STATISTICAL ANALYSIS PLAN

Study Title:	A Phase 3, Randomized, Double-Blind Study to Evaluate the Efficacy and Safety of Switching from Tenofovir Disoproxil Fumarate (TDF) 300 mg QD to Tenofovir Alafenamide (TAF) 25 QD in Subjects with Chronic Hepatitis B who are Virologically Suppressed
Name of Test Drug:	Tenofovir Alafenamide (TAF)
Study Number:	GS-US-320-4018
Protocol Version (Date):	Amendment 1 (21 December 2016)
Analysis Type:	Final Analysis
Analysis Plan Version:	1.0
Analysis Plan Date:	11 March 2020
Analysis Plan Author(s):	PPD

CONFIDENTIAL AND PROPRIETARY INFORMATION

TABLE OF CONTENTS

LIST OF IN-TEXT TABLES

LIST OF ABBREVIATIONS

AASLD	American Association for the Study of Liver Diseases
AE	adverse event
ALT	alanine aminotransferase
ANCOVA	analysis of covariance
ANOVA	analysis of variance
Anti-HBe	antibody to HBeAg
Anti-HBs	antibody to HBsAg
AST	aspartate aminotransferase
BLQ	below the limit of quantitation
BMD	bone mineral density
BMI	body mass index
bsAP	Bone specific alkaline phosphatase
CDER	Center for Drug Evaluation and Research
CFR	Code of Federal Regulations
CHB	chronic hepatitis B
CI	confidence interval
CK	creatine kinase
CKD	chronic kidney disease
CKD-EPI	Chronic Kidney Disease Epidemiology Collaboration formula for calculating glomerular filtration rate
CL_{CR}	creatinine clearance
CMH	Cochran-Mantel-Haenszel
CRF	case report form
CTX	c-type collagen sequence
CV	coefficient of variation
CSR	clinical study report
CTCAE	Common Toxicity Criteria for Adverse Events
DILI	drug-induced liver injury
DMC	data monitoring committee
DNA	deoxyribonucleic acid
DXA	dual-energy x-ray absorptiometry
ECG	electrocardiogram
eCRF	electronic case report form
ED	early discontinuation
eGFR	estimated glomerular filtration rate
ESDD	early study drug discontinuation
FAS	Full Analysis Set
FDA	Food and Drug Administration
$FEPO_4$	fractional excretion of filtered phosphate

GFR	glomerular filtration rate
Gilead	Gilead Sciences, Inc.
Hb	hemoglobin
HBeAb	hepatitis B e antibody
HBeAg	hepatitis B e antigen
HBsAb	hepatitis B surface antibody
HBsAg	hepatitis B surface antigen
HBV	hepatitis B virus
HCC	hepatocellular carcinoma
HCV	hepatitis C virus
HDL	high density lipoproptein
HDV	hepatitis D virus
HIV	human immunodeficiency virus
HLGT	high-level group term
HLT	high-level term
ICH	International Conference on Harmonization (of Technical Requirements for Registration of Pharmaceuticals for Human Use)
ITT	intent to treat
IWRS	interactive web response system
LDL	low density lipoprotein
LLN	lower limit of the normal range
LTT	lower-level term
LOQ	limit of quantitation
M = E	missing = excluded
M = F	missing = failure
MDRD	Modification of Diet in Renal Disease
MedDRA	Medical Dictionary for Regulatory Activities
MH	Mantel-Haenszel
OC	osteocalcin
P1NP	procollagen type 1 N-terminal propeptide
PBMC	peripheral blood mononuclear cell
PCR	polymerase chain reaction
PP	per protocol
PT	preferred term
PVE	Pharmacovigilance and Epidemiology
Q1, Q3	first quartile, third quartile
QD	once daily
RBP	retinol binding protein
SAE	serious adverse event
SAP	statistical analysis plan
SD	standard deviation

SI (units)	international system of units
SOC	system organ class
TAF	tenofovir alafenamide
TDF	tenofovir disoproxil fumarate (Viread®)
TFLs	tables, figures, and listings
TmP	tubular maximum reabsorption rate of phosphate
UACR	urine albumin to creatinine ratio
ULN	upper limit of normal
UPCR	urine protein to creatinine ratio
WHO	World Health Organization

1. INTRODUCTION

This statistical analysis plan (SAP) describes the statistical analysis methods and data presentations to be used in the tables, figures, and listings (TFLs) in the final clinical study report (CSR) for Study GS-US-320-4018. The final efficacy and safety analysis will be performed after all subjects have completed the study or prematurely discontinued the study. This SAP is based on the study protocol Amendment 1 dated 21 December 2016 and the electronic case report form (eCRF). The SAP will be finalized before database finalization for the final analysis. Any changes made after the finalization of the SAP will be documented in the CSR.

1.1. Study Objectives

The primary objectives of this study are as follows:

- To evaluate the efficacy of switching to tenofovir alafenamide (TAF) 25 mg once daily (QD) versus continuing tenofovir disoproxil fumarate (TDF) 300 mg QD in virologically suppressed subjects with chronic hepatitis B virus (HBV) as determined by the proportion of subjects with HBV DNA ≥ 20 IU/mL (as defined by the modified United States (US) Food and Drug Administration (FDA) defined snapshot algorithm) at Week 48

- To compare the safety and tolerability of switching to TAF 25 mg QD versus continuing TDF 300 mg QD in virologically suppressed subjects with chronic HBV at Week 48

The key secondary objectives of this study are as follows:

- To compare the safety of switching to TAF 25 mg QD versus continuing TDF 300 mg QD as determined by the percent change from baseline in hip and spine bone mineral density (BMD) at Week 48

- To compare the safety of switching to TAF 25 mg QD versus continuing TDF 300 mg QD as determined by the change from baseline in estimated creatinine clearance by Cockcroft-Gault method ($eGFR_{CG}$) at Week 48

Other secondary objectives of this study are as follows:

- To evaluate the proportion of subjects with HBV DNA < 20 IU/mL and target detected/not detected (ie, < lower limit of detection [LLOD]) at Week 48 in subjects who switched to TAF 25 mg QD versus those who continued TDF 300 mg QD

- To compare the safety of switching to TAF 25 mg QD for 96 weeks versus continuing TDF 300 mg QD for 48 weeks followed by open-label TAF for 48 weeks as determined by the percent change from baseline in hip and spine BMD at Week 96

- To compare the safety of switching to TAF 25 mg QD for 96 weeks versus continuing TDF 300 mg QD for 48 weeks followed by open-label TAF for 48 weeks as determined by the change from baseline in $eGFR_{CG}$ at Week 96

- To compare the serological response (loss of hepatitis B surface antigen [HBsAg] and seroconversion to antibody against hepatitis B surface antigen [anti-HBs], and loss of hepatitis B e antigen [HBeAg] and seroconversion to antibody against hepatitis B e antigen [anti-HBe] in HBeAg-positive subjects) of switching to TAF 25 mg QD versus continued TDF 300 mg QD at Week 48

- To compare biochemical response (normal alanine aminotransferase [ALT] and normalized ALT) of switching to TAF 25 mg QD versus continued TDF 300 mg QD at Week 48

- To compare the change in fibrosis as assessed by FibroTest® after switching to TAF 25 mg QD versus continued TDF 300 mg QD at Week 48

- To evaluate the comparative open-label efficacy (virologic, serological, and biochemical) and safety of switching to TAF 25 mg QD from Week 48 through Week 96 in subjects initially randomized to TAF 25 mg QD and in subjects sequentially treated with continued TDF 300 mg QD for 48 weeks and then switched to open-label TAF 25 mg QD at Week 96

1.2. Study Design

1.2.1. Design Configuration and Subject Population

This is a randomized, double-blind, multicenter, active-controlled study to evaluate the safety and efficacy of TAF 25 mg QD in virologically suppressed chronic hepatitis B (CHB) subjects who switch from TDF to TAF compared to continued TDF treatment.

Approximately 460 adult (about half ≥ 50 years old) subjects with CHB, who are currently virologically suppressed while receiving therapy with TDF, will be randomized in a 1:1 ratio to receive either TAF 25 mg QD and matched placebo of TDF 300 mg QD or Gilead-provided TDF 300 mg QD and matched placebo of TAF 25 mg QD.

1.2.2. Treatment Groups and Randomization

At randomization, subjects will be stratified by screening HBeAg status (HBeAg-negative vs. HBeAg-positive) and age (≥ 50 or < 50 years).

1.2.3. Key Eligibility Criteria

Subjects were to have met all eligibility criteria, including the following key eligibility criteria:

- Documented evidence of chronic HBV infection previously (eg, documented HBsAg positive for more than 6 months)

- Maintained on TDF 300 mg QD for at least 48 weeks, and as monotherapy for CHB for at least 24 weeks prior to screening and with viral suppression (HBV DNA < lower limit of quantitation [LOQ] by local laboratory assessment) for a minimum of 12 weeks prior to screening, and including a screening HBV DNA value of < 20 IU/mL (by a central laboratory)

- Estimated creatinine clearance ≥ 50 ml/min (using the Cockcroft-Gault method) based on serum creatinine and actual body weight as measured at screening

- Did not have co-infection with hepatitis C virus (HCV), HIV, or hepatitis D virus (HDV), evidence of hepatocellular carcinoma or recent (≤ 5 years) history of clinical hepatic decompensation, and abnormal hematological and biochemical parameters

1.2.4. Study Periods and Phases

The duration of double-blind treatment is 48 weeks. All subjects who complete 48 weeks of treatment are eligible for participation in the open label TAF 25-mg extension period for an additional 48 weeks (through Week 96). During the double-blind period only, subjects with a confirmed $eGFR_{CG}$ < 50 ml/min, and > 20% decline in estimated glomerular filtration rate ($eGFR_{CysC}$) by CKD-EPI (cystatin C) compared to baseline during the study, were required to undergo dose modification to every other day dosing of study drug. Subjects with confirmed creatinine clearance < 30 mL/min during the double-blind period of the study were to have the study drug discontinued.

1.2.5. Schedule of Assessments

Laboratory analyses (serum chemistry and liver tests, hematology, $eGFR_{CG}$, urinalysis, plasma HBV DNA levels), pregnancy testing [for females of childbearing potential]), vital signs, adverse events, and concomitant medications will be measured or assessed at screening baseline, Weeks 4, 8, 12, 24, 36, 48, 60 ,72, and 96 (and at early discontinuation [ED]). Fasting metabolic panels will be measured at baseline, and Weeks 12, 24, 36, 48, 60, 72, and at 96 (and at ED). HBV serology (HBsAg, HBeAg, reflex hepatitis B antibody [HBeAb], hepatitis B surface antibody [HBsAb], and quantitative HBsAg) will be conducted at screening, baseline, Weeks 12, 24, 36, 48, 60 72, and 96 (and at ED).

Dual energy x-ray absorptiometry (DXA) scans of the hip and spine will be performed at baseline and Weeks 24, 48, 72, and 96 (and at ED). Bone and renal biomarker testing will be performed at baseline and Weeks 4, 12, 24, 48, 72, and 96 (and at ED). FibroTest® and the Child Pugh Score will be performed at baseline, Week 48, and Week 96 (and at ED).

Complete physical examination will be performed at screening, baseline, and Weeks 24, 48, 72, and 96 (and at ED). Symptom-driven physical examinations including body weight assessment will be conducted at all other visits. Electrocardiogram (ECG) will be performed at screening, Week 48, and Week 96 (and at ED). Health-related quality of life questionnaires including Chronic Liver Disease Questionnaire (CLDQ), SF-36, and Work Productivity and Activity Impairment (WPAI) will be filled out at baseline, and Weeks 24, 48, and 96 (and at ED).

Follow-up assessments will occur every 4 weeks for 12 weeks for subjects who discontinue study drug due to HBsAg loss and for 24 weeks for subjects who permanently discontinue study drug for other reasons, and will include the following: concomitant medications, adverse events, body weight, vital signs, symptom driven physical examinations, laboratory analyses, and HBV serology.

1.3. Sample Size and Power

With respect to the primary efficacy endpoint of proportion of subjects with HBV DNA ≥ 20 IU/mL (the lower LOQ of the current polymerase chain reaction [PCR] assay), as determined by the modified US FDA-defined snapshot algorithm, at Week 48, when the sample sizes are 230 (TAF 25 mg arm) and 230 (TDF 300 mg arm), a two-group, large-sample normal approximation test of proportions with a one-sided 0.025 significance level will have 80% power to establish non-inferiority of TAF to TDF. It is assumed that both treatment arms will have subjects who have HBV DNA ≥ 20 IU/mL (as defined by the modified US FDA-defined snapshot algorithm) at a rate of 2.4% at Week 48 (based on the rates of suppressed subjects having HBV DNA ≥ 29 IU/mL, the lower LOQ of the assay used in those studies, after 48 weeks of treatment from the combined data from TDF studies GS-US-174-0102 and GS-US-174-0103 with the ratio of HBeAg- to HBeAg+ subjects assumed to be 2:1), with a non-inferiority margin of 4%.

The proposed sample size (n 230 for the TAF 25 mg arm, n 230 for the TDF 300 mg arm) also provides > 99% power to detect a 1.81% difference in the percentage change from baseline in hip BMD at Week 48 (assuming a 1.47% [SD 2.71%] change from baseline in the TAF 25 mg arm and 0.34% [SD 2.83%] change from baseline in the TDF 300 mg arm, with a two-sided α 0.05); a > 99% power to detect a 2.00% difference in the percentage change from baseline in spine BMD at Week 48 (assuming a 1.56% [SD 3.84%] change from baseline in the TAF 25 mg arm and 0.44% [SD 4.14%] change in the TDF 300 mg arm, with a two-sided α 0.05); a 97% power to detect a 4.9 mL/min difference in the change from baseline in $eGFR_{CG}$ at Week 48 (assuming a 1.3 mL/min [SD 13.2] change from baseline in the TAF 25 mg arm and 3.6 mL/min [SD 13.0] change from baseline in the TDF 300 mg arm, with a two-sided α 0.05). These assumptions were derived from an HIV switch study (GS-US-292-0109) due to the unavailability of stable and suppressed switch data from TAF HBV studies.

2. TYPE OF PLANNED ANALYSIS

2.1. Data Monitoring Committee (DMC) Analysis

An external multidisciplinary Data Monitoring Committee (DMC) will review the progress of the study and perform interim reviews of the safety data in order to protect subject welfare and preserve study integrity. To ensure the best interests of the participants, the DMC will recommend to the sponsor if the nature, frequency, and severity of adverse effects associated with the study treatment warrant the early termination of the study, the continuation of the study, or the continuation of the study with modifications.

The initial review was conducted after all subjects completed 24 weeks of the study. No additional meetings were requested by the DMC.

The DMC's role and responsibilities and the scope of analysis to be provided to the DMC are provided in a mutually agreed upon charter, which defines the DMC membership, meeting logistics, and meeting frequency.

2.2. Week 48 Analysis (Primary Analysis)

The Week 48 analysis was conducted after the last subject completed the Week 48 visit or prematurely discontinued study drug.

2.3. Final Analysis

The final analysis for the study will be conducted after all subjects have completed the study or prematurely discontinued the study, for all study data from double-blind phase to open-label phase to treatment-free follow-up.

This SAP described the plan for the final analysis.

3. GENERAL CONSIDERATIONS FOR DATA ANALYSES

Analysis results will be presented using descriptive statistics. For categorical variables, the number and percentage of subjects in each category will be presented; for continuous variables, the number of subjects (n), mean, standard deviation (SD) or standard error (SE), median, first quartile (Q1), third quartile (Q3), minimum, and maximum will be presented.

By-subject listings will be presented for all subjects in the Randomized Analysis Set and sorted by subject ID number, visit date, and time (if applicable). Data collected on log forms, such as adverse events (AEs), will be presented in chronological order within the subject. The treatment group to which subjects were randomized will be used in the listings. Age, sex at birth, race, and ethnicity will be included in the listings, as space permits.

3.1. Analysis Sets

Analysis sets define the subjects to be included in an analysis. Analysis sets and their definitions are provided in this section. The analysis set will be identified and included as a subtitle of each table, figure, and listing.

For each analysis set, the number and percentage of subjects who were included out of all randomized subjects will be summarized by treatment group and overall.

A listing of reasons for exclusion from analysis sets will be provided by subject.

3.1.1. Randomized Analysis Set

The Randomized Analysis Set includes all subjects who were randomized into the study. This is the primary analysis set for by-subject listings.

3.1.2. Safety Analysis Set

The Safety Analysis Set will include all randomized subjects who have received at least 1 dose of study drug. Subjects will be analyzed according to the treatment they actually received during the double-blind phase. The actual treatment received will differ from the randomized treatment only when their actual treatment differs from randomized treatment for the entire treatment duration during the double-blind phase. This is the primary analysis set for safety analyses.

3.1.3. Open-Label Safety Analysis Set

The Open-Label Safety Analysis Set includes all randomized subjects who have received at least 1 dose of open-label study drug. Subjects will be analyzed according to the treatment they actually received during the double-blind phase.

3.1.4. Treatment-Free Follow-up Safety Analysis Set

The treatment-free follow-up (TFFU) Safety Analysis Set includes all randomized subjects who entered the TFFU period. Subjects will be analyzed according to the treatment they actually received during the double-blind phase.

3.1.5. Full Analysis Set

The Full Analysis Set (FAS) will include all randomized subjects who have received at least 1 dose of study drug. Subjects will be analyzed according to the treatment to which they were randomized during the double-blind phase. This is the primary analysis set for efficacy analyses, unless otherwise specified.

3.1.6. Open-Label Full Analysis Set

The Open-Label FAS includes all randomized subjects who have received at least 1 dose of open-label study drug. Subjects will be analyzed according to the treatment to which they were randomized during the double-blind phase.

3.1.7. TFFU Full Analysis Set

The TFFU FAS includes all randomized subjects who entered the TFFU period. Subjects will be analyzed according to the treatment to which they were randomized during the double-blind phase.

3.1.8. Serologically Evaluable Full Analysis Set

3.1.8.1. Serologically Evaluable Full Analysis Set for HBeAg loss/seroconversion

The Serologically Evaluable Full Analysis Set for HBeAg loss/seroconversion will include all subjects who were randomized and had received at least 1 dose of study drug, and with HBeAg positive and HBeAb negative or missing values at baseline. Subjects will be analyzed according to the treatment they were randomized to during the double-blind phase.

3.1.8.2. Serologically Evaluable Full Analysis Set for HBsAg loss/seroconversion

The Serologically Evaluable Full Analysis Set for HBsAg loss/seroconversion will include all subjects who were randomized and had received at least 1 dose of study drug, and with HBsAg positive and HBsAb negative or missing values at baseline. Subjects will be analyzed according to the treatment they were randomized to during the double-blind phase.

3.1.9. Open-Label Serologically Evaluable Full Analysis Set

3.1.9.1. Open-Label Serologically Evaluable Full Analysis Set for HBeAg Loss/Seroconversion

The Open-Label Serologically Evaluable Full Analysis Set for HBeAg loss/seroconversion includes all subjects who were randomized and had received at least 1 dose of open-label study

drug, and with HBeAg positive and HBeAb negative or missing at open-label baseline. Subjects will be analyzed according to the treatment they were randomized to during the double-blind phase.

3.1.9.2. Open-Label Serologically Evaluable Full Analysis Set for HBsAg Loss/Seroconversion

The Open-Label Serologically Evaluable Full Analysis Set for HBsAg loss/seroconversion includes all subjects who were randomized and had received at least 1 dose of open-label study drug, and with HBsAg positive and HBsAb negative or missing at open-label baseline. Subjects will be analyzed according to the treatment they were randomized to during the double-blind phase.

3.1.10. TFFU Serologically Evaluable Full Analysis Set

3.1.10.1. TFFU Serologically Evaluable Full Analysis Set for HBeAg Loss/Seroconversion

The TFFU Serologically Evaluable Full Analysis Set for HBeAg loss/seroconversion includes all subjects who were randomized and entered the TFFU period, and with HBeAg positive and HBeAb negative or missing at TFFU baseline. Subjects will be analyzed according to the treatment they were randomized to during the double-blind phase.

3.1.10.2. TFFU Serologically Evaluable Full Analysis Set for HBsAg Loss/Seroconversion

The TFFU Serologically Evaluable Full Analysis Set for HBsAg loss/seroconversion includes all subjects who were randomized and entered the TFFU period, and with HBsAg positive and HBsAb negative or missing at TFFU baseline. Subjects will be analyzed according to the treatment they were randomized to during the double-blind phase.

3.1.11. DXA Analysis Set

3.1.11.1. Hip DXA Analysis Set

The Hip DXA Analysis Set will include all subjects who were randomized and had received at least 1 dose of study drug, and had nonmissing baseline hip BMD values. Subjects will be analyzed according to the treatment they actually received during the double-blind phase.

3.1.11.2. Spine DXA Analysis Set

The Spine DXA Analysis Set will include all subjects who were randomized and had received at least 1 dose of study drug, and had nonmissing baseline spine BMD values. Subjects will be analyzed according to the treatment they actually received during the double-blind phase.

3.1.12. Open-Label DXA Analysis Set

3.1.12.1. Open-Label Hip DXA Analysis Set

The Open-Label Hip DXA Analysis Set includes all subjects in the Open-Label Safety Analysis Set who had nonmissing open-label baseline hip BMD values. Subjects will be analyzed according to the treatment received in double-blind phase.

3.1.12.2. Open-Label Spine DXA Analysis Set

The Open-Label Spine DXA Analysis Set includes all subjects in the Open-Label Safety Analysis Set who had nonmissing open-label baseline spine BMD values. Subjects will be analyzed according to the treatment received in double-blind phase.

3.2. Subject Grouping

For analyses based on the FAS, Open-Label FAS, and TFFU FAS, including the serologically evaluable FASs, subjects will be grouped according to the treatment to which they were randomized to during the double-blind phase. For analyses based on the Hip and Spine DXA Analysis Sets, Safety Analysis Set, Open-Label Hip and Spine DXA Analysis Sets, Open-Label Safety Analysis Set, and TFFU Safety Analysis Set, subjects will be grouped according to the actual treatment received during the double-blind phase.

3.3. Strata and Covariates

Subjects will be randomly assigned to treatment groups via the interactive voice or web response system (IXRS) in a 1:1 ratio using a stratified randomization schedule. Stratification will be based on screening for HBeAg status (HBeAg-negative vs HBeAg-positive) and age ($\geq$ 50 or $<$ 50 years). For statistical analyses, the HBeAg status and age strata recorded in the clinical database will be reclassified using the corresponding baseline value and used for analysis. If the number of subjects in a stratum is too small, the stratum may be combined with other strata for analysis.

3.4. Missing Data and Outliers

3.4.1. Missing Data

A missing datum for a given study analysis window may be due to any of the following reasons:

- A visit occurring in the window but data were not collected or were unusable
- A visit not occurring in the window
- A subject permanently discontinuing from the study before reaching the window

In general, values for missing data will not be imputed, unless methods for handling missing data are specified.

For missing last dosing date of study drug, imputation rules are described in Section 3.6. The handling of missing or incomplete dates for AE onset is described in Section 7.1.6.2, and for concomitant medications in Section 7.7.

3.4.2. Outliers

Outliers will be identified during the data management and data analysis process, but no sensitivity analyses will be conducted. All data will be included in the data analysis.

3.5. Data Handling Conventions and Transformations

In general, age (in years) on the date of the first dose of study drug will be used for analyses and presentation in listings. If an enrolled subject was not dosed with any study drug, the randomization date will be used instead of the first dosing date of study drug. For screen failures, the date the informed consent was signed will be used for age calculation. If only the birth year is collected on the eCRF, "01 July" will be used for the unknown birth day and month for the purpose of age calculation. If only birth year and month are collected, "01" will be used for the unknown birth day.

Non-PK data that are continuous in nature but are less than the lower limit of quantitation (LOQ) or above the upper LOQ will be imputed as follows:

- A value that is 1 unit less than the LOQ will be used to calculate descriptive statistics if the datum is reported in the form of "< x" (where x is considered the LOQ). For example, if the values are reported as < 50 and < 5.0, values of 49 and 4.9, respectively, will be used to calculate summary statistics. An exception to this rule is any value reported as < 1 or < 0.1, etc. For values reported as < 1 or < 0.1, a value of 0.9 or 0.09, respectively, will be used to calculate summary statistics.

- A value that is 1 unit above the LOQ will be used to calculate descriptive statistics if the datum is reported in the form of "> x" (where x is considered the LOQ). Values with decimal points will follow the same logic as above.

- The LOQ will be used to calculate descriptive statistics if the datum is reported in the form of "≤ x" or "≥ x" (where x is considered the LOQ).

Logarithm (base 10) transformations will be applied to HBV DNA and quantitative HBsAg data, as appropriate, in efficacy analyses.

For HBV DNA, if the value in IU/mL is above the upper limit of quantification, the corresponding diluted value, if available, will be used. HBV DNA results of "<20 IU/mL HBV DNA detected" or "No HBV DNA detected" will be imputed as 19 IU/mL.

3.6. Analysis Visit Windows

3.6.1. Definition of Study Day 1 and Other Definitions

Study Day 1 is defined as the day when the first dose of blinded study drug was taken, as recorded on the Study Drug Administration eCRF.

Open-Label Study Day 1 is defined as the day when the first dose of the open-label study drug was taken, as recorded on the Study Drug Administration eCRF.

Study days are calculated relative to Study Day 1. For events that occurred on or after Study Day 1 date, study days are calculated as (visit date Study Day 1 + 1). For events that occurred prior to Study Day 1, study days are calculated as (visit date Study Day 1).

Open-Label Study days are calculated relative to Open-Label Study Day 1. For events that occurred on or after Open-Label Study Day 1, study days are calculated as (visit date Open-Label Study Day 1 + 1).

Follow-up days are for visits occurred during the treatment-free follow-up period and calculated as (visit date last dose date).

Last Dose Date of Blinded Study Drug is the latest non-missing end date of blinded study drug, recorded on the Study Drug Administration eCRF form with "Study Drug Permanently Withdrawn" box checked for subjects who prematurely discontinued blinded study drug or who completed blinded study drug according to the Blinded Study Drug Completion eCRF. If the last dose date of blinded study drug is missing (eg, due to lost to follow up) for subjects who prematurely discontinued blinded study drug or completed blinded study drug, or for subjects who are still on blinded study drug, the latest of non-missing blinded study drug start dates and end dates, the clinical visit dates, and the laboratory visit dates, excluding the dates during open-label treatment and 24-week treatment-free follow up, will be used to impute the last dose date of blinded study drug.

For subjects who prematurely discontinued blinded study drug or who completed blinded study drug but did not enter the open-label phase, the **Last Dose Date** is the same as Last Dose Date of Blinded Study Drug.

For subjects who completed blinded study drug and entered the open-label phase, the **Last Dose Date** is the latest non-missing end date of open-label study drug, recorded on the Study Drug Administration eCRF form with "Study Drug Permanently Withdrawn" box checked for subjects who prematurely discontinued open-label study drug or who completed open-label study drug according to Open-Label Study Drug Completion eCRF. If the last dose date of open-label study drug is missing (eg, due to lost to follow up) for subjects who prematurely discontinued open-label study drug or who completed open-label study drug, or for subjects who are still on open-label study drug, the latest of non-missing open-label study drug start dates and end dates, the clinical visit dates, and the laboratory visit dates, excluding the dates during 24-week treatment-free follow-up, will be used to impute the last dose date.

Last Study Date is the latest of nonmissing study drug (blinded or open-label) start dates and end dates, the clinical visit dates, and the laboratory visit dates, including the 24-week treatment-free follow-up visit date, for subjects who prematurely discontinued study or who completed study according to the Study Completion eCRF.

3.6.2. Analysis Windows

Subject visits might not occur on protocol specified days. Therefore, for the purpose of analysis, observations will be assigned to analysis windows.

The following windows (Table 3-1 to Table 3-6) apply only to baseline and postbaseline on-treatment assessments collected during the double-blind phase and open-label phase. For summaries and analyses, assessments will first be categorized into baseline and postbaseline on-treatment assessments occurring during the double-blind or open-label phase before applying analysis windows.

Baseline value for the double-blind phase is defined as the last nonmissing value obtained on or prior to Study Day 1. For DXA BMD, it is defined as the last value on or prior to Study Day 14.

For subjects who completed blinded study drug and entered the open-label phase, laboratory assessments that occurred during the period from Study Day 2 up to and including the minimum of the Last Dose Date of Blinded Study Drug + 3 days and Open-label Study Day 1, will be considered as postbaseline on-treatment assessments during the double-blind phase. If a subject prematurely discontinued blinded study drug or did not enter open-label phase after completion of blinded study drug, then postbaseline on-treatment assessments during the double-blind phase will be defined as assessments that occurred during the period from Study Day 2 up to and including the Last Dose Date of Blinded Study Drug + 3 days. DXA assessments that occurred during the period from Study Day 15 up to and including the Last Dose Date of Blinded Study Drug + 14 days will be considered as postbaseline on-treatment assessments during the double-blind phase.

Baseline value for the open-label phase is defined as the last nonmissing value obtained on or prior to Open-Label Study Day 1. For DXA BMD, it is defined as the last value on or prior to Open-Label Study Day 14.

For subjects who entered the open-label phase, postbaseline on-treatment assessments during the open-label phase will be defined as assessments that occurred during the period from Open-label Study Day 2 up to the Last Dose Date + 3 days. DXA assessments that occurred during the period from Open-label Study Day 15 up to the Last Dose Date + 14 days will be considered as postbaseline on-treatment assessments during the open-label phase.

Baseline value for the TFFU phase is defined as the last nonmissing value obtained on or prior to the Last Dose Date + 3 days.

The analysis windows for HBV DNA, hematology, serum chemistry and liver function tests, urinalysis, urine pregnancy test, $eGFR_{CG}$, bone biomarker serum parathyroid hormone (PTH), urine fractional excretion of filtered phosphate ($FEPO_4$), renal tubular maximum reabsorption rate of phosphate to the glomerular filtration rate (TmP/GFR), urine protein to creatinine ratio

(UPCR), urine albumin to creatinine ratio (UACR), weight, and vital sign assessments are presented in Table 3-1.

Table 3-1. Analysis Windows for HBV DNA, Hematology, Serum Chemistry and Liver Function Tests, Urinalysis, Urine Pregnancy Test, $eGFR_{CG}$, PTH, $FEPO_4$, TmP/GFR, UPCR, UACR, Weight, and Vital Sign Assessments

Visit ID	Nominal Day	Lower Limit	Upper Limit
Baseline			1
Week 4	28	2	42
Week 8	56	43	70
Week 12	84	71	126
Week 24	168	127	210
Week 36	252	211	294
Week 48	336	295	378
Week 60	420	379	462
Week 72	504	463	588
Week 96	672	589	840

The analysis windows for safety ECG and Child Pugh Score are presented in Table 3-2.

Table 3-2. Analysis Windows for Safety ECG and Child Pugh Score

Visit ID	Nominal Day	Lower Limit	Upper Limit
Baseline			1
Week 48	336	2	504
Week 96	672	505	840

The analysis windows for FibroTest and health-related quality of life questionnaires are presented in Table 3-3.

Table 3-3. Analysis Windows for Fibrotest and Health Related Quality of Life Questionnaires

Visit ID	Nominal Day	Lower Limit	Upper Limit
Baseline			1
Week 24	168	2	252
Week 48	336	253	504
Week 96	672	505	840

The analysis windows for BMD results from DXA are presented in Table 3-4.

Table 3-4. Analysis Windows for BMD Results from DXA

Visit ID	Nominal Day	Lower Limit	Upper Limit
Baseline			14
Week 24	168	15	252
Week 48	336	253	420
Week 72	504	421	588
Week 96	672	589	840

The analysis windows for fasting glucose, fasting lipid panel including total cholesterol, high density lipoprotein (HDL), direct low density lipoprotein (LDL), total cholesterol to HDL ratio, and triglycerides, serum HBsAg (quantitative) and HBV serology are presented in Table 3-5.

Table 3-5. Analysis Windows for Fasting Glucose, Fasting Lipid Panel, Serum HBsAg (Quantitative) and HBV Serology

Visit ID	Nominal Day	Lower Limit	Upper Limit
Baseline			1
Week 12	84	2	126
Week 24	168	127	210
Week 36	252	211	294
Week 48	336	295	378
Week 60	420	379	462
Week 72	504	463	588
Week 96	672	589	840

The analysis windows for fasting renal biomarkers including urine retinol binding protein (RBP), urine RBP to creatinine ratio, urine beta-2-microglobulin (B2M), urine B2M to creatinine ratio, and bone biomarkers including C-type collagen sequence (CTX), procollagen type 1 N-terminal propeptide (P1NP) are presented in Table 3-6.

Table 3-6. Analysis Windows for Fasting Renal and Bone Biomarkers

Visit ID	Nominal Day	Lower Limit	Upper Limit
Baseline			1
Week 4	28	2	56
Week 12	84	57	126
Week 24	168	127	252
Week 48	336	253	420

Visit ID	Nominal Day	Lower Limit	Upper Limit
Week 72	504	421	588
Week 96	672	589	840

Post-treatment assessments during the treatment free follow-up (TFFU) phase will be defined as assessments that occurred during the period after the Last Dose Date (LDD) + 3 days up to the Last Study Date.

The analysis windows for post-treatment assessments are presented in Table 3-7.

Table 3-7. Analysis Windows for Post-Treatment Assessments

Visit ID	Nominal Follow-Up Day	Lower Limit	Upper Limit
Follow-Up Week 4	28	LDD+4	42
Follow-Up Week 8	56	43	70
Follow-Up Week 12	84	71	98
Follow-Up Week 16	112	99	126
Follow-Up Week 20	140	127	154
Follow-Up Week 24	168	155	182

3.6.3. Selection of Data in the Event of Multiple Records in an Analysis Visit Window

Depending on the statistical analysis method, single values are required for each analysis window. For example, change from baseline by visit usually requires a single value, whereas a time-to-event analysis would not require one value per analysis window. When a single value is needed, the following rule(s) will be used.

For baseline of the double-blind, open-label, and TFFU phases, the last available record on or prior to the first dose of blinded study drug, open-label study drug, and last dose date + 4 days will be selected, respectively. For DXA BMD baseline, the last value on or prior to Study Day 14 and Open-Label Study Day 14 will be selected for the double-bind and open-label phases, respectively. If there are multiple records with the same time or no time recorded on the same day for numeric observations, the average will be computed for that day, except for HBV DNA and quantitative HBsAg the geometric mean will be taken instead. If there are multiple records with the same time or no time recorded on the same day for categorical observations, the most conservative value will be taken, eg, negative will be selected over positive for HBeAg and HBsAg, and positive will be selected over negative for HBeAb and HBsAb.

The following specified rules will be used for postbaseline visits:

- **ALT:** The largest value will be included in the analysis when 2 or more ALT values occur within the same visit window.

- **BMD:** The latest record in the window will be selected

- **HBV DNA (IU/mL):** The record on the latest day in the window will be selected. If there is more than 1 record on the selected day, the geometric mean will be taken.

- **Quantitative HBsAg (IU/mL):** The record closest to the nominal day for that visit will be selected. If there are 2 records equidistant from the nominal day, the latest will be selected. If there is more than 1 record on the selected day, the geometric mean will be taken.

- **Serology:** For HBeAg, HBeAb, HBsAg, and HBsAb, the record closest to the nominal day for that visit will be selected. If there are 2 records equidistant from the nominal day, the latest will be selected. If there is more than 1 record on the selected day, the most conservative value will be taken, ie, positive will be selected over negative for HBeAg and HBsAg, and negative will be selected over positive for HBeAb and HBsAb.

For all other laboratory parameters:

- If multiple valid non-missing **numeric** observations exist in a window, then records will be chosen as follows:

  The record closest to the nominal day for that visit will be selected. If there are 2 records equidistant from the nominal day, the latest will be selected. If there is more than 1 record on the selected day, the average will be taken.

- If multiple valid non-missing **categorical** observations (eg, safety ECG results) exist in a window, then records will be chosen as follows:

  The most conservative value (eg, abnormal will be selected over normal for safety ECG) within the window will be selected. In the event that 2 values within a window are of equal abnormality, the value collected nearest to the nominal date will be used.

4. SUBJECT DISPOSITION

4.1. Subject Enrollment and Disposition

Subject enrollment summaries were performed as part of the Week 48 analysis and will not be repeated in the final analysis. The randomization schedule used for the study will be provided as an appendix to the CSR.

The summary of subject disposition will be provided by treatment group and overall for all screened subjects. This summary will include the number of subjects screened, screen failure subjects who were not randomized, subjects who met all eligibility criteria and were not randomized, subjects in the Randomized Analysis Set, subjects randomized but never treated, subjects in the Safety Analysis Set, subjects in the FAS, subjects in the Open-Label Safety Analysis Set, subjects in the Open-Label FAS, subjects in the TFFU Safety Analysis Set, and subjects in the TFFU FAS.

In addition, the number and percentage of the subjects in the following categories will be summarized using the Safety Analysis Set:

Double-blind Study Drug Completion

- Prematurely discontinued double-blind study treatment (with summary of reasons for discontinuing treatment)
- Completed double-blind study treatment

Open-label Study Drug Completion

- Willing to enter open-label phase
- Prematurely discontinued open-label study treatment (with summary of reasons for discontinuing treatment)
- Completed open-label study treatment

Study Completion

- Entered 24-week treatment-free follow-up period
- Prematurely discontinued study (with summary of reasons for discontinuing study)
- Completed protocol-planned duration of the study
- Started another HBV therapy

No inferential statistics will be generated. A data listing of reasons for premature study drug/study discontinuation will be provided.

4.2. Extent of Study Drug Exposure and Adherence

Exposure data described below will be summarized for double-blind phase and open-label phase separately.

4.2.1. Duration of Exposure to Blinded Study Drug

Duration of exposure to blinded study drug will be defined as (last dose date of blinded study drug first dose date of blinded study drug + 1), regardless of temporary interruptions in study drug administration, and will be expressed in weeks (recorded to 1 decimal place, eg, 4.5 weeks).

Duration of exposure to blinded study drug will be summarized using descriptive statistics (sample size, mean, SD, median, Q1, Q3, minimum, and maximum) and as the number and percentage of subjects exposed for specified periods, eg, ≥ 4 weeks (28 days), ≥ 8 weeks (56 days), ≥ 12 weeks (84 days), ≥ 24 weeks (168 days), ≥ 36 weeks (252 days), ≥ 48 weeks (336 days).

Summaries will be provided by treatment group for subjects in the Safety Analysis Set.
No inferential statistics will be provided.

4.2.2. Adherence with Blinded Study Drug Regimen

Study drug regimen adherence will be computed based on tablet counts for the active drug only (eg, adherence in TAF group only includes TAF and not placebo for TDF). The numbers of tablets of study drug dispensed and returned are captured on study drug accountability forms.

Adherence (%) of study drug regimen will be calculated as follows:

$$\text{Adherence}\,(\%) \quad 100 \times \frac{\text{Number of tablets taken}}{\text{Number of tablets prescribed}}$$

$$100 \times \frac{\sum \text{No. of tablets taken at each dispensing period [1]}}{(\sum \text{No. of tablets prescribed at each dispensing period [2]}) + 1}$$

[1] Number of tablets taken at a distinct dispensing period for a study drug is calculated as the minimum of (a) the daily number of tablets prescribed for the study drug multiplied by (**the duration of treatment +1 day)** at the dispensing period of the same dispensing date, and (b) the number of tablets taken for the study drug (number of tablets dispensed minus the number of tablets returned). Total number of tablets taken is determined by summing the number of tablets taken from all evaluable dispensing periods.

[2] Number of tablets prescribed at a distinct dispensing period for a study drug is calculated as the daily number of tablets prescribed for the study drug multiplied by **the duration of treatment** at the dispensing period of the same dispensing date. Total number of tablets prescribed is determined by summing the number of tablets prescribed from all evaluable dispensing periods.

The duration of treatment at a dispensing period for a study drug is calculated as the minimum of (a) the last returned date of the same dispensing period for the study drug, (b) date of premature discontinuation of the study drug, and (c) **next dispensing date** of the study drug, minus dispensing date of the study drug.

The next dispensing date is the following dispensing date of the study drug regardless of the bottle return date.

For a record where the number of tablets returned was missing (with "Yes" answered for "Was bottle returned?" question), it is assumed the number of tablets returned was 0. If the number of tablets dispensed was missing or any study drug bottle was not returned or the bottle return status was unknown for the same dispensing date, then all records for the same dispensing date for that study drug will be excluded from both denominator and numerator calculation.

Adherence up to the Week 48 visit for blinded study drug will be calculated using all data from the entire dosing period up to the date of permanent discontinuation of the blinded study drug for subjects who prematurely discontinued blinded study drug, or completed blinded study drug.

The number and percentage of subjects who return at least 1 bottle and have calculable adherence during the study, descriptive statistics for adherence up to Week 48 visit for the blinded study drug regimen (sample size, mean, SD, median, Q1, Q3, minimum, and maximum) along with the number and percentage of subjects belonging to adherence categories (eg, < 80%, ≥ 80% to < 90%, ≥ 90% to < 95%, ≥ 95%) will be provided by treatment group for the Safety Analysis Set. No inferential statistics will be provided.

4.2.3. Duration of Exposure to Open-Label Study Drug

Duration of exposure to open-label study drug will be defined as (last dose date of open-label study drug first dose date of open-label study drug + 1), regardless of temporary interruptions in study drug administration, and will be expressed in weeks (recorded to 1 decimal place, eg, 4.5 weeks).

Duration of exposure to open-label study drug will be summarized using descriptive statistics (sample size, mean, SD, median, Q1, Q3, minimum, and maximum) and as the number and percentage of subjects exposed for specified periods, eg, ≥ 12 weeks (84 days), ≥ 24 weeks (168 days), ≥ 36 weeks (252 days), ≥ 48 weeks (336 days).

Summaries will be provided by treatment group for subjects in the Open-Label Safety Analysis Set. No inferential statistics will be provided.

4.2.4. Adherence with Open-Label Study Drug Regimen

Study drug regimen adherence will be computed based on tablet counts for open-label TAF. The numbers of tablets of study drug dispensed and returned are captured on study drug accountability forms.

Adherence (%) of study drug regimen will be calculated as follows:

$$\text{Adherence}\,(\%) \quad 100 \times \frac{\text{Number of tablets taken}}{\text{Number of tablets prescribed}}$$

$$100 \times \frac{\sum \text{No. of tablets taken at each dispensing period [1]}}{(\sum \text{No. of tablets prescribed at each dispensing period [2]}) + 1}$$

[1] Number of tablets taken at a distinct dispensing period for a study drug is calculated as the minimum of (a) the daily number of tablets prescribed for the study drug multiplied by (**the duration of treatment +1 day)** at the dispensing period of the same dispensing date, and (b) the number of tablets taken for the study drug (number of tablets dispensed minus the number of tablets returned). Total number of tablets taken is determined by summing the number of tablets taken from all evaluable dispensing periods.

[2] Number of tablets prescribed at a distinct dispensing period for a study drug is calculated as the daily number of tablets prescribed for the study drug multiplied by **the duration of treatment** at the dispensing period of the same dispensing date. Total number of tablets prescribed is determined by summing the number of tablets prescribed from all evaluable dispensing periods.

The duration of treatment at a dispensing period for a study drug is calculated as the minimum of (a) the last returned date of the same dispensing period for the study drug, (b) date of premature discontinuation of the study drug, and (c) **next dispensing date** of the study drug, minus dispensing date of the study drug.

The next dispensing date is the following dispensing date of the study drug regardless of the bottle return date.

For a record where the number of tablets returned was missing (with "Yes" answered for "Was bottle returned?" question), it is assumed the number of tablets returned was 0. If the number of tablets dispensed was missing or any study drug bottle was not returned or the bottle return status was unknown for the same dispensing date, then all records for the same dispensing date for that study drug will be excluded from both denominator and numerator calculation.

Adherence for open-label study drug will be calculated using all data from the entire open-label dosing period up to the date of permanent discontinuation of the open-label study drug for subjects who prematurely discontinued open-label study drug, or completed open-label study drug.

The number and percentage of subjects who return at least 1 bottle and have calculable adherence during the study, descriptive statistics for adherence for the open-label study drug regimen (sample size, mean, SD, median, Q1, Q3, minimum, and maximum) along with the number and percentage of subjects belonging to adherence categories (eg, $< 80\%$, $\geq 80\%$ to $< 90\%$, $\geq 90\%$ to $< 95\%$, $\geq 95\%$) will be provided by treatment group for the Safety Analysis Set. No inferential statistics will be provided.

4.3. Protocol Deviations

A by-subject listing will be provided for those subjects in the Randomized Analysis Set who did not meet at least 1 eligibility (inclusion or exclusion) criterion. The listing will present the eligibility criterion (or criteria if more than 1 deviation) that subjects did not meet and related comments, if collected.

Protocol deviations occurring after subjects entered the study are documented during routine monitoring. The number and percentage of subjects with important protocol deviations by deviation reason (eg, nonadherence to study drug, violation of select inclusion/exclusion criteria) will be summarized by treatment group for the Randomized Analysis Set. A by-subject listing will be provided for subjects with important protocol deviations.

A separate by-subject listing of subjects who received the incorrect study drug will also be provided.

5. BASELINE CHARACTERISTICS

5.1. Demographics and Baseline Characteristics

Subject demographic data (eg, age, sex, race, and ethnicity) and baseline characteristics (eg, body weight, height, body mass index [BMI], and Vitamin D) will be summarized by treatment group and overall using descriptive statistics (sample size, mean, SD, median, Q1, Q3, minimum, and maximum) for continuous data and by the number and percentage of subjects for categorical data. Age is calculated as age in years at the first dose of study drug. The summaries of demographic data and baseline subject characteristics will be provided for the Open-Label Safety Analysis Set.

In addition, the following baseline characteristics and medical history will be summarized:

- Age group (< 50 years, ≥ 50 years and further split into ≥ 50 to < 60 years and > 60 years)
- Region (Asia, Europe or North America)
- BMI categories (< 18.5 kg/m^2 [underweight], ≥ 18.5 - 25.0 kg/m^2 [normal], ≥ 25.0 - 30.0 kg/m^2 [overweight], and ≥ 30.0 kg/m^2 [obese])
- HBV DNA categories (< 20 IU/mL, 20 to < 69 IU/mL, ≥ 69 IU/mL)
- HBsAg (log_{10} IU/mL)
- ALT (U/L)
- ALT level based on central laboratory normal range (≤ ULN, > ULN - 5 × ULN, > 5 × ULN)
- ALT level based on 2018 AASLD normal range with the ULN as 25 U/L for females and 35 U/L for males (≤ ULN, > ULN - 5 × ULN, > 5 × ULN)
- HBeAg and HBeAb status (HBeAg positive, HBeAg negative/HBeAb negative, HBeAg negative/HBeAb positive)
- Years positive for HBV
- Cirrhosis history (yes, no, indeterminate/unknown)
- FibroTest score
- Fibrosis stage by FibroTest score (0 - 0.48, 0.49 - 0.74, 0.75 - 1)
- Estimated GFR by CG, CKD-EPI creatinine, and CKD-EPI Cystatin C methods
- Proteinuria by urinalysis (dipstick) (Grade 0, Grade 1, Grade 2, Grade 3)

- Clinical BMD status for hip and spine (normal, osteopenia, osteoporosis)
- Hip fracture and major osteoporotic fracture probabilities estimated using FRAX® (see Section 7.3.4)
- Previous experience with treatment other than TDF (yes, no)
- Previous interferon experience (yes, no)
- Medical history: diabetes mellitus (yes, no), hypertension (yes, no), cardiovascular disease (yes, no), and hyperlipidemia (yes, no)

Diabetes mellitus, hypertension, cardiovascular disease, and hyperlipidemia above are determined by medical history, adverse events, and concomitant medication data, which will be reviewed by the Gilead medical monitor before to data finalization for the final analysis.

5.2. Medical History

Medical history will be collected at screening for disease-specific and general conditions (ie, conditions not specific to the disease being studied). A listing of medical history data will be provided using the Randomized Analysis Set.

6. EFFICACY ANALYSES

For the final analysis, efficacy data will be summarized for (1) subjects in the FAS from baseline up to Week 96 and statistical comparisons will be performed, (2) subjects in the Open-Label FAS from open-label baseline up to Week 96 and no statistical comparisons will be performed, and (3) subjects in the TFFU FAS from TFFU baseline to the end of the study and no statistical comparisons will be performed. All efficacy data, including data collected during the 24-week TFFU period, will be listed.

6.1. Primary Efficacy Endpoint

6.1.1. Definition of the Primary Efficacy Endpoint

The primary efficacy endpoint is the proportion of subjects with HBV DNA ≥ 20 IU/mL (as determined by the modified US FDA-defined snapshot algorithm) at Week 48. The primary efficacy endpoint was analyzed in the Week 48 analysis and will not be re-analyzed in the final analysis. Details are provided in the Week 48 SAP.

6.2. Secondary Efficacy Endpoints

6.2.1. Definition of Secondary Efficacy Endpoints

The secondary efficacy endpoints are:

- The proportion of subjects with HBV DNA ≥ 20 IU/mL (as determined by the modified US FDA-defined snapshot algorithm) at Week 96
- The proportion of subjects with HBV DNA < 20 IU/mL and target detected/not detected (ie, < LLOD) at Weeks 48 and 96
- The proportion of subjects with HBeAg loss and proportion with seroconversion to anti-HBe at Weeks 48 and 96
- The proportion of subjects with HBsAg loss and proportion with seroconversion to anti-HBs at Weeks 48 and 96
- The proportion of subjects with normal ALT and proportion with normalized ALT (by central laboratory and AASLD criteria) at Weeks 48 and 96
- The change from baseline in fibrosis as assessed by FibroTest at Weeks 48 and 96

For the final analysis, the following definitions will be used:

- HBsAg loss is defined as HBsAg test result changes from HBsAg positive at baseline to HBsAg negative at a postbaseline visit with baseline HBsAb negative or missing.
- HBeAg loss is defined as HBeAg test result changes from HBeAg positive at baseline to HBeAg negative at a postbaseline visit with baseline HBeAb negative or missing.
- HBsAg seroconversion is defined as HBsAg loss and HBsAb test result changes from HBsAb negative or missing at baseline to HBsAb positive at a postbaseline visit.
- HBeAg seroconversion is defined as HBeAg loss and HBeAb test result changes from HBeAb negative or missing at baseline to HBeAb positive at a postbaseline visit.
- ALT normalization is defined as ALT > ULN (by central laboratory normal range or 2018 AASLD normal range) at baseline but within normal range at a postbaseline visit.

Both baseline and postbaseline borderline serology results will be imputed using the following rules:

- HBsAg and HBeAg borderline will be considered as HBsAg positive and HBeAg positive.
- HBsAb and HBeAb borderline will be considered as HBsAb negative and HBeAb negative.

For the above definitions, baseline will be replaced with open-label baseline for the analyses of subjects in the Open-Label FAS, and with TFFU baseline for the analyses of subjects in the TFFU FAS.

6.2.2. Analysis Methods for Secondary Efficacy Endpoints

The primary analysis of the proportion of subjects in the Full Analysis Set with HBV DNA ≥ 20 IU/mL (as determined by the modified US FDA-defined snapshot algorithm {U. S. Department of Health and Human Services 2015}) will be repeated for the Week 96 endpoint using the FAS. The Week 96 analysis window is defined as from Study Day 589 to Study Day 840, inclusive.

The analyses for the other secondary efficacy endpoints will be conducted using the FAS. All secondary efficacy endpoints involving proportions will be analyzed from baseline to Week 96 using both the Missing Failure (M F) approach and the Missing Excluded (M E) approach, unless otherwise specified. In the M E approach, all missing data will be excluded in the computation, and in the M F approach, all missing data will be treated as failures and included in both the numerator and denominator. P-values will be calculated using the CMH test stratified by baseline HBeAg status (HBeAg-positive or HBeAg-negative) and age (< 50 or ≥ 50 years), and the proportion difference between the 2 treatment groups and the associated 95% CIs will be calculated based on the stratum-adjusted MH proportion.

The change from baseline in $\log_{10}$ (HBsAg) (IU/mL), both overall and by baseline HBeAg status, and ALT will be summarized by visit using observed data (ie, missing data will be excluded). The p-value, differences in change from baseline in $\log_{10}$ (HBsAg) (IU/mL) and ALT between the 2 treatment groups and the associated 95% CI will be constructed using analysis of variance (ANOVA) models, including treatment group, baseline HBeAg status (HBeAg-positive or HBeAg-negative), and age (< 50 or ≥ 50 years) as fixed effects in the model. For the analysis of change from baseline in $\log_{10}$ (HBsAg) (IU/mL) by baseline HBeAg status, only treatment group and age (< 50 or ≥ 50 years) will be included as fixed effects in the model.

In addition, the proportion of subjects in the FAS with HBV DNA < 20 IU/mL, the proportion of subjects with normal ALT by the central laboratory and 2018 AASLD criteria (ULN is 35 U/L for males and 25 U/L for females) using the M E approach, and the mean change from baseline in $\log_{10}$ (HBsAg) (IU/mL) and ALT (U/L) will be plotted with 95% CI over time using observed data from baseline to Week 96.

Fibrosis assessed by FibroTest at Week 24, 48, and 96, and the change from baseline in FibroTest score will be summarized using descriptive statistics (sample size, mean, SD, median, Q1, Q3, minimum, and maximum) by treatment group for subjects in the FAS. The p-value, differences in change from baseline in FibroTest score between the 2 treatment groups and the associated 95% CI will be constructed using ANOVA models, including treatment group, baseline HBeAg status (HBeAg-positive or HBeAg-negative) and age (< 50 or ≥ 50 years) as fixed effects in the model. A shift table of fibrosis stage from baseline based on the FibroTest score at Week 24, 48, and 96 will be provided. Mean change from baseline in FibroTest score will be plotted with 95% CI over time using observed data from baseline to Week 96. A listing of fibrosis assessment by FibroTest will also be provided.

All the analyses of the secondary efficacy endpoints, using both the M F and the M E approach, and plots described above will be repeated for subjects in the Open-Label FAS from open-label baseline to Week 96. No statistical comparisons will be performed.

For subjects in the TFFU FAS, only HBV DNA, ALT, HBsAg, HBeAg loss/seroconversion, and HBsAg loss/seroconversion endpoints will be summarized by visit in the TFFU phase as described above only using the M E approach. No statistical comparisons will be performed.

6.3. Changes From Protocol-Specified Efficacy Analyses

FibroTest measurements were not specified to be collected at Week 24 in the study protocol but were collected in the study. FibroTest analyses were updated to include the Week 24 measurement.

7. SAFETY ANALYSES

For the final analysis, cumulative safety data (treatment-emergent AEs, treatment-emergent laboratory abnormalities, etc.) will be summarized for (1) subjects in the FAS during the double-blind phase, (2) subjects in the Open-Label Safety Analysis Set during the open-label phase, and (3) subjects in the TFFU Safety Analysis Set during the TFFU phase. Exposure data will be summarized for (1) and (2).

By-visit summary tables of safety measurements will be presented for the entire study treatment period (double-blind and open-label phase combined) based on subjects in the Safety Analysis Set from baseline to Week 96, and open-label phase based on subjects in the Open-Label Safety Analysis Set from open-label baseline to Week 96. Statistical comparisons will only be performed for the analysis of the entire study treatment period, unless otherwise specified. No by-visit summaries for safety measurements will be generated for the TFFU phase.

All safety data including data collected during the treatment-free follow-up period will be included in data listings.

7.1. Adverse Events

7.1.1. Adverse Event Dictionary

AEs will be coded using the Medical Dictionary for Regulatory Activities (MedDRA). System Organ Class (SOC), high-level group term (HLGT), high-level term (HLT), preferred term (PT), and lower-level term (LLT) will be attached to the clinical database.

7.1.2. Adverse Event Severity

Adverse events are graded by the investigator as Grade 1 (mild), Grade 2 (moderate), Grade 3 (severe) or Grade 4 (life threatening) according to toxicity criteria specified in Appendix 4 of the clinical study protocol. The severity grade of events for which the investigator did not record severity will be categorized as “missing” for tabular summaries and data listings, and will be considered the least severe for the purposes of sorting for data presentation.

7.1.3. Relationship of AEs to Study Drug

Related AEs are those for which the investigator answers “Yes” to the question “Related to Study Treatment?” in the eCRF. Events for which the investigator did not record relationship to study drug will be considered related to study drug for summary purpose. Data listings will show relationship as missing.

7.1.4. Relationship of AEs to Study Procedure

AEs for which ‘Yes’ is marked for question ‘Related to Study Procedures?’ in the eCRF will be identified and included in AE listing.

7.1.5. Serious AEs

Serious AEs (SAEs) are those identified as serious in the eCRF, where 'Yes' was marked for 'AE serious'. The clinical database will be reconciled with the SAE database (from the Pharmacovigilance and Epidemiology [PVE] Department) before database finalization.

7.1.6. Treatment-Emergent AEs

7.1.6.1. Definition of Treatment-Emergent

Treatment-emergent AEs will be separately defined for the double-blind phase and the open-label phase. Post-treatment-emergent AEs will be defined only for the TFFU phase.

Treatment-emergent AEs occurring during the double-blind phase are defined as:

- Any AE with onset date on or after the blinded study drug start date and no later than the minimum of the blinded study drug stop date + 3 days and the day prior to the first dose date of open-label study drug, if applicable, for those who discontinued blinded study drug permanently, or
- Any AE with onset date on or after the blinded study drug start date for those who are still on the blinded study drug, or
- Any AE leading to blinded study drug discontinuation.

Treatment-emergent AEs occurring during the open-label phase are defined as:

- Any AE with onset date on or after the open-label study drug start date and no later than the open-label study drug stop date + 3 days for those who discontinued open-label study drug permanently, or
- Any AE with onset date on or after the open-label study drug start date for those who are still on open-label study drug, or
- Any AE leading to open-label study drug discontinuation.

Post-treatment-emergent AEs occurring during the TFFU phase are defined as:

- Any AE with onset date after the last dose of study drug + 3 days

7.1.6.2. Incomplete Dates

If an AE onset date is incomplete or completely missing, the following rules will be used to determine if the AE is considered treatment emergent during the double-blind and open-label phases, or post-treatment-emergent during the TFFU phase:

Events with Missing Onset Day and/or Month

The event is treatment-emergent during the double-blind phase if the following criteria are met:

- The month and year (or year) of onset date is the same as or after the month and year (or year) of the first dose of the blinded study drug, and
- For those who discontinued the blinded study drug permanently only: the month and year (or year) of onset date is the same as or before the month and year (or year) of the date of the minimum of the blinded study drug stop date + 3 days and the day prior to the first dose date of open-label study drug, if applicable, and
- AE end date is as follows:

  The (complete) end date is on or after the first dose date of the blinded study drug, or

  The month and year (or year) of end date is the same or after the month and year (or year) of the first dose of the blinded study drug, or

  End date is completely missing

The event is treatment-emergent during the open-label phase if the following criteria are met:

- The month and year (or year) of onset date is the same as or after the month and year (or year) of the first dose of the open-label study drug, and
- For those who discontinued the blinded study drug permanently only: the month and year (or year) of onset date is the same as or before the month and year (or year) of the date of the last dose of the open-label study drug + 3 days, and
- AE end date is as follows:

  The (complete) end date is on or after the first dose date of the open-label study drug, or

  The month and year (or year) of end date is the same or after the month and year (or year) of the first dose of the open-label study drug, or

  End date is completely missing

The event is post-treatment-emergent during the TFFU phase if the following criteria are met:

- The month and year (or year) of onset date is the same as or after the month and year (or year) of the day after the last dose date + 3 days

Events with Completely Missing Onset Date

An AE with a completely missing onset date is defined as treatment-emergent AE during the double-blind phase if end date is as follows:

- The (complete) end date is on or after the first dose date of the blinded study drug, or
- The month and year (or year) of end date is the same or after the month and year (or year) of the first dose of the blinded study drug, or
- End date is completely missing

An AE with a completely missing onset date is defined as treatment-emergent AE during the open-label phase if end date is as follows:

- The (complete) end date is on or after the first dose date of the open-label study drug, or
- The month and year (or year) of end date is the same or after the month and year (or year) of the first dose of the open-label study drug, or
- End date is completely missing

The event is post-treatment-emergent during the TFFU phase if the following criteria are met:

- The (complete) end date is after the last dose date + 3 days, or
- The month and year (or year) of end date is the same or after the month and year (or year) of the day after the last dose date + 3 days, or
- End date is completely missing

7.1.7. Summaries of AEs and Deaths

AEs will be summarized for (1) subjects in the Safety Analysis Set during the double-blind phase, (2) subjects in the Open-Label Safety Analysis Set during the open-label phase, and (3) subjects in the TFFU Safety Analysis Set during the TFFU phase. The full set of AE summaries as described below will be generated for (1) and (2).

A brief summary of AEs (ie, the number and percentage of subjects) will be presented by treatment group for the following: any treatment-emergent AE, any Grade 3 or 4 treatment-emergent AE, any Grade 2, 3 or 4 treatment-emergent AE, any treatment-emergent study drug-related AE, any Grade 3 or 4 treatment-emergent study drug-related AE, any Grade 2, 3 or 4 treatment-emergent study drug-related AE, any treatment-emergent serious adverse event (SAE),

any treatment-emergent study drug-related SAE, any treatment-emergent AE leading to premature study drug discontinuation, any treatment-emergent AE leading to dose modification or study drug interruption, and any death.

Summaries (number and percentage of subjects) of AEs (by SOC, HLT [if specified below], and PT) will be provided by treatment group and overall as follows:

- All treatment-emergent AEs summarized by SOC, HLT, and PT
- Any Grade 3 or 4 treatment-emergent AEs
- Any Grade 2, 3, or 4 treatment-emergent AEs
- All treatment-emergent nonserious AEs occurring in at least 5% of subjects in any treatment group (this summary is generated per requirement for reporting in ClinicalTrials.gov)
- All treatment-emergent study drug-related AEs summarized by SOC, HLT, and PT
- Any Grade 3 or 4 treatment-emergent study drug-related AEs
- Any Grade 2, 3, or 4 treatment-emergent study drug-related AEs
- All treatment-emergent SAEs
- All treatment-emergent study drug-related SAEs
- All treatment-emergent AEs leading to premature discontinuation from study drug
- All treatment-emergent AEs leading to dose modification or study drug interruption

Multiple events will be counted once only per subject in each summary. For data presentation, SOC (and HLT) will be ordered alphabetically, with PT sorted by decreasing total frequency. For summaries by severity grade, the most severe event will be selected.

For post-treatment AEs during the TFFU phase, the overall summary and individual summaries will be provided only for the following:

- All post-treatment-emergent AEs
- Any Grade 3 and 4 post-treatment-emergent AEs
- All post-treatment-emergent SAEs

In addition to the by-treatment summaries, data listings will be provided for the following:

- All AEs

- Grade 3 and 4 AEs
- SAEs
- Study drug-related SAEs
- Deaths
- AEs leading to premature discontinuation of study drug
- AEs leading to dose modification or study drug interruption

7.2. Laboratory Evaluations

Laboratory data collected during the study will be analyzed and summarized using both quantitative and qualitative methods. Summaries of laboratory data will be provided for (1) the double-blind phase and open-label phase for subjects in the Safety Analysis Set from baseline to Week 96, and (2) summaries of laboratory data will be provided for the open-label phase only for subjects in the Open-Label Safety Analysis Set from open-label baseline to Week 96. Only laboratory abnormalities will be summarized for subjects in the TFFU Safety Analysis Set from TFFU baseline to the end of study.

The analyses will be based on values reported in conventional units. When values are below the LOQ, they will be listed as such, and the closest imputed value will be used for the purpose of calculating summary statistics as specified in Section 3.5.

A by-subject listing for laboratory test results will be provided by subject ID number and visit in chronological order for hematology, serum chemistry, urine chemistry, and urinalysis separately. Values falling outside of the relevant reference range and/or having a severity grade of 1 or higher on the Gilead Grading Scale for Severity of Adverse Events and Laboratory Abnormalities will be flagged in the data listings, as appropriate.

7.2.1. Summaries of Numeric Laboratory Results

Descriptive statistics (sample size, mean, SD, median, Q1, Q3, minimum, and maximum) will be provided by treatment group for each laboratory test during the double-blind phase and open-label phase for subjects in the Safety Analysis Set from baseline to Week 96 as follows:

- Baseline values
- Values at each postbaseline analysis window
- Change from baseline to each postbaseline analysis window
- Percentage change from baseline to each postbaseline analysis window (if specified)

A baseline laboratory value will be defined as the last measurement obtained on or prior to the date of first dose of blinded study drug. Change from baseline to a postbaseline visit will be defined as the visit value minus the baseline value.

Descriptive statistics (sample size, mean, SD, median, Q1, Q3, minimum, and maximum) will also be provided by treatment group for each laboratory test during the open-label phase only based on subjects in the Open-Label Safety Analysis Set from open-label baseline to Week 96, as follows:

- Open-label baseline values
- Values at each open-label postbaseline analysis window
- Change from open-label postbaseline to each open-label postbaseline analysis window
- Percentage change from open-label postbaseline to each open-label postbaseline analysis window (if specified)

An open-label baseline laboratory value will be defined as the last measurement obtained on or prior to the date of first dose of open-label study drug. Change from open-label baseline to an open-label postbaseline visit will be defined as the visit value minus the open-label baseline value.

Descriptive summaries for numeric laboratory results measured during the TFFU phase will not be provided.

Mean, median, Q1, Q3, minimum, and maximum values will be displayed to the reported number of digits; SD values will be displayed to the reported number of digits plus 1. In the case of multiple values in an analysis window, data will be selected for analysis as described in Section 3.6.2.

Median (Q1, Q3) change from baseline to Week 96 in selected safety endpoints over time, including the fasting lipid panel parameters and fasting glucose, for subjects in the Safety Analysis Set will be plotted by treatment group. This will be repeated for subjects in the Open-Label Safety Analysis Set from open-label baseline to Week 96.

7.2.1.1. Metabolic Assessments

For the lipid panel and glucose measurements, only those under fasting status will be summarized for subjects in the Safety Analysis Set from baseline to Week 96. P-values for the difference between the 2 treatment groups in baseline values and the change from baseline in fasting lipid and glucose data (including total cholesterol, triglycerides, LDL, HDL, total cholesterol to HDL ratio, and glucose) will be estimated from the 2-sided Wilcoxon rank sum test.

Fasting lipid data (including total cholesterol, LDL, HDL and triglycerides) will also be analyzed using the following National Cholesterol Education Program (NCEP) Adult Treatment Panel (ATP) III categories {National Cholesterol Education Program (NCEP) 2001}:

- **For total cholesterol (mg/dL):** < 200 (desirable), 200-239 (borderline high), and ≥ 240 (high)
- **For LDL (mg/dL):** < 100 (optimal), 100-129 (near optimal/above optimal), 130-159 (borderline high), 160-189 (high), and ≥ 190 (very high)
- **For HDL (mg/dL):** < 40 (low), 40-59 (normal), and ≥ 60 (high)
- **For triglycerides (mg/dL):** < 150 (normal), 150-199 (borderline high), 200-499 (high), and ≥ 500 (very high)

The number and proportion of subjects for the above categories of each lipid parameter will be summarized by its baseline category for each treatment group at each visit.

The summaries described above will be repeated for the subjects in the Open-Label Safety Analysis from open-label baseline to Week 96 using open-label baseline instead of baseline, and no statistical comparisons will be made.

7.2.1.2. Calcium Corrected for Albumin

Calcium corrected for albumin will be calculated and summarized by-visit for subjects in the Safety Analysis Set from baseline to Week 96, and separately for subjects in the Open-Label Analysis Set from open-label baseline to Week 96. The following formula will be used when both serum calcium and albumin results for a given blood draw are available and serum albumin value is < 4.0 g/dL.

Calcium corrected for albumin (mg/dL) serum calcium (mg/dL) + 0.8 × (4.0 albumin (g/dL))

When albumin value is ≥ 4.0 g/dL, the actual calcium results will be used. Toxicity grading for calcium will be applied based on the corrected values.

7.2.2. Graded Laboratory Values

The criteria specified in the protocol will be used to grade laboratory results as Grade 0, Grade 1 (mild), Grade 2 (moderate), Grade 3 (severe) or Grade 4 (life-threatening). Grade 0 includes all values that do not meet criteria for an abnormality of at least Grade 1. Some laboratory tests have criteria for both increased and decreased levels; analysis for each direction (ie, increased, decreased) will be presented separately.

For triglycerides, LDL, and total cholesterol, the protocol-specified toxicity grade scale is for fasting test values; non-fasting lipid results (or lipid results without known fasting status) will not be graded or summarized by toxicity grades.

7.2.2.1. Treatment-Emergent Laboratory Abnormalities

Treatment-emergent laboratory abnormalities occurring in the double-blind phase are defined as values that increase by at least 1 toxicity grade from baseline at any postbaseline visit up to and including the minimum of the blinded study drug stop date + 3 days and the first dose date of open-label study drug, if applicable, for those who discontinued blinded study drug permanently, or values that increase by at least 1 toxicity grade from baseline at any postbaseline visit for those who are still on double-blind study drug. If the relevant baseline laboratory value is missing, any laboratory abnormality of at least Grade 1 observed within the double-blind phase will be considered treatment emergent.

Treatment-emergent laboratory abnormalities occurring in the open-label phase are defined as values that increase by at least 1 toxicity grade from open-label baseline at any open-label postbaseline visit up to and including the last dose date of the open-label study drug + 3 days for those who discontinued open-label study drug permanently, or values that increase by at least 1 toxicity grade from open-label baseline at any open-label postbaseline visit for those who are still on open-label study drug. For the analyses of abnormalities occurring during open-label treatment, open-label baseline will be the last available record on or prior to Open-Label Study Day 1.

Post-treatment-emergent laboratory abnormalities occurring in the TFFU phase are defined as values that increase by at least 1 toxicity grade from TFFU baseline. For the analyses of abnormalities occurring during the TFFU phase, TFFU baseline will be the last available record on or prior to last dose date + 3 days.

Fasting glucose and nonfasting glucose (including glucose results without a known fasting status) are graded based on different grading scales as specified in the protocol. Treatment-emergent and post-treatment-emergent laboratory abnormalities will be summarized for fasting glucose. Since nonfasting glucose was not assessed at baseline, the maximum postbaseline grade will be summarized.

7.2.2.2. Treatment-Emergent Marked Laboratory Abnormalities

Treatment-emergent marked laboratory abnormalities occurring in the double-blind phase are defined as values that worsen by at least 3 grades from baseline at any postbaseline visit up to and including the minimum of the blinded study drug stop date + 3 days and the first dose date of open-label study drug, if applicable, for those who discontinued blinded study drug permanently, or values that worsen by at least 3 grades from baseline at any postbaseline visit for those who are still on blinded study drug. If relevant baseline laboratory data are missing, any laboratory abnormalities of at least Grade 3 or 4 observed within the double-blind time frame specified above will be considered as treatment-emergent marked laboratory abnormalities during the double-blind phase.

Treatment-emergent marked laboratory abnormalities occurring in the open-label phase are defined as values that worsen by at least 3 grades from open-label baseline at any open-label postbaseline visit up to and including the date of last dose of open-label study drug + 3 days for those who discontinued open-label study drug permanently, or values that worsen by at least

3 grades from open-label baseline at any open-label postbaseline visit for those who are still on open-label study drug.

Post-treatment-emergent marked laboratory abnormalities will not be summarized for the TFFU phase.

7.2.2.3. Summaries of Laboratory Abnormalities

The following summaries (number and percentage of subjects) of laboratory abnormalities will be provided by treatment group (subjects categorized according to most severe abnormality grade) for (1) subjects in the Safety Analysis Set during the double-blind phase, (2) subjects in the Open-Label Safety Analysis Set during the open-label phase, and (3) subjects in the TFFU Safety Analysis Set during the TFFU phase:

- Treatment-emergent laboratory abnormalities
- Treatment-emergent Grade 3 and 4 laboratory abnormalities
- Treatment-emergent marked laboratory abnormalities (this will not be summarized for (3))

For all summaries of laboratory abnormalities, the denominator is the number of subjects with any nonmissing postbaseline (or open-label postbaseline, or post-TFFU baseline) value in the given study period. By-subject listings of Grade 3 or 4 and graded laboratory abnormalities will be provided by subject ID number and visit in chronological order.

7.2.3. ALT Elevation

An ALT elevation is defined as serum ALT > 2 × open-label baseline value and > 10 × ULN, with or without associated symptoms. Confirmed ALT elevation (ALT flare) is defined as ALT elevations at 2 consecutive open-label postbaseline visits. All treatment-emergent ALT elevations including confirmed ALT elevations will be summarized for the open-label phase only. All treatment-emergent and nontreatment-emergent ALT elevations will be included in a listing.

7.3. Bone Safety Analyses

7.3.1. Bone Mineral Density (BMD)

Percentage change from baseline in hip BMD and spine BMD during the entire study treatment period will be summarized by treatment group and visit using descriptive statistics for subjects in the Hip and Spine DXA Analysis Sets. Percentage change from open-label baseline in hip BMD and spine BMD during the open-label phase will be summarized by treatment group and visit using descriptive statistics for subjects in the Open-Label Hip and Spine DXA Analysis Sets.

For each subject and visit, the clinical BMD status will be defined for hip and spine BMD based on the corrected t-score in Table 7-1.

Table 7-1. Normal, Osteopenia, and Osteoporosis as Defined by T-Score

Clinical Status	BMD T-Score
Normal	t-score ≥ -1.0
Osteopenia	-2.5 ≤ t-score < -1.0
Osteoporosis	t-score < -2.5

The number and percentage of subjects in each clinical BMD status (normal, osteopenia, and osteoporosis) will be summarized by visit and by baseline (or open-label baseline) clinical status for both hip and spine using observed data. Only for the entire study treatment period summary, the distribution of the clinical BMD status will be compared between the 2 treatment groups, adjusting for baseline status using rank analysis of covariance (ANCOVA) {LaVange 2008}.

The number and percentage of subjects in each category based on percentage change from baseline (or open-label baseline) in hip BMD and spine BMD (> 7% decrease, > 5% to ≤ 7% decrease, > 3% to ≤ 5% decrease, > 1% to ≤ 3% decrease, > 0 to ≤ 1% decrease, 0 to ≤ 1% increase, > 1% to ≤ 3% increase, > 3% to ≤ 5% increase, > 5% to ≤ 7% increase, > 7% increase) will be summarized by treatment group and visit using observed data. Only for the entire study treatment period summary, the distribution difference in these categories between the 2 treatment groups will be compared using the CMH test (row mean scores differ statistic).

Median (Q1, Q3) and mean (95% CIs) of percentage change from baseline to Week 96 in observed hip BMD and spine BMD over time will be plotted by treatment group for subjects in the corresponding Hip and Spine DXA Analysis Sets. This will be repeated using subjects in the Open-Label Hip and Spine DXA Analysis Sets from open-label baseline to Week 96.

7.3.2. Bone Biomarkers

Bone biomarkers include serum CTX, P1NP, and PTH.

Baseline, postbaseline, change from baseline, and percentage change from baseline in bone biomarkers will be summarized by treatment group and visit using descriptive statistics for subjects in the Safety Analysis Set from baseline to Week 96. Percentage change from baseline will be compared between the 2 treatment groups using the Wilcoxon rank sum test.

Open-label baseline, open-label postbaseline, change from open-label baseline, and percentage change from open-label baseline in bone biomarkers will be summarized by treatment group and visit using descriptive statistics for subjects in the Open-Label Safety Analysis Set from open-label baseline to Week 96. No statistical comparisons will be performed.

Median (Q1, Q3) percentage change from the baseline to Week 96 in each bone biomarker over time will be plotted by treatment group for subjects in the Safety Analysis Set. This will be repeated using subjects in the Open-Label Safety Analysis Set from open-label baseline to Week 96.

7.3.3. Fracture Events

The PTs for fracture events were defined based on HLGT of Fractures from the most current version of MedDRA at the time of the analysis (see Appendix 3). The number and percentage of subjects in the Open-Label Safety Analysis Set who experienced fracture events will be summarized by treatment group for the open-label phase. A data listing of fracture events will be provided.

7.3.4. Assessment of Fracture Probability

Fracture probabilities will be assessed using FRAX, a computer based algorithm developed by the World Health Organization (WHO; http://www.shef.ac.uk/FRAX).

The FRAX algorithm is based on individual patient models that integrate the risks associated with clinical risk factors as well as BMD at the femoral neck. The algorithm provides both the 10-year probability of hip fracture and the 10-year probability of a major osteoporotic fracture (clinical spine, forearm, hip, or shoulder fracture).

The FRAX model is constructed from real data in population-based cohorts around the world that have a limited age range. For an age below 40 or above 90 years, the tool will calculate the probability of fracture at the age of 40 or 90 years, respectively. Due to the age limitation, 2 sets of analyses of fracture probabilities will be performed.

In the first set of analysis, summaries of baseline and change from baseline to Week 96 in the 10-year probabilities of hip fracture, as well as major osteoporotic fracture, will be presented by treatment group and visit for subjects in the Safety Analysis Set aged between 40 and 90 years. Change from baseline in the 10-year fracture probabilities at each analysis visit will be compared between the 2 treatment groups using ANOVA, which includes treatment as a fixed effect.

In the second set of analysis, the above-specified analysis will be performed to include all subjects, where subjects in the Safety Analysis Set with an age below 40 or above 90 years will be treated as having an age of 40 or 90 years, respectively, in computing their fracture probabilities.

The analyses described above will be repeated for subjects in the Open-Label Safety Analysis Set from open-label baseline to Week 96, using open-label baseline in place of baseline, and no statistical comparisons will be performed.

Only data listings of fracture risk assessment questionnaire and FRAX fracture probabilities will be provided.

7.3.5. Bone Events

The PTs for bone events were defined based a search for bone disorder related adverse events reviewed by PVE using the most current version of MedDRA at the time of the analysis (see Appendix 2). The number and percentage of subjects in the Open-Label Safety Analysis Set who

experienced treatment-emergent bone events will be summarized by treatment group for the open-label phase. A data listing of bone events will be provided.

7.4. Renal Safety Analyses

7.4.1. Confirmed Renal Abnormalities

Confirmed renal abnormalities are defined as follows:

- Confirmed increase from open-label baseline in creatinine of at least 0.5 mg/dL or
- Confirmed creatinine clearance CL_{cr} by CG below 50 mL/min or
- Confirmed phosphorous < 2 mg/dL

Confirmed renal abnormalities are defined as renal abnormalities at 2 consecutive open-label postbaseline visits. Treatment-emergent confirmed renal abnormalities during the open-label phase will be summarized, where open-label baseline will be defined as the last available record on or prior to Open-Label Study Day 1. All confirmed renal abnormalities including those that occur during the open-label phase will be listed.

7.4.2. Serum Creatinine

For the renal safety analyses, observed creatinine values will be used.

For subjects in the Safety Analysis Set, the baseline and change from baseline in serum creatinine during the entire study period will be summarized using descriptive statistics. The baseline serum creatinine and change from baseline in serum creatinine will be compared between the 2 treatment groups using the two-sided Wilcoxon rank sum test.

For subjects in the Open-Label Safety Analysis Set, the open-label baseline and change from open-label baseline in serum creatinine during the open-label phase will be summarized using descriptive statistics. No statistical comparisons will be performed.

Median (Q1, Q3) and mean (95% CIs) of change from baseline to Week 96 in serum creatinine over time will be plotted by treatment group for subjects in the Safety Analysis Set. This will be repeated using subjects in the Open-Label Safety Analysis Set from open-label baseline to Week 96.

7.4.3. Estimated Glomerular Filtration Rate

Estimated glomerular filtration rate will be summarized for subjects in the Safety Analysis Set from baseline to Week 96.

The following formulae will be used to calculate eGFR:

- CG based:

  $eGFR_{CG}$ (mL/min) [(140 age (yrs)) × weight (kg) × (0.85 if female)] / (SCr (mg/dL) × 72),

  where weight is actual total body mass in kilograms, and SCr is serum creatinine.

- CKD-EPI Creatinine based:

  $eGFR_{CKD\ EPI,\ creatinine}$ (mL/min/1.73 m^2) 141 × $\min(SCr/\kappa,1)^{\alpha}$ × $\max(SCr/\kappa,1)^{1.209}$ × 0.993^{age} × 1.018 (if female) × 1.159 (if black),

  where κ is 0.7 for females and 0.9 for males, α is 0.329 for females and 0.411 for males, min indicates the minimum of SCr/κ or 1, and max indicates the maximum of SCr/κ or 1 {Levey 2009}.

- CKD-EPI Cystatin C based:

  $eGFR_{CKD\ EPI,\ cysC}$ (mL/min/1.73 m^2) 133 × $\min(SCys/0.8, 1)^{0.499}$ × $\max(SCys/0.8, 1)^{1.328}$ × 0.996^{age} [× 0.932 if female],

  where SCys is serum cystatin C.

P-values for the baseline, the change from baseline in $eGFR_{CG}$ and $eGFR_{CKD\ EPI,\ creatinine}$ will be estimated from the 2-sided Wilcoxon rank sum test to compare the 2 treatment groups. For $eGFR_{CKD\ EPI,\ cysC}$, only the baseline value will be summarized as the postbaseline tests were only performed as reflex tests.

The number and proportion of subjects with decrease from baseline of ≥ 25% and ≥ 50% in $eGFR_{CG}$ and $eGFR_{CKD\ EPI,\ creatinine}$ will be summarized by treatment groups for subjects in the Safety Analysis Set from baseline to Week 96. Statistical comparisons of the subject incidence rates between the 2 treatment groups during the entire study period will be performed using Fisher's exact test.

The number and proportion of subjects in each stage of chronic kidney disease (CKD) will be summarized by baseline stages of CKD at Weeks 12, 24, 36, 48, 60, 72, and 96 for the entire study treatment period based on subjects in the Safety Analysis Set. The distribution of the stages of CKD at these visits will be compared between the 2 treatment groups adjusting for entire study treatment period using rank ANCOVA {LaVange 2008}.

The stages of CKD are defined as follows:

- **Stage 1:** $eGFR_{CG}$ ≥ 90 mL/min
- **Stage 2:** $eGFR_{CG}$ ≥ 60 and < 90 mL/min
- **Stage 3:** $eGFR_{CG}$ ≥ 30 and < 60 mL/min
- **Stage 4:** $eGFR_{CG}$ < 30 mL/min

The summaries described above will be repeated for subjects in the Open-Label Safety Analysis Set from open-label baseline to Week 96, using open-label baseline in place of baseline, and no statistical comparisons will be performed.

Median (Q1, Q3) change from baseline in eGFR by CG and by CKD-EPI Creatinine over time will be plotted by treatment group for subjects in the Safety Analysis Set from baseline to Week 96. This will be repeated for subjects in the Open-Label Safety Analysis Set from open-label baseline to Week 96.

7.4.4. Treatment-Emergent Proteinuria (Dipstick)

Treatment-emergent proteinuria by urinalysis (dipstick) will be summarized by treatment group for the subjects in the Open-Label Safety Analysis Set during the open-label phase. A listing of subjects with treatment-emergent proteinuria will be provided.

7.4.5. Urine Creatinine, Urine RBP to Creatinine Ratio and Beta-2-Microglobulin to Creatinine Ratio

Baseline, postbaseline, change from baseline and percentage change from baseline in urine creatinine, urine RBP to creatinine ratio, and beta-2-microglobulin to creatinine ratio will be summarized by treatment group and visit using descriptive statistics during the double-blind and open-label phase. The difference in percentage change from baseline in these ratios between the 2 treatment groups will be tested using the Wilcoxon rank sum test only for the entire study period summary.

The summary described above will be repeated for subjects in the Open-Label Safety Analysis Set from open-label baseline to Week 96, using open-label baseline in place of baseline, and no statistical comparisons will be performed.

Median (Q1, Q3) percentage change from the baseline in the urine RBP to creatinine ratio, and beta-2-microglobulin to creatinine ratio over time will be plotted by treatment group for subjects in the Safety Analysis Set from baseline to Week 96. This will be repeated for subjects in the Open-Label Safety Analysis Set from open-label baseline to Week 96.

7.4.6. Proteinuria by Quantitative Assessment

Baseline, postbaseline, changes from baseline, and percentage change from baseline in UPCR and UACR will be summarized by treatment group and visit using descriptive statistics, for subjects in the Safety Analysis Set from baseline to Week 96. Percentage change from baseline will be compared between the 2 treatment groups using the Wilcoxon rank sum test only for the entire study period summary.

The number and proportion of subjects with UPCR $\leq$ 200 mg/g versus > 200 mg/g will be summarized by baseline category for each postbaseline visit during the entire study treatment period {KDIGO Guideline Development Staff 2013}.

The number and proportion of subjects with UACR < 30 mg/g versus ≥ 30 mg/g will be summarized by baseline category for each postbaseline visit during the entire study treatment period {KDIGO Guideline Development Staff 2013}.

The distribution of the UPCR and UACR categories during the entire study period will be compared between the 2 treatment groups adjusting for baseline categories using rank ANCOVA {LaVange 2008}, respectively.

The summaries described above will be repeated for subjects in the Open-Label Safety Analysis Set from open-label baseline to Week 96, using open-label baseline in place of baseline, and no statistical comparisons will be performed.

Median (Q1, Q3) percentage change from baseline in the 2 ratios over time will be plotted by treatment group for subjects in the Safety Analysis Set from baseline to Week 96. This will be repeated for subjects in the Open-Label Safety Analysis Set from open-label baseline to Week 96.

7.4.7. Other Renal Biomarkers

Other renal biomarkers include TmP/GFR and $FEPO_4$.

TmP/GFR based on serum creatinine {Barth 2000} will be calculated as follows:

$$TmP/GFR = TRP \times SPO_4 \quad if\ TRP \le 0.86$$
$$TmP/GFR = 0.3 \times TRP / \left[1 \quad (0.8 \times TRP)\right] \times SPO_4 \quad if\ TRP > 0.86$$

where TRP (tubular reabsoprtion of phosphate) is calculated by:

$$TRP \quad 1 \quad \frac{UPO_4}{SPO_4} \times \frac{SCr}{UCr}$$

where SCr is serum creatinine concentration (mg/dL), UPO_4 is urine phosphate concentration (mg/dL), SPO_4 is serum phosphate concentration, and UCr is urine creatinine concentration (mg/dL).

Urine $FEPO_4$ will be calculated as follows:

$$FEPO_4\ (\%) \quad (SCr \times UPO_4) / (SPO_4 \times UCr) \times 100\ (\%)$$

The baseline, postbaseline, and change from baseline in TmP/GFR and $FEPO_4$ will be summarized by treatment group and visit using descriptive statistics during the double-blind and open-label phase using the subjects in the Safety Analysis Set. Baseline and change from baseline will be compared between the 2 treatment groups using the Wilcoxon rank sum test only for the entire study period.

The summaries described above will be repeated for subjects in the Open-Label Safety Analysis Set from open-label baseline to Week 96, using open-label baseline in place of baseline, and no statistical comparisons will be performed.

Median (Q1, Q3) change from the baseline in TmP/GFR and $FEPO_4$ over time will be plotted by treatment group for subjects in the Safety Analysis Set from baseline to Week 96. This will be repeated for subjects in the Open-Label Safety Analysis Set from open-label baseline to Week 96.

7.5. Child-Pugh-Turcotte Score

The Child-Pugh-Turcotte (CPT) score is calculated as: the sum of the scores related to the 5 items in the table below (if any of the components are missing, the score is not calculated):

Measure	1 point	2 points	3 points
Total Bilirubin (mg/dL)	< 2	2 - 3	> 3
Serum Albumin (g/dL)	> 3.5	2.8 – 3.5	< 2.8
INR	< 1.7	1.7 – 2.3	> 2.3
Ascites	None	Slight/Moderate (medically controlled)	Severe (medically uncontrolled)
Hepatic Encephalopathy	None	Slight/Moderate (medically controlled)	Severe (medically uncontrolled)

For summary purposes, the score will be further classified into classes: Class A (5-6 points), Class B (7-9 points), and Class C (10-15 points). The number and proportion of subjects in each class of CPT score will be summarized in a shift table with baseline, Week 48, and Week 96 CPT class based on subjects in the Safety Analysis Set during the entire study treatment period. The distribution of the CPT class at these visits will be compared between the 2 treatment groups adjusting for baseline class using rank ANCOVA {LaVange 2008}.

The summaries described above will be repeated for subjects in the Open-Label Safety Analysis Set from open-label baseline to Week 96, using open-label baseline in place of baseline, and no statistical comparisons will be performed.

7.6. Body Weight

Body weight at each visit and change from baseline in body weight will be summarized using descriptive statistics (sample size, mean, SD, median, Q1, Q3, minimum, and maximum) by treatment group for each postbaseline analysis window, using subjects in the Safety Analysis Set

from baseline to Week 96. In the case of multiple values in an analysis window, data will be selected for analysis as described in Section 3.6.2.

This summary will be repeated for subjects in the Open-Label Safety Analysis Set from open-label baseline to Week 96.

7.7. Concomitant Medications

Concomitant medications (ie, medications other than study drug that are taken while receiving study drug) will be coded using the WHO Drug Dictionary. Use of concomitant medications for the open-label phase will be summarized (number and percentage of subjects) by treatment group and preferred name. Multiple drug use (by preferred name) will be counted only once per subject. The summary will be sorted by decreasing total frequency.

If the start or stop date of concomitant medications is incomplete, the month and year (or year alone if month is not recorded) of start or stop date will be used to determine if the medications are concomitant as follows. The medication is concomitant for the double-blind phase if the month and year of start or stop (or year of the start or stop) of the medication do not meet any of following criteria:

- The month and year of start of the medication is after the date of the last dose of blinded study drug
- The month and year of stop of the medication is before the date of the first dose of blinded study drug

The medication is concomitant for the open-label phase if the month and year of start or stop (or year of the start or stop) of the medication do not meet any of following criteria:

- The month and year of start of the medication is after the date of the last dose of open-label study drug
- The month and year of stop of the medication is before the date of the first dose of open-label study drug

If both the start and stop date of the medication are missing, the medication will be considered as concomitant during both double-blind and open-label phases.

If the start and stop date of the medications are not missing, and the start date is not after the last dose date of the blinded study drug and the stop date is not before the first dose date of the blinded study drug, or the medications are marked as ongoing and start date is on or before the last dose date of the blinded study drug, the medications are considered concomitant during the double-blind phase.

Similarly, if the start and stop date of the medications are not missing, and the start date is not after the last dose date of the open-label study drug and the stop date is not before the first dose date of the open-label study drug, or the medications are marked as ongoing and start date is on

or before the last dose date of the open-label study drug, the medications are considered concomitant during the open-label phase.

Summaries of concomitant medications will be provided for the open-label phase using the Safety Analysis Set. No inferential statistics will be provided. Subjects with any concomitant medication use will also be listed.

7.8. Electrocardiogram (ECG) Results

The number and percentage of subjects with an investigator's ECG assessment of normal, abnormal but not clinically significant, or abnormal and clinically significant will be summarized by treatment group for subjects in the Safety Analysis Set from baseline to Week 96. This summary will be repeated for subjects in the Open-Label Safety Analysis Set from open-label baseline to Week 96. No inferential statistics will be provided.

A by-subject listing of safety ECG results will be provided including treatment, assessment date and time, and ECG results.

7.9. Other Safety Measures

A data listing will be provided for subjects experiencing pregnancy during the study.

Listings of cirrhosis and hepatocellular carcinoma assessment results will be provided.

7.10. Changes From Protocol-Specified Safety Analyses

There are no deviations from the protocol-specified safety analyses.

8. PATIENT REPORTED OUTCOMES

The effect of treatment on health related quality of life questionnaires (via SF-36, CLDQ, and WPAI questionnaires) will be analyzed by visit for subjects in the Safety Analysis Set for the entire study treatment period (ie, from baseline up to Week 96), and separately for subjects in the Open-Label Safety Analysis Set for the open-label phase only (ie, from open-label baseline up to Week 96).

8.1. Definition of Patient Reported Outcomes

8.1.1. SF-36

The Short Form-36 (SF-36) includes 11 questions with multiple sub-questions that yield a total of 36 questions. The responses are rated on a 3-point, 5-point, or 6-point scale, and scoring are performed on 8 domains: Physical function (PF), Role physical (RP), Bodily Pain (BP), General health perceptions (GH), Vitality (VT), Social functioning (SF), Role emotional (RE), Mental health (MH).

Scale	Item	Abbreviated Item Content
Physical Functioning (PF)	3a	Vigorous activities, such as running, lifting heavy objects, or participating in strenuous sports
	3b	Moderate activities, such as moving a table, pushing a vacuum cleaner, bowling, or playing golf
	3c	Lifting or carrying groceries
	3d	Climbing several flights of stairs
	3e	Climbing one flight of stairs
	3f	Bending, kneeling, or stooping
	3g	Walking more than a mile
	3h	Walking several hundred yards
	3i	Walking one hundred yards
	3j	Bathing or dressing oneself
Role-Physical (RP)	4a	Cut down the amount of time one spent on work or other activities
	4b	Accomplished less than you would like
	4c	Limited in kind of work or other activities
	4d	Had difficulty performing work or other activities (e.g., it took extra effort)
Bodily Pain (BP)	7	Intensity of bodily pain
	8	Extent pain interfered with normal work

Scale	Item	Abbreviated Item Content
General Health (GH)	1	Is your health: excellent, very good, good, fair, poor
	11a	Seem to get sick a little easier than other people
	11b	As healthy as anybody I know
	11c	Expect my health to get worse
	11d	Health is excellent
Vitality (VT)	9a	Feel full of life
	9e	Have a lot of energy
	9g	Feel worn out
	9i	Feel tired
Social Functioning (SF)	6	Extent health problems interfered with normal social activities
	10	Frequency health problems interfered with social activities
Role-Emotional (RE)	5a	Cut down the amount of time spent on work or other activities
	5b	Accomplished less than you would like
	5c	Did work or other activities less carefully than usual
Mental Health (MH)	9b	Been very nervous
	9c	Felt so down in the dumps that nothing could cheer you up
	9d	Felt calm and peaceful
	9f	Felt downhearted and depressed
	9h	Been happy
Health Transition (HT)	2	How health is now compared to before

Two summary scales will be presented based on the domain scores:

- Physical component summary (PCS): combines PF, RP, BP, and GH
- Mental component summary (MCS): combines VT, SF, RE, and MH

The summary scores aggregate information from the 8 SF-36 domains in a way that captures 80% to 85% of the variance in the 8 domains.

The SF-36 physical component and the mental component scores will be calculated based on the 8 domain scores using the SF-36v2 scoring method (http://www.sf-36.org/demos/SF-36v2.html).

8.1.2. CLDQ

The Chronic Liver Disease Questionnaire (CLDQ) includes 29 questions which aim to survey how subjects felt during the last two weeks. The responses are rated on a 7-point scale, and

translated into 6 domain scores: fatigue, activity, emotional function, abdominal symptoms, systemic symptoms, and worry, and an overall CLDQ score. Specifically,

- Abdominal symptoms (AB) Mean of {R1, R5, R17}
- Fatigue (FA) Mean of {R2, R4, R8, R11, R13}
- Systemic (SY) Mean of {R3, R6, R21, R23, R27}
- Activity (AC) Mean of {R7, R9, R14}
- Emotion (EM) Mean of {R10, R12, R15, R16, R19, R20, R24, R26}
- Worry (WO) Mean of {R18, R22, R25, R28, R29}

where R# is the response to question number # (eg, R1 is the response to question 1).

The overall CLDQ score will then be calculated by taking the mean of 6 domain scores.

8.1.3. WPAI:SHP

The Work Productivity and Activity Impairment Questionnaire: Specific Health Hepatitis (WPAI:SHP) includes 6 questions that are summarized into a four WPAI:SHP scores. The first question is to determine the employment status. Questions 2 through 5 are applicable to employed patients only. Using subject responses to six questions, four scores are derived:

- percentage of absenteeism,
- percentage of presenteeism (reduced productivity while at work),
- an overall work impairment score that combines absenteeism and presenteeism and
- percentage of impairment in activities performed outside of work.

Greater scores indicate greater impairment
(http://www.reillyassociates.net/WPAI_Scoring.html).

Score	Source	Formula
Absenteeism	Question 2: During the past seven days, how many hours did you miss from work because of *{your health problems}*?	100 x [Q2/(Q2+Q4)]
	Question 4: During the past seven days, how many hours did you actually work?	
Presenteesism	Question 5: During the past seven days, how much did *{your health problems}* affect your productivity while you were working?	100 x [(Q5)/10]
Work Productivity Loss	Question 2: During the past seven days, how many hours did you miss from work because of *{your health problems}*?	100 x [Q2/(Q2+Q4) + {1 – Q2/(Q2+Q4)}x(Q5)/10]
	Question 4: During the past seven days, how many hours did you actually work?	
	Question 5: During the past seven days, how much did *{your health problems}* affect your productivity while you were working?	
Activity Impairment	Question 6: During the past seven days, how much did *{your health problems}* affect your ability to do your regular daily activities, other than work at a job?	100 x [(Q6)/10]

For the WPAI:SHP, four scores will be calculated based on the results.

8.2. Analysis Methods for Patient Reported Outcomes

Descriptive statistics for each of the SF-36 domains, and physical and component scores, CLDQ domain and overall scores, and the WPAI:SHP scores at baseline, Week 24, Week 48, Week 96, and change from baseline at Week 24, Week 48, and Week 96, will be presented by treatment group for subjects in the Safety Analysis Set from baseline to Week 96. Change from baseline for each score will be compared between treatment groups using a 2-sided Wilcoxon rank-sum test.

This will be repeated for subjects in the Open-Label Safety Analysis Set from open-label baseline to Week 96, using open-label baseline in place of baseline, and no statistical comparisons will be performed.

A by-subject listing for the scores from each of the questionnaires will be provided by subject ID number and visits in chronological order.

8.3. Changes From Protocol-Specified Patient Reported Outcomes Analyses

There are no deviations from the protocol-specified patient reported outcomes analyses.

9. REFERENCES

Barth JH, Jones RG, Payne RB. Calculation of renal tubular reabsorption of phosphate: the algorithm performs better than the nomogram. Ann Clin Biochem 2000;37 (Pt 1):79-81.

KDIGO Guideline Development Staff. KDIGO 2012 Clinical Practice Guideline for the Evaluation and Management of Chronic Kidney Disease. Kidney international. Supplement 2013;3 (1):v-150.

LaVange LM, Koch GG. Randomization-Based Nonparametric (ANCOVA). In: D'Agostino Sr. RB, Sullivan LM, Massaro JM, eds. Wiley Encyclopedia of Clinical Trials. John Wiley & Sons, Inc.; 2008: 31-8. vol 4).

Levey AS, Stevens LA, Schmid CH, Zhang YL, Castro AF, 3rd, Feldman HI, et al. A new equation to estimate glomerular filtration rate. Ann Intern Med 2009;150 (9):604-12.

National Cholesterol Education Program (NCEP). Detection, Evaluation, and Treatment of High Blood Cholesterol in Adults (Adult Treatment Panel III) Executive Summary. National Institute of Health May, 2001.

U. S. Department of Health and Human Services, Food and Drug Administration (FDA), Center for Drug Evaluation and Research (CDER). Human Immunodeficiency Virus-1 Infection: Developing Antiretroviral Drugs for Treatment. Guidance for Industry. Silver Spring, MD. November, 2015.

10. SOFTWARE

SAS® (SAS Institute Inc., Version 9.2, Cary, NC) is to be used for all programming of tables, listings, and figures.

nQuery Advisor® (Statistical Solutions Ltd., Version 6.0, Cork, Ireland) was used for the sample size and power calculation.

FRAX® (WHO Collaborating Center for Metabolic Bone Disease, University of Sheffield, UK) is to be used for the 10-year probabilities of hip fracture or a major osteoporotic fracture.

11. SAP REVISION

Revision Date (DD MMM YYYY)	Section	Summary of Revision	Reason for Revision

12. APPENDICES

Appendix 1. Study Procedures Table

Study Procedures	Screening (45 days)	Baseline (Day 1)	Blinded (± 3 days)						Open Label (± 7 days)				Follow Up[h]
			Week										
			4	8	12	24	36	48[k]	60	72	96	ED[g]	
Informed Consent	x												
Inclusion/Exclusion Criteria	x	x											
Medical History (including HBV disease and treatment history)	x	x											
Concomitant Medications	x	x	x	x	x	x	x	x	x	x	x	x	x
Adverse Events	x	x	x	x	x	x	x	x	x	x	x	x	x
Complete Physical Examination with weight and vital signs	x	x				x		x		x	x	x	
Height	x												
Body Weight	x	x	x	x	x	x	x	x	x	x	x	x	x
Vital Signs[a]	x	x	x	x	x	x	x	x	x	x	x	x	x
Symptom driven Physical Examination			x	x	x	x	x		x				x
Health Related Quality of Life (CLDQ, SF 36, WPAI)[m]		x				x		x			x	x	
Serum Chemistry and Liver Tests[b]	x	x	x	x	x	x	x	x	x	x	x	x	x
Hematology	x	x	x	x	x	x	x	x	x	x	x	x	x
Fasting metabolic panel[c]		x			x	x	x	x	x	x	x	x	
Pregnancy testing (women of child bearing potential only)	serum	urine (+ve test to be confirmed with serum)											
	x	x	x	x	x	x	x	x	x	x	x	x	
Estimated Creatinine Clearance by Cockcroft Gault method ($eGFR_{CG}$)	x	x	x	x	x	x	x	x	x	x	x	x	x
HBV serology (qualitative HBsAg and HBeAg) and quantitative HBsAg[i]	x	x			x	x	x	x	x	x	x	x	x

Study Procedures	Screening (45 days)	Baseline (Day 1)	Blinded (± 3 days)						Open Label (± 7 days)				Follow Up[h]
			Week										
			4	8	12	24	36	48[k]	60	72	96	ED[g]	
HCV, HDV, HIV Testing	x												
α fetoprotein (AFP)	x												
Urinalysis	x	x	x	x	x	x	x	x	x	x	x	x	x
Urine drug screen	x												
DXA scans (Hip & Spine)[d]	x					x		x		x	x	x	
ECG[e]	x							x			x	x	
Plasma HBV DNA levels	x	x	x	x	x	x	x	x	x	x	x	x	x
Randomization		x											
Study Drug Accountability			x	x	x	x	x	x	x	x	x		
Study Drug Dispensation		x	x	x	x	x	x	x	x	x			
FibroTest®		x						x			x	x	
Child Pugh Score[n]		x						x			x	x	
Serum Cystatin C[l]		x											
Fasting Blood for Bone Biomarker[f]		x	x		x	x		x		x	x	x	
Fasting Urine for Renal Biomarker[f]		x	x		x	x		x		x	x	x	
Fracture Risk Assessment (FRAX)		x											
Virology (Sequence analysis of HBV pol/RT for resistance surveillance)[j]		x	x	x	x	x	x	x	x	x	x	x	
Vitamin D		x											

Study Procedures	Screening (45 days)	Baseline (Day 1)	Blinded (± 3 days) Week 4	8	12	24	36	48[k]	Open Label (± 7 days) Week 60	72	96	ED[g]	Follow Up[h]
CCI													
Serum, plasma and urine for storage		x	x	x	x	x	x	x	x	x	x	x	x (serum only)

a Vital signs include blood pressure, pulse, respiration rate and temperature
b Serum chemistry and Liver Function Tests: alkaline phosphatase, AST, ALT, GGT, total bilirubin, direct and indirect bilirubin, total protein, albumin, LDH, CPK, bicarbonate, BUN, calcium, chloride, creatinine, glucose, phosphorus, magnesium, potassium, sodium, uric acid, and amylase (reflex lipase testing is performed in subjects with total amylase > 1.5 × ULN), and PTH. PTH analyzed at all visits except for Screening. At Baseline, Weeks 24, 48, 72, 96, and ED, analyses of glucose will be done as part of the fasting metabolic assessments and not as part of the chemistry panel. Liver function tests: PT/INR will be done at Screening, Week 48, 96 and ED and then as a reflex only test for ALT flares.
c Fasting glucose and lipid panel (total cholesterol, HDL, direct LDL, triglycerides). At Baseline, Weeks 12, 24, 36, 48, 60, 72, and 96/ED analyses of glucose will be done as part of the fasting metabolic assessments and not as part of the chemistry panel.
d The Baseline DXA can be performed at any time during the Screening period, but should be completed at least 14 days prior to the Baseline visit to ensure results are received prior to the first dose of study drug. The Week 24 DXA window is ± 14 days. The Week48 DXA window is 14 days only. DXA required for Early Discontinuation (ED) visit if not done within the last 12 weeks and should be done within ±14 days of the expected ED visit date
e Subjects must rest quietly in the supine position for a minimum of 5 minutes prior to the recording.
f Blood for selected bone biomarkers and urine for selected renal biomarkers will be collected in a fasted state. Required for ED visit if the last sample was collected > 12 weeks
g The Early Discontinuation (ED) visit should be performed within 72 hours of the last study drug dose (+ 3 days)
h Subjects who discontinue study drug due to HBsAg loss with confirmed seroconversion to anti HBs on or after the Week 24 visit, will be followed off treatment every 4 weeks for 12 weeks and then per the original study visit schedule through Week 96/ED. Subjects who have received at least one dose of study drug and permanently discontinue study drug for reasons other than HBsAg loss with confirmed seroconversion to anti HBs will be followed every 4 weeks for 24 weeks off treatment or up to initiation of appropriate, alternative HBV therapy, whichever occurs first. Initiation of appropriate, alternative HBV therapy is highly encouraged.
i HBeAb and HBsAb reflex testing will be performed as needed.
j Resistance sequence analysis may be performed at BL for subjects with HBV DNA ≥ 69 IU/mL and may be attempted for viremic subjects (HBV DNA ≥ 69 IU/mL) at Wks 48 and 96/ED. Phenotypic analysis will be performed for subjects that are subjected to sequence analysis. As it may not be known at the time of the visit whether a subject is viremic or if it will be their last study visit, a virology sample will be collected as each visit. In the event of unconfirmed virologic rebound (HBV DNA ≥ 20 IU/mL), subjects will be asked to return to the clinic for a scheduled or unscheduled blood draw. For virologic rebound occurring within the first 12 weeks of the study, the next scheduled visit will used for follow up. For virologic rebound occurring after Week 12, the subject will return for an unscheduled visit 2 3 weeks after the date of the original test that resulted with HBV DNA virologic rebound for confirmation of virologic rebound. At this follow up visit, a serum blood sample for resistance testing will be obtained. For unscheduled visits, the subject will be required to bring their supply of study drug with them and be assessed for adherence by pill count, and if necessary, the subject will be re counseled on adherence to study medication.
k The visit window for Week 48 visit is 14 days only.
l Serum Cystatin C will be done when $eGFR_{CG}$ falls < 50 ml/min during blinded period of study.
m Health Related Quality of Life surveys required for Early Discontinuation (ED) visit if not done within the last 24 weeks of the expected ED visit date
n Child Pugh Score will be assessed at Baseline (Day 1), Weeks 48 and 96. Assessment requires: total bilirubin, albumin, PT/INR, Ascites assessment, and Hepatic encephalopathy assessment

Appendix 2. Bone Events

The selected Preferred Terms of bone events based on MedDRA 22.1 are listed as follows:

Selected Preferred Terms
Abscess oral
Acetabulum fracture
Aneurysmal bone cyst
Ankle fracture
Arm amputation
Arthrogram
Arthrogram abnormal
Arthrogram normal
Arthroscopy
Arthroscopy abnormal
Arthroscopy normal
Aspiration bursa
Aspiration bursa abnormal
Aspiration bursa normal
Aspiration joint
Aspiration joint abnormal
Aspiration joint normal
Biopsy abdominal wall
Biopsy abdominal wall abnormal
Biopsy abdominal wall normal
Biopsy bone
Biopsy bone abnormal
Biopsy bone normal
Biopsy cartilage
Biopsy cartilage abnormal
Biopsy cartilage normal
Biopsy muscle
Biopsy muscle abnormal
Biopsy muscle normal
Bone cyst
Bone development abnormal
Bone disorder

Selected Preferred Terms
Bone formation increased
Bone graft
Bone hypertrophy
Bone lesion excision
Bone pain
Bone scan
Bone trimming
Bunion operation
Calcification metastatic
Calcium deficiency
Calcium intoxication
Calcium metabolism disorder
Callus formation delayed
Chondrocalcinosis pyrophosphate
Clavicle fracture
Closed fracture manipulation
Coccydynia
Complicated fracture
Compression fracture
Dental alveolar anomaly
Dental necrosis
Dislocation of vertebra
Dwarfism
Epiphyses delayed fusion
Epiphyses premature fusion
Epiphysiolysis
Epiphysitis
Exostosis
Exostosis of external ear canal
Exostosis of jaw
External fixation of fracture
Face and mouth X-ray
Face and mouth X-ray abnormal
Face and mouth X-ray normal
Facial bones fracture

Selected Preferred Terms
Femoral neck fracture
Femur fracture
Fibula fracture
Finger amputation
Finger repair operation
Flail chest
Foot amputation
Foot fracture
Forearm fracture
Fracture
Fracture delayed union
Fracture malunion
Fracture nonunion
Fracture of clavicle due to birth trauma
Fractured ischium
Fractured maxilla elevation
Fractured sacrum
Fractured skull depressed
Greenstick fracture
Groin pain
Hand amputation
Hand fracture
Hand repair operation
Hip arthroplasty
Hip disarticulation
Hip fracture
Humerus fracture
Hypercalcaemia
Hypercalcaemia of malignancy
Hypercalcaemic nephropathy
Hypercalciuria
Hyperparathyroidism
Hyperparathyroidism primary
Hyperparathyroidism secondary
Hyperparathyroidism tertiary

Selected Preferred Terms
Hyperphosphataemia
Hypertrophic osteoarthropathy
Hypocalcaemia
Hypochondroplasia
Hypoparathyroidism
Hypophosphataemia
Ilium fracture
Internal fixation of fracture
Jaw cyst
Jaw fracture
Jaw lesion excision
Knee arthroplasty
Kyphoscoliosis
Kyphosis
Leg amputation
Metacarpal excision
Metatarsal excision
Metatarsalgia
Microsurgery to hand
Multiple fractures
Open fracture
Open reduction of fracture
Open reduction of spinal fracture
Osteitis
Osteitis condensans
Osteitis deformans
Osteoarthropathy
Osteodystrophy
Osteolysis
Osteomalacia
Osteomyelitis
Osteomyelitis acute
Osteomyelitis blastomyces
Osteomyelitis chronic
Osteomyelitis drainage

Selected Preferred Terms
Osteomyelitis salmonella
Osteonecrosis
Osteoporosis
Osteoporosis postmenopausal
Osteoporotic fracture
Osteosclerosis
Osteosis
Osteotomy
Pain in jaw
Patella fracture
Patellectomy
Pathological fracture
Periodontal destruction
Periostitis
Periostitis hypertrophic
Perthes disease
Post-traumatic osteoporosis
Pseudohypoparathyroidism
Radius fracture
Removal of epiphyseal fixation
Removal of foreign body from joint
Removal of internal fixation
Renal rickets
Resorption bone increased
Rib fracture
Rickets
Scapula fracture
Senile osteoporosis
Skeletal survey
Skeletal survey abnormal
Skeletal survey normal
Skeletal traction
Skull fractured base
Skull X-ray
Skull X-ray abnormal

Selected Preferred Terms
Skull X-ray normal
Spinal compression fracture
Spinal fracture
Spinal X-ray
Spinal X-ray abnormal
Spinal X-ray normal
Sternal fracture
Stress fracture
Synovectomy
Tetany
Tibia fracture
Toe amputation
Toe operation
Tooth abscess
Ulna fracture
Wrist fracture
X-ray of pelvis and hip
X-ray of pelvis and hip abnormal
X-ray of pelvis and hip normal
Pseudarthrosis
Tooth infection
Synoviorthesis
Vertebral lesion
Osteopathic treatment
Ankle operation
Osteopenia
Fractured coccyx
Arthrectomy
Bone density decreased
Traumatic fracture
Knee operation
Shoulder arthroplasty
Williams syndrome
Pubic pain
Extraskeletal ossification

Selected Preferred Terms
Bone infarction
Osteoporosis prophylaxis
Cervical vertebral fracture
Lumbar vertebral fracture
Thoracic vertebral fracture
Temporomandibular joint surgery
Ankle arthroplasty
Melorheostosis
Rotator cuff repair
Synovial fluid white blood cells positive
Joint prosthesis user
Osteocalcin
Osteocalcin increased
Osteocalcin decreased
Bone densitometry
Hypocalciuria
Oral surgery
Hip surgery
Wrist surgery
Hyperphosphaturia
Biopsy chest wall
Biopsy ligament
Biopsy tendon
Chvostek's sign
Bone fistula
Calciphylaxis
Bone erosion
Bone marrow oedema
Osteorrhagia
Latent tetany
Baker's cyst excision
Bone cyst excision
Periprosthetic osteolysis
Osteoporosis circumscripta cranii
Comminuted fracture

Selected Preferred Terms
Epiphyseal surgery
Fracture displacement
Discogram
Ostectomy
Rheumatoid nodule removal
Sequestrectomy
Abscess jaw
Bone swelling
N-telopeptide urine
N-telopeptide urine normal
N-telopeptide urine abnormal
N-telopeptide urine increased
N-telopeptide urine decreased
Epiphyseal fracture
Congenital syphilitic osteochondritis
Maxillofacial operation
Vertebral column mass
Meniscus operation
Intervertebral disc injury
Sternal injury
Limb reattachment surgery
Meniscus removal
Bone formation decreased
Body height below normal
Body height decreased
Body height abnormal
Synovial fluid red blood cells positive
Synovial fluid crystal present
Fracture debridement
Bone electrostimulation therapy
Fracture reduction
Bone debridement
Joint arthroplasty
Biopsy chest wall abnormal
Biopsy chest wall normal

Selected Preferred Terms
Bone density increased
Bone callus excessive
Biopsy ligament abnormal
Biopsy tendon abnormal
Biopsy ligament normal
Biopsy tendon normal
Arthroscopic surgery
Spinal deformity
Bankart lesion
Bone metabolism disorder
Jaw operation
Fractured zygomatic arch elevation
Pyridinoline urine
Pyridinoline urine increased
Pyridinoline urine decreased
Erdheim-Chester disease
Epiphyseal disorder
Hyperphosphatasaemia
Foot operation
Pelvic fracture
Limb operation
Jaw disorder
Phosphorus metabolism disorder
Skeletal injury
Skull fracture
Spinal disorder
Upper limb fracture
X-ray limb
X-ray limb abnormal
X-ray limb normal
Lower limb fracture
Amputation
Arthrodesis
Bone lesion
Bone operation

Selected Preferred Terms
Bone scan abnormal
Bone scan normal
Fracture treatment
Shoulder operation
Osteoprotegerin
Osteoprotegerin decreased
Osteoprotegerin increased
Osteoprotegerin ligand
Osteoprotegerin ligand decreased
Talipes correction
Limb immobilisation
Hungry bone syndrome
High turnover osteopathy
Carpal collapse
Low turnover osteopathy
Synovial fluid white blood cells
Synovial fluid red blood cells
Synovial fluid cell count
Synovial fluid crystal
Vertebroplasty
Ultrasound joint
Bone fissure
Bone fragmentation
Staphylococcal osteomyelitis
Bone marrow oedema syndrome
Osteonecrosis of jaw
Candida osteomyelitis
Periarthritis calcarea
Joint resurfacing surgery
Epiphyseal injury
Discogram abnormal
Discogram normal
Oroantral fistula
Ligament operation
Interscapulothoracic amputation

Selected Preferred Terms
Osteomyelitis bacterial
Osteomyelitis viral
Osteomyelitis fungal
Deoxypyridinoline urine increased
Medial tibial stress syndrome
Vertebral wedging
Capsulorrhaphy
Craniotabes
Bone loss
Torus fracture
Avulsion fracture
Bone contusion
Impacted fracture
Hypoparathyroidism secondary
N-telopeptide
C-telopeptide
Eagle's syndrome
Patella replacement
Alveolar osteitis
Elbow operation
Radiation osteitis
Osteoradionecrosis
Tophus removal operation
Bone groove deepening
Inadequate osteointegration
Post transplant distal limb syndrome
Bone resorption test
Calcanectomy
Bone abscess
Familial hypocalciuric hypercalcaemia
Tartrate-resistant acid phosphatase
Pseudohyperphosphataemia
Bone density abnormal
Limb reconstructive surgery
Radiotherapy to joint

Selected Preferred Terms
Bone atrophy
Arthrolysis
C-telopeptide increased
Bone resorption test abnormal
Deoxypyridinoline urine
Bone formation test
Bone formation test abnormal
Intramedullary rod insertion
Periprosthetic fracture
Dent's disease
Bone graft removal
Acute phosphate nephropathy
Spinal column injury
Biopsy soft tissue
Os trigonum syndrome
Scapulothoracic dissociation
Pubis fracture
Cemento osseous dysplasia
Arthrotomy
Brown tumour
Bone hyperpigmentation
Cementoplasty
Bone decalcification
Atypical femur fracture
Exposed bone in jaw
Enostosis
Gorham's disease
Trousseau's sign
Chronic recurrent multifocal osteomyelitis
Limb amputation
Hypercalcitoninaemia
Spinal pain
Atypical fracture
Hypocalcaemic seizure
Orthopaedic procedure

Selected Preferred Terms
Tooth demineralisation
Chance fracture
Bone prosthesis insertion
Synovial biopsy
Intervertebral disc biopsy
Radial head dislocation
Osteochondral fracture
Sacroiliac fracture
Limb fracture
Spinal fusion fracture
Delayed spinal fusion
Radiolucency around implant
Alveolar bone resorption
Subperiosteal abscess
Epiphysiodesis
Maxillonasal dysplasia
Primary hypoparathyroidism
Dental cyst
Vertebral body replacement
Synovial biopsy abnormal
Incomplete spinal fusion
Tartrate-resistant acid phosphatase decreased
Resorption bone decreased
Osteonecrosis of external auditory canal
Surgical fixation of rib fracture
Pseudohypercalcaemia
Periosteal haematoma
Rachitic rosary
Periostosis
Craniofacial fracture
Hereditary hypophosphataemic rickets
Chronic kidney disease-mineral and bone disorder
Orthoroentgenogram
Gluteoplasty
Rotationplasty

Selected Preferred Terms
Costal cartilage fracture
Astragalectomy
Lisfranc fracture
Primary familial brain calcification
Bone metabolism biochemical marker increased
Jaw fistula
Neonatal hypocalcaemia
Cystic angiomatosis
Fracture blisters
Trapeziectomy
Serum procollagen type I N-terminal propeptide decreased
Parathyroid hyperplasia
Metaphyseal corner fracture
Fracture infection
Subchondral insufficiency fracture
Gastric mucosal calcinosis
Itai-itai disease
X-ray dental abnormal
Peri-spinal heterotopic ossification
Pseudofracture
Idiopathic condylar resorption
Removal of external fixation
Osteophyte fracture
Femoral derotation osteotomy
Familial isolated hyperparathyroidism
Malacoplakia of bone
Post procedural hypoparathyroidism
Tartrate-resistant acid phosphatase increased
Distraction osteogenesis
Cheilectomy
Bone sequestrum
Maisonneuve fracture
Hypophosphataemic osteomalacia
Stapes fracture
Scapholunate dissociation

Selected Preferred Terms
Anterior labroligamentous periosteal sleeve avulsion lesion
Degenerative bone disease
Sesamoidectomy
Computerised tomogram spine
Os odontoideum
Iliocostal friction syndrome
Mueller-Weiss syndrome
Metastatic bone disease prophylaxis
Pathological fracture prophylaxis
Musculoskeletal toxicity
Vertebral end plate inflammation
Kohler's disease
Computerised tomogram neck
Magnetic resonance imaging joint
Magnetic resonance imaging neck

Appendix 3. Fracture Events

The selected Preferred Terms of fracture events from HLGT of Fractures based on MedDRA 22.1 are listed as follows:

Selected Preferred Term
Acetabulum fracture
Ankle fracture
Clavicle fracture
Complicated fracture
Compression fracture
Facial bones fracture
Femoral neck fracture
Femur fracture
Fibula fracture
Flail chest
Foot fracture
Forearm fracture
Fracture
Fracture delayed union
Fracture malunion
Fracture nonunion
Fracture of clavicle due to birth trauma
Fractured ischium
Fractured sacrum
Fractured skull depressed
Greenstick fracture
Hand fracture
Hip fracture
Humerus fracture
Ilium fracture
Jaw fracture
Multiple fractures
Open fracture
Osteoporotic fracture
Patella fracture
Pathological fracture
Radius fracture

Selected Preferred Term
Rib fracture
Scapula fracture
Skull fractured base
Spinal compression fracture
Spinal fracture
Sternal fracture
Stress fracture
Tibia fracture
Ulna fracture
Wrist fracture
Pseudarthrosis
Fractured coccyx
Traumatic fracture
Cervical vertebral fracture
Lumbar vertebral fracture
Thoracic vertebral fracture
Comminuted fracture
Fracture displacement
Epiphyseal fracture
Pelvic fracture
Skull fracture
Upper limb fracture
Lower limb fracture
Bone fissure
Bone fragmentation
Torus fracture
Avulsion fracture
Impacted fracture
Periprosthetic fracture
Scapulothoracic dissociation
Pubis fracture
Atypical femur fracture
Atypical fracture
Chance fracture
Osteochondral fracture

Selected Preferred Term
Sacroiliac fracture
Limb fracture
Spinal fusion fracture
Craniofacial fracture
Costal cartilage fracture
Lisfranc fracture
Fracture blisters
Metaphyseal corner fracture
Fracture infection
Subchondral insufficiency fracture
Pseudofracture
Osteophyte fracture
Maisonneuve fracture

SAP GS-US-320-4018 Final v1.0

ELECTRONIC SIGNATURES

Signed by	Meaning of Signature	Server Date (dd-MMM-yyyy hh:mm:ss)
PPD	Biostatistics eSigned	11-Mar-2020 23:29:44