exclusion from PP population:
   1. Inclusion Criterion #2: A BMD T-score above the required level.
   2. Exclusion Criterion #2: A BMD T-score below the required limit.
   3. Exclusion Criterion #1: Presence of abnormalities of the lumbar spine that would prohibit assessment of spinal bone mineral density.
   4. Exclusion Criterion #3: Unevaluable hip BMD or subjects who have undergone bilateral hip replacement.
   5. Exclusion Criterion #4: Fragility fracture within the 12 months prior to signing of informed consent.
   6. Exclusion Criterion #5: History of severe vertebral fracture or > 2 moderate vertebral fractures.
   7. Exclusion Criteria #6, #8, and #10: Diseases and treatments affecting the bone strength.
   8. Inclusion Criterion #5: Laboratory tests affecting the bone strength
   9. Exclusion Criteria #18, #19, #20, #21, #22, #23, #24, #25, and #27. Prior treatment with bone acting agents within the specified time period.

B. Deviations of study drug administration that will result in exclusion from PP population:
   1. The blinding of the study drug was broken during the subject's participation in this study before database lock.
   2. Duration of study drug exposure is less than 3 months (90 days)
   3. Compliance to study drug administration < 80%

C. Deviations of efficacy measurements that will result in exclusion from PP population:
   1. Lumbar spine BMD is missing at baseline
   2. No post-baseline lumbar spine BMD measurements recorded

D. Other significant protocol deviations
   1. Any protocol deviation with potential major impact on the primary efficacy endpoint will exclude a subject from the PP population. This will include subjects taking any prohibited medications during study.

Appendix B. Criteria for Potentially Clinically Significant Laboratory Values

Lab Parameter	Conventional Unit		SI Unit	
	Low	**High**	**Low**	**High**
Hematology				
Absolute Eosinophils		>5000 cells/mm^3		>5.00 10^9/L
Absolute Lymphocytes	≤499 cells/mm^3		≤0.499 10^9/L	
Absolute Neutrophils	≤999 cells/mm^3		≤0.999 10^9/L	
Hemoglobin	≤10.4 g/dL (male)	change from baseline ≥2.1 g/dL	≤104 g/L (male)	change from baseline ≥21 g/L
Platelets	≤99000 cells/mm^3		≤99 10^9/L	
White Blood Cells	≤1499 cells/mm^3	≥20001 cells/mm^3	≤1.499 10^9/L	≥20.0 10^9/L
Coagulation				
Activated Partial Thromboplastin Time		≥1.41*ULN		≥1.41*ULN
Serum Chemistry				
Alanine Aminotransferase		≥5.1*ULN		≥5.1*ULN
Albumin	<2.5 g/dL		<25 g/L	
Alkaline Phosphatase IFCC		≥3.1*ULN		≥3.1*ULN
Aspartate Aminotransferase		≥5.1*ULN		≥5.1*ULN
Bilirubin Total		≥1.51*ULN (with ALT and/or AST>ULN) ≥2.0*ULN (with normal liver function tests)		≥1.51*ULN (with ALT and/or AST>ULN) ≥2.0*ULN (with normal liver function tests)
Calcium	≤7.4 mg/dL	≥11.6 mg/dL	≤1.85 mmol/L	≥2.9 mmol/L
Cholesterol Total		>226 mg/dL		>5.8534 mmol/L
Creatine Kinase		≥3.1*ULN		≥3.1*ULN
Creatinine		≥2.1 mg/dL		≥ 185.64 umol/L
Glucose	≤54 mg/dL	>200 mg/dL (random)	≤ 2.997 mmol/L	> 11.1 mmol/L
Potassium	≤3.2 mEq/L	≥5.5 mEq/L	≤3.2 mmol/L	≥5.5 mmol/L
Protein Total	<5 g/dL		<50 g/L	
Sodium	≤129 mEq/L	≥148 mEq/L	≤129 mmol/L	≥148 mmol/L
Urine				
Glucose		2+ (≥100 mg/dL)		2+ (≥5.55 mmol/L)
Protein		2+ (≥100 mg/dL)		2+ (≥ 1.0 gm/L)
Blood (microscopic)		>50 rbc/hpf		>50 rbc/hpf
Source: *FDA Guidance for Industry 2007, Toxicity Grading Scale for Healthy Adult and Adolescent Volunteers Enrolled in Preventive Vaccine Clinical Trials*				

Appendix C. Criteria for Potentially Clinically Significant Vital Signs Values

Parameter	Unit	Lower Criteria	Upper Criteria
Heart Rate	bpm	< 45	>100 >130
Supine systolic blood pressure	mm Hg	< 80	>155
Supine diastolic blood pressure	mmHg		>100
Standing systolic blood pressure	mmHg	<80	>155
Standing diastolic blood pressure	mmHg		>100
Body temperature	°C		≥39.0

Appendix D. Conventions for Imputing Missing Dates

1. Adverse event (AE) start (onset) date imputation

AE start (onset) date is separated into 3 different data fields, including day, month, and year. Completely missing AE start date (i.e., missing day, missing month, and missing year) will not be imputed. Partial AE start date will be imputed according to treatment start date as follows in a sequential fashion.

- If AE year is missing, then no imputation
- If AE month and day are missing and
  - if AE year < Treatment year then AE month/day = '01JUL'
  - if AE year = Treatment year then AE month/day = treatment start date+1
  - if AE year > Treatment year then AE month/day = '01JAN'
- If AE day is missing and
  - if AE year < Treatment year then AE day = '15'
  - if AE year = Treatment year and
    - if AE month < Treatment month then AE day = '15'
    - if AE month = Treatment month then AE date = treatment start date+1
    - if AE month > Treatment month then AE day = '01'
  - if AE year > Treatment year, then AE day = '01'
- If AE start date is still missing after the above imputation and AE end date is nonmissing then AE start date = AE end date.
- If AE start date is still missing after the above imputation, then count this AE as a treatment-emergent AE.

2. Concomitant medication (CM) start and end date imputation

The imputation of missing medication start and/or end dates is performed for the following reasons:

- To determine if a medication is a prior medication or a concomitant medication.
- To determine if treatment with the medication satisfies any of the medication-related exclusion criteria #18, #19, #20, #21, #22, #23, #24, #25, and #27 for subject eligibility.

The medication start and end dates are separated into 3 different data fields, including day, month, and year. Completely missing medication start and/or end date (i.e., missing day, missing month, and missing year) will not be imputed. Partially missing medication start/end date will be imputed according to treatment start date. The rules on how to determine if a medication is a prior medication only or both prior and concomitant medication and to determine if subjects violated any protocol specified medication-related exclusion criteria are specified below.

The prior and concomitant medications data listings will display all data as collected; imputed values will not be shown.

- **Medication Start Date Imputation**
  a) If medication start year is missing, then no imputation
  b) If medication start year is reported while medication start month and day are missing, then medication start month/day = '01JAN'

c) If medication start year and month are reported while medication start day is missing, then medication start day = '01'

- **Medication End Date Imputation**

a) If medication start date and medication end date are completely missing; or medication end date is completely missing while medication start date is either a complete date or partially missing and the observed/imputed medication start date < treatment start date, then no imputation:
   - If Ongoing = Yes, the medication is considered to be both prior and concomitant medication
   - If Ongoing = No or missing, the medication is considered to be a prior medication only
   - Subject will be considered as violating the medication-related exclusion criteria if the medication is one of the prohibited medications according to the medication-related exclusion criteria.

b) If medication start date is completely missing while medication end date is partially missing; or medication end date is partially missing while medication start date is either a complete date or partially missing and the observed/imputed medication start date < treatment start date, then the missing medication end date will be imputed as follows:

	MMM Missing	MMM < TRT Month	MMM = TRT Month	MMM > TRT Month
YYYY < TRT Year	31DEC	Last day of MMM	Last day of MMM	Last day of MMM
YYYY = TRT Year	TRTDT *	Last day of MMM	TRTDT *	01MMM
YYYY > TRT Year	31JAN	Last day of MMM	Last day of MMM	Last day of MMM

* If Ongoing = Yes, then no imputation.

Where: YYYY = Medication end year; MMM = Medication end month; TRT Year = Treatment start year; TRT Month = Treatment start month; and TRTDT = Treatment start date.

   - The observed and/or imputed medication end date will be used to determine if a medication is a prior medication only; or is both a prior and a concomitant medication.
   - If the medication is one of the prohibited medications according to the medication-related exclusion criteria, the observed and/or imputed medication end date will also be used to determine if a subject meets any medication-related exclusion criteria.

c) If the medication start date is either a complete date or partially missing with the observed/imputed medication start date >= treatment start date, then no imputation for medication end date, and:
   - The medication is considered to be a concomitant medication only.
   - Subject doesn't meet any medication-related exclusion criteria.

3. **Other date imputations**

Any event with a partial date that requires imputation but has no specific imputation rule will be imputed as follows:

- If the day is missing, then impute the day as '01'
- If the month is missing, then impute the month as 'JUL'
- If the imputation of a start date causes the start date to be post the non-missing end date, then the start date will be imputed as the end date.
- If the imputation of a start date causes the start date to be prior to the start of treatment (but the event should happen post-baseline), then the start date will be imputed as "treatment start date + 1."

Any date cannot be imputed past the date of death. When this happens, the date will be imputed as the date of death.

Signature Page for RAD-CLIN-002815 v1.0

Approval	[redacted] 25-Aug-2021 13:52:49 GMT+0000

Approval	[redacted] 25-Aug-2021 14:24:02 GMT+0000

Approval	[redacted] 25-Aug-2021 14:28:48 GMT+0000

Signature Page for RAD-CLIN-002815 v1.0