NOVARTIS

Clinical Development

LJC242

CLJC242A2201J / NCT03517540

A randomized, double-blind, multicenter study to assess the safety, tolerability, and efficacy of a combination treatment of tropifexor (LJN452) and cenicriviroc (CVC) in adult patients with nonalcoholic steatohepatitis (NASH) and liver fibrosis (TANDEM)

Statistical Analysis Plan (SAP)

Author: [redacted]

Document type: SAP Documentation

Document status: Final 2.0

Release date: November 11, 2020

Number of pages: 33

Document History – Changes compared to previous final version of SAP

Date	Time point	Reason for update	Outcome for update	Section and title impacted (Current)
20 August 2018	Prior to DB lock	Creation of final version	N/A – first version	N/A – first version
10 July 2020	Prior to DB lock	Creation of Amendment 1	Added several supportive analyses for the secondary efficacy for the impact of COVID-19. Added AE summary by the onset time of COVID-19 pandemic in different countries and regions.	Section 2.7.3, 2.7.4 Section 2.8.1
			Added treatment extension and corresponding actions taken for analyses.	Section 1.1, 2.4.1
11 November 2020	Prior to DB lock	Creation of Amendment 2	Added several supportive analyses for the secondary efficacy.	Section 2.7
			Added analyses based on paired biopsy for secondary efficacy endpoints [redacted]	Section 2.7, 2.13

Table of contents

Table of contents 3
List of abbreviations 5
1 Introduction 7
1.1 Study design 7
1.2 Study objectives and endpoints 7
2 Statistical methods 10
2.1 Data analysis general information 10
2.2 Analysis sets 11
2.3 Patient disposition, demographics and other baseline characteristics 12
2.4 Treatments (study treatment, rescue medication, concomitant therapies, compliance) 12
2.4.1 Study treatment / compliance 12
2.4.2 Prior, concomitant, and post therapies 13
2.5 Analysis of the primary objective 13
2.5.1 Primary endpoint 13
2.5.2 Statistical hypothesis, model, and method of analysis 14
2.5.3 Handling of missing values/censoring/discontinuations 14
2.5.4 Supportive analyses 14
2.6 Analysis of the key secondary objective 14
2.7 Analysis of secondary efficacy objective(s) 15
2.7.1 Secondary endpoints 15
2.7.2 Statistical hypothesis, model, and method of analysis 15
2.7.3 Handling of missing values/censoring/discontinuations 16
2.7.4 Supportive Analyses 16
2.8 Safety analyses 17
2.8.1 Adverse events (AEs) 17
2.8.2 Deaths 18
2.8.3 Laboratory data 19
2.8.4 Other safety data 19
[redacted] 19
[redacted] 20
[redacted] 20
[redacted] 20
[redacted] 20
2.14 Interim analysis 22
3 Sample size calculation 22

3.1 Power considerations with given sample size for safety assessment 22
3.2 Power considerations with given sample size for efficacy assessment 24
4 Change to protocol specified analyses 26
5 Appendix 26
5.1 Imputation rules 26
5.1.1 Study drug 26
5.1.2 AE date imputation 26
5.1.3 Concomitant medication date imputation 27
5.1.4 Other imputations 29
5.2 AEs coding/grading 29
5.3 Laboratory parameters derivations 29
5.3.1 Laboratory parameters 29
[redacted] 30
[redacted] 30
[redacted] 30
5.3.5 Derived baseline characteristics 31
[redacted] 31
5.4 Statistical models 31
5.4.1 Primary analysis 31
5.4.2 Key secondary analysis 31
5.5 Not applicable. Rule of exclusion criteria of analysis sets 31
5.6 Other statistical aspects 32
6 Reference 33

List of abbreviations

AE	Adverse event
ALP	Alkaline phosphatase
ALT	Alanine transaminase
ANCOVA	Analysis of covariance
[redacted]	[redacted]
APTT	Activated partial thromboplastin time
AST	Aspartate transaminase
ATC	Anatomical Therapeutic Chemical classification
[redacted]	[redacted]
[redacted]	[redacted]
CI	Confidence interval
CRN	Clinical Research Network
CSR	Clinical Study report
CV	Coefficient of variation
CVC	Cenicriviroc
DMC	Data Monitoring Committee
eGFR	Estimated glomerular filtration rate
[redacted]	[redacted]
EOS	End of study
EOT	End of treatment
FAS	Full analysis Set
[redacted]	[redacted]
GGT	Gamma glutamyl transferase
[redacted]	[redacted]
[redacted]	[redacted]
INR	International Normalized Ratio
ITT	Intent-to-treat
[redacted]	[redacted]
MedDRA	Medical Dictionary for Regulatory Activities
[redacted]	[redacted]
[redacted]	[redacted]
NA	Not applicable
NAFLD	Non-alcoholic fatty liver disease
[redacted]	[redacted]
NASH	Non-alcoholic steatohepatitis
OCA	Obeticholic acid
[redacted]	[redacted]
[redacted]	[redacted]
[redacted]	[redacted]
[redacted]	[redacted]
[redacted]	[redacted]
PRO	Patient-reported Outcomes

PT	Preferred term / Prothrombin time
RAN	Randomized set
SAE	Serious adverse event
SAF	Safety set
SAP	Statistical Analysis Plan
SCR	Screened set
SD	Standard deviation
SMQ	Standardized MedDRA Query
SOC	System Organ Class
TFLS	Tables, Figures and Listings Shells
ULN	Upper limit of normal range
[redacted]	[redacted]
WHO	World Health Organization

1 Introduction

The purpose of Statistical Analysis Plan (SAP) is to describe the implementation of the statistical analysis planned in the protocol. It addresses the analyses intended for the final clinical study report.

Reference documents: Clinical Trial Protocol version 03, 7 May 2020.

1.1 Study design

This study is a 48-week, randomized, double-blind, multicenter study that consists of a screening period, a treatment period starting from randomization on Day 1 and running to Week 48, and a follow up period of 4 weeks after the last dose of study treatment. The total study duration is up to 60 weeks.

The study population are adult male and female patients with histologic evidence of NASH and fibrosis (stage 2 or 3 as per NASH CRN histological score, F2/F3); see Inclusion and Exclusion criteria in the protocol for details.

At baseline, approximately 200 patients whose eligibility is confirmed will be randomized in a 1:1:1:1 ratio, resulting in approximately 50 patients randomized in each of the four arms:

Arm A: tropifexor 140 µg, once-daily

Arm B: CVC 150 mg, once-daily

Arm C: tropifexor 140 µg + CVC 150 mg, once-daily

Arm D: tropifexor 90 µg + CVC 150 mg, once-daily.

[redacted]

The primary analysis will be performed when all randomized patients have completed or discontinued the trial participation. No interim analyses are planned.

Due to the COVID-19 pandemic, it may not be possible for patients to return in time for the Week-48 visit to perform scheduled assessments including biopsies, and according to drug dispensing plan in Section 5.5.2 of the protocol, the Week-48 visit and associated assessments can be delayed up to approximately 12 weeks. Since the primary objective of the study is to determine if there is a safe combination between one of the chosen doses of tropifexor and 150 mg of cenicriviroc, all assessment results obtained, including those from eligible patients after Week 48, will be included in analyses as they are with no special plan to account for the treatment extension.

1.2 Study objectives and endpoints

Table 1-1 Objectives and related endpoints

Objective(s)	Endpoint(s)
Primary Objective(s)	**Endpoint(s) for primary objective(s)**

Objective(s)	Endpoint(s)
To evaluate the safety and tolerability of tropifexor + CVC in patients with NASH and fibrosis (stage 2 or 3 as per NASH CRN histological score, F2/F3) by monitoring adverse events, vital signs and laboratory values during 48 weeks of treatment as compared to monotherapy with each of tropifexor and CVC	Occurrence of adverse events, serious adverse events, adverse events resulting in discontinuation of study treatment, adverse events of special interest and changes in vital signs and laboratory values over 48 weeks of treatment
Secondary Objective(s)	**Endpoint(s) for secondary objective(s)**
To characterize the efficacy of tropifexor + CVC in patients with NASH with fibrosis stage F2/F3 as assessed by histological improvement after 48 weeks of treatment compared to monotherapies (tropifexor and CVC) compared to baseline biopsy	Proportion of patients who have at least a one-point improvement in fibrosis Proportion of patients with resolution of steatohepatitis

Objective(s)	Endpoint(s)
[redacted]	

Objective(s)	Endpoint(s)
[redacted]	

2 Statistical methods

2.1 Data analysis general information

Analyses for the CSR will be performed by Novartis using SAS 9.4 or newer. Other statistical software and programming languages, such as R, may be used as well, in the newest version available in the validated environment. Data exploration will be aided by specific tools focusing on interactive visualization, e.g. an R Shiny App. [redacted] Unless mentioned otherwise in the respective section of this SAP and the TFLS document, formal tables, figures and listings will therefore be specified ad hoc for these.

Summary tables will be presented by treatment group and analysis visit (as applicable) using descriptive statistics. These include absolute and relative frequencies for categorical variables. Continuous variables will be summarized by arithmetic mean, standard deviation, minimum, maximum, median and first and third quartile per default. Where indicated, geometric mean and coefficient of variation will also be displayed, and the ratio instead of percentage change.

The data analysis for the CSR will be performed after the clinical database lock, when all patients have completed or discontinued the study, and will therefore include all collected data.

Data cutoffs for safety DMC analyses will occur approximately every 6 months after start of randomization.

Centers, which are expected to be small in this study, will be pooled for all analyses. No analyses by center and no stratification by center are planned upfront. In case of some unexpectedly large centers (>15 randomized subjects), effects in these centers will be explored ad hoc (e.g., using a ShinyApp). The randomization stratification factor [redacted] will not be used as a covariate in statistical analyses, because it represents an optional assessment and not a subject baseline characteristic.

General definitions

The term "Study treatment" refers to the treatment assigned to or received by the study subject, which is either tropifexor 140 µg (once-daily), CVC 150 mg (once-daily), tropifexor 140 µg + CVC 150 mg (once-daily), or tropifexor 90 µg + CVC 150 mg (once-daily).

The term "Study drug" refers to tropifexor, CVC, or placebo.

Date of first administration of study treatment: Date of first administration as recorded in the CRF (Drug Administration Record page)

Date of last administration of study treatment: Date of last administration as recorded in the CRF (Drug Administration Record page)

Study day: Study day is calculated from the date of first administration of study treatment, which is defined as Day 1.

Baseline: Generally, baseline is defined as the last assessment before date and time of first administration of study treatment; if only the date is available, the last assessment before or at the date of first administration of study drug will be used. For transaminases (ALT, AST, and GGT) and bilirubin, the baseline value will be calculated as the mean of the last two assessments before first administration of the study drug, which are usually those taken at the Screening 1 and Baseline visits (Screening 2 and Baseline if a test was performed during Screening 2 visit).

Last contact: Date of last data point entered in the database.

Treatment period: Period from Baseline visit to End-of-treatment visit (included).

Follow-up period: Period from End-of-treatment visit (not included) to End-of-Study visit (included).

2.2 Analysis sets

- Screened set (SCR) – All patients who signed the informed consent.
- Randomized set (RAN) – All patients who received a randomization number, regardless of receiving trial medication.
- Full analysis set (FAS) – All patients to whom study treatment has been assigned*. Following the intent-to-treat (ITT) principle, patients are analyzed according to the treatment they have been assigned to at randomization.
- Safety set (SAF) - All patients who received at least one dose of study treatment and have at least one post-baseline safety assessment. Of note, the statement that a patient had no adverse events also constitutes a safety assessment. Patients will be analyzed according to the treatment received.

[redacted]

[redacted]

** Excluding patients who were mis-randomized and did not take any study drug. Mis-randomized patients are those who were not qualified for randomization, but were inadvertently randomized into the study.*

The number of patients in each analysis set will be presented by treatment group and overall for the screened set.

Subgroup of interest

No subgroups of particular interest are defined in the protocol. Due to the relatively small sample size per group, subgroup analyses will only be performed within the data exploration framework.

2.3 Patient disposition, demographics and other baseline characteristics

Demographic variables and other baseline characteristics will be summarized for the FAS. Descriptive statistics (mean, median, standard deviation, minimum, and maximum) will be presented for continuous variables for each treatment group and for all patients (total). The number and percentage of patients in each category will be presented for categorical variables for each treatment group and all patients (total). In addition, all relevant medical history and protocol solicited medical history will be summarized by treatment group.

Patient disposition

The number of subjects in each analysis set will be presented overall for the screened set and by treatment group for the randomized set.

The number and percentage of subjects in the randomized set who completed or discontinued each study period (screening, treatment, and follow-up), and the reason for discontinuation will be presented for each treatment group and all subjects.

The frequency (%) of subjects with CSR reportable protocol deviations as well as the criteria leading to exclusion from analysis sets will be presented in separate tables for the randomized set.

2.4 Treatments (study treatment, rescue medication, concomitant therapies, compliance)

2.4.1 Study treatment / compliance

Overall duration (in weeks) for the double-blind investigational treatment will be computed for each randomized patient as follows:

(date of last administration of study treatment – date of first administration of study treatment + 1) / 7

If a patient was randomized, but did not receive any dose of randomized double-blind study treatment (protocol deviation) then the exposure is set to 0.

The duration of exposure to study treatment (in weeks) will be summarized for the SAF using descriptive statistics and additionally by duration category in steps of "week" (+/- one-day time window included):

- ≥ 0 week- < 12 weeks
- ≥ 12 week- < 24 weeks
- ≥ 24 weeks - < 36 weeks
- ≥ 36 weeks - < 48 weeks
- ≥ 48 weeks - < 52 weeks
- ≥ 52 weeks - < 56 weeks
- ≥ 56 weeks

The cutoff of 52 weeks and 56 weeks is to account for the extended treatment for patients who are unable to perform the Week-48 visit as scheduled.

2.4.2 Prior, concomitant, and post therapies

Medications will be identified using the WHO dictionary including ATC code and presented for the SAF. Prior medications are defined as any medications taken prior to the first administration of study drug (regardless of whether they are stopped or continued after randomization). Concomitant medications and significant non-drug therapies are defined as those used during the double-blind period. Prior and concomitant medications will be summarized by treatment group in separate tables. Medications will be presented in alphabetical order, by ATC codes and grouped by anatomical main group (the 1st level of the ATC code, e.g. "Cardiovascular system", C), and by the first subgroup, e.g. "Lipid modifying agents" (C10). Tables will also show the overall number and percentage of subjects receiving at least one drug of a particular ATC code and at least one drug in a particular anatomical main group and first subgroup.

Concomitant medications that were prohibited as per protocol and given between start and end of treatment period as well as significant non-drug therapies, and medications that are permitted only if the dose is stable within 25 percent of the baseline dose (Table 5-4 in the protocol) will be provided in separate tables.

2.5 Analysis of the primary objective

2.5.1 Primary endpoint

The primary study objective is to evaluate the safety and tolerability of tropifexor + CVC in patients with NASH and fibrosis (stage 2 or 3 as per NASH CRN histological score, F2/F3) by monitoring adverse events, vital signs and laboratory values during treatment as compared to monotherapy with each of tropifexor and CVC. The related endpoints are (over treatment):

- occurrence of adverse events,
- occurrence of serious adverse events,
- occurrence of adverse events resulting in discontinuation of study treatment,

- occurrence of adverse events of special interest,
- changes in vital signs and laboratory values.

Only treatment emergent adverse events will be considered for the analysis, which are defined as those starting on or after first day of application of study treatment and before EOS visit or 30 days after EOT visit, whichever comes earlier.

2.5.2 Statistical hypothesis, model, and method of analysis

There are no pre-specified hypotheses and statistical models in this study. The methods to analyze the primary safety variables are outlined in Table 2-1.

Table 2-1 Primary variables and methods of analysis

Variable	Method of analysis
Occurrence of adverse events	Summary table of absolute and relative frequency, overall and by preferred term
Occurrence of serious adverse events	Summary table of absolute and relative frequency, overall and by preferred term
Occurrence of adverse events resulting in discontinuation or dose reduction of study treatment	Summary table of absolute and relative frequency, overall and by preferred term
Occurrence of adverse events of special interest	Summary table of absolute and relative frequency, by type of AE as the risks are described for both tropifexor and CVC (risk definition for tropifexor are those that appear in the tropifexor IB or CVC IB, DSPP and eCRS) together with any other events identified during the course of this or other studies
Changes in vital signs	Descriptive statistics by visit
Changes in laboratory data	Descriptive statistics by visit

The statistical analysis of these variables is described in Section 2.8, along with the analysis of other safety variables.

2.5.3 Handling of missing values/censoring/discontinuations

Imputation of incomplete adverse events start and end dates follow standard conventions and are described in detail in Section 5.1.2.

Missing follow-up assessments for laboratory and vital signs will not be imputed.

2.5.4 Supportive analyses

Supportive analyses are added for evaluation of the impact of COVID-19 on study results. See Section 2.8.1.

2.6 Analysis of the key secondary objective

There is no key secondary objective defined in the protocol.

2.7 Analysis of secondary efficacy objective(s)

2.7.1 Secondary endpoints

There is a single secondary objective in this study which is to characterize the efficacy of tropifexor + CVC in patients with NASH with fibrosis stage F2/F3 as assessed by histological improvement after 48 weeks of treatment compared to monotherapies (tropifexor and CVC) relative to baseline biopsy. There are two endpoints that will be used to evaluate this objective:

- Proportion of patients who have at least a one-point improvement in fibrosis
- Proportion of patients with resolution of steatohepatitis

The determination of fibrosis improvement will be based on NASH CRN staging (see Table 2-2). Only main stages (0, 1, 2, 3, and 4) will be considered. For example, a change from 1c to 1b or 1a will not be counted as a one-point change.

Table 2-2 Fibrosis score according to NASH CRN

Fibrosis score	Histopathology
0	none
1a	mild perisinusoidal fibrosis
1b	moderate perisinusoidal fibrosis
1c	portal/periportal fibrosis only
2	perisinusoidal and periportal fibrosis
3	bridging fibrosis
4	cirrhosis

Results of the paired biopsy readings will be used for all biopsy-based outputs, yet not retrospectively for re-evaluation of patient eligibility at screening, which is performed based on real-time biopsy reading.

Resolution of steatohepatitis will be determined as diagnostic category “not NAFLD” or “NAFLD, not NASH” as provided in the central biopsy report.

2.7.2 Statistical hypothesis, model, and method of analysis

The two secondary endpoints are incorporated in two estimands.

The first estimand is defined as below:

- Treatment: The randomized treatment (tropifexor 140 µg once-daily, CVC 150 mg once-daily, tropifexor 140 µg + CVC 150 mg once-daily, and tropifexor 90 µg + CVC 150 mg once-daily) taken for 48 weeks, further defined in Section 5.2 in the protocol.
- Population: Full analysis set.
- Endpoint: Proportion of patients who have at least a one-point improvement in fibrosis.
- Summary measure: Difference in the proportion of patients on the different tropifexor + CVC regimens who achieve at least a one-point improvement in fibrosis at Week 48 biopsy compared to tropifexor and CVC monotherapy patients.

The second estimand is the same as the first one except

- Endpoint: Proportion of patients with resolution of steatohepatitis

- Summary measure: Difference in the proportion of patients on the different tropifexor + CVC regimens who achieve resolution of steatohepatitis at Week 48 relative to baseline compared to tropifexor and CVC monotherapy patients.

Both estimands will be evaluated using a Cochran-Mantel-Haenszel test controlling for baseline fibrosis stage (F2/F3). Risk differences and odds ratios with 95% confidence intervals will provide supporting information for both estimands.

2.7.3 Handling of missing values/censoring/discontinuations

Missing baseline data (if applicable) for the secondary endpoints will not be imputed.

Missing follow-up data for the secondary endpoints will be handled as follows:

If a patient stayed on randomized study treatment for less than 24 weeks, missing data will not be imputed, and if data are available (which is not expected as per protocol) they will not be included in the analysis.

If a patient stayed on randomized study treatment for at least 24 weeks, but follow-up data are not available, data will be imputed by multiple inputation (see Section 5.1.4).

If a patient stayed on randomized study treatment for at least 24 weeks, and follow-up data are available, data will be used as observed, even if treatment was discontinued prior to Week 48.

Due to the COVID-19 pandemic, it was not possible for some patients to return in time for the Week-48 visit to perform a liver biopsy. Instead, it was possible for additional treatment to be dispensed (see drug dispensing plan in Section 5.5.2 in the protocol), and the Week-48 liver biopsy could be delayed up to approximately 12 weeks. Since the purpose of the secondary efficacy analysis is to evaluate the potential histological benefit of study treatment, all biopsies obtained, including those from eligible patients after Week 48, will be included in the analysis as Week-48 biopsies.

2.7.4 Supportive Analyses

Since that some Week-48 biopsies may not be obtained as originally scheduled, the same endpoints as in Section 2.7.2 will be re-evaluated. Summary tables will be provided for corresponding summary measures by treatment group and by whether the patients received extended treatment.

In addition, the same endpoints as in Section 2.7.2 will be re-evaluated for patients with at least 10% weight loss at Week 48 compared to Baseline. Summary tables will be provided for corresponding summary measures by treatment group and by population with or without at least 10% weight loss at Week 48 compared to Baseline.

Analyses of proportion of patients who have at least one-point improvement in fibrosis described in Section 2.7.2 and corresponding supportive analyses described in Section 2.7.4 will be repeated based on modified ISHAK scoring system.

Analyses of both secondary efficacy endpoints described in Section 2.7.2 and corresponding supportive analyses described in Section 2.7.4 will be repeated based on real-time biopsy reading.

Qualitative results of comparison between biopsies obtained during Screening and Week 48 based on paired biopsy readings will be summarized by treatment group into three categories (better, same, [redacted]) for following parameters: steatosis, [redacted] NASH CRN fibrosis, and modified ISHAK score.

2.8 Safety analyses

The analysis of safety parameters will be performed in the Safety Analysis Set.

2.8.1 Adverse events (AEs)

Treatment emergent adverse events (events started after the first dose of study treatment or events present prior to the first dose of study treatment but increased in severity based on preferred term, until EOS visit or 30 days after EOT visit, whichever comes earlier) will be summarized. AEs will be summarized by presenting, for each treatment group, the number and percentage of patients having experienced

- Any adverse event (AE),
- Any serious adverse event (SAE),
- Any adverse event by primary system organ class (SOC),
- Any adverse event by preferred term (PT),
- Any adverse event by severity,
- Any adverse event possibly related to study treatment (investigator assessment),
- Any adverse event resulting in discontinuation of study treatment,
- Any adverse events of special interest for tropifexor or CVC treatment,
- Any adverse events by onset time before and after the start date of COVID-19 pandemic by country/region,
- Any serious adverse events by onset time before and after the start date of COVID-19 pandemic by country/region.

If a patient reported more than one adverse event with the same preferred term, the adverse event with the greatest severity will be presented. If a patient reported more than one adverse event within the same primary system organ class, the patient will be counted only once with the greatest severity at the system organ class level, where applicable.

The start date of COVID-19 for different countries or regions is defined in Table 5-1 of "COVID-19 (Coronavirus) Guidance - Start and End Dates by Region for Sensitivity Analyses" Version 1.0 released on May 19, 2020. As of the time of finalization of this SAP, the start (and end) dates are listed below:

Table 2-3 Defined Dates

Region/Country	Start Date	End Date
Italy	23-Feb-2020	End date has not yet been defined
Rest of the World	01-Mar-2020	End date has not yet been defined

Tabulations of adverse events, SAE, and adverse events of special interest broken down by time of onset according to the following intervals will be provided as well:

- Starting within 4 weeks after first administration of study treatment,
- Starting between 4 weeks and 12 weeks after first administration of study treatment,
- Starting more than 12 weeks after first administration of study treatment.

For the legal requirements of ClinicalTrials.gov and EudraCT, two required tables on treatment emergent adverse events which are not serious adverse events with an incidence greater than or equal to 5% in any treatment group and on treatment emergent serious adverse events and SAE suspected to be related to study treatment will be provided by system organ class and preferred term on the safety set population.

If for a same patient, several consecutive AEs (irrespective of study treatment causality, seriousness and severity) occurred with the same SOC and PT:

- *a single occurrence will be counted if there is ≤ 1 day gap between the end date of the preceding AE and the start date of the consecutive AE*
- *more than one occurrence will be counted if there is > 1 day gap between the end date of the preceding AE and the start date of the consecutive AE*

For occurrence, the presence of at least one SAE / SAE suspected to be related to study treatment / non SAE has to be checked in a block e.g., among AE's in a ≤ 1 day gap block, if at least one SAE is occurring, then one occurrence is calculated for that SAE.

The number of deaths resulting from SAEs suspected to be related to study treatment and SAEs irrespective of study treatment relationship will be provided by SOC and PT.

Adverse events of special interest / grouping of AEs

Adverse events of special interest are the risks as described for both tropifexor and CVC (those that appear in tropifexor Development Safety Profiling Plan or CVC Investigator's Brochure, respectively). These will be summarized with absolute and relative frequency overall and by type of risk / AE.

Relative and absolute frequencies of patients with liver and renal events (see protocol Appendix 2 and 3, respectivley) will also be provided.

2.8.2 Deaths

A separate summary and listing will be provided for deaths, if they occur during the course of the study.

2.8.3 Laboratory data

The summary of safety laboratory evaluations will be presented for the groups of laboratory tests (e.g., hematology, clinical chemistry). Descriptive summary statistics for the change from baseline to each study visit will be presented. These descriptive summaries will be presented by test group, laboratory test, and treatment group. Change from baseline will only be summarized for patients with both baseline and post baseline values. Shift tables based on the normal laboratory ranges will be provided. For the shift tables, the normal laboratory ranges will be used to evaluate whether a particular laboratory test value was normal, low, or high for each visit value relative to whether or not the baseline value was normal, low, or high. These summaries will be presented by laboratory test category and treatment group.

The number and percentage of patients with new clinically notable laboratory results after baseline will be presented. Clinically notable laboratory results are listed in Section 5.3.1, for those parameters where ranges are available.

Safety laboratory parameters which are at the same time part of the efficacy analyses will be included in the safety tables as well.

2.8.4 Other safety data

2.8.4.1 ECG and cardiac imaging data

ECG data will be summarized by treatment and visit (post-baseline ECGs are performed at Week 24 and Week 48).

The Fridericia QT correction formula (QTcF) will be used for clinical decisions and for analyses.

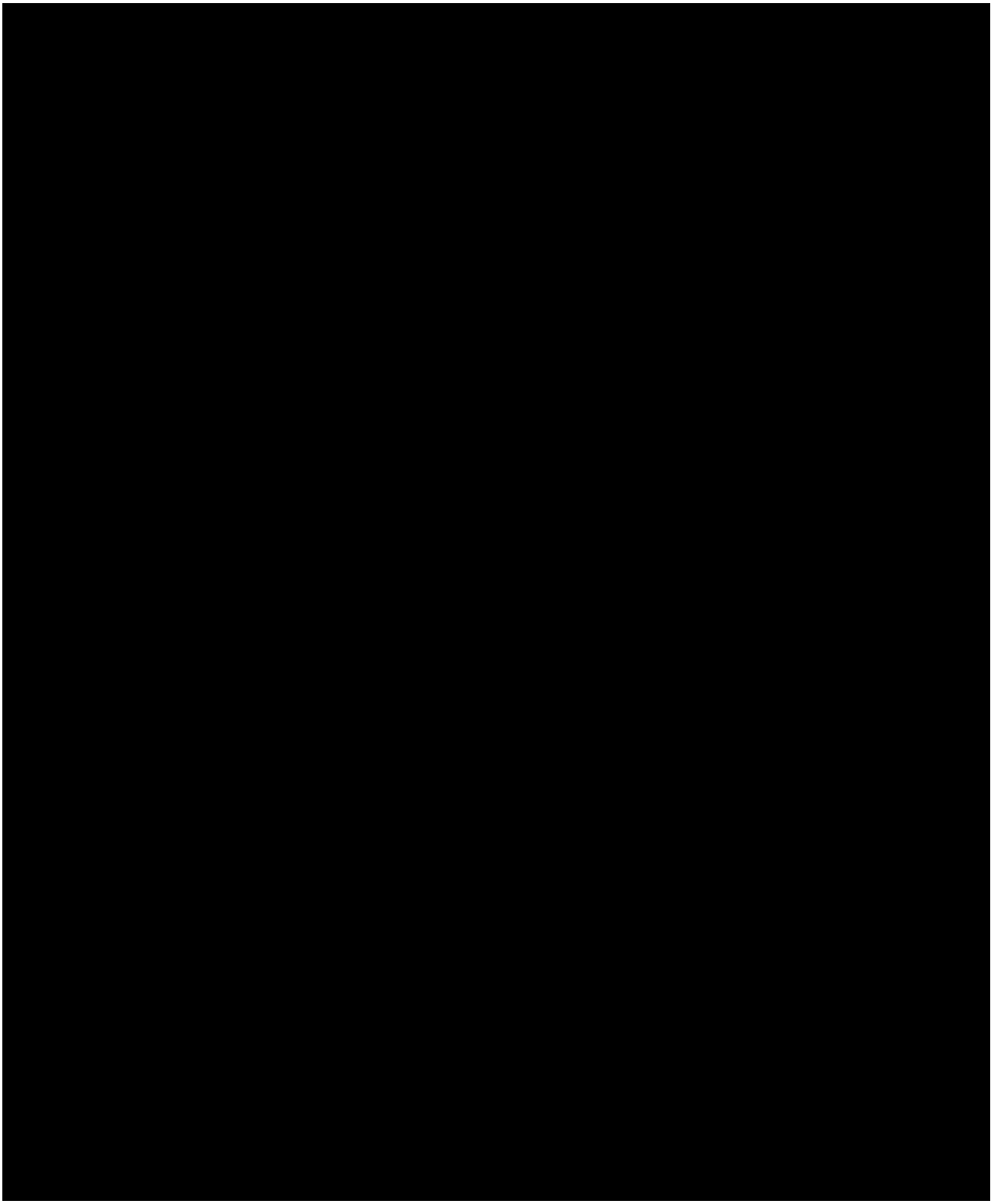
Notable QTcF values and changes from baseline will be summarized at Week 24 and Week 48. A notable value is defined as a QTcF interval of greater than 450 ms for both male and female. Those with QTcF interval greater than 500 ms will also be presented. The categories used for the change (increase) in QTcF are: ≤ 0 ms, > 0 and ≤ 30 ms, > 30 to ≤ 60 ms and > 60 ms.

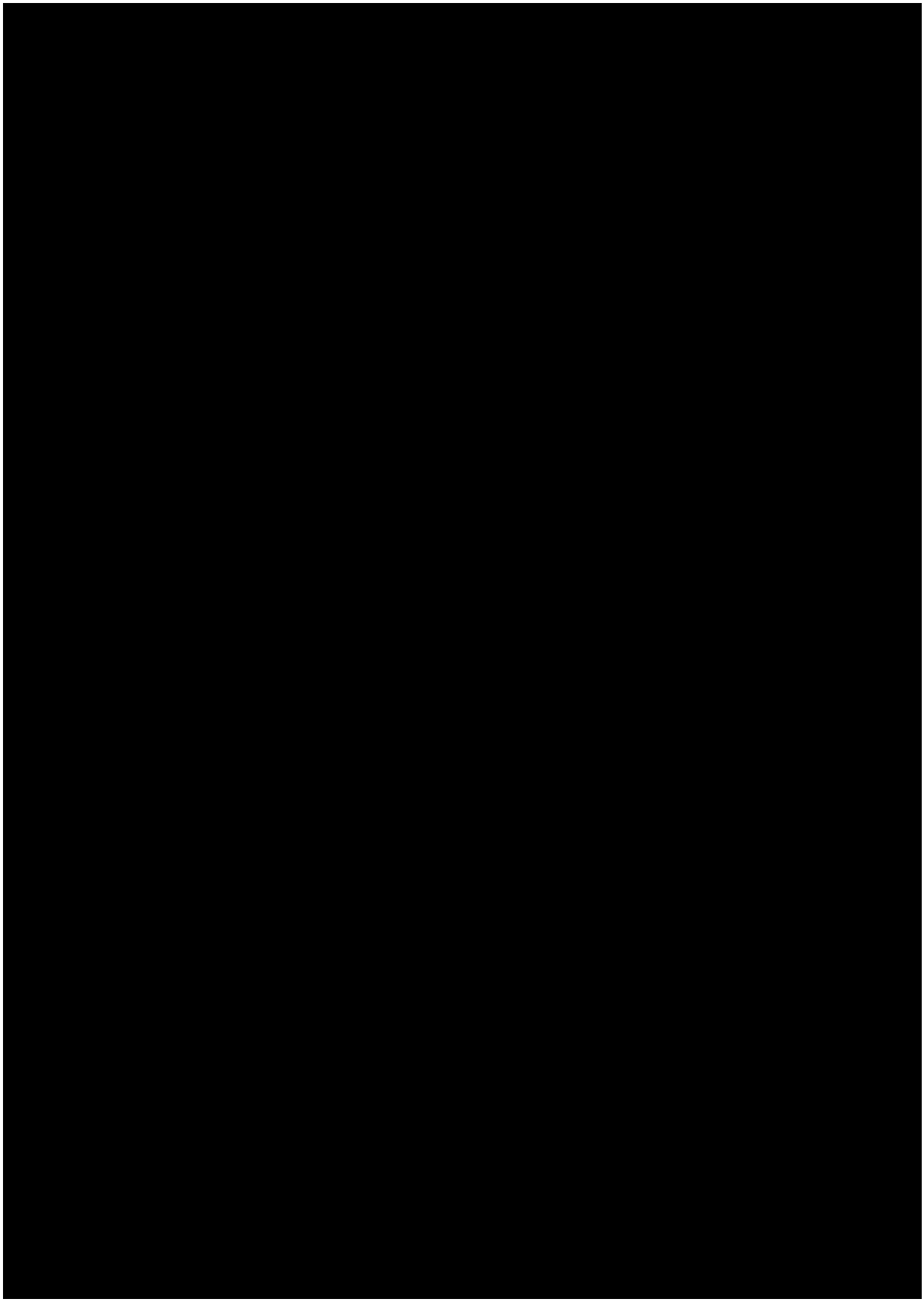
2.8.4.2 Vital signs

Analysis of the vital sign measurements using summary statistics for the change from baseline for each post-baseline visit will be performed. These descriptive summaries will be presented by vital sign parameter and treatment group. [redacted].

Change from baseline will only be summarized for patients with both baseline and post-baseline values. Patients with notable vital signs as defined in protocol Appendix 1 will be listed.

[redacted]





2.14 Interim analysis

No interim analysis is planned for the study.

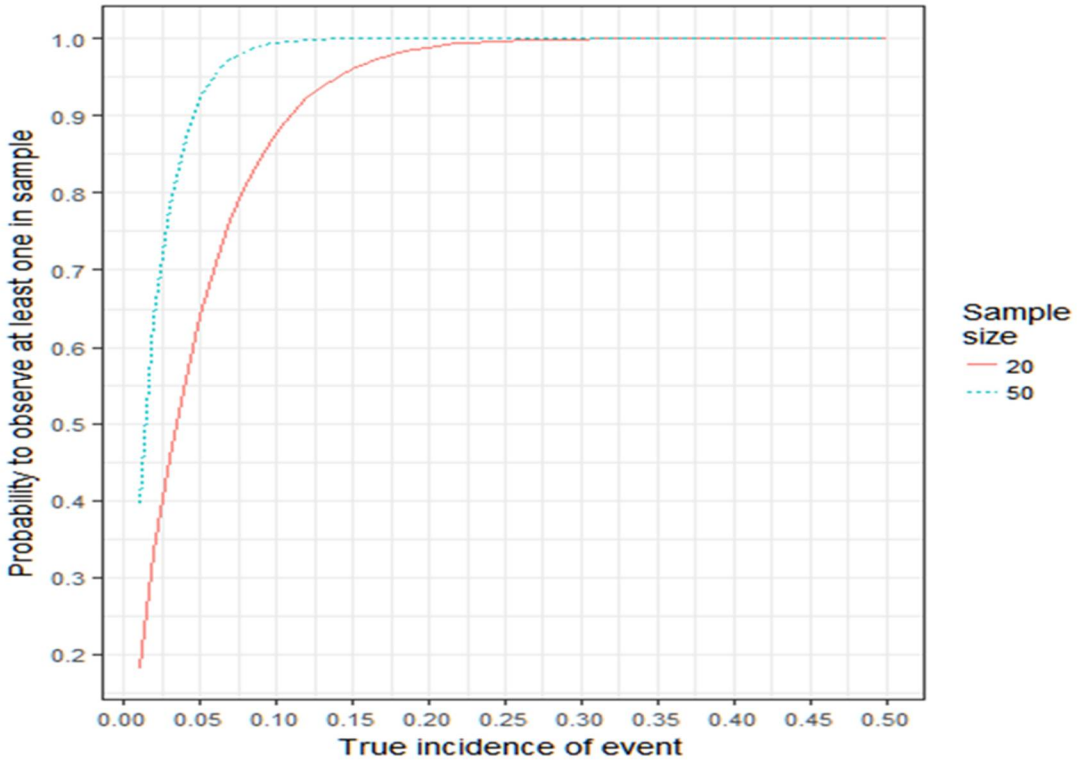
3 Sample size calculation

The primary objective of the study is to determine if there is a safe combination between one of the chosen doses of tropifexor and 150 mg of cenicriviroc. However, the assessment will be made based on the whole safety profile and not on quantitatively formulated hypotheses for distinct parameters. Therefore, the sample size is based on the feasibility with respect to expected speed of enrollment and duration of the study, not on formal statistical criteria.

3.1 Power considerations with given sample size for safety assessment

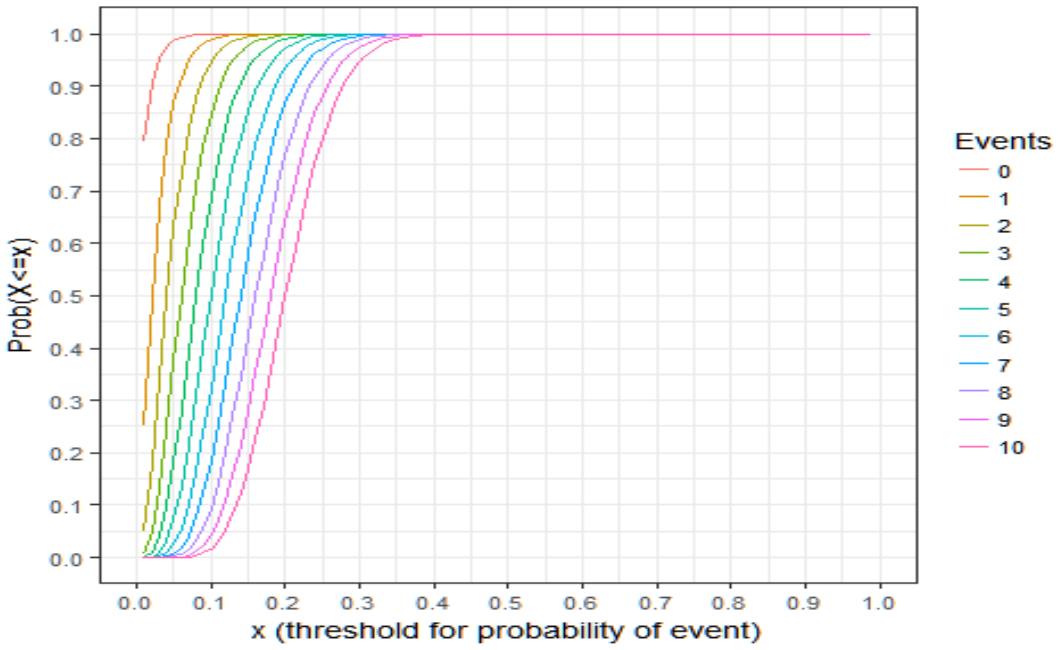
Events with a true incidence of 30% and above are likely to be observed (almost 100% probability) in a group of 50 patients (size of each treatment group). Events with true incidences below 10% down to 3% are still very likely to be observed, while events are observed with less than 50% probability only if the true incidence is less than about 2.5% (Figure 3-1). It is noteworthy, however, that a single patient constitutes 2% in a sample of 50.

Figure 3-1 Binomial probability to observe an event with given sample size



Probabilities of the incidence being below a certain threshold are plotted in Figure 3-2 for a sample size of 50 patients when the event is observed in 0, 1, 2, ...,10 patients (assuming a beta distribution with prior shape parameters 0.33, 0.33).

Figure 3-2 Predictions for probability of event based on observed number occurring in a group of 50 subjects



For example, if 0 events are observed, the probability (Prob) that the incidence is ≤ 5% would be 98.7%. If an event is observed in one patient, the probability that the incidence is ≤ 5% would be 86.6%. Similarly, if an event is observed in 10 patients (one fifth), the probability that the incidence is ≤ 20% would be approximately 50% (calculated using R function pbeta).

3.2 Power considerations with given sample size for efficacy assessment

To evaluate the effectiveness of tropifexor + CVC with respect to the proportion of patients who have at least a one point improvement in fibrosis at Week 48 compared to baseline (i.e, the response rate), one needs to consider the effect size that one is able to detect with respect to both tropifexor and CVC as monotherapy.

- For CVC, based on information from the CENTAUR Phase 2b study, Friedman (2017), it can be assumed that the monotherapy response rate at Week 48 in F2/F3 patients would be approximately 35%
- For tropifexor, a response rate of 42% is assumed (best-case scenario) based on the interpolation of expected Week 72 response rate and assuming linear improvement over time
- For the combined effect of tropifexor + CVC it will be assumed that 75% of the efficacy for CVC will be added to the effect of tropifexor (i.e. a response rate of 42 + 0.75*35 = 69%)

For comparisons between tropifexor + CVC and CVC monotherapy a 2-group continuity corrected χ^2 test of proportions with type I error rate of 0.10 (2-sided, no adjustment for multiple comparisons), a sample size of 50 per group results in a power of 95% (nQuery Advisor 7.0).

For comparisons between tropifexor + CVC and tropifexor monotherapy a 2-group continuity corrected χ^2 test of proportions with type I error rate of 0.10 (2-sided, no adjustment for multiple comparisons), a sample size of 50 per group results in a power of 81% (nQuery Advisor 7.0).

Further scenarios assuming different response rates for tropifexor and the magnitude of effect added to tropifexor by CVC are described in Table 3-1.

Table 3-1 Power for detecting a treatment difference between tropifexor + CVC and tropifexor and CVC monotherapy

Assumed tropifexor response rate	Assumed response rate of tropifexor + CVC	Power N = 50/arm vs. tropifexor monotherapy	Power N = 50/arm vs. CVC monotherapy
Current target 42% response rate at Week 48 (50% improvement over OCA)	½ effect of CVC added to tropifexor ((60%)	48%	75%
	¾ effect of CVC added to tropifexor (69%)	81%	95%
	Full effect of CVC added to tropifexor (77%)	96%	99%
Response rate between target and OCA at Week 48 (35%)	½ effect of CVC added to tropifexor (53%)	48%	48%
	¾ effect of CVC added to tropifexor (62%)	81%	81%
	Full effect of CVC added to tropifexor (70%)	96%	96%
Worst case: tropifexor same effect as OCA 28% at Week 48	½ effect of CVC added to tropifexor (46%)	50%	23%
	¾ effect of CVC added to tropifexor (55%)	82%	56%
	Full effect of CVC added to tropifexor (63%)	96%	81%

*Power calculated assuming a two-sided Type I error rate of 0.10

4 Change to protocol specified analyses

- AE and SAE summary by the onset time with respect to COVID-19 pandemic in different countries and regions are added.
- Paired-biopsy reading will be used for all biopsy-based outputs.
- NAS-based definition of resolution of steatohepatitis is added in the supportive analyses for secondary efficacy endpoint [redacted]

5 Appendix

5.1 Imputation rules

5.1.1 Study drug

The date of the end-of-treatment visit (EOT, Week 48) will be used if the date of last administration of study treatment is missing and the date of randomization will be used if the date of first administration is missing.

5.1.2 AE date imputation

Adverse Event Start Date Imputation (#IMPUTAEV)

This algorithm is expressed in the Variable Source Derivation column as **#IMPUTAEV(*event*)** where *event* is the partial start date of the adverse event.

The following table explains the notation used in the logic matrix. Please note that **missing start dates** will not be imputed.

	Day	Month	Year
Partial Adverse Event Start Date	Not used	MON	YYYY
Treatment Start Date	Not used	TRTM	TRTY

The following matrix explains the logic behind the imputation.

	MON MISSING	MON < TRTM	MON = TRTM	MON > TRTM
YYYY MISSING	**NC** Uncertain	**NC** Uncertain	**NC** Uncertain	**NC** Uncertain
YYYY < TRTY	**(D)** Before Treatment Start	**(C)** Before Treatment Start	**(C)** Before Treatment Start	**(C)** Before Treatment Start
YYYY = TRTY	**(B)** Uncertain	**(C)** Before Treatment Start	**(B)** Uncertain	**(A)** After Treatment Start
YYYY > TRTY	**(E)** After Treatment Start	**(A)** After Treatment Start	**(A)** After Treatment Start	**(A)** After Treatment Start

The following table is the legend to the logic matrix.

If AE end date is complete and AE end date < Treatment start date then AE start reference = min (informed consent date, earliest visit date from SV)

Else if AE end date is partial or AE is ongoing then AE start reference = treatment start date

Relationship	Date Imputation
Before AE Start reference	Partial date indicates AE start date **prior** to AE start reference
After AE Start reference	Partial date indicates AE start date **after** AE start reference
Uncertain	Partial date **insufficient to determine** relationship of AE start date to AE start reference
Imputation Calculation	
NC / Blank	No convention
(A)	01MONYYYY
(B)	Treatment start date+1
(C)	max(15MONYYYY, the start date of the screening period +1 day)
(D)	max(01JULYYYY, the start date of the screening period +1 day)
(E)	01JANYYYY
Complete date	No date imputation

Adverse Event End Date Imputation

Imputed date = date part of original date, if complete date

If a patient is not randomized, then the AE end date is the date of completion of pre-randomization period.

If a patient is randomized:

If AE end date, month, and year are missing or just month is missing, then the AE end date is set to the study completion/discontinuation visit date.

If AE day is missing, then it is set to minimum of (treatment end day, last day of the month).

If imputed AE end date is less than the AE start date, use the AE start date as the imputed AE end date.

Impute Date Flag:

If year of the imputed date <> YYYY then date flag = Y

else if month of the imputed date <> MON then date flag = M

else if day of the imputed date <> day of original date then date_flag = D

else date flag = null

5.1.3 Concomitant medication date imputation

This algorithm is used when *event* is the partial start date of the concomitant medication or non-drug therapy/procedure.

The following table explains the notation used in the logic matrix. Please note that **completely missing start dates** will not be imputed. Also note that imputation of a start date must not result in a date later than the end date. In such case, start date will be max(01-MMM-YYYY, Treatment start date) if only the day is missing and max(01-JAN-YYYY, Treatment start date) if day and month are missing.

	Day	Month	Year
Partial CMD Start Date	Not used	MON	YYYY
Treatment Start Date	Not used	TRTM	TRTY

The following matrix explains the logic behind the imputation.

	MON MISSING	MON < TRTM	MON = TRTM	MON > TRTM
YYYY MISSING	**(C2)** Uncertain	**(C1)** Uncertain	**(C1)** Uncertain	**(C1)** Uncertain
YYYY < TRTY	**(D)** Before Treatment Start	**(A)** Before Treatment Start	**(A)** Before Treatment Start	**(A)** Before Treatment Start
YYYY = TRTY	**(C2)** Uncertain	**(A)** Before Treatment Start	**(C2)** Uncertain	**(B)** After Treatment Start
YYYY > TRTY	**(E)** After Treatment Start	**(B)** After Treatment Start	**(B)** After Treatment Start	**(B)** After Treatment Start

The following table is the legend to the logic matrix.

Relationship	**Date Imputation**
Before Treatment Start	Partial date indicates CMD start date **prior** to Treatment Start Date
After Treatment Start	Partial date indicates CMD start date **after** Treatment Start Date
Uncertain	Partial date **insufficient to determine** relationship of CMD start date relative to Treatment Start Date
Imputation Calculation	
(A)	15MONYYYY
(B)	01MONYYYY
(C1 or C2)	IF relative reference start = before THEN Treatment Start Date -1 day ELSE IF relative reference start = ' ' THEN Treatment Start Date +1 day
(D)	01JULYYYY
(E)	01JANYYYY

Concomitant Medication End Date Imputation:

If not ongoing then

Imputed date = date part of CMENDTC, if complete date

Imputed date = min(reference end date, DEC 31) , if month is missing, (C2, D, E)

Imputed date = min(reference end date, last day of the Month) , if day is missing. (A, B, C1)

Concomitant Medication Date Flag:

If not a complete date then

Y - If year of the imputed date <> YYYY else

M – If month of the imputed date <> MON else

D

5.1.3.1 Prior therapies date imputation

Same as concomitant medication date imputation (as applicable).

5.1.3.2 Post therapies date imputation

Not applicable.

5.1.4 Other imputations

A multiple imputation approach will be applied to missing follow-up biopsies.

For subjects who do not have a Week 48 liver biopsy or not remain on their randomized study treatment for at least 24 weeks (even if a Week 48 biopsy was obtained), the biopsy based outcomes will be imputed by multiple imputation (MI). It is assumed that biopsy results are missing at random (MAR). Available results for subjects who discontinue treatment before 24 weeks are set to missing for this analysis. Imputation will be based on available results from subjects who did not discontinue study treatment, thus addressing the question what the results would have been if subjects had not discontinued. The imputation model will take into account the treatment group, the baseline fibrosis stage (categorical), [redacted], the baseline ALT, [redacted] and change in ALT, [redacted] from baseline to Week 12 and to Week 24 (continuous). As there is only one post-baseline biopsy assessment, an arbitrary missing pattern is assumed and a fully conditional specification (FCS) method with logistic regression is applied. The randomization seed is 4572201 and the number of imputations will be set to 50.

5.2 AEs coding/grading

AEs are coded using the MedDRA dictionary, and severity is graded mild/moderate/severe.

5.3 Laboratory parameters derivations

5.3.1 Laboratory parameters

Liver laboratory events will be classified in two ways:

- Events as defined in Appendix 2 of the protocol (Table 14-1). Only the laboratory results criteria will be considered. Confirmation of the results is not taken into account. For Table 14-1, criteria of ALT and AST are evaluated as either-or relationship, while those of baseline and treatment-emergent assessments as both-and relationship.
- Events as defined in the internal Novartis “Analysis plan for liver safety data”, Table 6-1.

Renal laboratory events will be classified in two ways:

- Events as defined in Appendix 3 of the protocol (Table 15-1). Only the laboratory results criteria will be considered. Confirmation of the results is not taken into account.
- Events as defined in the internal Novartis “Analysis plan for renal safety data”,

Criteria for other laboratory parameters:

Table 5-1 Notable criteria for other laboratory parameters

Parameter	Threshold value	Unit
Albumin	<32	g/L
Hemoglobin	<70	g/L
Hemoglobin	>200	g/L
White blood cell count	<2.0	10^9/L
White blood cell count	>35.0	10^9/L
Platelets	<50	10^9/L
Platelets	>1000	10^9/L
Prothrombin Time INR	>4.0	
PT	>40.0	sec
APTT	>80.0	sec
Sodium	<120	mmol/L
Sodium	>160	mmol/L
Potassium	<3.0	mmol/L
Potassium	>6.0	mmol/L
Glucose	<2.2	mmol/L
Glucose	>27.8	mmol/L
Calcium	<1.50	mmol/L
Calcium	>3.00	mmol/L
Phosphate	<0.29	mmol/L
Creatinine	>177	µmol/L
Calculated eGFR	<60	mL/min

5.3.5 Derived baseline characteristics

Use of lipid reducing drugs (e.g. statins)

Concomitant use of statins is determined as follows:

If drug belonging to ATC code C10 was used either starting prior to randomization and ongoing at randomization, or starting during the treatment epoch, concomitant use of statins is considered Yes, otherwise No.

5.4 Statistical models

5.4.1 Primary analysis

Not applicable.

5.4.2 Key secondary analysis

5.5 Not applicable. Rule of exclusion criteria of analysis sets

Table 5-2 Criteria for exclusion from analysis sets

Deviation ID	Description of Deviation	Exclusion from
OTH10	Patient was rescreened but did not sign a new ICF	All analysis sets
TRT05	No drug taken after randomization	SAF
INCL01	ICF not signed	All analysis sets

5.6 Other statistical aspects

Secondary efficacy endpoints

Multiple imputation is adopted to impute missing Week 48 biopsy results for eligible patients. See Section 5.1.4 for more details.

CMH test is adopted for inference controlled for the baseline fibrosis stage (F2/F3). In order to pool results from multile imputation, test statistic of CMH test is transformed using Wilson-Hilferty transformation

$$\mathrm{ST}_{\mathrm{WH\_CMH}^{(m)}} = \frac{\sqrt[3]{\mathrm{CMH}^{(m)}/df} - (1 - \frac{2}{9 * df})}{\sqrt[2]{2/(9 * df)}},$$

Where $\mathrm{CMH}^{(m)}$ is the CMH test statistic computed from the m-th imputed dataset, df is the degrees of freedom accosiated with the test statsitic, and $\mathrm{ST}_{\mathrm{WH\_CMH}^{(m)}}$ follows approximately standard normal distribution. Consequently, the test results obtained from muliple imputation datasets can be combined by specifying the model effects and standard error in PROC MIANALYZE in SAS and the final one-side *p*-value for CMH test can be obtained.

Odds ratio and relative risk obtained from PROC FREQ also need to be transformed to normal distribution before pooling. Both statistics are log-normally distributed and hence a simple log transformation would transform them to normal and similar procedure to the above can be followed to obtain the pooled results. Finally, exponential tranformation is implemented to tranform the statistics and their CI back to the original scale.

Repeated measurements

For inference of repeated measurements, PROC MIXED with REML option is employed for MMRM model. The model will include treatment group, visit, interaction between treatment group and visit, and baseline value of the variable of interest and its interaction with visit. The model will adopt unstructured covariance matrix and the degrees-of-freedom method of Kenward and Roger. The difference in the variable of interest between treatment groups and its CI will be calculated using LSMEANS statement.

Crude incidence and related risk estimates

For n patients each at risk to experience a certain event with probability π, the crude incidence is estimated as p=x/n, where x is the number of patients with the event.

Odds ratio and $100(1-\alpha)\%$ confidence interval

For an investigational drug group with n_1 patients at risk, independent from placebo with n_0 patients at risk, of whom x_1 and x_0 experience a certain event with probability π_1 and π_0 respectively, the odds ratio is estimated as $\frac{p_1/(1-p_1)}{p_0/(1-p_0)}$ with $p_1 = x_1/n_1$ and $p_0 = x_0/n_0$. A conditional exact $100(1-\alpha)\%$ confidence interval will be obtained by using the SAS procedure PROC FREQ with statement EXACT OR or EXACT COMOR. In case that exact method fails to produce valid results, asymptotic approach should be adopted. EXACT

statement is not to be used in combination with multiple imputation, in which case asymptotic approach is adopted.

Risk difference and $100(1-\alpha)\%$ confidence interval

For an investigational drug group with n_1 patients at risk, independent from placebo with n_0 patients at risk, of whom $x_1$and x_0 experience a certain event, the risk difference is estimated as $p_1 - p_0$ with $p_1 = x_1/n_1$ and $p_0 = x_0/n_0$. Exact unconditional confidence limits for the risk difference will be obtained with SAS procedure PROC FREQ and option RISKDIFF in the TABLES statement, specifying the RISKDIFF option also in the EXACT statement. In case that exact method fails to produce valid results, asymptotic approach should be adopted. EXACT statement is not to be used in combination with multiple imputation, in which case asymptotic approach is adopted.

Geometric mean and coefficient of variation

The geometric mean (where applicable) will be presented for the baseline values, absolute post-dose values and for the ratio to baseline (or pre-dose, respectively) values. The geometric mean of the ratio to baseline will be presented in terms of % change from baseline and will be calculated as follows: (exp (mean of the log-transformed ratio to baseline values) -1)*100.

The Coefficient of Variation (CV) will be calculated for the baseline values, the absolute post-dose values and the ratio to baseline (or pre-dose, respectively) values.

6 Reference

[redacted]

van Ginkel, Joost R., and Pieter M. Kroonenberg. "Analysis of variance of multiply imputed data." *Multivariate behavioral research* 49.1 (2014): 78-91.