STATISTICAL ANALYSIS PLAN

TITLE PAGE

Study Number TVB009-IMB-30085

A Randomized, Double-Blind, Multinational, Multicenter Study to Compare Efficacy, Safety, and Immunogenicity of TVB-009P and Denosumab (PROLIA®) in Patients with Postmenopausal Osteoporosis

Final Version 3: April 18, 2023

TABLE OF CONTENTS **PAGE**

DECLARATION

I, the undersigned, declare that I have prepared the statistical analysis plan along with TLF mockups and that to the best of my knowledge this document is internally consistent with protocol and scientifically rational.

Prepared by:

[redacted]

[redacted]
Statistician

AUTHORIZATION: I, the undersigned, declare that I have reviewed the statistical analysis plan along with TLF mock-ups and that to the best of my knowledge the document accurately reflects the protocol objectives.

Authorized by:

[redacted]

[redacted]

REVISION HISTORY

Version	Date	Author	Reasons
Final 1	December 5, 2022	[redacted]	The first approved version
Final 2	February 22, 2023	[redacted]	Clarification that BMD assessments must be done by machines of the same manufacturer
Final 3	April 18, 2023	[redacted]	Inclusion of the analysis of fractures identified by the central reader of X-rays

LIST OF ABBREVIATIONS

Abbreviation or special term	Explanation
ADA	Anti-Drug Antibodies
AE	Adverse Event
ANCOVA	Analysis of Covariance
ATC	Anatomic-therapeutic-chemical
$AUC_{0\text{-}tau}$	Area under the curve over the dosing period
BLA	Biologics License Application
BLQ	Below Limit of Quantification
BMD	Bone Mineral Density
BMI	Body Mass Index
C_{2weeks}	Serum concentration at 2 weeks post-dose
CI	Confidence Interval
C_{max}	Maximum observed concentration
COVID-19	Coronavirus disease 2019
CS	Clinically Significant
CSR	Clinical Study Report
C_{trough}	Serum concentration before next dose
DXA	Dual-energy X-ray Absorptiometry
ECG	Electrocardiogram
eCRF	Electronic Case Report Form
EOS	End of Study
ET	Early Termination
EU	European Union
HEENT	Head, Eyes, Ears, Nose, and Throat
IgG2	Immunoglobulin G2
IMP	Investigational Medicinal Product
ITT	Intent-to-Treat

Abbreviation or special term	Explanation
LLOQ	Low Limit of Quantification
LS	Least Squares
LS-BMD	Lumbar Spine-Bone Mineral Density
MAA	Marketing Authorisation Application
mAb	Monoclonal Antibody
MAR	Missing at Random
MCMC	Markov Chain Monte Carlo
MedDRA	Medical Dictionary for Regulatory Activities
mITT	Modified Intent-to-Treat
MNAR	Missing Not at Random
NCS	Not Clinically Significant
NRS	Numeric Rating Scale
P1NP	Procollagen Type 1 N Propeptide
PK	Pharmacokinetic
PMO	Postmenopausal Osteoporosis
PP	Per Protocol
PT	Preferred Term
RANKL	Receptor Activator of Nuclear Factor Kappa-B Ligand
RTSM	Randomization and Trial Supply Management
SAP	Statistical Analysis Plan
sc	Subcutaneous
SD	Standard Deviation
sCTX-1	C-telopeptide Cross-Link of Type 1 Collagen
SOC	System Organ Class
TEAE	Treatment-Emergent Adverse Event
TITT	Transition Intent-to-Treat
TmITT	Transition Modified Intent-to-Treat
US	United States

Abbreviation or special term	Explanation
WHO-DRL	World Health Organization Drug Reference List

1 INTRODUCTION

This statistical analysis plan (SAP) describes the statistical methods and data handling methods to be followed during the final reporting and analyses of data collected for the study Protocol TVB009-IMB-30085 "A Randomized, Double-Blind, Multinational, Multicenter Study to Compare Efficacy, Safety, and Immunogenicity of TVB-009P and Denosumab (PROLIA®) in Patients with Postmenopausal Osteoporosis".

This SAP should be read in conjunction with the study protocol. This version of the plan has been developed using the protocol with Amendment 2 dated 29 Jun 2021.

TVB-009 is a fully human immunoglobulin G2 (IgG2)/kappa monoclonal antibody (mAb) directed against receptor activator of nuclear factor kappa-B ligand (RANKL), which is being developed by Teva as a biosimilar candidate to PROLIA and XGEVA with the active substance denosumab. Denosumab is currently licensed as PROLIA and XGEVA, both being first approved in the US in 2010, and in the EU in 2010 (PROLIA) and 2011 (XGEVA).

2 STUDY DETAILS

2.1 Study Objectives and Endpoints

The primary and secondary study objectives and endpoints are:

Objectives	Endpoints
The **primary objective** of the study is to demonstrate that there are no clinically meaningful differences in efficacy between TVB-009P and PROLIA US administered subcutaneously (sc) in patients with postmenopausal osteoporosis.	The primary efficacy endpoint is: • the percent change from baseline in lumbar spine-bone mineral density (LS-BMD) at week 52 based on centrally assessed dual-energy X-ray absorptiometry (DXA) measurements • The co-primary efficacy endpoint for the European Union (EU) submission is the percent change from baseline in serum C-telopeptide cross-link of type 1 collagen (sCTX-1) at week 26. For the EU submission, this endpoint is regarded as co-primary. For the United States (US) submission, this endpoint is regarded as secondary
A **secondary objective** is to compare further efficacy and pharmacodynamic parameters between TVB-009P and PROLIA US	The secondary efficacy and pharmacodynamic endpoints are: • percent change from baseline in LS-BMD at week 26 based on centrally assessed DXA measurements • percent change from baseline in femoral neck bone mineral density (BMD) by DXA at week 26 and at week 52 • percent change from baseline in total hip BMD by DXA at week 26 and at week 52 • percent change from baseline in sCTX-1 at all time points • sCTX-1 suppression at week 4 • percent change from baseline in procollagen type 1 N propeptide (P1NP) at week 26 and week 52 • incidence of fractures up to week 52
A **secondary objective** of the study is to compare efficacy and pharmacodynamic parameters	The pharmacodynamic/efficacy endpoints in the transition period are:

Objectives	Endpoints
between TVB-009P and PROLIA US after a single transition from PROLIA US to TVB-009P	• percent change from week 52 in LS-BMD by DXA at week 78 • difference between percent change from baseline in sCTX-1 between week 52 and week 78 • percent change from week 52 in femoral neck BMD by DXA at week 78 • percent change from week 52 in total hip BMD by DXA at week 78 • difference between percent change from baseline in P1NP between week 52 and week 78 • incidence of fractures up to week 78
A **secondary objective** of this study is to compare the safety and tolerability, including device-related events, between TVB-009P and PROLIA US throughout the study. A **secondary objective** of this study is to compare the safety and tolerability between TVB-009P and PROLIA US after a single transition from PROLIA US to TVB-009P, including device-related events.	The safety and tolerability endpoints are: • adverse events (and the number of patients who withdraw from the study due to adverse events) • vital signs • laboratory tests (hematology, serum chemistry, and urinalysis) • electrocardiogram (ECG) • local tolerability at injection site • use of concomitant medications • device-related adverse events and malfunctions
A **secondary objective** of this study is to assess the immunogenicity of TVB-009P in comparison with PROLIA US throughout the study. A **secondary objective** of this study is to assess the immunogenicity of TVB- 009P in comparison with PROLIA US after a single transition from PROLIA US to TVB-009P.	The immunogenicity endpoint is: • incidence of patients with confirmed positive anti-drug antibody (ADA) sample For confirmed positive samples, the ADA titer and the neutralizing potential will be tested. The effect of positive immunogenicity findings on pharmacokinetics, efficacy, and safety will be assessed if applicable
Explorative objectives and endpoints	
An **exploratory objective** of this study is to compare pharmacokinetics	The pharmacokinetic endpoints are: • serum concentration before next dose (C_{trough}),

Objectives	Endpoints
between TVB-009P and PROLIA US throughout the study	before second and third dose, and 6 months after third dose • serum concentration at 2 weeks post-dose (C_{2weeks}) Pharmacokinetic endpoints derived from population pharmacokinetics analysis are: • area under the curve over the dosing period (AUC_{0-tau}) after first dose • C_{trough} before second dose • maximum observed concentration (C_{max}) after first dose

2.2 Primary Estimand

The primary estimand assesses clinically meaningful differences in effectiveness in patients with Postmenopausal Osteoporosis (PMO), based on surrogate measures, focusing on the causal effects attributable to the Investigational Medicinal Product (IMP).

The primary estimand is the difference in mean percent change in lumbar spine-bone mineral density (LS-BMD) from baseline at week 52 between TVB-009P and PROLIA US treatment arms, regardless of intercurrent events in the target population of patients with PMO who receive at least one dose of IMP and have both a baseline and at least 1 post-baseline assessment of LS-BMD. For the submission in the US, this estimand is regarded as primary and for the submission in the European Union (EU), this estimand is regarded as co-primary.

The co-primary estimand for the EU filing is the difference in mean percent change in sCTX-1from baseline at week 26, between TVB-009P and PROLIA US treatment arms, regardless of intercurrent events in the target population of patients with PMO. The same analysis set will be used for both estimands and therefore patients who terminate before week 26 will not be included in the sCTX-1 analysis. For the EU submission, this estimand is regarded as co-primary. For the US submission, this estimand is regarded as secondary.

2.3 Study Design

This is a randomized, double-blind, multinational, multicenter study to demonstrate similar efficacy and safety of TVB-009P compared to PROLIA US administered in 3 sc doses of 60 mg every 26 weeks (3 injections) in patients with PMO. This study will consist of a screening period (up to 4 weeks) and a 52-week double-blind main treatment period, followed by a 26-week double-blind transition period.

The total duration of the study for each patient is approximately 82 weeks.

Screening will take place within 4 weeks before the first dose. After informed consent is obtained, eligibility criteria will be reviewed, and screening evaluations will be performed.

At baseline, patients will be randomized in a 1:1 ratio to receive the first 2 doses of TVB-009P or PROLIA US ("main treatment period"). The first dose of TVB-009P or PROLIA US will be administered following randomization. The second dose will be administered 26 weeks after the first dose.

At week 52 (26 weeks after the second dose and prior to receiving their third dose), patients in the PROLIA US arm will be re-randomized 1:1 to either continue with a third dose of PROLIA US or transition to TVB-009P and receive a single dose of TVB-009P in the transition period to assess primarily immunogenicity and safety after a transition from PROLIA US to TVB-009P. All patients in the TVB-009P group will continue treatment with a third dose of TVB-009P. All patients who do not terminate the study before the third dose will be followed for 26 weeks after the third dose of study drug.

Final procedures and assessments will be performed at the end of study (EOS) visit at the end of the 78-week study period. Patients who withdraw from the study before completing the 78-week study period will have early termination (ET) procedures and assessments performed at their final visit. A patient who is randomized but does not complete all treatment periods will not be replaced. In case the assessment of LS-BMD is missing or not evaluable from more than 10% of the patients at the week 26 visit, the sponsor may decide to continue the recruitment of patients in order to increase the number of evaluable patients.

The EOS is defined as the last visit of the last patient of the transition period.

The assessments and procedures performed during each study visit are detailed in Table 3 in section 7.1.

2.4 Determination of Sample Size

[redacted]

[redacted]

2.5 Randomization and Blinding

This is a randomized, double-blind study. Patients and investigators will remain blinded to IMP assignment during the study.

At baseline, patients will be randomized in a 1:1 ratio to receive the first 2 doses of TVB-009P or PROLIA US ("main treatment period"). The randomization will be stratified by region (US/non-US) and any use of previous bisphosphonates (yes/no). At week 52, prior to receiving their third dose of study medication, patients in the PROLIA US treatment group will be re-randomized in a 1:1 ratio to receive a third dose of PROLIA US or switch to TVB-009P ("transition period") to receive a single dose of TVB-009P. The re-randomization will be stratified by region (US/non-US) and any use of previous bisphosphonates (yes/no). All patients in the TVB-009P group will receive a third dose of TVB-009P at week 52.

During the main treatment period, the sponsor, investigators (and other site staff involved in study assessments) and patients will be blinded to the treatment assignment of all patients. To maintain blinding of the patients and investigators (and other site staff involved in study assessments) to the treatment assignment of all patients during the entire study, and maintain blinding of the sponsor until the main treatment period is unblinded, the re-randomization process will be performed for all patients, including the patients in the TVB-009P group (although only patients in the PROLIA US group will actually be re-randomized while patients in the TVB-009P group will continue to receive

TVB-009P in the transition period). The randomization will be implemented using Randomization and Trial Supply Management (RTSM).

Before database lock (DBL), staff responsible for efficacy and safety analysis, pharmacokinetic and immunogenicity bioanalysis, population pharmacokinetic analysis, and/or pharmacokinetic/pharmacodynamic analysis will not have access to the patient treatment randomization. After last patient last visit and DBL of the main treatment period, the sponsor will unblind the treatments for the analysis of the main treatment period (up to and including week 52; not including third IMP dose and assessments following the third dose). A full CSR will be prepared with all data related to the first and second doses up to the week 52 end of main treatment phase procedures. Only after completion of the study (after week 78) and DBL will the transition period be fully unblinded and analyzed. The results will be reported separately in a CSR addendum, including any updates to the safety analysis of the main treatment period.

3 DATA ANALYSIS CONSIDERATION

3.1 General Principles

The statistical analyses will be performed by [redacted] with approval of the Sponsor, using SAS Version 9.4 (or higher). All tables, figures and listings will be produced in landscape format.

In general, all data will be listed by subject and visit/time point where appropriate. The summary tables will be stratified by, or have columns corresponding to, treatment groups (see section 3.4).

The total number of subjects in the treatment group (N) under the specified analysis set will be displayed in the header of summary tables.

Data will be summarized using descriptive statistics for continuous variables. Unless otherwise specified, descriptive statistics will include number of subjects, mean, standard deviation, minimum, median and maximum. Number of subjects with missing values will also be displayed, but only if non-zero. If the number of subjects (N) is less than or equal to 5, only N, median, minimum and maximum will be displayed and N/A will be shown for the mean and SD. The minimum and maximum statistics will be presented to the same number of decimal places as the original data. The mean and median will be presented to one more decimal place than the original data. The standard deviation will be presented to two more decimal places than the original data.

In summary tables of categorical variables, counts and percentages will be displayed. The count [n] indicates the actual number of subjects in a particular category, which should always be less than or equal to the total number of subjects in the respective study group [M]. Percentage will be obtained by: % = n/M*100. Unless otherwise specified, all percentages will be expressed to one decimal place.

Baseline will be defined as the last assessment, scheduled or not, prior to the first dose of the study drug, unless otherwise specified.

Any post-baseline BMD assessment by DXA must be made by a machine of the same manufacturer as the baseline assessment of the subject. Post-baseline BMD assessments made by a machine of a different manufacturer (if any) will be considered invalid and treated as missing (as if not done at all) for all analyses and for inclusion in analysis populations.

Unscheduled and early termination (ET) visits will be mapped to the closest scheduled visit in the period according to the allowed time window specified in in Table 3 in section 7.1 when no assessment from the schedule visit exists.

In by-visit summaries, only data collected on scheduled visits/timepoints will be summarized (including unscheduled visits remapped to scheduled ones per the previous paragraph). Data from unscheduled assessments that were not remapped to scheduled visits will be included in listings and may be used in determination of baseline if applicable.

Relative days will be calculated relative to date of first dose of the study drug. Relative days will be calculated as follows only when the full assessment date is known (i.e., partial dates will have missing relative days).

For assessment on or after the day of first dose of the study drug:

Relative Day = Date of Assessment – Date of First Dose of Study Drug+1.

For assessment before the day of first dose of the study drug:

Relative Day = Date of Assessment – Date of First Dose of Study Drug.

Additionally transition period days will be calculated for data collected in the transition period as

Transition Period Day = Date of Assessments – Date of Transition Period Dose + 1

Transition period days will be presented in listings of transition period data.

All dates will be displayed in DDMMMYYYY format.

3.2 Coding Dictionaries Used

Medical history and adverse events will be coded using Medical Dictionary for Regulatory Activities (MedDRA) version 23.1.

Prior and concomitant medication will be coded using World Health Organisation-Drug Reference List (WHO-DRL) version September 2020 and by Anatomical Therapeutic Chemical (ATC) (the highest level available).

3.3 Analysis Sets

3.3.1 Intent-to-Treat Analysis Set

The intent-to-treat (ITT) analysis set will include all randomized patients.

In the ITT analysis set, treatment will be assigned based on the treatment to which patients were randomized, regardless of which treatment they actually received.

3.3.2 Modified Intent-to-Treat Analysis Set

The modified intent-to-treat (mITT) analysis set will include all randomized patients who received at least 1 dose of IMP and had at least 1 post-baseline evaluation of LS-BMD.

Patients who withdraw from the study prior to week 26 will not have a post-baseline LS-BMD measurement and will therefore not be included in the mITT analysis set.

In the mITT analysis set, treatment will be assigned based on the treatment to which patients were randomized, regardless of which treatment they actually received.

3.3.3 Transition Intent-to-treat Analysis Set

The transition intent-to-treat (TITT) analysis set will include all patients re-randomized in the transition period.

In the TITT analysis set, treatment will be assigned based on the treatment to which patients were randomized in the main and transition treatment periods, regardless of which treatment they actually received.

3.3.4 Transition Modified Intent to Treat Analysis Set

The transition modified intent-to-treat (TmITT) analysis set will include all patients who received the third dose of IMP and had an end-of-study evaluation of LS-BMD.

In the TmITT analysis set, treatment will be assigned based on the treatment to which patients were randomized in the main treatment period and in the transition period, regardless of which treatment they actually received.

3.3.5 Safety Analysis Set

The safety analysis set will include all randomized patients who received at least 1 dose of IMP.

In the safety analysis set, treatment will be assigned based on the treatment patients actually received, regardless of the treatment to which they were randomized, unless otherwise specified.

Treatment assignment in the main treatment period in case of potential mixed actual dosing will be discussed on a case-by-case basis and documented prior to the database lock and unblinding for analysis.

3.3.6 Transition Safety Analysis Set

The transition safety analysis set will include all patients who received the third dose of the IMP.

In the transition safety analysis set, treatment will be assigned based upon the treatment patients actually received during and transition period, regardless of the treatment to which they were randomized.

3.3.7 Per-Protocol Analysis Set

The per-protocol (PP) analysis set is a subset of the mITT analysis set including only patients who received treatment with IMP at the baseline and week 26 visit as randomized and completed the main treatment period without any major protocol deviations that may impact the similarity assessments of LS-BMD and sCTX-1.

The exclusion of patients from the PP analysis set will be discussed on a case-by-case basis and documented prior to the database lock and unblinding for analysis.

3.4 Treatment Groups

Summaries of the baseline characteristics and assessments in the main treatment period will be presented by the following treatment groups:

- TVB-009P
- PROLIA US

Summaries of assessments in the transition period will be presented by the following treatment groups, based on the subject's treatment in the main treatment period and transition period:

- TVB-009P/TVB-009P
- PROLIA US/PROLIA US
- PROLIA US/TVB-009P

Summaries of assessments in the overall treatment period are planned for the safety analyses only and will include only subjects that stayed on the same treatment throughout the study. They will therefore include the following treatment groups:

- TVB-009P/TVB-009P
- PROLIA US/PROLIA US

Total treatment group can be included where appropriate.

3.5 Handling Withdrawals and Missing Data

Data imputation is planned for the primary endpoint analysis, see section 4.6.3 for details. The missing assessments of LS-BMD or sCTX-1 (either due to withdrawal from the study or for other reasons) are assumed to be missing at random (MAR) in this study.

The mITT analysis set will be used for the primary analysis. This analysis set includes only patients that had at least 1 post-baseline evaluation of LS-BMD (see Section 3.3.2), i.e., reached week 26; the drop-out rate prior to week 26 is assumed to be low, [redacted]. Underlying assumptions are that the drop-out rate prior to week 26 will be similar in both treatment groups, and that drop-outs are not related to efficacy. As both treatment groups are active, and improvement in the underlying disease is not expected to be felt by the patient, these assumptions are considered reasonable. The assumption that drop-out rates are comparable between the treatment groups will be assessed using descriptive statistics (see section 4.1).

In the primary analysis, missing LS-BMD at Week 52 will be imputed based on available assessments using multiple imputation under the MAR assumption (see Section 4.6.3.1); this is a conservative approach for similarity testing, as missing data will be imputed within each treatment group separately.

The imputation methodology is reliable as long as the missing LS-BMD rate (i.e. percentage of patients with missing LS-BMD) at Week 52 is low. Sensitivity and supplementary analyses for missing data in the primary analysis are presented in section 4.6.4 and section 4.6.5.

Since the mITT analysis set includes only patients that reached at least week 26, missing sCTX-1 at week 26 is expected to be negligible and will not be imputed.

For all the other variables, only the observed data from the patients will be used in the statistical analyses, i.e., there is no plan to estimate missing data, unless otherwise specified.

3.6 Multiple Comparisons and Multiplicity

Formal statistical analysis of similarity will be performed only on the following endpoints:

(1) percent change from baseline to week 52 in LS-BMD

(2) percent change from baseline in sCTX-1 at week 26

For the US filing, percent change from baseline to week 52 in LS-BMD is the primary endpoint. Percent change from baseline in sCTX-1 at week 26 is a secondary endpoint. For this submission, the percent change from baseline in sCTX-1 at week 26 will be tested only if the primary endpoint was found to be statistically similar (i.e., met the similarity criteria).

For the EU submission, percent change from baseline in LS-BMD at week 52 and percent change from baseline in sCTX-1 at week 26 are co-primary endpoints. For this submission, success will be declared if similarity is demonstrated for both co-primary endpoints.

The analysis of all other endpoints will be descriptive, therefore no control for multiplicity is needed.

3.7 Sequence of Final Analyses

Upon completion of the main treatment period (up to and including week 52) the sponsor will unblind the main treatment period (up to and including week 52; not including third IMP dosing and

assessments following the third dosing), while keeping the investigators (and other site staff involved in study assessments) and the patients blinded (see section 2.5). A CSR that contains the results of the analysis of the main treatment period (up to and including week 52) will be submitted with the filing of the Biologics License Application (BLA) and/or Marketing Authorisation Application (MAA). The results of the analysis of the transition period (up to and including week 78) will be included in a CSR addendum.

4 ANALYSIS METHODS

4.1 Disposition

The number of patients screened, patients screened but not randomized, and patients randomized in the main treatment period and in the transition period will be presented by treatment group.

For the main treatment period, the summary will include number and percentage of patients who were randomized (ITT analysis set), included in the mITT, safety, and PP analysis sets, patients that withdrew up to week 26, and patients that completed the main treatment period. Patients that withdrew will also be summarized by reason for withdrawal as recorded in the disposition eCRF. Patient lost to follow-up will be further summarized by country. The percentages will be based on the number of randomized patients.

For the transition period, the summary will include number and percentage of patients randomized (TITT analysis set), patients in the TmITT, and transition safety analysis sets, patients that withdraw during the transition period, and patients that complete the study. Patients that withdrew will also be summarized by reason for withdrawal as recorded in the disposition eCRF. The percentages will be based on the number of patients who were re-randomized for the transition period.

Number of subjects who discontinued due to reasons related to COVID-19 pandemic will be provided separately, as applicable.

If more than 10% of the patients withdraw from the study before the end of the main treatment period, Kaplan-Meier curves for the number of days until study discontinuation will be plotted by treatment group for the main and transition period, as applicable.

All disposition information will be listed. Also, a separate listing will present the date of informed consent and inclusion/exclusion criteria violated, if any.

4.2 Demography and Baseline Characteristics

Patient demographic and baseline characteristics will include age, race, ethnicity, baseline height, weight, and BMI, baseline LS-BMD, baseline sCTX-[redacted] presence of prior fractures (yes/no) and whether the patient is taking

calcium and vitamin D supplements. Demographic and baseline characteristics will be summarized using descriptive statistics by treatment group based on the safety, ITT, and mITT analysis sets.

For continuous variables descriptive statistics will be provided. For categorical variables patient counts and percentages will be provided.

All demographic and baseline characteristics will be listed.

4.3 Medical History

Medical history will be summarized by the number and percentage of subjects by MedDRA system organ class, preferred term, and treatment group for the safety analysis set.

Subjects will be counted only once for each applicable preferred term, and each applicable system organ class. System organ classes and preferred terms will be presented in alphabetical order.

All medical history information will be listed.

4.4 Protocol Deviations

Protocol deviations will be recorded in the eCRF during the study and identified programmatically.

Each protocol deviation will be classified as minor or major. Major protocol deviations will exclude the subject from the PP analysis set. Specific deviation and their severities are defined in the separate Protocol Deviations Specification document. The exclusionary status of each deviation will be reviewed and approved by the sponsor prior to the database lock.

All major protocol deviations will be summarized by deviation category and treatment group. This analysis will be performed for the ITT analysis set. All classification of protocol deviations will be performed prior to database lock.

Deviations related to the COVID-19 pandemic will be identified.

All protocol deviations will be listed.

4.5 Study Drug Exposure and Treatment Compliance

Number and percentage of subjects receiving 1 or 2 doses in the main treatment period will be presented by treatment group for the ITT analysis set. Similarly, number and percentage of subjects receiving the dose in the transition period will be presented by treatment group for the TITT analysis set.

Number and percentage of subjects receiving each dose in full, at least 50% and less than 50% will also be presented.

All details of study drug administration captured on the eCRF will be listed.

4.6 Efficacy Analysis

The mITT analysis set will be used as the primary analysis set for efficacy in the main treatment period.

Supplementary analysis will be carried out using the ITT and PP analysis sets.

The TmITT analysis set will be used as the primary analysis set for efficacy in the transition period.

4.6.1 Primary Endpoint

The primary efficacy endpoint for the US filing is the percent change from baseline in LS-BMD at Week 52, mean difference between TVB-009P and PROLIA US treatment groups.

The co-primary endpoints for the EU filing are: (1) percent change from baseline in LS-BMD at week 52, mean difference between TVB-009P and PROLIA US treatment groups; (2) percent change from baseline in sCTX-1 at week 26, mean difference between TVB-009P and PROLIA US treatment groups.

4.6.2 Secondary Endpoints

4.6.2.1 Secondary Efficacy and Pharmacodynamic Endpoints in the Main Treatment Period

The secondary efficacy and pharmacodynamic endpoints in the main treatment period are:

- percent change from baseline in LS-BMD at week 26 by DXA
- percent change from baseline in femoral neck BMD by DXA at week 26 and at week 52
- percent change from baseline in total hip BMD by DXA at week 26 and at week 52
- percent change from baseline in sCTX-1 at all time points
- sCTX-1 suppression at week 4 (defined as sCTX-1 level below the limit of quantification)
- percent change from baseline in P1NP at week 26 and week 52
- incidence of fractures up to week 52

The comparisons will be made between the TVB-009P and PROLIA US groups.

4.6.2.2 Secondary Efficacy and Pharmacodynamic Endpoints in the Transition Period

The secondary efficacy and pharmacodynamic endpoints in the transition period are:

- percent change from week 52 in LS-BMD by DXA at week 78

- difference between percent change from baseline in sCTX-1 between week 52 and week 78
- percent change from week 52 in femoral neck BMD by DXA at week 78
- percent change from week 52 in total hip BMD by DXA at week 78
- difference between percent change from baseline in P1NP between week 52 and week 78
- incidence of fractures up to week 78

The comparisons will be made between the PROLIA US/PROLIA US and PROLIA US/TVB-009P groups.

4.6.3 Primary Efficacy Analysis

The primary analysis of LS-BMD percent change from baseline at week 52 will be an analysis of covariance (ANCOVA) model [redacted] Similarity will be demonstrated if the 95% CI for the LS mean difference between TVB-009P and PROLIA US falls entirely within the similarity margin of ±1.45%.

The primary analysis of sCTX-1 percent change from baseline at week 26 will be an ANCOVA model [redacted] Similarity will be demonstrated if the 95% CI for the LS mean difference between TVB-009P and PROLIA US falls entirely within the similarity margin of ±20%.

Missing LS-BMD at Week 52 will be imputed as described in the section 4.6.3.1.

Very few missing sCTX-1 assessments at Week 26 are expected in the mITT analysis set, and therefore not imputation for missing values of this endpoint will be performed. However, sCTX-1 values that are below the limit of quantification (BLQ) will be imputed as the low limit of quantification (LLOQ).

The primary analysis will be based on the mITT analysis set.

4.6.3.1 Missing LS-BMD Multiple Imputation

Missing LS-BMD percent change from baseline at Week 52, due to early discontinuation or any other reason, will be imputed as follows.

The imputation will be performed on all subjects in the mITT analysis set.

The imputation will proceed in two steps [redacted] It is expected that very few subjects will have intermittent missing value at Week 26. [redacted] The resulting dataset with monotone missing pattern will be used in the next step, in which trailing missing values will be imputed using the monotone regression predictive mean matching multiple imputation method (Heitjan and Little 1991, Schenker and Taylor 1996), for each treatment arm separatel [redacted]

The resulting complete, imputed datasets will each be analyzed using the model specified in the section 4.6.3, and the resulting statistics combined using methodology provided by Rubin (1987) and Little and Rubin (2002). See section 7.2.1 for sample SAS code.

4.6.4 Sensitivity Analysis for the Primary Analysis

To assess the robustness of the primary efficacy analysis using the same estimand, sensitivity analyses will include:

Sensitivity analyses for the statistical model:

- primary model, but including only treatment group as a covariate, will be performed for both LS-BMD and sCTX-1 endpoints. For LS-BMD endpoint this analysis will be based on the same multiple-imputed dataset as the primary analysis.

Sensitivity analysis for BLQ imputation:

For the primary analysis of sCTX-1 percent change from baseline at week 26, values BLQ will be imputed as the LLOQ (0.033 ng/mL). Sensitivity analyses for this imputation will be performed as follows:

- two-dimensional tipping-point: for the analysis of sCTX-1 percent change from baseline at week 26, sensitivity analysis for the BLQ imputation will be conducted by repeating the analysis using different imputed values for patients in the TVB-009P group and the PROLIA US group, in the range of 0 to LLOQ (0.033 ng/mL). Values of 0, 0.011, 0.022 and 0.033 will be used for each treatment groups separately, for a total of 16 combinations.

4.6.5 Supplementary Analysis for the Primary Analysis

Supplementary analyses to evaluate effect of assumptions on missing data on the results:

- two-dimensional tipping-point (missing not at random [MNAR]): in order to assess the sensitivity of the primary analysis to the MAR assumption on missing data for patients that have at least 1 post-baseline value, supplementary analysis for the primary analysis will be conducted using multiple imputation under MNAR assumption. In this analysis, missing LS-BMD percent change from baseline at week 52 will be imputed using the same predictive mean matching multiple imputation method as in the primary analysis, adjusted under MNAR assumption:

  - the percent change from baseline in LS-BMD at week 52 in patients randomized to TVB-009P with missing LS-BMD at week 52 will be imputed assuming the treatment effect is worsened by δ_1 compared to the patients who have no missing value (where $\delta_1 = 0$ to 2% or estimated treatment effect the TVB-009P group, whichever is higher, it steps of 0.5%);

  - the percent change from baseline in LS-BMD at week 52 in patients randomized to PROLIA US with missing LS-BMD at week 52 will be imputed assuming the treatment effect is worsened by δ_2 compared to the patients who have no missing value (where $\delta_2 = 0$ to 2% or estimated treatment effect the PROLIA US group, whichever is higher, in steps of 0.5%);

The resulting complete, imputed datasets will each be analyzed using the same model as the primary analysis model, and the resulting statistics combined using methodology provided by Rubin (1987) and Little and Rubin (2002). See section 7.2.3 for sample SAS code.

- to further alleviate the concern on the uncertainty introduced by missing data, the following 2 separate 1-sided tests of alpha=0.05 with missing data imputed under the corresponding null using a multiple imputation method will be conducted:

  - in the first test missing values for the TVB-009P group will be imputed assuming the treatment effect is worsened (i.e. LS-BMD percent change from baseline decreased) by the equivalence margin value of 1.45%, while the missing values for the PROLIA US group are imputed without penalization. Non-inferiority will then be tested by checking that the lower 95% confidence limit for the mean difference TVB-009P – PROLIA US is greater than the margin value of -1.45%.

  - in the second test missing values for the TVB-009P group will be imputed assuming the treatment effect is improved (i.e. LS-BMD percent change from baseline increased) by

the equivalence margin value of 1.45%, while the missing values for the PROLIA US group are imputed without penalization. Non-superiority will then be tested by checking that the upper 95% confidence limit for the mean difference TVB-009P – PROLIA US is less than the margin value of 1.45%.

The resulting complete, imputed datasets will each be analyzed using the same model as the primary analysis model, and the resulting statistics combined using methodology provided by Rubin (1987) and Little and Rubin (2002). See section 7.2.4 for sample SAS code.

- primary analysis repeated on the ITT analysis set, with the same multiple imputation approach as for the primary analysis but repeated on the entire ITT analysis set.

- Two-dimensional tipping-point (MNAR) repeated on the ITT analysis set

Other analyses:

- primary model repeated for the PP analysis set without imputations

- primary model repeated for the mITT analysis set, excluding patients who received IMP not as randomized at baseline and/or at week 26

4.6.6 Secondary Efficacy Analysis

No formal hypothesis testing is planned for the secondary efficacy endpoints. Analysis will be descriptive in nature.

4.6.6.1 Continuous endpoints

For continuous endpoints such as percent change from baseline in LS-BMD, femoral neck BMD, total hip BMD, sCTX-1, and P1NP at weeks 26 and 52 (as applicable) descriptive statistics will be presented by treatment group and visit.

4.6.6.2 Binary endpoints

For binary endpoints such as sCTX-1 suppression or incidence of fractures, number and percentage of subjects achieving the endpoint will be presented by treatment group. For descriptive purposes, 95% CIs for the differences in percentages between treatment groups will be presented.

Vertebral fractures reported by sites in the eCRF, vertebral fractures identified by the central reader of X-rays (new fractures not found at Screening) and non-vertebral fractures reported by sites in the eCRF will be summarized separately. Vertebral fractures identified by the central reader and non-vertebral fractures will also be summarized together. Subjects with vertebral fractures identified by the central reader will also be presented by the highest fracture grade (Grade 1, 2 or 3).

4.6.7 Efficacy Analysis in the Transition Period

Descriptive statistics will be presented by the treatment groups to which the patients were assigned in the main treatment and transition periods (TVB-009P/TVB-009P, PROLIA US/PROLIA US and PROLIA US/TVB-009P).

For the binary endpoint of incidence of fractures, number and percentage of subjects with a fracture in the transition period will be presented by treatment group. For descriptive purposes, 95% CIs for the differences in percentages between the PROLIA US/TVB-009P and PROLIA US/PROLIA US treatment groups will be presented.

Vertebral fractures reported by sites in the eCRF, vertebral fractures identified by the central reader of X-rays (new fractures in the transition period, not found at Week 52) and non-vertebral fractures reported by sites in the eCRF will be summarized separately. Vertebral fractures identified by the central reader and non-vertebral fractures will also be summarized together. Subjects with vertebral fractures identified by the central reader will also be presented by the highest fracture grade (Grade 1, 2 or 3).

The efficacy analyses in the transition period will be based on the TmITT analysis sets.

All efficacy analyses in the transition period are considered descriptive and no formal hypothesis testing is planned.

4.7 Safety Analysis

Safety analyses will be performed on the safety and transition safety analyses sets.

The safety of TVB-009P and PROLIA US will be assessed throughout the study by evaluating adverse events, clinical laboratory test results, vital signs measurements, ECG, physical examination results, local tolerability, and concomitant medication usage.

4.7.1 Safety Analysis in the Main Treatment Period

Safety analyses in the main treatment period will be performed on the safety analysis set. Summaries will be presented by treatment group (TVB-009P and PROLIA US) and for all patients.

All safety variables at the Week 52 Visit that are assessed prior to IMP administration will be considered as occurring during the main treatment period.

4.7.2 Safety Analysis in the Transition Period

Safety analyses in the transition period will be performed on the transition safety analysis set. Summaries will be presented by the treatment groups to which the patients were assigned in the main treatment and transition periods (TVB-009P/TVB-009P, PROLIA US/PROLIA US and PROLIA US/TVB-009P) and for all patients.

4.7.3 Safety Analysis in the Overall Treatment Period

Safety analyses in the overall treatment period will be performed on the safety analysis set. The analyses will include only patients in the TVB-009P/TVB-009P and PROLIA US/PROLIA US treatment groups and for all patients.

The analyses will be similar to the analyses of the main treatment period. Summaries will be presented by treatment group and for all patients included in the analysis.

4.7.4 Adverse Events

Adverse Events will be coded using the Medical Dictionary of Regulatory Activities (MedDRA) AE coding system for purposes of summarization.

Only Treatment Emergent Adverse Events (TEAE) will be used for the summary analysis. An AE will be considered as treatment-emergent if the date of onset is on or after the first study drug administration date. AEs with unknown start dates will be counted as treatment-emergent unless the AE resolution date is prior to the first study drug administration date. If the AE start date is partially missing, the AE will be considered treatment-emergent, unless the month and year (when available) rule out the possibility that the event occurred post start of the study drug.

All TEAEs will be assigned to either main treatment period or transition period based on the date of onset. In case a TEAE has a partial start date so that is ambiguous whether the AE started in the main treatment period or transition period, it will be assigned to the main treatment period.

In summaries of TEAEs a subject experiencing the same AE (with the same preferred term) multiple times within the same study period will only be counted once for that preferred term and study period. Similarly, if a subject experiences multiple AEs within the same system organ class in the same study period, that subject will be counted only once in that system organ class for that study period. When summarizing AEs by severity, only the most severe occurrence within the preferred term or system organ class and study period will be used. Similarly, when summarizing AEs by relationship to study drug, only the most related occurrence within the preferred term or system organ class and study period will be selected for displays in summary tables.

AEs will be summarized for the main treatment period, transition period and the overall treatment period.

An overall summary will include, by study period and by treatment group and overall, the number and percentage of subjects reporting at least 1 TEAE, as well as the number of TAEs, in the following categories:

- Any TEAE
- Treatment-related TEAE

- Device-related TEAE
- Serious TEAE
- TEAE leading to discontinuation of the study drug
- TEAE leading to death.

The following TEAE frequency tables will be prepared summarizing the overall number of TEAEs, the number and percentage of subjects reporting at least one TEAE by MedDRA System Organ Class (SOC) and preferred term (PT), by treatment group and by study period:

- All TEAEs
- Serious TEAEs
- Treatment-related TEAEs
- Device-related TEAEs
- TEAEs leading to discontinuation of the IMP
- TEAEs in subjects with COVID-19, as applicable
- TEAEs by severity
- Treatment-related TEAEs by severity

The summaries by SOC and PT will be ordered alphabetically by SOC and PT.

Separately all TEAEs will be presented by PT only, sorted in descending order of frequency in the Total treatment group.

All information pertaining to adverse events noted during the study will be listed by subject, detailing verbatim, preferred term, system organ class, start date, stop date, severity, outcome, action taken and causal relationship to the study drug. Separate listings will be prepared for serious AEs and AEs leading to discontinuation of the IMP.

4.7.5 Laboratory Evaluations

Laboratory testing will be performed according to the schedule of assessments (see section 7.1).

Values and changes from baseline in hematology, chemistry (including vitamin D) and urinalysis results will be summarized descriptively by visit and treatment group for the main treatment period and overall treatment period. Changes from Week 52 will be presented for the transition period.

Additionally, numeric hematology, chemistry (including vitamin D) and urinalysis results will be classified as Low (below the reference range), Normal (within the reference range) or High (above

the reference range). Categorical Urinalysis results will be classified as Normal or Abnormal. Shifts among these categories between baseline and last available post-baseline assessment will be provided for the main treatment period and overall treatment period for the safety analysis set. Similarly shifts from Week 52 to the last available assessment in the transition period will be provided for the transition safety analysis set.

Potentially clinically significant values will be summarized and listed separately.

Number and percentage of subjects with potentially clinically significant abnormal values will be summarized for the main treatment period, transition period and overall treatment period using the criteria specified in Table 1.

Table 1: Criteria for Potentially Clinically Significant Laboratory Values

Test		Criterion / value
Serum chemistry		
Alanine aminotransferase (ALT)		≥3 x ULN
Aspartate aminotransferase (AST)		≥3 x ULN
Alkaline phosphatase		≥3x ULN
Gamma-glutamyl transpeptidase (GGT)		≥3x ULN
Lactate dehydrogenase (LDH)		≥3x ULN
Creatinine		≥177 µmol/L
Uric acid	Men	≥625 µmol/L
	Women	≥506 µmol/L
Bilirubin (total)		≥34.2 µmol/L
Hematology		
Hematocrit	Men	<0.37 L/L
	Women	<0.32 L/L
Hemoglobin	Men	≤115 g/L
	Women	≤95 g/L
White blood cell (WBC) counts		≤3 x 10^9/L ≥20 x 10^9/L
Eosinophils		≥10%
Platelet counts		≤75 x 10^9/L ≥700 x 10^9/L

ULN=upper limit of normal range

All results will be listed. A separate listing of subjects who tested positive for COVID-19, if any, will be provided.

4.7.6 Vital Signs

Vital signs will be collected according to the schedule of assessments (see section 6.1). Pulse rate, blood pressure (systolic/diastolic), and respiratory rate will be recorded. Weight will be measured as part of the physical examination.

Vital signs (including weight) and their changes from baseline will be summarized descriptively by visit and treatment group for the main treatment period and overall treatment period . Changes from Week 52 (visit 9) will be presented for the transition period.

Vital signs will be classified as Normal, Abnormal not clinically significant or Abnormal clinically significant by the investigator, and this classification will be captured in the eCRF. Number and percentage of subjects in each category will be presented by visit and treatment group for the main treatment period and transition period. Shifts among these categories between baseline and last available post-baseline assessment will be provided for the main treatment period and overall treatment period for the safety analysis set. Similarly shifts from Week 52 to the last available assessment in the transition period will be provided for the transition safety analysis set.

Number and percentage of subjects with potentially clinically significant abnormal values will be summarized using the criteria specified in Table 2 for the main treatment period, transition period and overall treatment period. These summaries will include all post-baseline values (including scheduled, unscheduled, and early termination time points/visits). Note that in order to qualify as potentially clinically significant abnormal, a value needs to meet both criteria below: i.e., have a value beyond the criterion value and a change of at least the magnitude specified in the change relative to baseline column. For the transition period, summaries Week 52 will serve as baseline.

Table 2: Criteria for Potentially Clinically Significant Vital Signs

Vital Sign	Criterion value	Change relative to baseline
Pulse	≥120 bpm	Increase of ≥15
	≤50 bpm	Decrease of ≥15
Systolic blood pressure	≥180 mm Hg	Increase of ≥20
	≤90 mm Hg	Decrease of ≥20

Table 2: Criteria for Potentially Clinically Significant Vital Signs

Vital Sign	Criterion value	Change relative to baseline
Diastolic blood pressure	≥105 mm Hg	Increase of ≥15
	≤50 mm Hg	Decrease of ≥15

All vital signs will be listed.

4.7.7 Electrocardiogram

ECG will be performed according to the schedule of assessments (see section 6.1). Standard ECGs parameters will be recorded, and the ECG will be interpreted locally by the principal investigator (or qualified physician). Only ECG interpretation (Normal, Abnormal Not Clinically Significant or Abnormal Clinically Significant) will be entered into the study database and used in analysis.

Overall interpretation will be summarized categorically by visit and treatment group.

Shifts in overall interpretation from baseline to the last available post-baseline assessment will be tabulated.

All results will be listed.

4.7.8 Physical Examination

Physical examination will be performed according to the schedule of assessments (see section 6.1).

The physical examination will include, at a minimum, examination of head, eyes, ears, nose, and throat [HEENT], oral cavity, chest, cardiovascular system, abdomen, and skin.

Each system will be classified as Normal, Abnormal Not Clinically Significant (NCS) or Abnormal Clinically Significant (CS).

Number and percentage of subjects with each assessment result will be tabulated by body system, visit and treatment group, for the main treatment period using the safety analysis set and for the transition period using the transition safety analysis set.

All results will be listed.

4.7.9 Assessment of Local Tolerability and Pain

Local tolerability will be assessed according to the schedule of assessments (see section 6.1).

Number and percentage of subjects with each severity level of injection site tolerability signs will be presented by visit/timepoint and treatment group.

Patient-reported pain NRS will be summarized descriptively by visit/timepoint and treatment group.

The summary will be presented for the main treatment period using the safety analysis set and for the transition period using the transition safety analysis set.

4.7.10 Prior and Concomitant Medications

Prior medications are defined as medications that were taken prior to subject's first dose of the IMP. Concomitant medications are defined as medications taken after the first dose of the IMP, i.e. either stopped after the first dose of the IMP or ongoing.

A concomitant medication will be associated with the main treatment period, if it is taken during that period, i.e. prior to the first transition period treatment. A concomitant medication will be associated with the transition period if it is taken after the first transition period treatment. Concomitant medications taken both prior to and after the first transition period treatment will be associated with both periods.

Prior medications will be summarized for the ITT analysis set. Concomitant medications will be summarized separately for the main treatment period for the safety analysis set and for the transition period for the transition period analysis set.

In all cases prior and concomitant medications will be summarized by ATC class (highest level available), WHO Drug Dictionary preferred name and treatment group. One subject will be counted once for each applicable preferred name and ATC class. ATC classes and preferred names will be presented alphabetically.

All prior and concomitant medications will be presented in separate listings.

4.8 Pharmacokinetic Analysis

PK concentrations will be summarized descriptively by visit and treatment group for the safety analysis set in the main treatment period and transition safety analysis set in the transition period.

The following exploratory pharmacokinetic parameters will be summarized using descriptive statistics:

- serum concentration before next dose (C_{trough}), before second and third dose, and 6 months after third dose
- serum concentration at 2 weeks postdose (C_{2weeks})

Individual data will be listed.

Details of the pharmacokinetic analysis of the area under the curve over the dosing period ($AUC_{0\text{-}tau}$) after first dose derived from population pharmacokinetics analysis, C_{trough} before second dose derived from population pharmacokinetics analysis, and maximum observed concentration (C_{max}) after first dose derived from population pharmacokinetic analysis will be given in a separate TVB-009P population pharmacokinetic statistical analysis plan. The results of the population pharmacokinetic analysis will be provided in a separate report.

4.9 Immunogenicity Analysis

If more than 5 patients develop ADA anytime post baseline, the incidence of ADA positive will be summarized by treatment group, and ADA positive/negative, titer level, and neutralizing ADA positive/negative will be summarized at each visit using descriptive statistics. In addition, the incidence of ADA positive anytime within the period will be summarized.

Results will be summarized in the main treatment period for the safety analysis set and in the transition period for the transition safety analysis set.

Results of immunogenicity assessment will be listed.

The effect of positive immunogenicity findings on efficacy and safety may be investigated, if applicable.

5 CHANGES TO ANALYSES SPECIFIED IN PROTOCOL

There are no changes to the analyses specified in the protocol.

6 REFERENCES

1. Study Protocol: “A Randomized, Double-Blind, Multinational, Multicenter Study to Compare Efficacy, Safety, and Immunogenicity of TVB-009P and Denosumab (PROLIA®) in Patients with Postmenopausal Osteoporosis”
2. Heitjan, F., and Little, R. J. A. (1991). “Multiple Imputation for the Fatal Accident Reporting System.” *Journal of the Royal Statistical Society, Series C* 40:13–29.
3. Little RJA, Rubin DB. Statistical Analysis with Missing Data, Second Edition. 2002, New York: John Wiley & Sons.
4. Rubin DB. Multiple Imputation for Nonresponse in Surveys. 1987, New York: John Wiley & Sons.
5. Schenker, N., and Taylor, J. M. G. (1996). “Partially Parametric Techniques for Multiple Imputation.” *Computational Statistics and Data Analysis* 22:425–446.

7 APPENDICES

7.1 Study Schedule

Table 3: Study Procedures and Assessments

Study period	Screening	Baseline	Main treatment period								Transition period		
Visit number	V1	V2	V3[a]	V4[a]	V5[a]	V6[a]	V7	V8[a]	V9[b] EOM	V9[b] SOT	V10[a]	V11[a]	V12 EOS/ET
Day/week and allowed time windows	Up to 4 weeks before V2	Day 1	Day 15 ±3 days	Week 4 ±3 days	Week 8 ±5 days	Week 12 ±7 days	Week 26 ±7 days	Week 39 ±14 days	Week 52 ±14 days		Week 54 2 weeks ±3 days after V9	Week 65 13 weeks ±14 days after V9	Week 78 26 weeks ±14 days after V9
Procedures and assessments													
Informed consent	X												
Inclusion and exclusion criteria	X	X											
Demographics/medical history	X												
Vitamin D and calcium dispensing[c]	X												
Pregnancy test (urine dipstick)		X[d]											
COVID-19 viral test[e]	X												
Randomization		X								X[f]			
IMP administration[g]		X[h]					X[h]			X[h]			
Prior medication and treatment history	X	X											

Study period	**Screen-ing**	**Base-line**	**Main treatment period**								**Transition period**		
Visit number	**V1**	**V2**	**V3[a]**	**V4[a]**	**V5[a]**	**V6[a]**	**V7**	**V8[a]**	**V9[b]**		**V10[a]**	**V11[a]**	**V12**
									EOM	**SOT**			**EOS/ET**
Day/week and allowed time windows	**Up to 4 weeks before V2**	**Day 1**	**Day 15 ±3 days**	**Week 4 ±3 days**	**Week 8 ±5 days**	**Week 12 ±7 days**	**Week 26 ±7 days**	**Week 39 ±14 days**	**Week 52 ±14 days**		**Week 54 2 weeks ±3 days after V9**	**Week 65 13 weeks ±14 days after V9**	**Week 78 26 weeks ±14 days after V9**
Procedures and assessments													
Local tolerability at the injection site		X[i]					X[i]			X[i]			
Clinical laboratory tests sampling (serum chemistry, hematology, and urinalysis)[j]	X[k,l]	X		X		X[l]	X[l]	X	X[l]			X	X[l]
Immunogenicity sampling (serum ADA)[m]		X[n]	X	X	X	X	X[n]	X	X[n]		X	X	X
Physical examination, including height and weight[o]	X						X		X				X
12-lead ECG	X	X					X		X				X
Vital signs measurement[p]	X	X	X	X	X	X	X	X	X		X	X	X
LS-BMD by DXA	X						X		X				X[q]
Total hip and femoral neck BMD by DXA[r]	X						X		X				X[q]
Lateral spine X-ray	X								X				X[q]

Study period	**Screen-ing**	**Base-line**	**Main treatment period**								**Transition period**		
Visit number	**V1**	**V2**	**V3[a]**	**V4[a]**	**V5[a]**	**V6[a]**	**V7**	**V8[a]**	**V9[b]**		**V10[a]**	**V11[a]**	**V12**
									EOM	**SOT**			**EOS/ET**
Day/week and allowed time windows	**Up to 4 weeks before V2**	**Day 1**	**Day 15 ±3 days**	**Week 4 ±3 days**	**Week 8 ±5 days**	**Week 12 ±7 days**	**Week 26 ±7 days**	**Week 39 ±14 days**	**Week 52 ±14 days**		**Week 54 2 weeks ±3 days after V9**	**Week 65 13 weeks ±14 days after V9**	**Week 78 26 weeks ±14 days after V9**
Procedures and assessments													
Pharmacokinetics sampling (serum concentration of IMP)		X	X	X	X	X	X	X	X		X	X	X
Pharmacodynamics sampling (sCTX-1 and P1NP)[s]		X	X	X	X	X	X	X	X		X	X	X
Adverse events inquiry	X	X	X	X	X	X	X	X	X		X	X	X
Concomitant medication inquiry	X	X	X	X	X	X	X	X	X		X	X	X
COVID-19 inquiry	X	X	X	X	X	X	X	X	X		X	X	X
Inform of study restrictions and compliance requirements	X	X	X	X	X	X	X	X	X		X	X	

[a] Visits 3, 4, 5, 6, 8, 10, and 11 may be conducted at the patient's home should circumstances dictate.
[b] End of main treatment period is after all assessments (except for local tolerance at the injection site) have been performed at visit 9. Transition period begins with IMP administration.
[c] Vitamin D and calcium will be dispensed as necessary throughout the study to facilitate daily supplementation from screening to the end of the study.
[d] Prior to randomization.
[e] COVID-19 test can be performed at any time the patient displays symptoms of the disease.

[f] Patients in the PROLIA US treatment group will be re-randomized prior to receiving their third dose in a 1:1 ratio to receive an additional dose of PROLIA US or switch to TVB-009P and receive a single dose of TVB-009P in the transition period. Patients who were initially randomized to TVB-009P will continue to receive TVB-009P; however, to maintain blinding, the randomization process will be performed for all patients (although only patients in the PROLIA US arm will actually be re-randomized).
[g] All the assessments scheduled for visits 2, 7, and 9, with the exception of local tolerability, will need to be completed before IMP administration.
[h] All device-related adverse events, malfunctions etc will be recorded; their impact relative to the safety and tolerability of the IMP will be evaluated.
[i] Local tolerability at the injection site will be assessed 1 hour ±10 minutes after IMP administration.
[j] Coagulation (INR, PT and PPT), cholesterol (LDL, HDL, HDL/total), triglycerides and urinalysis will only be conducted at visits 2, 7, and 9. Blood sampling for lipid profile should be taken in the morning hours after overnight fasting. Samples should be collected consistently at the same time of the day for an individual patient at all visits. Vigorous exercise should be avoided the day prior to sampling
[k] Additional laboratory parameters, such as FSH and E2 (as applicable), estimated glomerular filtration rate (eGFR) (using the Modification of Diet in Renal Disease [MDRD] formula), will also be assessed at screening.
[l] Vitamin D levels will be assessed at screening and visits 6, 7, 9, 12, and ET/EOS. Vitamin D and/or calcium supplements will be provided. A patient with a screening adjusted calcium level outside the normal range or serum 25 (OH) vitamin D level ≤20 ng/mL may be rescreened once to re-evaluate calcium and vitamin D level post repletion.
[m] If any severe hypersensitivity reaction (eg, anaphylaxis), serious adverse or immunogenicity-related adverse event is observed, additional sample(s) will be collected for immunogenicity assessment as close to onset of the event as possible, at resolution of the event, and 30 days following the event onset, if possible.
[n] When a number of assessments are to be conducted at the same time point, the immunogenicity blood sample should be taken after other assessments but before IMP administration.
[o] Physical examination will include, at a minimum, head, eyes, ears, nose, and throat (HEENT), oral cavity, chest, cardiovascular system, abdomen, and skin. Oral examination will be performed to exclude risks for osteonecrosis of the jaw. If required by local authorities, a mandible radiograph may be carried out at screening.
[p] Vital signs includes pulse rate, blood pressure (systolic/diastolic), and respiratory rate. Before blood pressure and pulse are measured, the patient must rest in a supine or seated position for at least 5 minutes. The same position and arm should be used each time vital signs are measured for a given patient. For any abnormal vital signs value, the measurement should be repeated as soon as possible. In the event of suspected severe hypersensitivity (including anaphylaxis), vital signs, including oxygen saturation and respiration rate, will be measured.
[q] In case of early termination, LS-BMD DXA, total hip and femoral neck BMD DXA, and lateral spine X-ray will be assessed only in patients who terminate the study between the week 26 and week 52 visits. The minimum time between two DXA scans should be 3 months. These assessments will not be performed for patients who terminate the study earlier than week 26 or during the transition period.
[r] These scans will be unilateral only. For each patient, the same hip should be scanned throughout the study.
[s] Blood sampling for sCTX-1 and P1NP assessment should be taken in the morning hours after overnight fasting. Samples should be collected consistently at the same time of the day for an individual patient at all visits. Vigorous exercise should be avoided the day prior to sampling.

ADA=anti-drug antibody; BMD=bone mineral density; COVID-19=coronavirus 2019; DXA=dual-energy X-ray absorptiometry; ECG=electrocardiogram; E2=estradiol; EOM=end of main treatment period; EOS=end of study; ET=early termination; FSH=follicle stimulating hormone; HDL=high density lipoprotein; IMP=investigational medicinal product; INR=international normalized ratio; LDL=low density lipoprotein; LS-BMD=lumbar spine-bone mineral density; P1NP=procollagen type 1 N propeptide; PK=pharmacokinetic; PPT=partial prothrombin; PT=prothrombin; sCTX-1=serum C-telopeptide cross-link of type 1 collagen; SOT=start of transition period; V=visit.