STATISTICAL ANALYSIS PLAN

Study Title:	A Phase 3, Randomized, Double-Blind, Placebo-Controlled Study Evaluating the Safety and Efficacy of Selonsertib in Subjects with Nonalcoholic Steatohepatitis (NASH) and Bridging (F3) Fibrosis
Name of Test Drug:	Selonsertib
Study Number:	GS-US-384-1943
Protocol Version (Date):	Amendment 2 (31 May 2018)
Analysis Type:	Interim and Final Analysis
Analysis Plan Version:	Version 1.0
Analysis Plan Date:	23 January 2019
Analysis Plan Author(s):	PPD PPD

CONFIDENTIAL AND PROPRIETARY INFORMATION

TABLE OF CONTENTS

LIST OF IN-TEXT TABLES

LIST OF IN-TEXT FIGURES

LIST OF ABBREVIATIONS AND DEFINITION OF TERMS

AE	adverse event
ALP	alkaline phosphatase
ALT	alanine aminotransferase
ANCOVA	analysis of covariance
ARFI	acoustic radiation force impulse
APRI	AST to platelet ratio index
AST	aspartate aminotransferase
α-SMA	α-smooth muscle actin
BLQ	below the limit of quantitation
BMI	body mass index
BP	blood pressure
CI	confidence interval
CLDQ	chronic liver disease questionnaire
CK-18	cytokeratin-18
CMH	Cochran-Mantel-Haenszel
CP	Child-Pugh
CRF	case report form
CRN	clinical research network
CRP	C-reactive protein
CSR	clinical study report
CTCAE	Common Toxicity Criteria for Adverse Events
DILI	drug-induced liver injury
DMC	data monitoring committee
eCRF	electronic case report form
ECG	electrocardiogram
EFS	event-free survival
ELF™	enhanced liver fibrosis
ET	early termination
EQ-5D	EuroQol five dimensions
FAS	Full Analysis Set
FIB-4	fibrosis-4
Gilead	Gilead Sciences
GGT	gamma glutamyl transferase
HbA1c	Hemoglobin A1c
HCC	hepatocellular carcinoma
HCV	hepatitis C virus
HE	hepatic encephalopathy
Hgb	hemoglobin
HDL	high density lipoprotein

HLGT	high-level group term
HLT	high-level term
HOMA-IR	homeostasis model assessment of insulin resistance
HR	hazard ratio
HRQoL	Health Related Quality of Life
HRUQ	Health Resource Utilization Questionnaires
ID	identification
IFG	impaired fasting glucose
INR	international normalized ratio
ITT	intent to treat
IWRS	interactive web response system
KM	Kaplan-Meier
KMMI	Kaplan-Meier multiple imputation
LDL	low density lipoprotein
LLT	lower-level term
LOCF	last observation carried forward
LOQ	limit of quantitation
MAR	missing at random
MedDRA	Medical Dictionary for Regulatory Activities
MELD	model for end-stage liver disease
MH	Mantel-Haenszel
MI	multiple imputation
MRE	magnetic resonance elastography
NAFLD	non-alcoholic fatty liver disease
NAS	NAFLD activity score
NASH	non-alcoholic steatohepatitis
NFS	NAFLD fibrosis score
NRI	non-responder imputation
OL	open-label
PK	Pharmacokinetic
PIIINP	procollagen III N-terminal propeptide
PT	preferred term
PTM	placebo-to-match
PVE	Pharmacovigilance and Epidemiology
Q1, Q3	first quartile, third quartile
QoL	quality of life
SAE	serious adverse event
SAF	steatosis, activity, and fibrosis
SAP	statistical analysis plan
SD	standard deviation
SEL	selonsertib

SF-36	Short Form (36) Health Survey
SI (units)	international system of units
SWE	shear-wave elastography
SOC	system organ class
SVR	sustained virologic response
TEAE	treatment-emergent adverse event
TFLs	tables, figures, and listings
TIMP1	tissue inhibitor of metalloproteinase
VLDL	very low density lipoprotein
ULN	upper limit of normal
WBC	white blood cell count
WHO	World Health Organization
WPAI	work productivity and activity impairment questionnaire

1. INTRODUCTION

This statistical analysis plan (SAP) describes the statistical analysis methods and data presentations to be used in tables, figures and listings (TFLs) for the interim analysis at Week 48 and final analysis at Week 240 for Study GS-US-384-1943.

This analysis plan is based on the study protocol Amendment 2 dated 31 May 2018 and the electronic case report form (eCRF). The SAP will be finalized before breaking the blind for the interim analysis. Any changes made after the finalization of the SAP will be documented in the clinical study report (CSR).

1.1. Study Objectives

The primary objective of this study is as follows:

- To evaluate whether selonsertib (SEL) can cause fibrosis regression and reduce progression to cirrhosis and associated complications in subjects with NASH and bridging (F3) fibrosis

The secondary objective of this study is as follows:

- To assess the safety and tolerability of SEL in subjects with NASH and bridging (F3) fibrosis

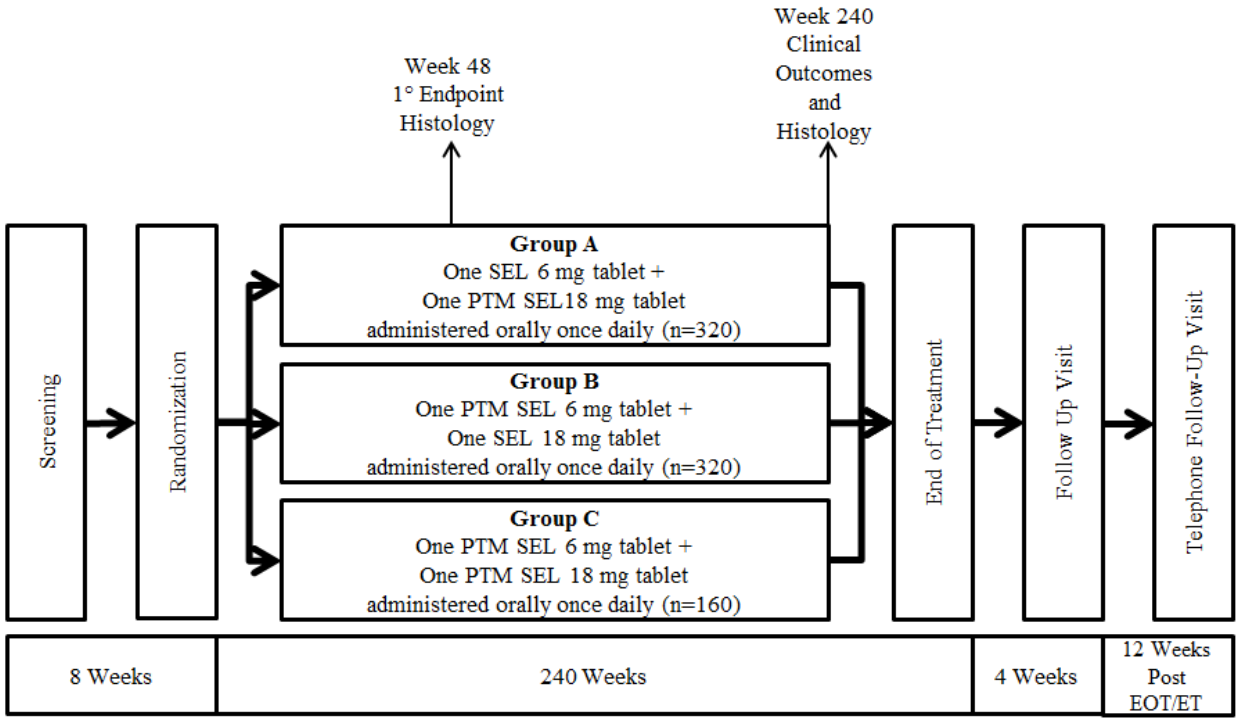
The exploratory objectives of this study are as follows:

1.2. Study Design

This is a Phase 3, randomized, double-blind, placebo-controlled study evaluating the safety and efficacy of SEL in subjects with NASH and bridging (F3) fibrosis.

The overall study design is presented graphically in Figure 1-1.

Figure 1-1. Overall Study Design



- EOT = end of treatment; ET = early termination; PTM = Placebo-to-Match

Subjects meeting the study's entry criteria were randomly assigned in a 2:2:1 ratio to 1 of 3 treatment groups: Group A, Group B and Group C, respectively, as shown in Figure 1-1. Randomization was stratified by the presence or absence of diabetes mellitus (determined by medical history or based on Screening laboratory values if previously undiagnosed [ie, hemoglobin A1c (HbA1c) ≥ 6.5% or fasting plasma glucose ≥ 126 mg/dL]) and by Enhanced Liver Fibrosis (ELF™) test score ≥ 9.76 or < 9.76 during Screening. Study drugs will be administered for up to a total of 240 weeks.

Subjects who experience a clinical event (except all-cause mortality and liver transplantation) prior to completing the Week 240 Visit of the Randomized Phase will be offered the option to rollover into an Open-Label (OL) Phase for a total treatment duration of 240 weeks inclusive of the Randomized Phase. Rollover into the OL Phase of the study must occur within 60 days of confirmation of the event. Subjects starting the OL Phase of the study will complete the same study procedures as during the Randomized Phase of the study, CCI [redacted]

The schedule of assessments is provided as an appendix to this analysis plan (Appendix 1).

1.3. Sample Size and Power

With a sample size of 320 subjects in each active treatment arm and 160 subjects in the placebo arm, the study has 94% power to detect a difference in the proportion of subjects with a ≥ 1-stage improvement in fibrosis without worsening of NASH of 15% or more at Week 48 at a significance level of 0.025 (2-sided), assuming the proportion of subjects that meet the endpoint in the placebo arm is 12%.

With regard to the clinical efficacy endpoint, subjects will be followed for a period of up to 240 weeks. The estimate of the expected event rate is 30% in untreated patients, producing an expected EFS rate of 70% at 240 weeks in the placebo arm. SEL is expected to improve the EFS rate to 83.7% (expected event rate of 16.3%) compared with placebo (hazard ratio (HR), 0.50). We also expect a 20% overall dropout rate. Given the above assumptions, for comparison of the SEL arms versus the placebo arm, together with a total sample size of 800 subjects (2:2:1 ratio), and assuming the log-rank test statistic is evaluated using a 2-sided 0.025 significance level, the study will have 85% power to evaluate the superiority of each SEL arm versus placebo with respect to EFS.

Therefore, the overall power for the trial, assuming independence between the primary efficacy endpoint and clinical efficacy endpoint, is 80% (94% × 85%). It should be noted that this is a lower bound estimate as these 2 endpoints are expected to be positively correlated.

2. TYPE OF PLANNED ANALYSIS

2.1. DMC Interim Analyses

An external Data Monitoring Committee (DMC) that consists of 3 hepatologists and a PhD statistician will review the progress of the study, perform interim reviews of safety data, and provide recommendations to Gilead whether the nature, frequency, and severity of adverse events (AEs) associated with study treatment warrant the early termination of the study in the best interests of the participants, whether the study should continue as planned, or whether the study should continue with modifications.

The initial meeting of the DMC will occur after 50 subjects have completed their Week 4 visit and approximately every 6 months thereafter to monitor the study for safety events.

The DMC's specific activities will be defined by a mutually agreed upon charter, which will define the DMC's membership, conduct, and meeting schedule.

While the DMC will be asked to advise Gilead regarding future conduct of the study, including possible early study termination, Gilead retains final decision-making authority on all aspects of the study.

2.2. Interim (Week 48) Analysis

When all subjects have completed the Randomized Phase Week 48 visit, or have a clinical event before Randomized Phase Week 48 visit, or have prematurely discontinued from the study, superiority of each SEL arm relative to the placebo arm will be assessed with respect to the primary histology efficacy endpoint.

2.3. Final (Week 240) Analysis

After all subjects have completed the study, or have prematurely discontinued from the study, outstanding data queries have been resolved or adjudicated as unresolvable, and the data have been cleaned and finalized, the study blind will be broken and the final analysis of the data will be performed. The final analysis will assess the superiority of each SEL arm relative to placebo with respect to the primary clinical efficacy endpoint of EFS.

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), median, first quartile (Q1), third quartile (Q3), minimum, and maximum will be presented.

By-subject listings will be presented for all subjects in the All Randomized Analysis Set and sorted by subject identification (ID) number, visit date, and time (if applicable). Data collected on log forms, such as AEs, will be presented in chronological order within subject. The treatment group to which subjects were initially assigned 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. Subjects included in each analysis set will be determined before the study blind is broken for analysis. The analysis set will be identified and included as a subtitle of each table, figure, and listing. Subjects who have been randomized but never dosed will also be summarized.

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

3.1.1. All Randomized Analysis Set

All Randomized Analysis Set includes all subjects who were randomized in the study.

3.1.2. Full Analysis Set

The Full Analysis Set (FAS) includes all subjects who were randomized in the study and received at least 1 dose of study drug. This is the primary analysis set for efficacy analyses.

3.1.3. Safety Analysis Set

The Safety Analysis Set includes all subjects who took at least 1 dose of study drug. This is the primary analysis set for safety analyses.

Safety analysis set for the OL Phase (OL Safety Analysis Set) includes all subjects who were enrolled into the OL Phase and received at least 1 dose of OL study drug. This is the primary analysis set for OL Phase safety analyses.

3.1.4. Pharmacokinetic Analysis Set

The PK Analysis Set will include all randomized subjects who took at least 1 dose of study drug and had at least 1 nonmissing postdose concentration of SEL (and/or its metabolite GS-607509) reported by the bioanalytical laboratory. This is the primary analysis set for all PK analyses.

3.2. Subject Grouping

For analyses based on the FAS, subjects will be grouped according to the treatment to which they were randomized. For analyses based on the Safety Analysis Set, subjects will be grouped according to actual treatment received. The actual treatment received will differ from the randomized treatment only when the actual treatment differs from the randomized treatment for the entire treatment duration. In this case, the actual treatment received is defined as the treatment received for the entire treatment duration.

For safety analyses, subjects will be grouped into the following:

Randomized Phase (only Randomized Phase data will be summarized for these groups)

1) SEL 18 mg: This group includes all subjects who received SEL 18 mg in the Randomized Phase.

2) SEL 6 mg: This group includes all subjects who received SEL 6 mg in the Randomized Phase.

3) SEL Pooled: This group includes all subjects who received SEL 18 mg or SEL 6 mg in the Randomized Phase.

4) Placebo: This group includes all subjects who received only placebo in the Randomized Phase.

Open-Label Phase (only OL Phase data will be summarized for this group):

OL SEL 18 mg: This group includes all subjects who received OL study drug.

3.3. Strata and Covariates

An Interactive Web Response System (IWRS) will be used for centralized randomization and treatment assignment. Randomization will be stratified by

- Presence or absence of diabetes mellitus (as determined by medical history or based on Screening laboratory values if previously undiagnosed [i.e. HbA1c ≥ 6.5% or fasting plasma glucose ≥ 126 mg/dL])

- ELF score ≥ 9.76 or < 9.76 during Screening.

If there are discrepancies in stratification factor values between the IWRS and the clinical database, the baseline values recorded in the clinical database will be used for analyses.

Efficacy endpoints will be evaluated using stratification factors as covariates for analyses, as specified in Section 6.

3.4. Examination of Subject Subgroups

The primary and key secondary efficacy endpoints as defined in Section 6 will be examined using the following subgroups:

- Subjects with diabetes at baseline
- Subjects without diabetes at baseline
- Subjects with ELF test score ≥ 9.76 at baseline
- Subjects with ELF test score < 9.76 at baseline
- Subjects with body weight loss ≥ 5% at Week 48/Week 240 (for Week 48/240 analysis, respectively)
- Subjects with body weight loss < 5% or body weight gain at Week 48/Week 240 (for Week 48/240 analysis, respectively)
- Subjects treated with Vitamin E at baseline
- Subjects not treated with Vitamin E at baseline
- Subjects with a history of sustained virologic response (SVR) to therapy for chronic hepatitis C virus (HCV)
- Subjects without a history of SVR to therapy for chronic HCV.

For the comparison of the treatment differences in each subgroup, point estimate and 95% confidence intervals (CIs) based on the primary analysis approach will be presented.

3.5. Multiple Testing

When the primary histology efficacy endpoint is compared between each of the SEL arms and the placebo arm at Week 48, it will be tested at a 2-sided 0.025 significance level to control for an overall Type I error rate of 0.05 by Bonferroni adjustment.

The testing strategy for the clinical efficacy endpoint (EFS) is described below and is presented graphically in Figure 3-1.

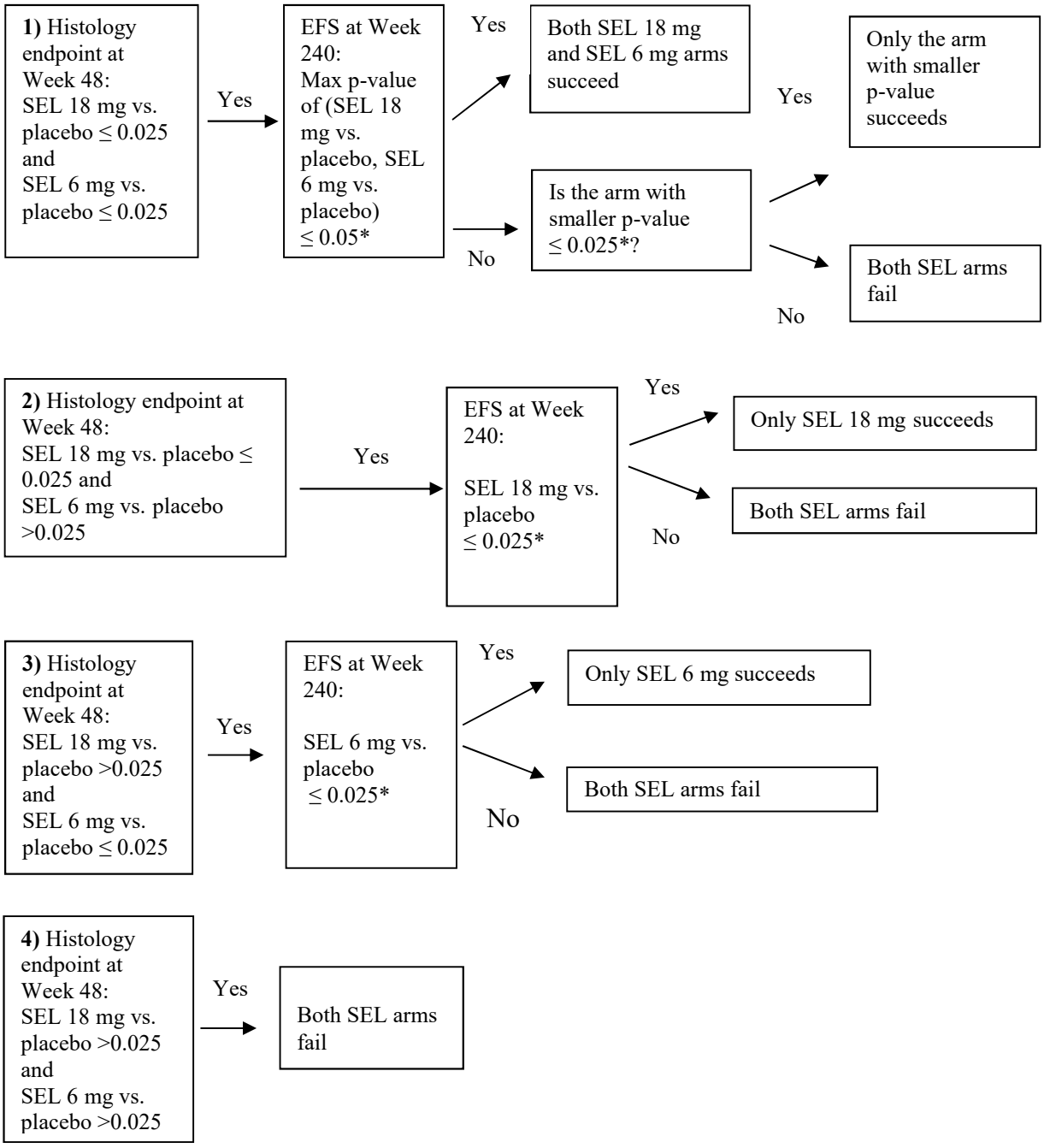
1) If the primary histology efficacy endpoint is significant in both SEL arms, the EFS endpoint will be tested for both SEL arms against placebo at Week 240 at a 2-sided 0.05 significance level by the Hochberg procedure {Hochberg 1988}. If the larger p-value of SEL18 mg versus placebo and SEL 6 mg versus placebo is less than or equal to 0.05, the EFS endpoint for both SEL arms will be a success; if the larger p-value is > 0.05, the treatment arm with the larger p-value fails, and the smaller p-value will be compared with the significance level of 0.025 to determine the success of the treatment arm.

2) If the primary histology efficacy endpoint is significant in the SEL 18 mg arm but not the SEL 6 mg arm, the EFS endpoint will be tested for SEL 18 mg against placebo at Week 240 at a 2-sided 0.025 significance level.

3) If the primary histology efficacy endpoint is significant in the SEL 6 mg arm but not the SEL18 mg arm, the EFS endpoint will be tested for SEL 6 mg against placebo at Week 240 at a 2-sided 0.025 significance level.

4) If the primary histology efficacy endpoint is not significant in either the SEL 18 mg or SEL 6 mg arms, the trial will be terminated.

Figure 3-1. Multiple Testing Flow Chart



* Nominal alpha will be adjusted using the Haybittle-Peto approach

Although no formal statistical inference on EFS is planned for the Week 48 Interim Analysis, the actual nominal significance level for the EFS endpoint at the Week 240 Analysis will be adjusted using the Haybittle-Peto approach {Haybittle 1971}. The adjustment will be based on the number of events at both the Week 48 and Week 240 analyses, assuming an allocation of a 2-sided significance level of 0.001 at the Week 48 Interim Analysis.

3.6. Missing Data and Outliers

3.6.1. Missing Data

For missing last dosing date of study drug, imputation rules are described in Section 4.2.1. The handling of missing or incomplete dates for AE onset is described in Section 7.1.5.3, and for prior and concomitant medications in Section 7.4.

The following imputation approaches will be applied to efficacy endpoints as specified:

1) Non-Responder Imputation (NRI)

   For purpose of analysis at a defined time point, when NRI is used, all subjects with missing values will be analyzed as not meeting the corresponding endpoint. The NRI method will be applied to subjects with or without baseline value.

2) Multiple Imputation (MI):

   The MI procedure replaces each missing data point of the binary primary histology efficacy endpoint (defined in Section 6.1.1) with a set of plausible values that represent the uncertainty about the right value to impute under the missing at random (MAR) assumption. The imputation model used to impute missing data points will be based on the logistic regression with the treatment arms and the 2 stratification factors as covariates. A total of 50 imputed datasets will be generated by the imputation model. That is, each missing data point will be imputed 50 times. Each imputed data set will be analyzed by the method for the primary analysis as specified in Section 6.1.3. The results from these 50 imputed data sets will be combined using Rubin's rule {Rubin 1987}.

3) Tipping Point Analysis (TPA):

   If any significant result in favor of SEL 18 mg or SEL 6 mg is detected in the primary analysis of the binary primary histology efficacy endpoint (described in Section 6.1.3), a delta-adjusting pattern-mixture approach for TPA {Ratitch 2013} will be conducted to assess the robustness of the primary analysis results under a missing not at random assumption for SEL arms. Specifically, it is assumed that a systematic difference exists between the conditional distributions of the missing and observed data in the SEL arms. To reflect such a systematic difference, a shift parameter δ will be applied to the imputation model (ie, the logistic regression with the treatment arms and the 2 stratifications factors as covariates) when the missing data points in the SEL arms are imputed. For each value of δ, multiply imputed data sets will be generated, each data set will be analyzed by the same method for the primary analysis, and analysis results will then be combined by Rubin's rule. Thus, by

varying the values of δ, the impact from missing data on the analysis results will be examined to identify tipping points (ie, the values at which conclusions from statistical inference change from being significant to being insignificant for SEL arm(s)).

4) Last Observation Carried Forward (LOCF)

For purpose of analysis at a defined time point, if the value is missing, the last observed value prior to the time point will be used to impute the missing value. Refer to Section 3.8.2 for the handling of the data to be included in efficacy analysis. The LOCF approach will be applied to those subjects who have a baseline value. No imputation will be performed if the value at baseline is missing.

5) Observed Case (OC)

Only subjects with observed Baseline and Week 48/240 values will be included at the corresponding Week 48/240 analysis.

6) Kaplan-Meier Multiple Imputation (KMMI)

Sensitivity analysis will be performed under the KMMI strategy {Zhao 2014} if there are significant differences in the Kaplan-Meier (KM) estimates of survival rates between SEL 18 mg or SEL 6 mg and placebo. This method introduces a sensitivity parameter, θ, defined as the fixed hazard ratio of a subject who dropped out having the event after being censored at a certain centering time relative to subjects who remain on treatment having the event after that censoring time. The MAR-like assumption (ie, noninformative independent censoring) corresponds to $\theta = 1$. $\theta > 1$ implies a higher hazard after censoring for subjects who drop out compared with their counterparts who remain on treatment. By varying the values of θ, the KMMI strategy addresses departures from the MAR-like assumption when unobserved time-to-event data are multiply imputed. The KMMI approach can be summarized as follows:

i). Calculate KM estimates of survival rates for each treatment group within each stratum separately.

ii). Obtain the survival estimate at each censoring time by linear interpolation of the KM survival rates corresponding to the 2 event times immediately before and after. The KM survival function may be extrapolated to calculate the survival estimates at censoring time which is after the event time of the last observed event up to the maximum follow-up time (ie, 252 weeks [or 58 months]).

iii). Apply each value of θ to impute the unobserved event time for a dropout subject who will experience the event after the time being censored for a total of 50 times.

After the multiple imputation, the pairwise logrank test will be performed to compare the survival rates between SEL 18 mg or SEL 6 mg and placebo on each complete data set (described in Section 6.2.3). Then, by combining the normalized chi-square statistics using Rubin's rule, the impact of departures from noninformative censoring can be investigated by

the combined treatment effect inference, which is a function of θ. Thus, the robustness of the significant results can be assessed.

The NRI approach will be used for the primary analysis of the primary histology efficacy endpoint and secondary endpoints evaluated at Week 48. LOCF analysis and OC analysis will be carried out as sensitivity analyses for these endpoints at Week 48 using the same analysis method as defined in Section 6.1.3. "Last observation" includes both baseline and postbaseline records, and only the records eligible for efficacy analysis will be carried forward.

The MI and TPA approaches will be used for the sensitivity analyses for the primary histology efficacy endpoint. KMMI approach will be used for the sensitivity analysis for the clinical efficacy endpoint.

The NRI approach will be used for the primary analysis for the secondary endpoints evaluated at Week 240. The LOCF and OC approaches will be used for sensitivity analysis of these endpoints.

CCI [redacted]

In the subgroup analysis by body weight change from baseline, missing Week 48/Week 240 (for the Week 48/240 analysis, respectively) body weight will be imputed using the LOCF approach to determine which subgroup a subject belongs to.

3.6.2. Outliers

Outliers will be identified during the data management and data analysis process. All data, including outliers, will be included in analyses.

3.7. 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 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 for calculation of 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 for calculation of 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 for calculation of summary statistics.

- A value that is 1 unit above the LOQ will be used for calculation of 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 for calculation of descriptive statistics if the datum is reported in the form of "≤ x" or "≥ x" (where x is considered the LOQ).

When any ELF component (hyaluronic acid, procollagen III N-terminal propeptide [PIIINP] and tissue inhibitor of metalloproteinase 1 [TIMP1]) is less than the lower LOQ or above the upper LOQ, the ELF component will be imputed based on the rules described above, and the ELF test score will be calculated based on the imputed value of the component(s) as defined in Appendix 7.

Natural logarithm transformation will be used for plasma concentrations and analysis of PK parameters. Plasma concentration values that are below the limit of quantitation (BLQ) will be presented as "BLQ" in the concentration data listing. Values that are BLQ will be treated as 0 at predose time points and one-half the value of the LOQ at postbaseline time points.

The following conventions will be used for the presentation of summary and order statistics:

- If at least 1 subject has a concentration value of BLQ for the time point, the minimum value will be displayed as "BLQ."

- If more than 25% of the subjects have a concentration data value of BLQ for a given time point, the minimum and Q1 values will be displayed as "BLQ."

- If more than 50% of the subjects have a concentration data value of BLQ for a given time point, the minimum, Q1, and median values will be displayed as "BLQ."

- If more than 75% of the subjects have a concentration data value of BLQ for a given time point, the minimum, Q1, median, and Q3 values will be displayed as "BLQ."

- If all subjects have concentration data values of BLQ for a given time point, all order statistics (minimum, Q1, median, Q3, and maximum) will be displayed as "BLQ."

3.8. Analysis Visit Windows

3.8.1. Definition of Study Day

Study day/OL study day will be calculated from the first dosing date of study drug in the corresponding phase (Randomized or OL) and derived as follows:

- For postdose study days: Assessment Date – First Dosing Date + 1
- For days prior to the first dose: Assessment Date – First Dosing Date

Study Day 1/OL Study Day 1 is the day of first dose of study drug administration in the corresponding phase.

3.8.2. Analysis Visit Windows

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

On-treatment visit windows will be calculated from Study Day 1/OL Study Day 1 for selected efficacy measures, vital signs, elastography, and safety laboratory data.

Selected safety and efficacy data, unless otherwise specified, collected up to and including the last dosing date in the Randomized Phase + 30 days and before OL Study Day 1 for the Randomized Phase, whichever is earlier, or OL last dosing date + 30 days for the OL Phase, will be mapped according to the following analysis windows unless the nominal visit name is Follow-Up or OL Follow-Up.

All biopsy data and elastography data collected will be mapped to Randomized Phase visit windows unless on or after OL Study Day 1. HRQoL questionnaires and HRUQ will not be windowed, and will be summarized by nominal visit. For those endpoints, missing (OL) Baseline/Day 1 values will not be imputed.

The analysis windows for select measures are provided in Table 3-1 to Table 3-6.

Table 3-1. Analysis Visit Windows for Chemistry, Hematology, Coagulation, CP/MELD scores, Vital Signs, Body Weight, and Stool Frequency

Analysis Visit	Nominal Day	Lower Limit	Upper Limit
Baseline/OL-Baseline	1	(none)	1
Week 1/OL-Week 1	8	2	18
Week 4/OL-Week 4	29	19	42
Week 8/OL-Week 8	57	43	70
Week 12/OL-Week 12	85	71	98
Week 16/OL-Week 16	113	99	126
Week 20/OL-Week 20	141	127	154
Week 24/OL-Week 24	169	155	182
Week 28/OL-Week 28	197	183	210
Week 32/OL-Week 32	225	211	238
Week 36/OL-Week 36	253	239	266
Week 40/OL-Week 40	281	267	294
Week 44/OL-Week 44	309	295	322
Week 48/OL-Week 48	337	323	378
Week 60/OL-Week 60	421	379	462
Week 72/OL-Week 72	505	463	546
Week 84/OL-Week 84	589	547	630
Week 96/OL-Week 96	673	631	714
Week 108/OL-Week 108	757	715	798
Week 120/OL-Week 120	841	799	882
Week 132/OL-Week 132	925	883	966
Week 144/OL-Week 144	1009	967	1050
Week 156/OL-Week 156	1093	1051	1134
Week 168/OL-Week 168	1177	1135	1218
Week 180/OL-Week 180	1261	1219	1302
Week 192/OL-Week 192	1345	1303	1386
Week 204/OL-Week 204	1429	1387	1470
Week 216/OL-Week 216	1513	1471	1554
Week 228/OL-Week 228	1597	1555	1638
Week 240/OL-Week 240	1681	1639	≥1681

Table 3-2. Analysis Visit Windows for Waist Circumference, Insulin, Homeostasis Model Assessment of Insulin Resistance (HOMA-IR), Lipids and Hemoglobin A1c

Analysis Visit	Nominal Day	Lower Limit	Upper Limit
Baseline/OL-Baseline	1	(none)	1
Week 4/OL-Week 4	29	2	56
Week 12/OL-Week 12	85	57	126
Week 24/OL-Week 24	169	127	252
Week 48/OL-Week 48	337	253	420
Week 72/OL-Week 72	505	421	588
Week 96/OL-Week 96	673	589	756
Week 120/OL-Week 120	841	757	924
Week 144/OL-Week 144	1009	925	1092
Week 168/OL-Week 168	1177	1093	1260
Week 192/OL-Week 192	1345	1261	1428
Week 216/OL-Week 216	1513	1429	1596
Week 240/OL-Week 240	1681	1597	≥1681

Table 3-3. Analysis Visit Windows for ELF test score and components, FibroSure/FibroTest, α_2-Macroglobin, Haptoglobin, Apolipoprotein A1, C-Reactive Protein (CRP) and Total Bile Acids

Analysis Visit	Nominal Day	Lower Limit	Upper Limit
Baseline/OL-Baseline	1	(none)	1
Week 12/OL-Week 12	85	2	126
Week 24/OL-Week 24	169	127	252
Week 48/OL-Week 48	337	253	420
Week 72/OL-Week 72	505	421	588
Week 96/OL-Week 96	673	589	756
Week 120/OL-Week 120	841	757	924
Week 144/OL-Week 144	1009	925	1092
Week 168/OL-Week 168	1177	1093	1260
Week 192/OL-Week 192	1345	1261	1428
Week 216/OL-Week 216	1513	1429	1596
Week 240/OL-Week 240	1681	1597	≥1681

Table 3-4. Analysis Visit Windows for Elastography and Pooled Cohort Risk and Score

Analysis Visit	Nominal Day	Lower Limit	Upper Limit
Baseline	1	(none)	1 (28*)
Week 24	169	2(29*)	252
Week 48	337	253	420
Week 72	505	421	588
Week 96	673	589	840
Week 144	1009	841	1176
Week 192	1345	1177	1512
Week 240	1681	1513	≥1681

* for Elastography.

Table 3-5. Analysis Visit Windows for Cytokeratin-18 (CK-18) M30 and CK-18 M65

Analysis Visit	Nominal Day	Lower Limit	Upper Limit
Baseline	1	(none)	1
Week 24	169	2	252
Week 48	337	253	504
Week 96	673	505	840
Week 144	1009	841	1176
Week 192	1345	1177	1512
Week 240	1681	1513	≥1681

Table 3-6. Analysis Visit Windows for Biopsy Data

Analysis Visit	Nominal Day	Lower Limit	Upper Limit
Baseline	1	(none)	1
Week 48	337	169	504
Week 240	1681	505	≥1681

Data relating to unscheduled visits and early termination visits may be assigned to a particular visit or time point based on the visit windows. The following conventions will be followed:

- An unscheduled visit prior to the first dosing of study drug may be included in the calculation of the baseline value, if applicable.
- Unscheduled visits after the first dose of study drug will be included in determining the maximum postbaseline toxicity grade.
- Data collected on a follow-up visit will be summarized as a separate visit, and labeled "Follow-up Visit" for each phase.

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

Depending on the statistical analysis method, single values may be 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 1 value per analysis window.

If multiple, valid, nonmissing, continuous measurements exist in an analysis window, records will be chosen based on the following rules if a single value is needed:

- In general, the baseline value will be the last nonmissing value on or prior to the first dosing date of study drug in the corresponding phase. If multiple measurements occur on the same day, the last nonmissing value prior to the time of first dosing of study drug will be considered as the baseline value. If these multiple measurements occur at the same time or the time is not available, the average of these measurements (for continuous data) will be considered as the baseline value. Baseline values for liver tests (ALT, AST, total bilirubin and direct bilirubin) in the Randomized Phase will be determined by averaging the values obtained between and including Screening and Day 1; baseline values for liver tests (ALT, AST, total bilirubin, and direct bilirubin) in the OL Phase will be determined by averaging the values obtained between OL Study Day - 55 and OL Study Day 1 (inclusive); baseline values for elastography will be determined as the last nonmissing values on or prior to (OL) Day 28. Missing (OL) Baseline/Day 1 values of HRQoL questionnaires and HRUQ will not be imputed.
- For postbaseline visits:

    1) The record closest to the nominal day for that visit will be selected.
    2) If there are 2 records that are equidistant from the nominal day, the later record will be selected.
    3) In general, if there is more than 1 record on the selected day, the average will be taken. For serum creatinine, if both enzymatic and regular creatinine are collected from the same blood sample and are analyzable, regular creatinine will be selected for analysis.

If multiple, valid, nonmissing, categorical measurements exist in an analysis window, records will be chosen based on the following rules if a single value is needed:

- For baseline, the last available record on or prior to the date of the first dose of study drug will be selected in the corresponding phase. If there are multiple records on the same day, the value with the lowest severity will be selected (eg, normal will be selected over abnormal for safety electrocardiogram [ECG] findings).

- For postbaseline visits, if there are multiple records on the same day, the value with the worst severity within the window will be selected (eg, abnormal will be selected over normal for safety ECG findings).

Postbaseline parameters used to define histology endpoints will be selected in each visit window according to the following rules:

- Histology endpoints including a ≥ 1-stage improvement in fibrosis without worsening of NASH, progression to cirrhosis, ≥ 1-stage improvement in fibrosis, NASH resolution without worsening of fibrosis, NAS and elements, SAF and elements will be derived based on parameters collected from 1 biopsy sampling date if the following 4 parameters are available: NASH CRN fibrosis stage, lobular inflammation, hepatocellular ballooning, and steatosis grades.

- When the set of the 4 parameters (NASH CRN fibrosis stage, lobular inflammation, hepatocellular ballooning, and steatosis grades) is available for multiple biopsy samples obtained on different dates in 1 visit window, the set with the biopsy sample collection date closest to the nominal day will be chosen for analysis.. When 2 samples are equidistant from the nominal day, the later set will be selected.

- When the set of the 4 parameters is not complete for any biopsy sample from a single day, the histology endpoints will be derived based on the nonmissing value of each individual parameter closest to the nominal day for that visit window. When 2 records are equidistant from the nominal day, the later record will be selected.

- For Ishak fibrosis stage, the nonmissing value from the biopsy sample collection date closest to the nominal day for the visit window will be selected regardless of the missing status of other parameters. When 2 records are equidistant from the nominal day, the later record will be selected.

- When an endpoint cannot be derived based on all biopsy samples in a visit window, the histology endpoint will be considered missing, and imputation methods listed in Section 3.6.1 will apply.

Liver stiffness by transient elastography data in each analysis visit window will be chosen based on the following rules:

- For baseline, measurements by XL probe will be selected for analysis if available; otherwise measurements by M probe will be selected.

- For postbaseline visits, measurements by the same probe type (XL or M) selected for the subject at baseline will be selected in each analysis visit window. If no measurement by the same probe type as baseline is available, the analysis value for the corresponding postbaseline visit will be considered missing. If there are multiple postbaseline records by the same probe type as baseline, the rules to choose postbaseline continuous measurements as described above will apply.

4. SUBJECT DISPOSITION

4.1. Subject Enrollment and Disposition

A summary of subject enrollment will be provided by treatment group and overall for each country and investigator within country. The summary will present the number and percentage of subjects enrolled. For each column, the denominator for the percentage calculation will be the total number of subjects analyzed for that column.

A similar enrollment table will be provided by randomization stratum. The denominator for the percentage of subjects in the stratum will be the total number of randomized subjects. If there are discrepancies in the value used for stratification assignment between the IWRS and the clinical database, the value collected in the clinical database will be used for the summary. A listing of subjects with discrepancies in the value used for stratification assignment between the IWRS and the clinical database at the time of data finalization will be provided.

The randomization schedule used for the study will be provided as a by-subject listing for the CSR.

A summary of subject disposition will be provided by treatment group. This summary will present the number of subjects screened, the number of subjects randomized, and the number of subjects in each of the categories (as applicable) listed below:

- Safety Analysis Set
- Full Analysis Set
- Continuing randomized study drug (for interim analysis)
- Completed randomized study drug
- Did not complete randomized study drug with reasons for premature discontinuation of study drug in the Randomized Phase
- OL Safety Analysis Set
- Continuing OL study drug (for interim analysis)
- Completed OL study drug
- Did not complete OL study drug with reasons for premature discontinuation of study drug in the OL Phase
- Continuing study (for interim analysis)

- Completed study
- Did not complete study with reasons for premature discontinuation of study

For the Randomized Phase, 2 additional categories will be summarized under "Completed study drug":

- Completed study drug due to protocol-specified clinical events
- Completed 240 weeks of study drug

For the status of study drug and study completion and reasons for premature discontinuation, the number and percentage of subjects in each category will be provided. The denominator for the percentage calculation will be the total number of subjects in the Safety Analysis Set corresponding to that column in the corresponding study phase. In addition, a flowchart will be provided to depict the disposition by study phase.

In the Week 48 interim analysis, subject disposition will be summarized by study phase in 2 tables. The table for the Randomized Phase will be provided by treatment group and include items listed above, except for those based on the OL Safety Analysis Set and the study completion status. The table for the OL Phase will be provided on the OL Safety Analysis Set and will include study drug completion status. The flowcharts also will be provided by study phase.

The following by-subject listings will be provided by subject ID number in ascending order to support the above summary tables:

- Reasons for premature study drug or study discontinuation
- Reasons for screen failure (will be provided by screening ID number in ascending order)
- Lot number and bottle ID

4.2. Extent of Study Drug Exposure and Adherence

Extent of exposure to study drug will be examined by assessing the total duration of exposure to study drug and the level of adherence to the study drug specified in the protocol.

4.2.1. Duration of Exposure to Study Drug

For each study phase, duration of exposure to study drug will be defined as last dosing date minus first dosing date plus 1 and then divided by 7, regardless of any temporary interruptions in study drug administration, and will be expressed in weeks using up to 1 decimal place (eg, 4.5 weeks). For the Randomized Phase, first dosing date is the earlier date between the 2 drug types; last dosing date is the later date between the 2 drug types. If the last study drug dosing date is missing, the latest date among the study drug end date, clinical visit date, laboratory sample collection date, and vital signs assessment date that occurred during the corresponding phase will

be used for the final analysis. At the time of an interim analysis, the missing last dosing date will be imputed by the data snapshot date (or data cut date, if applicable) for subjects who are still on treatment.

The duration of exposure to study drug will be summarized using descriptive statistics and the number (ie, cumulative counts) and percentage of subjects exposed through the following time periods: 1 day, 1 week, 4 weeks, 12 weeks, 24 weeks, 36 weeks, 48 weeks, and every 24 weeks thereafter. Summaries will be provided according to subject grouping defined in Section 3.2.

No formal statistical testing is planned.

4.2.2. Adherence to Study Drug

The total number of tablets administered will be summarized using descriptive statistics.

The presumed total number of tablets administered to a subject will be determined by the data collected on the drug accountability eCRF using the following formula:

$$\text{Total Number of Doses Administered} = \left(\sum \text{No. of Doses Dispensed}\right) - \left(\sum \text{No. of Doses Returned}\right)$$

The level of on-treatment adherence to study drug regimen will be determined by the total amount of study drug administered relative to the total amount of study drug expected to be administered in each treatment phase during a subject's actual on-treatment period based on the study drug regimen.

The level of on-treatment adherence will be expressed as a percentage using the following formula:

$$\text{On-Treatment Adherence (\%)} = \left(\frac{\text{Total Amount of Study Drug Administered}}{\text{Study Drug Expected to be Administered on Treatment}}\right) \text{x } 100$$

4.2.2.1. Randomized Phase

Study drug expected to be administered will be calculated as treatment duration in days (one tablet per day for each drug type) as defined in Section 4.2.1. The adherence rates of SEL 18 mg tablet for SEL 18 mg treatment group and SEL 6 mg tablet for SEL 6 mg treatment group, and the average of the adherence rates between PTM 18 mg and PTM 6 mg tablets for placebo group will be summarized by treatment group (1, 2 and 4) as defined in Section 3.2.

4.2.2.2. Open-Label Phase

On-treatment adherence will be calculated for OL SEL 18 mg. Study drug expected to be administered will be calculated as treatment duration in days (1 tablet per day). Adherence rates will be summarized for the OL SEL 18 mg group as defined in Section 3.2.

At interim analyses, if the subject is ongoing, the number of tablets administered and expected to be administered on treatment will be calculated based on study drug accountability data.

No formal statistical testing is planned. Descriptive statistics for the level of on-treatment adherence with the number and percentage of subjects belonging to adherence categories ($< 75\%$, ≥ 75 to $< 90\%$, $\geq 90\%$) will be provided.

By-subject listing of study drug administration and drug accountability will be provided separately by subject ID number (in ascending order) and visit (in chronological order).

4.3. Protocol Deviations

Subjects who did not meet the eligibility criteria for study entry, but enrolled in the study will be summarized. No exemptions will be granted for this study. The summary will present the number and percentage of subjects who did not meet at least 1 eligibility criterion and the number of subjects who did not meet specific criteria by treatment group based on the All Randomized Analysis Set. A by-subject listing will be provided for those subjects 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 will be summarized by treatment group in the Randomized and OL Phases for the All Randomized Analysis Set and OL Safety Analysis Set respectively. A by-subject listing will be provided for those subjects with any important protocol deviation.

5. BASELINE CHARACTERISTICS

5.1. Demographics

Subject demographic variables (ie, age, sex, race, and ethnicity) will be summarized by treatment group and overall for each study phase using descriptive statistics for age, and using number and percentage of subjects for sex, race, and ethnicity. Age in the Randomized/OL Phase summary is calculated in years at the date of first study drug administration in the Randomized/OL Phase. The summary of demographic data will be provided for the Safety Analysis Set and OL Safety Analysis Set, respectively.

A by-subject demographic listing, including the informed consent date, will be provided by subject ID number in ascending order for the All Randomized Analysis Set.

5.2. Other Baseline Characteristics

For the Randomized Phase, baseline characteristics include:

- Fibrosis stage score (fibrosis stages according to the NASH CRN classification and Ishak method will be summarized separately)
- Hepatic collagen content
- NAS
- Steatosis grade
- Lobular inflammation grade
- Hepatocellular ballooning grade
- ALT
- AST
- GGT
- ALP
- Total bilirubin
- Direct bilirubin
- INR
- Platelets

- MELD score
- Child-Pugh (CP) score
- Height
- Body weight
- Body mass index (BMI) as a continuous variable and as categories (18 to < 25 kg/m^2, 25 to <30 kg/m^2 and ≥ 30 kg/m^2)
- Diabetes mellitus (absence or presence)
- Fasting glucose
- HbA1c
- ELF test score
- Fibrosis-4 (FIB-4) index
- NAFLD Fibrosis Score (NFS)
- Liver stiffness by transient elastography
- Use of Vitamin E
- History of SVR to therapy for chronic HCV

Use of Vitamin E is defined as treatment with Vitamin E on Day 1. Sustained virologic response is defined as HCV antibody positive and HCV virus negative at baseline.

For the OL Phase, baseline characteristics include:

- ALT
- AST
- GGT
- ALP
- Total bilirubin
- INR
- Platelets

- Body weight
- BMI as a continuous variable and for categories (< 18 kg/m^2[if applicable], 18 to < 25 kg/m^2, 25 to <30 kg/m^2 and ≥ 30 kg/m^2)
- Fasting glucose
- HbA1c

All baseline characteristics will be summarized by treatment group and overall for the Safety Analysis Set and OL Safety Analysis Set as defined in Section 3.2. Descriptive statistics for continuous variables and number and percentage of subjects for categorical variables will be used. No formal statistical testing is planned.

A by-subject listing of baseline characteristics will be provided by subject ID number in ascending order for the All Randomized Analysis Set.

5.3. Medical History

General medical history data will be collected at Screening. General medical history data will be coded using the current version of 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 provided in the medical history dataset.

6. EFFICACY ANALYSES

Efficacy analyses will include data collected in the Randomized Phase only. All data on or after the first dose of OL study drug will be excluded from efficacy analyses.

6.1. Primary Histology Efficacy Endpoint

6.1.1. Definition of Primary Histology Efficacy Endpoint

The primary histology efficacy endpoint at Week 48 is the proportion of subjects who achieve a ≥ 1-stage improvement in fibrosis (according to the NASH CRN classification) without worsening of NASH (defined as a ≥ 1 point increase in hepatocellular ballooning or lobular inflammation).

6.1.2. Statistical Hypothesis for Primary Histology Efficacy Endpoint

The statistical hypotheses to be tested can be stated as:

$$H_{s0i}: \delta_{si} = 0 \text{ versus } H_{sAi}: \delta_{si} \neq 0\,,\ i = 1 \text{ or } 2$$

Where δ_{s1} (δ_{s2}) is the difference in the proportion of subjects who achieve a ≥ 1-stage improvement in fibrosis without worsening of NASH at Week 48 between SEL 18 mg (or SEL 6 mg) and placebo. The hypotheses will be tested in the FAS with α = 0.025 for the comparison between each SEL arm and the placebo arm. The study will evaluate if each SEL arm performs better (ie, has a higher proportion of subjects who achieve a ≥ 1-stage improvement in fibrosis without worsening of NASH) than the placebo arm.

6.1.3. Analysis of Primary Histology Efficacy Endpoint

A stratified Mantel-Haenszel (MH) test will be used to compare the differences in proportions of subjects who achieve a ≥ 1-stage improvement in fibrosis without worsening of NASH at Week 48 between SEL 18 mg (SEL 6 mg) and placebo, adjusting for the presence or absence of diabetes mellitus at baseline and ELF test score ≥ 9.76 or < 9.76 at baseline. Subjects with missing data regarding fibrosis stage and NAS elements used to define worsening of NASH (hepatocellular ballooning and lobular inflammation) at Week 48 will be analyzed as treatment failures for this endpoint. The 2-sided 95% CIs for the difference in proportions between SEL 18 mg (SEL 6 mg) and placebo will be constructed based on stratum-adjusted MH proportions as follows {Koch 1989}:

$$p_A - p_B \pm Z_{(1-\alpha/2)} * SE(p_A - p_B),$$

where

- $(p_A - p_B) = \frac{\sum w_h d_h}{\sum w_h}$, is the stratum-adjusted MH proportion difference, where $d_h = p_{Ah} - p_{Bh}$ is the difference in the proportion of subjects who achieved a ≥ 1-stage improvement in fibrosis without worsening of NASH between SEL 18 mg (SEL 6 mg) and placebo in stratum h.

- $w_h = \frac{n_{Ah} n_{Bh}}{n_{Ah} + n_{Bh}}$, is the weight based on the harmonic mean of sample size per treatment group for each stratum where n_{Ah} and n_{Bh} are the sample sizes of SEL 18 mg (SEL 6 mg) and placebo in stratum h.

- $SE(p_A - p_B) = \sqrt{\frac{\sum w_h^2 \left[\frac{p^*_{Ah}(1-p^*_{Ah})}{n_{Ah}-1} + \frac{p^*_{Bh}(1-p^*_{Bh})}{n_{Bh}-1} \right]}{(\sum w_h)^2}}$, where

  $p^*_{Ah} = \frac{m_{Ah} + 0.5}{n_{Ah} + 1}$ and $p^*_{Bh} = \frac{m_{Bh} + 0.5}{n_{Bh} + 1}$, and

  m_{Ah} and m_{Bh} are the number of subjects who achieved a ≥ 1-stage improvement in fibrosis without worsening of NASH in SEL 18 mg (SEL 6 mg) and placebo in stratum h.

- $\alpha = 0.05$ for the calculation of 95% CI

- $Z_{(1-\alpha/2)} = Z_{0.975} = 1.96$ is the 97.5th percentile of the normal distribution

- Z score = $\frac{(p_A - p_B)}{SE_{(p_A - p_B)}}$

If the computed lower confidence bound is less than –1, the lower bound is defined as –1. If the computed upper confidence bound is greater than 1, then the upper bound is defined as 1.

The point estimates, 95% CIs, and corresponding p-values based on Z tests for the differences in proportions, and the 95% CIs for proportion of responders in each treatment group based on the Clopper-Pearson method will be calculated.

A logistic regression analysis will also be performed to assess the odds ratio between each of the SEL arms and the placebo arm, adjusting for stratification factors. The point estimate of the odds ratio, 95% CIs of odds ratio, and corresponding p-values for each SEL arm versus the placebo arm will be reported.

Forest plots of the treatment differences in the proportion of subjects who achieved primary histology endpoint will be generated for the subgroup analyses as described in Section 3.4 and sensitivity analyses as described in Section 3.6.1, respectively.

6.2. Clinical Efficacy Endpoint

6.2.1. Definition of Clinical Efficacy Endpoint

A composite of clinical events that constitute the clinical efficacy endpoint include:

1. Progression to cirrhosis as defined by a liver biopsy showing F4 fibrosis according to the NASH CRN classification as assessed by the central reader

2. Events of hepatic decompensation, including:

    a) Clinically apparent ascites requiring treatment

    b) Hepatic encephalopathy (HE) of Grade 2 or above (according to the West Haven criteria defined in Appendix 5 of the protocol) requiring treatment

    c) Portal hypertension-related upper gastrointestinal bleeding identified by endoscopy and requiring hospitalization, including events of bleeding from esophageal varices, gastric varices, and portal hypertensive gastropathy

3. Liver transplantation or qualification for liver transplantation (defined as MELD score ≥ 15 on at least 2 consecutive occasions at least 4 weeks apart)

4. All-cause mortality

Each of these clinical events (except histologic progression to cirrhosis, all-cause mortality and liver transplantation) will require confirmation by a Hepatic Events Adjudication Committee. The onset date of the first clinical event (biopsy showing F4 fibrosis, death, liver transplantation or adjudication confirmed event) will be used in the analysis.

6.2.2. Statistical Hypothesis for Clinical Efficacy Endpoint

The null statistical hypotheses to be tested can be stated as the survival and hazard functions being identical between SEL 18 mg (SEL 6 mg) and placebo.

6.2.3. Analysis of Clinical Efficacy Endpoint

Differences in EFS between SEL 18 mg (SEL 6 mg) and placebo will be assessed using the log-rank test stratified by presence or absence of diabetes mellitus and ELF test score ≥ 9.76 or < 9.76 at baseline.

Time (in months) to the first clinical event/censoring date will be calculated as: [first clinical event date or censoring date - randomization date + 1] / 30.4375.

Censoring date will be the later date between the last visit date and latest AE onset date in the Randomized Phase for subjects who have prematurely discontinued from the Randomized Phase, or who have completed the Randomized Phase. For the Week 48 Interim Analysis, the data cut date will be used as the censoring date for ongoing subjects in the Randomized Phase.

For the purpose of analysis, if the onset date of a clinical event or an AE is incomplete, the first day of the month (if date is missing) or the first day of the year (if date and month are missing) will be used to derive the clinical event or censoring date. If such imputation leads to an imputed event date that is earlier than the randomization date, then the randomization date should be used to impute the incomplete event date.

In the following code, *Months* is the number of months from randomization date to the first clinical event or to the censoring date for subjects who don't have clinical events. *Status* is a binary variable with 1 representing no clinical event observed and 0 representing a clinical event observed.

```
proc lifetest data=efs_2groups;
        time Months*Status(1);
        strata BL_Diabetes BL_ELF_category / group=treatment;
run;
```

A proportional hazards model will be used to calculate the point estimate and 95% CIs for the hazard ratio between SEL 18 mg (SEL 6 mg) and placebo controlling for stratification factors. Proportional hazard assumption will be evaluated.

```
proc phreg data=efs_3groups;
        class treatment BL_Diabetes BL_ELF_category;
        model Months*Status(1) =  treatment / RL;
        strata BL_Diabetes BL_ELF_category;
run;
```

For each treatment group, the KM estimates of EFS probability at 3-month intervals (ie, at Month 3, 6, 9, etc) and 95% CIs will be provided. The 95% CI will be calculated based on Greenwood's formula and log-log transformation of the survival function.

Median, Q1 and Q3 of the EFS will be provided for each treatment group. The 95% CI for median EFS based on Brookmeyer-Crowley method {Brookmeyer 1982} of inverting a generalization of the sign test for censored data will also be provided for each treatment group.

Kaplan-Meier curves will be plotted for the clinical EFS by treatment arm. A by-subject listing of clinical events will be provided.

A sensitivity analysis of EFS will be performed where mortality includes only liver-related death. All deaths will be reviewed by the Hepatic Event Adjudication Committee to determine if they are liver-related.

In addition, a sensitivity analysis for the clinical efficacy endpoint will be performed by using the KMMI approach as described in Section 3.6.1 if the survival functions of SEL 18 mg (SEL 6 mg) and placebo are demonstrated to be significantly different (in favor of SEL 18 mg or SEL 6 mg).

Forest plots of treatment effect in the clinical efficacy endpoints, assessed by the hazard ratios between each SEL arm and placebo arm, will be generated for the Week 48 interim analysis and the Week 240 final analysis. Forest plots for subgroup analyses will be generated in the final analysis at Week 240.

6.3. Secondary Efficacy Endpoints

6.3.1. Definition of Secondary Efficacy Endpoints

The secondary endpoints of this study are as follows:

1) Proportion of subjects who have progression to cirrhosis by Week 48

2) Proportion of subjects who have a ≥ 1-stage improvement in fibrosis (according to the NASH CRN classification) without worsening of NASH at Week 240

3) Proportion of subjects who have a ≥ 1-stage improvement in fibrosis (according to the NASH CRN classification) at Week 48 and Week 240

4) Proportion of subjects who have NASH resolution (defined as lobular inflammation of 0 or 1 from at least 1 at baseline and hepatocellular ballooning reduced to 0 from at least 1 at baseline; both criteria are necessary conditions) without worsening of fibrosis (according to the NASH CRN classification) at Week 48 and Week 240

6.3.2. Analysis of the Secondary Efficacy Endpoints

For the secondary efficacy endpoints, the stratified MH test defined for the primary histology efficacy endpoint in Section 6.1.3 will be used to compare the differences in proportions between SEL 18 mg (SEL 6 mg) and placebo, adjusting for the presence or absence of diabetes mellitus at baseline and ELF test score ≥ 9.76 or < 9.76 at baseline. The point estimates, 95% CIs and corresponding p-values based on a Z test for the differences in proportions, and 95% CIs for proportion of responder in each treatment group based on the Clopper-Pearson method will be calculated.

A logistic regression analysis will also be performed to assess the odds ratio between each of the SEL arms and the placebo arm, adjusting for stratification factors. The point estimates of odds ratios, 95% CIs of odds ratios, and corresponding p-values for each SEL arm versus placebo arm will be reported.

Forest plots of the treatment differences in proportion of subjects who achieved a response for each secondary efficacy endpoint will be generated for subgroup analyses as described in Section 3.4 and sensitivity analyses as described in Section 3.6.1, respectively.

The point estimates and 95% CIs based on stratified MH tests of NRI data will be provided in the subgroup analysis for the primary histology endpoint and secondary endpoints. The point estimates and 95% CIs based on a proportional hazard model will be provided for the subgroup analysis for the primary clinical efficacy endpoint. When the grouping variable is identical to one of the stratification factors (ie, subgroups by baseline diabetes status or baseline ELF category), this stratification factor will not be adjusted in the modeling.

6.4. Exploratory Efficacy Endpoints

6.4.1. Definition of Exploratory Efficacy Endpoints

CCI [redacted]

CCI [redacted]

CCI [redacted]

CCI [redacted]

6.4.2. Analysis Methods for Exploratory Efficacy Endpoints

CCI [redacted]

CCI

[redacted]

6.5. Change From Protocol-Specified Efficacy Analyses

The analyses for the objective to determine whether noninvasive measures of fibrosis can predict histological regression and/or progression of fibrosis and clinical complications of cirrhosis will be included in the statistical analysis plan for noninvasive tests.

7. SAFETY ANALYSES

Subjects will be grouped as defined in Section 3.2.

7.1. Adverse Events and Deaths

7.1.1. Adverse Event Dictionary

Clinical and laboratory AEs will be coded using the current version of MedDRA. SOC, HLGT, HLT, PT, and LLT will be provided in the AE dataset.

7.1.2. Adverse Event Severity

Adverse events are graded by the investigator as Grade 1, 2, 3, 4, or 5 according to Common Terminology Criteria for Adverse Events (CTCAE). The severity grade of events for which the investigator did not record severity will be categorized as "missing" for tabular summaries and data listings. The most severe grade will be considered in data summary if a subject experienced multiple events under the same category. The missing category will be listed last in summary presentation when applicable.

7.1.3. Relationship of Adverse Events to Study Drug

Related AEs are those for which the investigator selected "Related" on the AE CRF to the question of "Related to Study Treatment." Relatedness will always default to the investigator's choice, not that of the medical monitor. Events for which the investigator did not record relationship to study drug will be considered related to study drug for summary purposes. However, by-subject data listings will show the relationship as missing.

7.1.4. Serious Adverse Events

Serious adverse events (SAEs) will be identified and captured as SAEs if AEs met the definitions of SAEs that were specified in the study protocol. Serious adverse events captured and stored in the clinical database will be reconciled with the SAE database from Gilead Pharmacovigilance and Epidemiology before data finalization.

7.1.5. Treatment-Emergent Adverse Events

7.1.5.1. Definition of Treatment-Emergent Adverse Events

Treatment-emergent adverse events (TEAEs) for the Randomized/OL Phase are defined as 1 or both of the following:

- Any AEs with an onset date on or after the study drug start date and no later than 30 days after permanent discontinuation of study drug in the Randomized Phase (and before the first dosing date in the OL Phase)/ OL Phase.

- Any AEs leading to premature discontinuation of study drug in the Randomized/OL Phase.

7.1.5.2. Definition of Adverse Events > 30 Days Posttreatment

Adverse events > 30 days posttreatment for the Randomized/OL Phase are defined as any AEs with an onset date later than 30 days after permanent discontinuation of study drug in the Randomized Phase (and before the first dosing date in the OL Phase)/ OL Phase.

7.1.5.3. Incomplete Dates

If the onset date of an AE is incomplete and the AE stop date is not prior to the first dosing date of study drug, then the month and year (or year alone if month is not recorded) of onset determine whether an AE is treatment emergent in the corresponding treatment phase. The event is considered treatment emergent if both of the following 2 criteria are met:

- The AE onset is the same as or after the month and year (or year) of the first dosing date of study drug, and
- The AE onset date is the same as or before the month and year (or year) of the date corresponding to 30 days after the date of the last dose of study drug in the Randomized Phase (and before the first dosing date in the OL Phase)/OL Phase.

An AE with completely missing onset and stop dates, or with the onset date missing and a stop date later than the first dosing date of study drug, will be considered to be treatment emergent. In addition, an AE with the onset date missing and incomplete stop date with the same or later month and year (or year alone if month is not recorded) as the first dosing date of study drug will be considered treatment emergent. In above cases if an AE is considered treatment-emergent, it will not be considered > 30 days posttreatment.

7.1.6. Summaries of Adverse Events and Deaths

Treatment-emergent adverse events and AEs > 30 days posttreatment will be summarized for the Safety Analysis Set/OL Safety Analysis Set according to the subject grouping defined in Section 3.2.

The number and percentage of subjects who experienced at least 1 TEAE will be provided and summarized by SOC, HLT and PT. For other AEs described below, summaries will be provided by SOC and PT:

- TEAEs by severity grade
- TEAEs of Grade 3 or higher (by maximum severity)
- TEAEs of Grade 2 or higher
- All TE treatment-related AEs
- TE Treatment-related AEs of Grade 3 or higher (by maximum severity)

- TE Treatment-related AEs of Grade 2 or higher
- All TE SAEs
- All TE treatment-related SAEs
- All TE treatment-related AEs by severity grade
- Death
- All TEAEs leading to discontinuation of study drug In the Randomized Phase, TEAEs leading to discontinuation of study drug include 2 categories: protocol-specified clinical events, and TEAEs leading to premature discontinuation of study drug. The latter will be provided in a separate summary table.

A brief, high-level summary of AEs described above (except for the summary by severity grade) will be provided by treatment group and by the number and percentage of subjects who experienced the above AEs. All deaths observed in the study and subjects who completed randomized study drug due to protocol-specified clinical events will be also included in this summary.

Multiple events will be counted only once per subject in each summary. Adverse events will be summarized and listed first in alphabetic order of SOC and HLT within SOC (if applicable), and then in descending order of PTs by frequency in the (OL) SEL 18 mg group within each SOC or HLT in the Randomized/OL Phase. For summaries by severity grade, the most severe grade will be used for those AEs that occurred more than once in an individual subject in the Randomized/OL Phase.

In addition to the above summary tables, the following tables will be summarized by PT only, in descending order of frequency in the (OL) SEL 18 mg group for the Randomized/OL Phase:

- All TEAEs
- TEAEs of Grade 3 or higher (by maximum severity)
- TEAEs of Grade 2 or higher
- TE treatment-related AEs
- TE Treatment-related AEs of Grade 3 or higher (by maximum severity)
- TE Treatment-related AEs of Grade 2 or higher
- All TESAEs
- All TE treatment-related SAEs

The following tables for AEs > 30 days posttreatment will be summarized by PT only, in descending order of frequency in the (OL) SEL 18 mg group for Randomized/OL Phase:

- All AEs > 30 days posttreatment
- Treatment-related AEs > 30 days posttreatment
- SAEs > 30 days posttreatment

A brief, high-level summary of AEs > 30 days posttreatment described above will be provided by treatment group and by the number and percentage of subjects who experienced the above AEs.

In addition, data listings will be provided for the following:

- All AEs, indicating whether the event is treatment emergent
- All AEs of Grade 3 or higher
- SAEs
- Deaths
- AEs leading to discontinuation of study drug

7.1.7. Additional Analysis of Adverse Events

7.1.7.1. Cardiovascular Events

Proportion of subjects with adjudication-confirmed cardiovascular events will be summarized by treatment group for the Randomized and OL Phase. Multiple events will be counted only once per subject in each summary for each phase. All adjudicated cardiovascular events will be listed for All Randomized Analysis set.

Time to the first cardiovascular event will be assessed in Randomized Phase. Time (in months) to the first cardiovascular event/censoring date will be calculated as: (first cardiovascular event date or censoring date − randomization date + 1) / 30.4375. Censoring date will be the latest date among the last visit date in the Randomized Phase, the latest AE onset date in the Randomized Phase, the day prior to OL Study Day 1 for subjects discontinued from the Randomized Phase, and the data cut date for ongoing subjects in the Randomized Phase (for the Week 48 Interim Analysis only).

For the purpose of analysis, if the onset date of a cardiovascular event or an AE is incomplete, the first day of the month (if date is missing) or the first day of the year (if date and month are missing) will be used to derive the first cardiovascular event/censoring date. If such imputation leads to an imputed event date that is earlier than the randomization date, then the randomization date should be used to impute the incomplete event date.

Proportional hazards model will be used to calculate the point estimate and 95% CI for the hazard ratio between SEL 18 mg (SEL 6 mg) and placebo controlling for stratification factors.

For each treatment group, KM estimates of cardiovascular event probability at 3-month intervals (ie, at Month 3, 6, 9, etc) and 95% CIs will be provided. The 95% CIs will be calculated based on Greenwood's formula and log-log transformation of the survival function.

Median, Q1 and Q3 of the time to the first cardiovascular event will be provided for each treatment group. The 95% CI for median EFS based on Brookmeyer-Crowley method {Brookmeyer 1982} of inverting a generalization of the sign test for censored data will also be provided for each treatment group.

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 the (OL) Safety Analysis Set as appropriate and will include data collected up to the last dose of any study drug in the Randomized/OL Phase plus 30 days for subjects who have permanently discontinued study drug in the Randomized (and before the first dosing date in the OL Phase)/OL Phase. The analysis 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 3.7.

A by-subject listing for laboratory test results will be provided by subject ID number and visit in chronological order for hematology, serum chemistry, and coagulation separately. Values falling outside of the relevant reference range and/or having a severity grade of 1 or higher on the CTCAE severity grade as described in Appendix 2 will be flagged in the data listings, as appropriate.

No formal statistical testing is planned.

7.2.1. Summaries of Numeric Laboratory Results

Alanine aminotransferase, AST, albumin, ALP, glucose, total bilirubin, direct bilirubin, GGT, HOMA-IR, HbA1c, and INR in the Randomized Phase will be summarized as efficacy endpoints. Descriptive statistics will be provided by treatment group for ALT, AST, albumin, ALP, glucose, total bilirubin, direct bilirubin, GGT, HOMA-IR, HbA1c, and INR in the OL Phase, and creatinine, creatinine clearance (Cockroft-Gault), white blood cells (WBCs), neutrophils, lymphocytes, hemoglobin, and platelets in the Randomized/OL Phase as follows:

- Baseline value
- Value at each postbaseline visit
- Change from baseline at each postbaseline visit

Change from baseline at a postbaseline visit will be defined as the visit value minus the baseline value in the corresponding study phase. The 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.

Baseline and postbaseline CP score in the OL Phase will be summarized by counts of subjects and percentages at each category level. Change from baseline will be summarized in the following 3 ordered categories: improvement, no change, and worsening. Worsening refers to an increase in the score, and improvement refers to a decrease in the score.

Median (Q1, Q3) of the observed values for WBC, neutrophils, lymphocytes, hemoglobin, and platelets will be plotted using a line plot by treatment group (1, 2, and 4) and visit for the Randomized Phase.

In the case of multiple values in an analysis window, data will be selected for analysis as described in Section 3.8.3.

7.2.2. Graded Laboratory Values

The CTCAE Version 5.0 will be used for assigning toxicity grades (0 to 4) to laboratory results for analysis. Grade 0 includes all values that do not meet the criteria for an abnormality of at least Grade 1. For laboratory tests with criteria for both increased and decreased levels, analyses for each direction (i.e., increased, decreased) will be presented separately.

For the baseline ALT, AST, total bilirubin and direct bilirubin toxicity grades of each phase, the CTCAE version 5.0 will be used to assign grades to the derived average values.

7.2.2.1. Treatment-Emergent Laboratory Abnormalities

Treatment-emergent laboratory abnormalities are defined as values that increase by at least 1 toxicity grade from baseline at any postbaseline time point in the corresponding study phase, up to and including the date of the last dose of study drug plus 30 days for subjects who permanently discontinued study drug (and before OL Study Day 1 for abnormalities in the Randomized Phase), or the last available date in the database snapshot for subjects who were still on treatment at the time of an interim analysis. If the relevant baseline laboratory value is missing, any abnormality of at least Grade 1 observed within the time frame specified above will be considered treatment emergent.

7.2.2.2. Treatment-Emergent Marked Laboratory Abnormalities

Treatment-emergent marked laboratory abnormalities are defined as values that increase from baseline by at least 3 toxicity grades at any postbaseline time point, up to and including the date of the last dose of study drug in the corresponding study phase plus 30 days for subjects who permanently discontinued study drug (and before OL Study Day 1 for abnormalities in the Randomized Phase) or the last available date in the database snapshot for subjects who were still on treatment at the time of an interim analysis. If the relevant baseline laboratory value is

missing, any Grade 3 or above values observed within the timeframe specified above will be considered treatment-emergent marked abnormality.

7.2.2.3. Summaries of Laboratory Abnormalities

Laboratory data that are categorical will be summarized using the number and percentage of subjects in the study with the given response at baseline and each scheduled postbaseline visit for the Randomized/OL Phase.

The following summaries (number and percentage of subjects) for treatment-emergent laboratory abnormalities will be provided by laboratory test and treatment group; subjects will be categorized according to the most severe postbaseline abnormality grade for a given laboratory test:

- Graded laboratory abnormalities
- Grade 3 or 4 laboratory abnormalities
- Marked laboratory abnormalities

For all summaries of laboratory abnormalities, the denominator is the number of subjects with nonmissing postbaseline values up to 30 days after the last dosing date (and before OL Study Day 1 for the Randomized Phase) of the reporting phase.

A by-subject listing of treatment-emergent Grade 3 or 4 laboratory abnormalities will be provided by subject ID number and visit in chronological order. This listing will include all test results that were collected throughout the study for the laboratory test of interest, with all applicable severity grades or abnormal flags displayed.

7.2.3. Liver-related Laboratory Evaluations

The summary will be provided by 3 subgroups of subjects according to their baseline ALT/AST level in the Randomized/OL Phase:

- Normal
- Elevated (> 1 to $< 5 \times$ upper limit of normal [ULN])
- Elevated ($\geq 5 \times$ ULN)

Liver-related abnormalities after initial study drug dosing will be examined and summarized using the number and percentage of subjects who were reported to have the following laboratory test values for postbaseline measurements:

For subjects with normal baseline ALT/AST:

- Subjects meeting criteria for close observation
    - ALT/AST > 3 × ULN
- Subjects meeting criteria for holding study drug
    - ALT/AST > 8 × ULN
    - ALT/AST > 5 × ULN for 2 weeks
    - ALT/AST > 3 × ULN and total bilirubin > 2 × ULN (or direct bilirubin 2 × baseline in subjects with Gilbert's syndrome)
    - ALT/AST > 3 × ULN and INR > 1.5 (if not on anticoagulation)

For subjects with baseline ALT/AST between 1 and 5 × ULN

- Subjects meeting criteria for close observation
    - ALT/AST > 2 × Baseline or > 300 U/L
- Subjects meeting criteria for holding study drug
    - ALT/AST > 8 × Baseline or > 500 U/L
    - ALT/AST > 3 × Baseline or > 300 U/L, and total bilirubin > 2 × ULN (or direct bilirubin 2 x Baseline in subjects with Gilbert's syndrome)
    - ALT/AST > 3 × Baseline and INR > 1.5 (if not on anticoagulation)

For subjects with baseline ALT/AST ≥ 5 × ULN

- Subjects meeting criteria for close observation
    - ALT/AST > 2 × Baseline
- Subjects meeting criteria for holding study drug
    - ALT/AST > 500 U/L

The above summary will include data from all postbaseline visits up to 30 days after the last dose of study drug for subjects who permanently discontinued study drug (and before OL Study Day 1 for abnormalities in the Randomized Phase) in the reporting phase, or the last available date in the database snapshot for subjects who were still on treatment at the time of an interim analysis. The denominator is the number of subjects in the (OL) Safety Analysis Set who have nonmissing

postbaseline values of all relevant tests at the same postbaseline visit date in the summarizing group. A listing of subjects meeting any of the criteria above will be provided.

7.2.4. Drug-Induced Liver Injury (DILI)

Due to the challenge of recognizing and diagnosing DILI in subjects with pre-existing hepatic dysfunction, a DILI Adjudication Committee will review potential cases of DILI identified based on laboratory parameters listed in Section 7.2.3. Proportion of subjects with DILI events determined by the Adjudication Committee, including confirmed DILI events (or worsening of hepatic function attributable to study drug could not be excluded) and insufficient data, will be summarized by treatment group for the Randomized and OL Phase. Multiple events will be counted only once per subject for each phase. All adjudicated DILI events will be listed for All Randomized Analysis set.

7.3. Body Weight and Vital Signs

Descriptive statistics will be provided by treatment group for body weight, waist circumference, BMI and vital signs (systolic and diastolic blood pressure [mmHg], pulse [bpm], respiratory rate [breath/min] and temperature [°C]) in the Randomized/OL Phase as follows:

- Baseline value
- Value at each postbaseline visit
- Change from baseline at each postbaseline visit
- Percent change from baseline at each postbaseline visit (only for body weight)

A baseline value in the Randomized/OL Phase will be defined as the last available value collected on or prior to the date/time of first dose of study drug in the corresponding treatment phase. Change from baseline to a postbaseline visit will be defined as the postbaseline value minus the baseline value. Vital signs measured at unscheduled visits will be included for the baseline value selection. CCI [redacted]

In the case of multiple values in an analysis window, data will be selected for analysis as described in Section 3.8.3. No formal statistical testing is planned.

A by-subject listing of vital signs will be provided by subject ID number and visit in chronological order. A by-subject listing of body weight, height, BMI and waist circumference will be provided.

7.4. Prior and Concomitant Medications

Medications collected at Screening and during the study will be coded using the current version of the World Health Organization (WHO) Drug dictionary.

7.4.1. Prior Medications

Prior medications are defined as any medication taken before a subject took the first study drug in the Randomized Phase.

Prior medications will be summarized by preferred name using the number and percentage of subjects for each treatment group and overall. A subject reporting the same medication more than once will be counted only once when calculating the number and percentage of subjects who received that medication. The summary will be ordered by preferred name in order of descending overall frequency. For drugs with the same frequency, sorting will be done alphabetically.

For the purposes of analysis, any medication with a start date prior to the first dosing date of study drug will be included in the prior medication summary regardless of when the stop date is. If a partial start date is entered the medication will be considered prior unless the month and year (if day is missing) or year (if day and month are missing) of the start date are after the first dosing date. Medications with a completely missing start date will be included in the prior medication summary, unless otherwise specified.

Summaries will be based on the Safety Analysis Set. No formal statistical testing is planned.

7.4.2. Concomitant Medications

Concomitant medications are defined as medications taken while a subject took study drug. Use of concomitant medications will be summarized by preferred name using the number and percentage of subjects for each treatment group in the Randomized/OL Phase. A subject reporting the same medication more than once will be counted only once when calculating the number and percentage of subjects who received that medication. The summary will be ordered by preferred name in descending frequency in the (OL) SEL 18 mg group for the Randomized/OL Phase. For drugs with the same frequency, sorting will be done alphabetically.

For the purposes of analysis, any medication with a start date prior to or on the first dosing date of study drug and continued to take after the first dosing date, or started after the first dosing date but prior to or on the last dosing date of study drug will be considered concomitant medications. Medications started and stopped on the same day as the first dosing date or the last dosing date of study drug will also be considered concomitant. Medications with a stop date prior to the date of first dosing date of study drug or a start date after the last dosing date of study drug will be excluded from the concomitant medication summary. If a partial stop date is entered, any medication with the month and year (if day is missing) or year (if day and month are missing) prior to the date of first study drug administration will be excluded from the concomitant medication summary. If a partial start date is entered, any medication with the month and year (if day is missing) or year (if day and month are missing) after the study drug stop date will be excluded from the concomitant medication summary. Medications with completely missing start and stop dates will be included in the concomitant medication summary, unless otherwise specified. Summaries will be based on the (OL) Safety Analysis Set. No formal statistical testing is planned.

All prior and concomitant medications (other than per-protocol study drugs) will be provided in a by-subject listing sorted by subject ID number and administration date in chronological order.

7.5. Electrocardiogram Results

Electrocardiogram (ECG) data will not be presented in the CSR since ECGs were not assessed in this study other than as part of the screening process for potential new subjects.

7.6. Other Safety Measures

7.6.1. Stool Frequency

Stool frequency subscore from the partial Mayo Score will be summarized by visit and treatment group. Change from baseline will be summarized in the following 3 ordered categories: increased, no change, and decreased. A baseline value in the Randomized/OL Phase will be defined as the last available value collected on or prior to the date/time of first dose of study drug in the corresponding treatment phase.

Table 7-1. Stool Frequency Subscore of Partial Mayo Score

Domain	Description	Score
Stool Frequency (based on the past 3 days)	Normal number of stools	0
	1-2 stools more than normal	1
	3-4 stools more than normal	2
	5 or more stools more than normal	3

7.7. Changes From Protocol-Specified Safety Analyses

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

8. PHARMACOKINETIC ANALYSES

A single PK blood sample will be collected during the Randomized Phase at Week 1, 4, 12, 24, and 48 visits in all subjects and during the OL Phase at all visits only in subjects with severe hepatic impairment (Child-Pugh Class C) in combination with renal impairment (eGFR < 30 mL/min)

CCI [redacted]

Concentrations of SEL and GS-607509 in plasma will be determined using validated bioanalytical assays. A population PK model will be developed to characterize the PK of SEL and its metabolite (as applicable). Data from this study (single PK and PK substudy) will be combined with data from other studies in a meta-population analysis using nonlinear mixed-effects modeling techniques. Details of the population PK analysis will be provided in a separate population PK analysis report.

The following listing will be provided for both single and intensive PK samples:

- PK sampling details by subject and analyte including actual dosing time and actual draw time, calculated time postdose of sample collection, differences in scheduled and actual draw times, sample age, and sample concentration

9. REFERENCES

Brookmeyer R, Crowley J. A confidence interval for the median survival time. Biometrics 1982;38:29-41.

Goff DC, Jr., Lloyd-Jones DM, Bennett G, Coady S, D'Agostino RB, Gibbons R, et al. 2013 ACC/AHA guideline on the assessment of cardiovascular risk: a report of the American College of Cardiology/American Heart Association Task Force on Practice Guidelines. Circulation 2014;129 (25 Suppl 2):S49-73.

Haybittle JL. Repeated assessment of results in clinical trials of cancer treatment. Br J Radiol 1971;44 (526):793-7.

Hochberg Y. A sharper Bonferroni procedure for multiple tests of significance. Biometrika 1988;75 (4):800-2.

Koch GG, Carr GJ, Amara IA, Stokes ME, Uryniak TJ. Categorical Data Analysis. Chapter 13 in Berry, D.A. (ed.). Statistical Methodology in the Pharmaceutical Sciences. New York: Marcel Dekker, Inc., 1989:pp. 414-21.

Ratitch B, O'Kelly M, Tosiello R. Missing data in clinical trials: from clinical assumptions to statistical analysis using pattern mixture models. Pharm Stat 2013;12 (6):337-47.

Rubin DB. Multiple Imputation for Nonresponse in Surveys. New York, NY: John Wiley & Sons, Inc; 1987.

Zhao Y, Herring AH, Zhou H, Ali MW, Koch GG. A Multiple Imputation Method for Sensitivity Analyses of Time-to-Event Data with Possibly Informative Censoring [Author Manuscript]. J Biopharm Stat 2014;24 (2):229-53.

10. SOFTWARE

SAS® Software Version 9.4. SAS Institute Inc., Cary, NC, USA.

nQuery +nTerim Version 4.0. Statistical Solutions, Cork, Ireland.

EAST Version 6. Cytel. Cambridge, MA, USA.

11. SAP REVISION

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

12. APPENDICES

Appendix 1. Study Procedures Table – Screening to Week 48

Assessments	Screening	Day 1[a]	On-treatment Visits												
			Week 1	Week 4	Week 8	Week 12	Week 16	Week 20	Week 24	Week 28	Week 32	Week 36	Week 40	Week 44	Week 48
Clinical Assessments															
Written Informed Consent	X														
Determine Eligibility	X														
Medical History	X														
Physical Examination	X	X[b]	X[b]	X[b]	X[b]	X[b]	X[b]	X[b]	X[b]	X[b]	X[b]	X[b]	X[b]	X[b]	X[b]
Vital Signs including Body Weight	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X
Height	X														
Waist Circumference	X	X		X		X			X						X
12-lead ECG	X														
CP/MELD Scores[c]	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X
Assess Ascites and Hepatic Encephalopathy	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X
Liver Biopsy[d]	X														X
Abdominal Ultrasound (OL Phase only)		X							X						X
Elastography (if available)	X[e]	X[f]							X						X

Assessments	Screening	Day 1[a]	On-treatment Visits												
			Week 1	Week 4	Week 8	Week 12	Week 16	Week 20	Week 24	Week 28	Week 32	Week 36	Week 40	Week 44	Week 48
Health Related Quality of Life Questionnaires[g]		X				X			X						X
Health Resource Utilization Questionnaires[g]		X				X			X						X
Stool Frequency Assessment		X	X	X	X	X	X	X	X	X	X	X	X	X	X
Lifestyle Modification Counseling		X	X	X	X	X	X	X	X	X	X	X	X	X	X
Adverse Events	X[h]	X	X	X	X	X	X	X	X	X	X	X	X	X	X
Concomitant Medications	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X
Dispense Study Drug		X		X	X	X	X	X	X	X	X	X	X	X	X
Review of Study Drug Dosing Compliance (Pill Count)			X	X	X	X	X	X	X	X	X	X	X	X	X
Laboratory Assessments															
Subject Fasting	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X
Chemistry[i], Hematology, Coagulation	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X
Insulin and Lipids		X		X		X			X						X
HbA1c	X	X		X		X			X						X
eGFR	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X
Pregnancy Testing[j]	X	X		X	X	X	X	X	X	X	X	X	X	X	X

Assessments	Screening	Day 1[a]	On-treatment Visits												
			Week 1	Week 4	Week 8	Week 12	Week 16	Week 20	Week 24	Week 28	Week 32	Week 36	Week 40	Week 44	Week 48
Single PK Sampling		X[l]	X[k]	X[k]	X[l]	X[k]	X[l]	X[l]	X[k]	X[l]	X[l]	X[l]	X[l]	X[l]	X[k]
Blood Collection (Biomarkers)	X	X				X			X						X
Urine Collection		X				X			X						X
Urine Drug Screen	X														
HIV-1, HBV, HCV Serology	X														
CCI [redacted] CCI															
Approximate amount of blood drawn (mL)[n]	70	65	15	15	12	60	12	12	60	12	12	12	12	12	65

Study Procedures Table – Week 60 to Follow Up

Assessments	On-treatment Visits										Follow Up[o]	Telephone Follow-Up[p]
	Week 60	Week 72	Week 84	Week 96	Week 108	Week 120	Week 132	Week 144	Week 156	ET[o]		
		Week 168	Week 180	Week 192	Week 204	Week 216	Week 228	Week 240 (EOT)				
Physical Examination	X[b]	X[b]	X[b]	X[b]	X[b]	X[b]	X[b]	X[b]	X[b]	X[b]	X[b]	
Vital Signs including Body Weight	X	X	X	X	X	X	X	X	X	X	X	
Waist Circumference		X		X		X		X		X		
Liver Biopsy[d]								X[q]		X[r]		
CP/MELD scores[c]	X	X	X	X	X	X	X	X	X	X		
Assess Ascites and Hepatic Encephalopathy	X	X	X	X	X	X	X	X	X	X		
Abdominal Ultrasound (OL Phase only)		X		X		X		X		X		
Elastography (if available)				X				X		X		
Health Related Quality of Life Questionnaires[g]		X		X		X		X		X		
Health Resource Utilization Questionnaires[g]		X		X		X		X		X		
Stool Frequency Assessment	X	X	X	X	X	X	X	X	X	X	X	
Lifestyle Modification Counseling	X	X	X	X	X	X	X	X	X	X		
Adverse Events	X	X	X	X	X	X	X	X	X	X	X	X
Concomitant Medications	X	X	X	X	X	X	X	X	X	X	X	
Dispense Study Drug	X	X	X	X	X	X	X	X[s]	X			
Review of Study Drug Dosing Compliance (Pill Count)	X	X	X	X	X	X	X	X	X	X		

Assessments	On-treatment Visits										Follow Up[o]	Telephone Follow-Up[p]
	Week 60	Week 72 / Week 168	Week 84 / Week 180	Week 96 / Week 192	Week 108 / Week 204	Week 120 / Week 216	Week 132 / Week 228	Week 144 / Week 240 (EOT)	Week 156	ET[o]		
Laboratory Assessments												
Subject Fasting	X	X	X	X	X	X	X	X	X	X	X	
Chemistry[i], Hematology, Coagulation	X	X	X	X	X	X	X	X	X	X	X	
Insulin and Lipids		X		X		X		X		X		
HbA1c		X		X		X		X		X		
eGFR	X	X	X	X	X	X	X	X	X	X	X	
Pregnancy Testing[j]	X	X	X	X	X	X	X	X	X	X	X	
Single PK sampling[l]	X	X	X	X	X	X	X	X	X	X[t]		
Blood Collection (Biomarkers)				X				X		X		
Urine Collection				X				X		X		
CCI												
Approximate amount of blood drawn (mL)[n]	12	15	12	65	12	15	12	65	12	65	10	

a Subjects starting the OL Phase of the study will complete the same study procedures as during the Randomized Phase of the study, starting with the Day 1 visit
b Symptom-driven PE
c CP score to be calculated only at Screening
d Not to be performed during the OL Phase of the study
e CCI
f CCI
g For subjects with questionnaires available at Day 1
h Adverse Events reporting during Screening is limited to serious adverse events and adverse events related to study procedures
i Additional testing: digoxin level at Day 1 and Week 1 and as needed in subjects taking digoxin (refer to Section 5.4 of the protocol) and CPK level for subjects in close observation for DILI (refer to Section 7.5 of the protocol)

j Females of childbearing potential only (refer to Appendix 3 of the protocol). Serum pregnancy test at Screening and urine pregnancy test at Day 1 and every 4 weeks thereafter. Starting at the Week 48 visit, urine pregnancy testing kits will be provided for home testing every 4 weeks, between in-clinic study visits. All females of childbearing potential will be contacted every 4 weeks and asked to report the result of the urine pregnancy tests. FSH testing as per Appendix 3 of the protocol
k To be performed during the Randomized Phase in all subjects and during the OL Phase, only in subjects with severe hepatic impairment (Child-Pugh Class C) in combination with renal impairment (eGFR < 30 mL/min)
l To be performed during the OL Phase, only in subjects with severe hepatic impairment (Child-Pugh Class C) in combination with renal impairment (eGFR < 30 mL/min)
m CCI [redacted]
n For specific blood volumes at each visit, please refer to the current approved ICF
o Subjects prematurely discontinuing from the study should complete an ET visit within 30 days of last dose, Follow-Up visit 4 weeks later and a Telephone Follow-Up visit after the ET visit
p To be performed 12 weeks after the Week 240/EOT or ET visit. At the discretion of the investigator, an unscheduled visit may be completed if the subject reports abnormal or concerning symptoms
q Not to be performed at the Week 144 visit
r To be performed at the discretion of the investigator
s Not to be performed at the Week 240/EOT visits
t To be collected in all subjects that ET in Randomized or OL Phase

Appendix 2. CTCAE Grade for Laboratory Parameters

CTCAE 5.0	CTCAE Grade				
CTCAE Term	**Grade 1**	**Grade 2**	**Grade 3**	**Grade 4**	**Grade 5**
Anemia	Hemoglobin (Hgb) <LLN - 10.0 g/dL; <LLN - 6.2 mmol/L; <LLN - 100 g/L	Hgb <10.0 - 8.0 g/dL; <6.2 - 4.9 mmol/L; <100 - 80g/L	Hgb <8.0 g/dL; <4.9 mmol/L; <80 g/L; transfusion indicated	Life-threatening consequences; urgent intervention indicated	Death
Activated partial thromboplastin time prolonged	>ULN - 1.5 x ULN	>1.5 - 2.5 x ULN	>2.5 x ULN; bleeding	-	-
Alanine aminotransferase increased	>ULN - 3.0 x ULN if baseline was normal; 1.5 - 3.0 x baseline if baseline was abnormal	>3.0 - 5.0 x ULN if baseline was normal; >3.0 - 5.0 x baseline if baseline was abnormal	>5.0 - 20.0 x ULN if baseline was normal; >5.0 - 20.0 x baseline if baseline was abnormal	>20.0 x ULN if baseline was normal; >20.0 x baseline if baseline was abnormal	-
Alkaline phosphatase increased	>ULN - 2.5 x ULN if baseline was normal; 2.0 - 2.5 x baseline if baseline was abnormal	>2.5 - 5.0 x ULN if baseline was normal; >2.5 - 5.0 x baseline if baseline was abnormal	>5.0 - 20.0 x ULN if baseline was normal; >5.0 - 20.0 x baseline if baseline was abnormal	>20.0 x ULN if baseline was normal; >20.0 x baseline if baseline was abnormal	-
Aspartate aminotransferase increased	>ULN - 3.0 x ULN if baseline was normal; 1.5 - 3.0 x baseline if baseline was abnormal	>3.0 - 5.0 x ULN if baseline was normal; >3.0 - 5.0 x baseline if baseline was abnormal	>5.0 - 20.0 x ULN if baseline was normal; >5.0 - 20.0 x baseline if baseline was abnormal	>20.0 x ULN if baseline was normal; >20.0 x baseline if baseline was abnormal	-
Blood bicarbonate decreased	<LLN and no intervention initiated	-	-	-	-
Blood bilirubin increased	>ULN - 1.5 x ULN if baseline was normal; > 1.0 - 1.5 x baseline if baseline was abnormal	>1.5 - 3.0 x ULN if baseline was normal; >1.5 - 3.0 x baseline if baseline was abnormal	>3.0 - 10.0 x ULN if baseline was normal; >3.0 - 10.0 x baseline if baseline was abnormal	>10.0 x ULN if baseline was normal; >10.0 x baseline if baseline was abnormal	-
Cholesterol high	>ULN - 300 mg/dL; >ULN - 7.75 mmol/L	>300 - 400 mg/dL; >7.75 - 10.34 mmol/L	>400 - 500 mg/dL; >10.34 - 12.92 mmol/L	>500 mg/dL; >12.92 mmol/L	-
CPK increased	>ULN - 2.5 x ULN	>2.5 x ULN - 5 x ULN	>5 x ULN - 10 x ULN	>10 x ULN	-

CTCAE 5.0	CTCAE Grade				
CTCAE Term	**Grade 1**	**Grade 2**	**Grade 3**	**Grade 4**	**Grade 5**
Creatinine increased	>ULN - 1.5 x ULN	>1.5 - 3.0 x baseline; >1.5 - 3.0 x ULN	>3.0 x baseline; >3.0 - 6.0 x ULN	>6.0 x ULN	-
GGT increased	>ULN - 2.5 x ULN if baseline was normal; 2.0 - 2.5 x baseline if baseline was abnormal	>2.5 - 5.0 x ULN if baseline was normal; >2.5 - 5.0 x baseline if baseline was abnormal	>5.0 - 20.0 x ULN if baseline was normal; >5.0 - 20.0 x baseline if baseline was abnormal	>20.0 x ULN if baseline was normal; >20.0 x baseline if baseline was abnormal	-
Haptoglobin decreased	<LLN	-	-	-	-
Hemoglobin increased	Increase in >0 - 2 g/dL	Increase in >2 - 4 g/dL	Increase in >4 g/dL	-	-
INR increased	>1.2 - 1.5; >1 - 1.5 x baseline if on anticoagulation; monitoring only indicated	>1.5 - 2.5; >1.5 - 2.5 x baseline if on anticoagulation; dose adjustment indicated	>2.5; >2.5 x baseline if on anticoagulation; bleeding	-	-
Lipase increased	>ULN - 1.5 x ULN	>1.5 - 2.0 x ULN; >2.0 - 5.0 x ULN and asymptomatic	>2.0 - 5.0 x ULN with signs or symptoms; >5.0 x ULN and asymptomatic	>5.0 x ULN and with signs or symptoms	-
Lymphocyte count decreased	<LLN - 800/mm3; <LLN - 0.8 x 10e9/L	<800 - 500/mm3; <0.8 - 0.5 x 10e9 /L	<500 - 200/mm3; <0.5 - 0.2 x 10e9 /L	<200/mm3; <0.2 x 10e9 /L	-
Lymphocyte count increased	-	>4000/mm3 - 20,000/mm3	>20,000/mm3	-	-
Neutrophil count decreased	<LLN - 1500/mm3; <LLN - 1.5 x 10e9 /L	<1500 - 1000/mm3; <1.5 - 1.0 x 10e9 /L	<1000 - 500/mm3; <1.0 - 0.5 x 10e9 /L	<500/mm3; <0.5 x 10e9 /L	-
Platelet count decreased	<LLN - 75,000/mm3; <LLN - 75.0 x 10e9 /L	<75,000 - 50,000/mm3; <75.0 - 50.0 x 10e9 /L	<50,000 - 25,000/mm3; <50.0 - 25.0 x 10e9 /L	<25,000/mm3; <25.0 x 10e9 /L	-
Serum amylase increased	>ULN - 1.5 x ULN	>1.5 - 2.0 x ULN; >2.0 - 5.0 x ULN and asymptomatic	>2.0 - 5.0 x ULN with signs or symptoms; >5.0 x ULN and asymptomatic	>5.0 x ULN and with signs or symptoms	-
White blood cell decreased	<LLN - 3000/mm3; <LLN - 3.0 x 10e9 /L	<3000 - 2000/mm3; <3.0 - 2.0 x 10e9 /L	<2000 - 1000/mm3; <2.0 - 1.0 x 10e9 /L	<1000/mm3; <1.0 x 10e9 /L	-

CTCAE 5.0	CTCAE Grade				
CTCAE Term	**Grade 1**	**Grade 2**	**Grade 3**	**Grade 4**	**Grade 5**
Hypercalcemia	Corrected serum calcium of >ULN - 11.5 mg/dL; >ULN - 2.9 mmol/L; Ionized calcium >ULN - 1.5 mmol/L	Corrected serum calcium of >11.5 - 12.5 mg/dL; >2.9 - 3.1 mmol/L; Ionized calcium >1.5 - 1.6 mmol/L; symptomatic	Corrected serum calcium of >12.5 - 13.5 mg/dL; >3.1 - 3.4 mmol/L; Ionized calcium >1.6 - 1.8 mmol/L; hospitalization indicated	Corrected serum calcium of >13.5 mg/dL; >3.4 mmol/L; Ionized calcium >1.8 mmol/L; life-threatening consequences	Death
Hyperkalemia	>ULN - 5.5 mmol/L	>5.5 - 6.0 mmol/L; intervention initiated	>6.0 - 7.0 mmol/L; hospitalization indicated	>7.0 mmol/L; life-threatening consequences	Death
Hypermagnesemia	>ULN - 3.0 mg/dL; >ULN - 1.23 mmol/L	-	>3.0 - 8.0 mg/dL; >1.23 - 3.30 mmol/L	>8.0 mg/dL; >3.30 mmol/L; life-threatening consequences	Death
Hypernatremia	>ULN - 150 mmol/L	>150 - 155 mmol/L; intervention initiated	>155 - 160 mmol/L; hospitalization indicated	>160 mmol/L; life-threatening consequences	Death
Hypertriglyceridemia	150 mg/dL - 300 mg/dL; 1.71 mmol/L - 3.42 mmol/L	>300 mg/dL - 500 mg/dL; >3.42 mmol/L - 5.7 mmol/L	>500 mg/dL - 1000 mg/dL; >5.7 mmol/L - 11.4 mmol/L	>1000 mg/dL; >11.4 mmol/L; life-threatening consequences	Death
Hyperuricemia	>ULN without physiologic consequences	-	>ULN with physiologic consequences	Life-threatening consequences	Death
Hypoalbuminemia	<LLN - 3 g/dL; <LLN - 30 g/L	<3 - 2 g/dL; <30 - 20 g/L	<2 g/dL; <20 g/L	Life-threatening consequences; urgent intervention indicated	Death
Hypocalcemia	Corrected serum calcium of <LLN - 8.0 mg/dL; <LLN - 2.0 mmol/L; Ionized calcium <LLN - 1.0 mmol/L	Corrected serum calcium of <8.0 - 7.0 mg/dL; <2.0 - 1.75 mmol/L; Ionized calcium <1.0 - 0.9 mmol/L; symptomatic	Corrected serum calcium of <7.0 - 6.0 mg/dL; <1.75 - 1.5 mmol/L; Ionized calcium <0.9 - 0.8 mmol/L; hospitalization indicated	Corrected serum calcium of <6.0 mg/dL; <1.5 mmol/L; Ionized calcium <0.8 mmol/L; life-threatening consequences	Death

CTCAE 5.0	CTCAE Grade				
CTCAE Term	**Grade 1**	**Grade 2**	**Grade 3**	**Grade 4**	**Grade 5**
Hypoglycemia	<LLN - 55 mg/dL; <LLN - 3.0 mmol/L	<55 - 40 mg/dL; <3.0 - 2.2 mmol/L	<40 - 30 mg/dL; <2.2 - 1.7 mmol/L	<30 mg/dL; <1.7 mmol/L; life-threatening consequences; seizures	Death
Hypokalemia	<LLN - 3.0 mmol/L	Symptomatic with <LLN - 3.0 mmol/L; intervention indicated	<3.0 - 2.5 mmol/L; hospitalization indicated	<2.5 mmol/L; life-threatening consequences	Death
Hypomagnesemia	<LLN - 1.2 mg/dL; <LLN - 0.5 mmol/L	<1.2 - 0.9 mg/dL; <0.5 - 0.4 mmol/L	<0.9 - 0.7 mg/dL; <0.4 - 0.3 mmol/L	<0.7 mg/dL; <0.3 mmol/L; life-threatening consequences	Death
Hyponatremia	<LLN - 130 mmol/L	125-129 mmol/L and asymptomatic	125-129 mmol/L symptomatic; 120-124 mmol/L regardless of symptoms	<120 mmol/L; life-threatening consequences	Death

* Since anticoagulation medication is **NOT** captured in lab data, the condition for anticoagulation medication is ignored in the grade derivation.
a Note: Refer to Common Terminology Criteria for Adverse Events (CTCAE) Version 5.0, which can be found at https://evs.nci nih.gov/ftp1/CTCAE/CTCAE_5.0

Appendix 3. Pooled Cohort Risk and Score Calculation

No.	Parameters	Coefficient			
		Female		Male	
		White or Other	African America	White or Other	African America
1	ln Age (y)	–29.799	17.114	12.344	2.469
2	ln Age, Squared	4.884	N/A	N/A	N/A
3	ln Total Cholesterol (mg/dL)	13.540	0.940	11.853	0.302
4	ln Age × ln Total Cholesterol	–3.114	N/A	–2.664	N/A
5	ln HDL Cholesterol (mg/dL)	–13.578	–18.920	–7.990	–0.307
6	ln Age × ln HDL Cholesterol	3.149	4.475	1.769	N/A
7	ln Treated Systolic BP	2.019	29.291	1.797	1.916
8	ln Age × ln Treated Systolic BP	N/A	–6.432	N/A	N/A
9	ln Untreated Systolic BP	1.957	27.820	1.764	1.809
10	ln Age × ln Untreated Systolic BP	N/A	–6.087	N/A	N/A
11	Current Smoker (1=Yes, 0=No)	7.574	0.691	7.837	0.549
12	ln Age × Current Smoker	–1.665	N/A	–1.795	N/A
13	Diabetes (1=Yes, 0=No)	0.661	0.874	0.658	0.645
A	Baseline Survival	0.9665	0.9533	0.9144	0.8954
B	Individual Sum = ∑(coefficient×parameter)	For each individual patient, individual sum is summation of product of each parameter and its corresponding coefficient for that sex-race group for all parameters listed above (#1 - #13). Note: N/A indicates that the specific coefficient is not available thus the corresponding parameter will not be included in calculation.			
C	Population Mean	–29.18	86.61	61.18	19.54
D	**Pooled Cohort SCORE**	= B – C = Individual Sum – Population Mean			
E	**Pooled Cohort Risk** = Estimated probability of a first hard ASCVD* event within 10 years	**= 1 - A^exp(B-C)** = 1 - Baseline Survival ^ exp (Individual Sum- Population Mean)			

* ASCVD = atherosclerotic cardiovascular disease; BP = blood pressure; Ln = natural logarithm.

- Age is the actual age when the laboratory parameters are collected for the visit. Smoker is collected at Screening/Baseline of the Randomized Phase.

- If a subject had diabetes at baseline, the diabetic status will be yes for all the postbaseline visits. If the subject does not have diabetes at baseline, the diabetic status will be determined based on the AE start date. If the AE start date is on or prior to the laboratory date of a specific visit, then the diabetic status will be yes for that visit and later visits. If the day of the AE start date is missing, it will be imputed using the 1st day of the month.

- Treatment status of hypertensives will be determined by the usage of anti-hypertensive medications on the dates when blood pressures are collected. Anti-hypertensive medication with a start date prior to or on the date of blood pressures taken and a stop date on or after it. Systolic BP will be considered as treated Systolic BP.

- Total Cholesterol and HDL Cholesterol need to be measured from the same blood draw. Systolic BP record closest to the date when Total Cholesterol and HDL Cholesterol are measured will be used for the risk and score calculation.

- Round Pooled Cohort Risk Score to 0.001.

Appendix 4. Liver Function Prognostic Scores

- **Child-Pugh Score Calculation**

	1	2	3
HE	**None** No encephalopathy and not on any treatment for hepatic encephalopathy	**Medication-Controlled** Subject is lethargic, may have moderate confusion Subject is receiving medical therapy for HE	**Medication-Refractory** Marked confusion/incoherent, rousable but sleeping or comatose
Ascites	**None** No ascites and not on treatment for ascites	**Mild/Moderate** Cross sectional imaging showing ascites Abdominal distension Medication for ascites	**Severe (diuretic-refractory)** Visible clinically
Total Bilirubin (mg/dL)	< 2	2-3	> 3
Albumin (g/dL)	> 3.5	2.8-3.5	< 2.8
INR	< 1.7	1.7-2.3	> 2.3

There are 5 components for CP score as seen in the first column. Each will be assigned a subscore ranging between 1 and 3 based on the matching condition as described in the corresponding row and the value in the header row of the corresponding column. For example, when total bilirubin value is 2.5 mg/dL, the total bilirubin subscore will be 2.

CP score is obtained by adding the subscore for each parameter, ranging between 5 and 15.

For Screening, HE and ascites closest to laboratory collection date will be used; for Day 1 and postbaseline visits, HE and ascites will be from the same day as laboratory data collected. If baseline CP score is missing due to unevaluable laboratory data, baseline CP score will be calculated using the last laboratory values collected prior to (OL) Day 1.

- **MELD Score Calculation**

MELD score = 3.78 [Ln total bilirubin (mg/dL)] + 11.2 [Ln INR] + 9.57 [Ln serum creatinine (mg/dL)] + 6.43. If the serum creatinine, the total bilirubin or the INR value is < 1.00 mg/dL, the calculation will use 1.00 as the test value. If the serum creatinine is > 4.00 mg/dL or subjects on dialysis, the calculation will use 4.00 as the serum creatinine value.

Appendix 5. Health Related QoL Score Calculation

- **SF-36**

Scoring of the SF-36 scales will be performed as described in Chapter 6 of the SF-36 Health Survey Manual and Interpretation Guide, Version 2. Summary will be done for 8 domains of the SF-36 (physical functioning, role-physical, bodily pain, general health, vitality, social functioning, role-emotional, and mental health), and for the physical component score and mental component summary.

- **CLDQ-NAFLD**

CLDQ-NAFLD scores are calculated using subject responses to 36 questions in the questionnaire. If Ri is the score for the patient's response to the item i, for i=1, 2, …., 36 then the 6 domain scores are calculated as follows:

- Abdominal = Mean of {R1, R5, R17}
- Fatigue = Mean of {R2, R4, R8, R11, R13, R35}
- Systemic = Mean of {R3, R6, R21, R23, R27, R36}
- Activity = Mean of {R7, R9, R14, R30, R31}
- Emotion = Mean of {R10, R12, R15, R16, R19, R20, R24, R26, R34}
- Worry = Mean of {R18, R22, R25, R28, R29, R32, R33}

Here "Mean" is the average of nonmissing items (SAS mean function). Each score is calculated only if at least half of corresponding items are not missing. Otherwise, the score will be missing.

Overall CLDQ score is calculated by taking the mean of 6 domain scores {abdominal, fatigue, systemic, activity, emotion, worry}. Overall CLDQ score will be summarized.

- **WPAI: NASH**

The response to Question 1 of this questionnaire provides the binary endpoint whether or not the subject had been in a paid employment during the week prior to assessment.

If the subject had been in a paid employment (Response to Q1 is "Yes") at the visit when questionnaire was given, then following three scores are derived:

- Percent work time missed = 100 × Q2 / (Q2 + Q4)
- Percent impairment while working = 100 × Q5 / 10

- Percent overall work impairment =

$$100 \times \left[\frac{Q2}{(Q2+Q4)} + \left(1 - \frac{Q2}{Q2+Q4)} \right) \times \frac{Q5}{10} \right]$$

Question 6 is applicable to all subjects:

— Percent activity impairment = 100 × Q6 / 10.

— Percent overall work impairment and percent activity impairment will be summarized.

- **EQ-5D**

Scoring of EQ-5D will be performed as described in EQ-5D-5L user guide. Summary will be done for 5 dimensions in descriptive system (mobility, self-care, usual activities, pain/discomfort and anxiety/depression) and the EQ Visual Analogue Scale.

- **HRUQ**

The number of visits/days for the following questions will be summarized.

— Outpatient non-emergent clinic visits (number of visits)

— Outpatient non-emergent clinic visits related to disease (number of visits)

— Outpatient emergent department visits (number of visits)

— Outpatient emergent department visits related to disease (number of visits)

— Outpatient surgeries and procedures (number of visits)

— Outpatient surgeries and procedures related to disease (number of visits)

— Inpatient surgeries and procedures (number of visits)

— Inpatient surgeries and procedures related to disease (number of visits)

— Number of days hospitalized (number of days)

— Number of days hospitalized related to disease (number of days)

Appendix 6. NAFLD Activity Score (NAS) and Steatosis, Activity, and Fibrosis (SAF) Score Calculation

NAS score is calculated as the sum of steatosis, lobular inflammation and hepatocellular ballooning. SAF score is calculated as the sum of steatosis (same as NAS calculation), activity and fibrosis (according to the NASH CRN classification) subscores. Activity subscore is defined as sum of lobular inflammation and hepatocellular ballooning, where both grade 2 and 3 of lobular inflammation will be treated as 2.

Appendix 7. Noninvasive markers for Fibrosis

- **Enhanced Liver Fibrosis (ELF)**

ELF test score = 2.278 + 0.851× ln (hyaluronic acid) + 0.751× ln (PIIINP) + 0.394 × ln (TIMP1);

Note: All ELF test score components (hyaluronic acid, PIIINP and TIMP1) need to be measured from the same blood draw. ELF test score will only be calculated when ELF test score is missing and components are less than the LOQ or above the upper LOQ. Individual components are to be imputed per data handling conventions in Section 3.7, and ELF test score will be calculated based on the imputed values of components.

- **Fibrosis-4 (FIB-4) Index Calculation:**

FIB-4 Index = round ((age × AST) / (platelet × sqrt (ALT)), 0.01);

- **AST to Platelet Ratio Index (APRI) Calculation:**

APRI = round (AST/ASTULN × 100 / platelet, 0.1);

Note: for FIB-4 index and APRI calculation, the laboratory parameters need to be measured from the same blood draw. Age should be the actual age at the date when laboratory values are taken.

- **NAFLD fibrosis score (NFS) Calculation:**

NFS = -1.675 + 0.037 × age (years) + 0.094 × BMI (kg/m^2) + 1.13 × IFG (impaired fasting glucose) / pre-diabetes or diabetes (yes = 1,no = 0) + 0.99 × AST / ALT ratio - 0.013 × platelet ($\times 10^9/L$) - 0.66 × albumin (g/dL). Keep 3 decimal places.

The laboratory parameters need to be measured from the same blood draw. The last non-missing BMI on or prior to laboratory date should be used. Age should be the actual age at the date when laboratory values are taken. Status of pre-diabetes/diabetes should also be decided on the laboratory date. If a subject had pre-diabetes/diabetes at baseline, the pre-diabetic/diabetic status will be yes for all the postbaseline visits. If the subject does not have pre-diabetic/diabetes at baseline, the diabetic status will be determined based on the start date AEs of pre-diabetes and diabetes, and the collection date when fasting glucose is greater than 100 (IFG). If the AE start date or the fasting glucose collection date is on or prior to the laboratory date of a specific visit, then the diabetic/pre-diabetic status will be yes for that visit and later visits. If the day of the AE start date is missing, it will be imputed using the 1st day of the month.

Appendix 8. SAS Program for Tipping Point Analysis for Binary Endpoint

The following **%tp_binary** macro generates multiple imputed data and a set of the shift parameters that adjust the imputed values will be examined.

```
/*-------------------------------------------------------------------------------------------------------------------*/
/*--- Delta-Adjusting Method for Tipping Point Analysis for Binary Endpoint                                       ---*/
/*--- Generate imputed data set for specified shift parameters                                                    ---*/
/*--- data= input data set                                                                                        ---*/
/*--- smin= min shift parameter for active drug                                                                   ---*/
/*--- smax= max shift parameter for active drug                                                                   ---*/
/*--- sinc= increment of the shift parameter for active drug                                                      ---*/
/*--- trt= treatment group indicator                                                                              ---*/
/*--- out= output imputed data set                                                                                ---*/
/*-------------------------------------------------------------------------------------------------------------------*/

%macro tp_binary( data=, smin=, smax=, sinc=, trt=, out=);
          data &out;
            set _null_;
          run;

          /*------------ # of shift values ------------*/
          %let ncase= %sysevalf( (&smax-&smin)/&sinc, ceil );

          /*------- Imputed data for each shift -------*/
          %do jc=0 %to &ncase;

          %let sj= %sysevalf( &smin + &jc * &sinc);

                    proc mi data=&data seed=20190000 nimpute=50 out=outmi;
                              var trt01pn strat1V strat2V pe48;
                              class trt01pn strat1V strat2V pe48;
                              monotone logistic (pe48 / link=logit);
                              mnar adjust( pe48(event='1') / adjustobs=(trt01pn="&trt") shift= &sj);
                    run;

          data outmi;
            set outmi;
          Shift_Trt= &sj;
          run;

          data &out;
            set &out outmi;
          run;

          %end;
%end;
%mend tp_binary;
```

SAP GS-US-384-1943

ELECTRONIC SIGNATURES

Signed by	Meaning of Signature	Server Date (dd-MMM-yyyy hh:mm:ss)
PPD	Clinical eSigned	23-Jan-2019 21:31:31
PPD	Biostatistics eSigned	25-Jan-2019 02:05:16