
# Statistical Analysis Plan

| | |
|---|---|
| **Protocol Title:** | A Multicenter, Randomized, Double-Blind, Placebo-Controlled, Parallel-Group, Phase 2a Clinical Trial to Evaluate the Safety, Tolerability, Efficacy, and Pharmacokinetics of Orally Administered TERN-101 Tablets in Adult Patients with Presumed Non-Cirrhotic Non-Alcoholic Steatohepatitis (NASH) |
| **Compound:** | TERN-101 |
| **US IND Number:** | [redacted] |
| **ClinicalTrials.gov Identifier:** | NCT04328077 |
| **Protocol Version and Date:** | Amendment 1, 22 May 2020 |
| **Phase:** | 2a |
| **Methodology:** | Randomized, Double-Blind, Placebo-Controlled, Parallel-Group |
| **Sponsor:** | Terns, Inc.<br>1065 E. Hillsdale Blvd. Suite 100<br>Foster City, California 94404 |
| **Analysis Plan Date:** | 19 May 2021 |
| **Analysis Plan Version:** | Final |

**CONFIDENTIAL**

## TABLE OF CONTENTS

## LIST OF TABLES



## LIST OF FIGURES

# 1. INTRODUCTION

This statistical analysis plan (SAP) document outlines the statistical methods to be implemented during the analyses of Study Protocol TERN101-2001. The purpose of this plan is to provide specific guidelines from which the analyses will proceed. Any deviations from these guidelines will be documented in the clinical study report (CSR).

This SAP supersedes the statistical considerations identified in the protocol; where considerations are substantially different, they are identified in Section 11. Any post-hoc or unplanned analyses, or significant changes from the planned analysis in this SAP performed to provide results for inclusion in the CSR but not included in this SAP, will be clearly identified in the CSR. Changes to planned analyses do not require an updated SAP but will be included in the CSR if significant.

# 2. INFORMATION FROM THE STUDY PROTOCOL

## 2.1. Study Objectives

Primary Objective:

- To evaluate safety and tolerability of 3 doses of orally administered TERN-101 versus placebo for 12 weeks in non-cirrhotic presumed NASH patients, with clinical or histological NASH diagnosis

Secondary Objectives:

- To evaluate the change in alanine aminotransferase (ALT) with 3 doses of TERN-101 for 12 weeks versus placebo in non-cirrhotic presumed NASH patients
- To evaluate the pharmacokinetics (PK) of 3 doses of TERN-101 in non-cirrhotic presumed NASH patients

Exploratory Objectives:

- To assess the effect of 3 doses of TERN-101 on liver fat content (LFC) and additional indicators of NASH disease activity
- To evaluate farnesoid X receptor (FXR) target engagement of 3 doses of TERN-101 in non-cirrhotic presumed NASH patients
- To evaluate the effect of 3 doses of TERN-101 on biomarkers of liver fibrosis

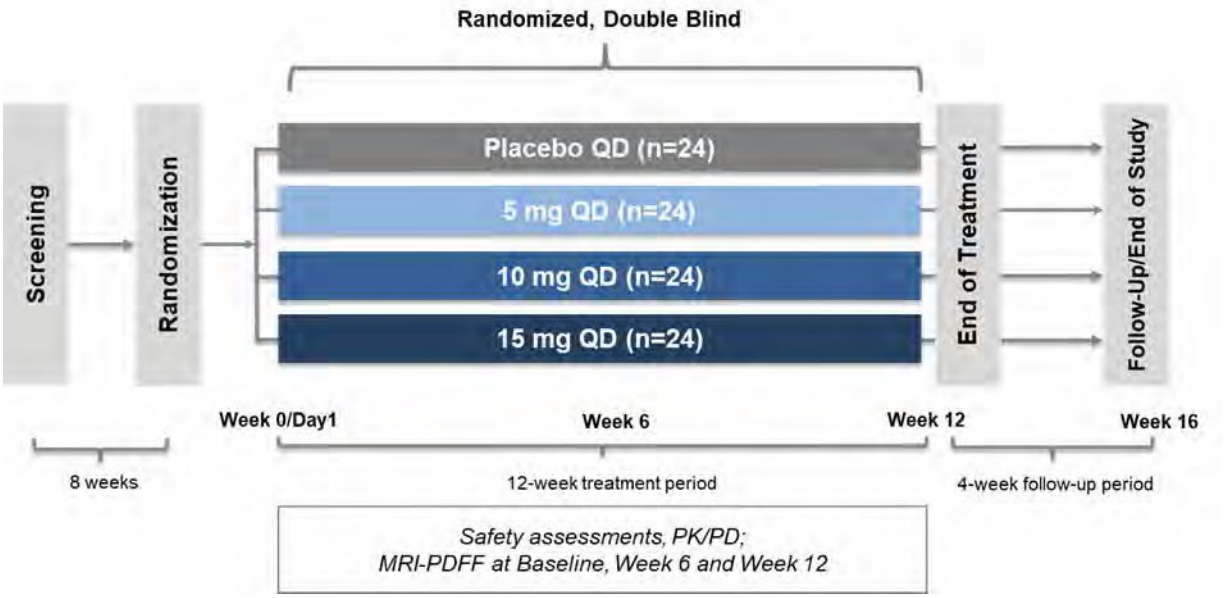
## 2.2. Study Design

This study is a randomized, double-blind, placebo-controlled, parallel-group, Phase 2a study to evaluate the safety, tolerability, efficacy, and PK of various doses of TERN-101 in non-cirrhotic presumed non-alcoholic steatohepatitis (NASH) patients. TERN-101 is a highly selective and potent small molecule agonist of farnesoid X receptor (FXR) in development for the treatment of patients with NASH, a serious and advanced form of non-alcoholic fatty liver disease (NALFD). The study aims to characterize the effects of varying doses of TERN-101 in a NAFLD patient

population with presumed NASH. Currently, there are no approved therapies for the treatment of NASH.

Figure 1 describes the overall study design. The total study duration will be approximately 24 weeks, consisting of an 8-week Screening Period, a 12-week Treatment Period and a 4-week Follow-up Period.

**Figure 1: Overall Study Design**




**Abbreviations:** MRI-PDFF = magnetic resonance imaging proton density fat fraction; PK/PD = pharmacokinetics/pharmacodynamics; QD = once a day

Of the approximately 96 patients randomized, 24 patients will take part in an intensive PK and pharmacodynamic (PD) collection after the first dose and after the last dose of study drug. This includes 6 patients at each TERN-101 dose level, and 6 patients in the placebo group. Patients who are not participating in the PK/PD sub-study will have trough PK/PD sampling only.

### 2.2.1. Study Procedures

The overall schedule of activities is outlined in Section 1.3 of the study protocol. See study protocol for complete details.

### 2.2.2. Study Population

This study includes non-cirrhotic presumed NASH patients. Inclusion and exclusion criteria are provided in Sections 5.1 and 5.2, respectively, of the study protocol.

### 2.2.3. Study Drug

Study drug is defined as any investigational intervention(s), marketed product(s), or placebo intended to be administered to a study patient according to the study protocol.

**Table 1: Study Drug**

[redacted]

### 2.2.4. Treatment Assignment, Blinding, and Randomization Methodology

Patients will be randomized to TERN-101 or placebo using an Interactive Web Response System (IWRS). Approximately 96 patients will be randomized at an overall ratio of 1:1:1:1 into 4 groups receiving 5 mg (n = 24), 10 mg (n = 24), 15 mg (n = 24) TERN-101 tablet, or matching placebo (n = 24). Randomization will ensure 6 patients at each TERN-101 dose level and 6 patients in the placebo group are assigned to the PK/PD sub-study.

Investigators, patients, all study personnel, and the Sponsor will remain blinded to each patient's assigned dose group throughout the course of the study. The randomization codes will be provided to bioanalytics to allow for the exclusion of placebo patients from the PK assay.

Bioanalytics will not disclose the randomization code or the results of their measurements until the trial is officially unblinded.

Emergency unblinding procedures are outlined in Section 6.4.1 of the study protocol and in the unblinding plan.

## 2.3. Study Endpoints

### 2.3.1. Primary Endpoints

The primary endpoints are the patient incidence of treatment emergent adverse events (TEAEs), defined in Section 9.2, and patient incidence of treatment emergent clinical safety laboratory abnormalities, defined in Section 9.3.1.

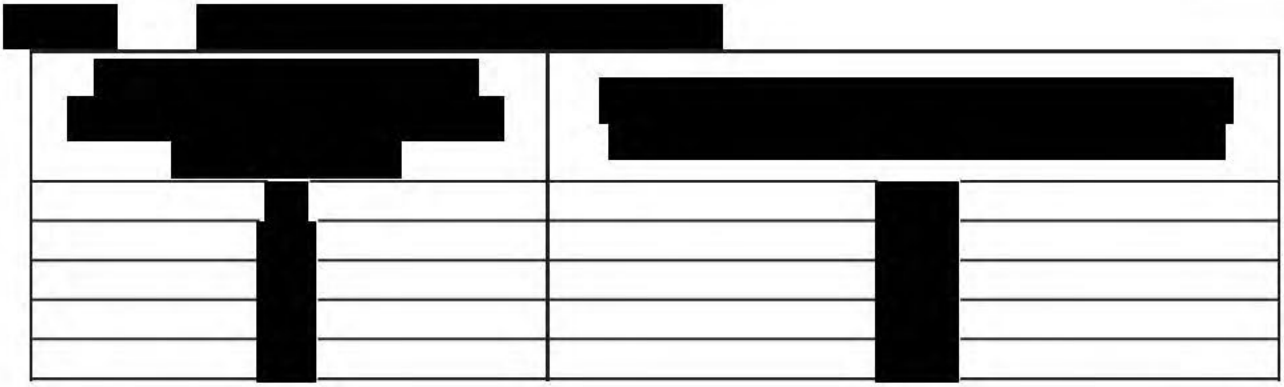
### 2.3.2. Secondary Endpoints

- Percent change from baseline in ALT levels at Week 12
- Plasma PK for TERN-101 and its main metabolite TRN-000971

### 2.3.3. Exploratory Endpoints

- Absolute change and relative change from baseline in LFC by Magnetic Resonance Imaging (MRI) Proton Density Fat Fraction (PDFF)
- Proportion of patients achieving a $\geq$ 30% relative decline from baseline in LFC by MRI-PDFF
- Absolute change and relative change from baseline in corrected T1 (cTI) by MRI
- Change and percent change from baseline in select liver function tests: ALT, aspartate transaminase (AST), and gamma-glutamyl transpeptidase (GGT)
- Change and percent change from baseline in plasma PD biomarkers including fibroblast growth factor-19 (FGF19), 7α-hydroxy-4-cholesten-3-one (7αC4), [redacted]
- Change and percent change from baseline in NASH and fibrosis blood biomarkers including cytokeratin-18 (CK-18) (M30 and M65), procollagen III n-terminal propeptide (PIIINP), tissue inhibitor of metalloproteinases-1 (TIMP-1), hyaluronic acid (HA), pro-peptide of type III collagen (Pro-C3) and calculated Enhanced Liver Fibrosis (ELF), Fibrosis-4 (FIB-4) and Nonalcoholic Fatty Liver Disease (NAFLD) fibrosis scores
- [redacted]

[redacted]



[redacted]

# 4. GENERAL STATISTICAL METHODS

## 4.1. Reporting Conventions

Individual patient data obtained from electronic case report forms (eCRFs), central laboratories, external sources, and any derived data will be presented in data listings by patient. The primary data source will be used for all analyses. All data listings that contain an evaluation date will contain a relative study day. Pre-treatment and on-treatment study days are numbered relative to the day of the first dose of study drug which is designated as Day 1. The preceding day is Day 1, the day before that is Day 2, etc.

All outputs will be incorporated into Microsoft Word rich text format (.rtf) files, sorted and labeled according to the International Council for Harmonisation (ICH) E3 guideline, E3

Structure and Content of Clinical Study Reports, and formatted to the appropriate page size(s), font type, and font size according to Food and Drug Administration (FDA) guidance of Portable Document Format Specifications.

For categorical variables, summary tabulations of the number and percentage of patients within each category of the parameter will be presented. Percentage calculations will be based on non-missing data, unless otherwise specified. If there are missing values, the number missing will be presented, but without a percentage. Percentages are rounded to 1 decimal place, unless otherwise specified.

For frequency counts of categorical variables, categories whose counts are zero will be displayed for the sake of completeness. For example, if none of the patients discontinue due to "Lost to Follow-up," this reason will be included in the table with a count of 0. Percentages will be presented to 1 decimal place, with the exception of 0, which will be presented without percent, and 100, which will be presented without decimal places. Values less than 0.1% will be presented as "<0.1%." Values less than 100% but greater than 99.9% will be presented as ">99.9%."

For continuous variables, the number of patients, mean, standard deviation (SD), median, 25$^{th}$ (Q1) and 75$^{th}$ (Q3) percentiles, minimum, and maximum values will be presented. The precision of summary statistics, unless otherwise specified will be as follows: mean, median, Q1, and Q3 to 1 more decimal place than the raw data, SD to 2 decimal places more than the raw data, and minimum and maximum to the same decimal places as the raw data. In general, the number of decimal places should not exceed 3 decimal places unless appropriate.

For log-normal data [redacted] the geometric mean and geometric coefficient of variation (CV%) will also be presented. The precision of descriptive statistics will be as follows: mean, median, Q1, and Q3 to 1 more decimal place than the raw data, SD to 2 decimal places more than the raw data, the minimum and maximum to the same number of decimal places as the raw data, geometric mean to 3 significant figures, and geometric CV% to 1 decimal place. [redacted]

Confidence intervals (CIs) will be provided and will be rounded to 1 decimal place, unless otherwise specified, in the table and listing shell. A full set of summary statistics will only be presented if there are 3 or more values available. If there are less than 3 values, only the min, max, and N will be presented. The other summary statistics will be denoted as not calculated (NC).

For tables where rounding is required, rounding will be done to the nearest round-off unit. For example, when rounding to the nearest integer, values ≥XX.5 will be rounded up to XX+1 (e.g., 97.5 will round up to 98), while values <XX.5 will be rounded down to XX (e.g., 97.4 will round down to 97).

All statistical tests comparing groups will be conducted at the 2-sided, 0.05 level of significance, unless otherwise specified. Summary statistics for each treatment group will be presented, as well as two-sided 90% and 95% CIs comparing groups will be provided.

Other general programming specifications are provided in Appendix 1.

### 4.2. Computing Environments

All descriptive statistical analyses will be performed using SAS software Version 9.4 or higher, unless otherwise noted.

Plasma PK [redacted] for TERN-101 will be estimated using non-compartmental methods with Phoenix WinNonlin® Version 8.3 or higher.

Medical history and AEs will be coded using Medical Dictionary for Regulatory Activities (MedDRA) version 23.0. Concomitant medications will be coded using World Health Organization Drug Dictionary (WHODD), B2 Enhanced March 2020.

Study Data Tabulation Model (SDTM) and Analysis Data Model (ADaM) datasets will be prepared using Clinical Data Interchange Standards Consortium (CDISC) SDTM Implementation Guide (IG) Version 3.3, ADaM Version 2.1, and CDISC ADaM IG Version 1.2 including most current occurrence and time to event IGs.

### 4.3. Partial Dates

Imputation of partial adverse event and concomitant medication dates are specified in Appendix 2, respectively.

All data recorded on the case report form will be included in data listings.

### 4.4. Data Conventions

The precision of original measurements will be maintained in summaries, when possible.

Quantitative safety laboratory tests containing less than (<) and greater than (>) symbols are test results that are below and above quantifiable limits, respectively. In order to retain these values for analysis purpose, the following will be imputed and stored within the analysis datasets:

- For values with <, the imputed value will be the numeric portion × 0.9.
- For values with >, the imputed value will be the numeric portion × 1.1.

PK and PD imputation of values below the limitation of quantitation (BLQ) are specified in Section 8.9.1 and Section 8.10.1, respectively.

Variables with a non-normal distribution that impacts the interpretation or validity of the planned analysis may have a data transformation applied (e.g., ln, $\log_{10}$).

### 4.5. Standard Calculations

Variables requiring calculation will be derived using the following formulas:

- Days – A duration expressed in days between one date (*date1*) and another later date (*date2*) will be calculated using the following formulas:

  duration in days = date2 – date1 + 1, where date1 ≥first administration date

  duration in days = date2 – date1, where date1 <first administration date

- Months – A duration expressed in months is calculated as the number of days divided by 30.4375
- Years – A duration expressed in years between one date (*date1*) and another date (*date2*) is calculated using the following formulas:

  duration in years = (date2 – date1 + 1)/365.25, where date1 ≥first administration date

  duration in years = (date2 – date1)/365.25, where date1<first administration date
- Body Mass Index (BMI) – BMI is calculated using height (cm) and weight (kg) using the following formula:

  $BMI\ (kg/m^2) = weight\ (kg) / ([height\ (cm)/100]^2)$
- Change (CHG; equivalent to absolute change for liver fat content [LFC]) – Change will be calculated as:

  Change = later value – earlier (i.e. baseline) value
- Percent change (PCHG; equivalent to relative change for LFC) – Percent change will be calculated as:

  Percent change = ([Change] / earlier [i.e. baseline] value) × 100

## 4.6. Treatments

Table 4 presents how the dose groups will be presented on TFLs as treatment groups, including the order.

**Table 4: Treatment Group Labels and Ordering**

| Treatment Group Label | Order on TFLs |
|---|---|
| Placebo | 1 |
| 5 mg | 2 |
| 10 mg | 3 |
| 15 mg | 4 |
| All TERN-101 | 5 |
| Overall (where specified) | 6 |

Screen failures will be presented on by-patient listings where data is available.

## 4.7. Visits

### 4.7.1. Windows

Each visit will be denoted by its visit "Week". The first dose day is denoted as Day 1. In data listings, the relative study day of all dates from first dose will be presented.

In the event of unscheduled visits or early treatment termination (ET) assessments, these will be reassigned to a scheduled visit for analysis purposes according to Table 5 for on treatment assessments. If multiple assessments occur within a single visit window, after reassignment of unscheduled visits and ET assessments, the assessment closest to the target day of the visit window will be used in the analysis. If there is a tie, the later assessment will be used in the analysis.

**Table 5: Treatment Period Visit Windows**

| Target Scheduled Visit | Target Study Day [a] | Analysis Window Study Day [a] | |
|---|---|---|---|
| | | Low | High |
| Baseline[b] | 1 | See Section 4.7.2 | |
| Week 2 | 15 | 2 | 22 |
| Week 4 | 29 | 23 | 36 |
| Week 6 | 43 | 37 | 50 |
| Week 8 | 57 | 51 | 71 |
| Week 12[c] | 85 | 72 | 92[c] |

[a] Study day will be calculated from first dose of study drug.
[b] Baseline is defined in Section 4.7.2.
[c] LFC and cTI by MRI will have no upper limit on Week 12.

The follow up visit (Week 16) will not be windowed. As described in Section 8.7.1, sensitivity analyses will be conducted including only Week 16 assessments which are at least 22 days from last dose.

### 4.7.2. Definition of Baselines

The baseline value for statistical analyses of quantitative laboratory parameters is defined as the mean of all available evaluations, including any unscheduled or repeat assessments, prior to the first administration of study drug, unless otherwise specified. If there is only one evaluation prior to the first administration of study drug, then the available data from this evaluation will be used as the baseline value.

For glucose and lipid parameters, only assessments taken while fasting for 8 hours will be included in the baseline derivation.

The baseline value for analyses of qualitative parameters (e.g., normal/abnormal) is defined as the last evaluation prior to the first administration of study drug.

The baseline value for analyses of weight, body mass index (BMI), and electrocardiogram parameters is defined as the last evaluation prior to the first administration of study drug.

### 4.7.3. Definition of End of Treatment

The End of Treatment value will be defined as the Week 12 value, either scheduled or windowed per Table 5. If a Week 12 value is not available, the last non-missing on or before the last dose date will be used.

## 5. ANALYSIS SETS

Table 6 defines the analysis sets to be used.

**Table 6: Analysis Sets**

| Analysis Sets | Description |
|---|---|
| Screened | All patients who sign the informed consent form (ICF). |
| Randomized | All patients who are randomized into any one of treatment groups. Treatment assignment will be based on the randomized treatment. |
| [redacted] | [redacted] |
| Pharmacodynamic (PD) | All randomized patients who received at least 1 dose of study drug (TERN-101 or placebo) and for whom at least one post-baseline assessment of FGF19, 7αC4, [redacted] is available. Treatment assignment will be based on the treatment actually received. The PD analysis set will include patients with trough concentrations and with intensive PD sampling as part of the PK/PD substudy. |
| Safety | All patients who received at least 1 dose of study drug. Treatment assignment will be based on the treatment actually received. |
| Efficacy | All randomized patients who received at least 1 dose of study drug. Treatment assignment will be based on the randomized treatment. |
| Per Protocol (PP) | All randomized patients who have completed the study without any protocol deviations affecting the interpretation of the efficacy, PK, or PD results. Treatment assignment will be based on the treatment actually received. |

| Analysis Sets | Description |
|---|---|
| Treatment Completer | All patients who complete 12 weeks of treatment with at least 80% compliance based on study drug accountability. Treatment assignment will be based on the randomized treatment. |

# 7. STUDY POPULATION

## 7.1. Patient Disposition

Patient disposition will be summarized for the Enrolled Analysis Set by treatment group and overall total. The summary will include:

- Number of patients screened.
- Number of screen failures, i.e., screened but not randomized, with reasons for screen failure. The denominator for percentage of screen failures will be the number of patients screened. The denominator for percentages for reasons for screen failures will be the number of screen failures.
- Number randomized. The denominator for percentage of randomized patients will be the number of patients screened.
- Number randomized and not treated. The denominator for percentage of randomized and not treated patients will be the number of patients randomized.

- Number in each analysis set. The denominator for percentage in each analysis set will be the number of randomized patients. In addition, the number of patients randomized to the PK/PD substudy will be presented.
- Number who discontinued treatment early and reason(s) for discontinuation of treatment. The denominator for percentages will be the number of randomized and treated patients (Efficacy Analysis Set).
- Number who discontinued from study prior to completing the study and reason(s) for discontinuation. The denominator for percentages will be the number of randomized and treated patients (Efficacy Analysis Set).

A by-patient data listing of study disposition information including the reasons for treatment and/or study termination will be presented. Withdrawal of consent details will also be included, if applicable.

A by-patient data listing including the reasons for exclusion from each analysis set will be presented.

### 7.2. Demographics and Baseline Characteristics

Demographic variables will include the following:

- Age at informed consent including the subgroup of <65 years vs ≥65 years
- Sex
- Race
- Ethnicity

Other baseline characteristics will include the following:

- Weight (kg)
- Height (cm)
- BMI (kg/m$^2$) including the following subgroup: <30, ≥30 - <35, ≥35 - <40, ≥40 kg/m$^2$
- Waist circumference (cm)
- Baseline liver stiffness by transient elastography (kPa)
- Baseline Controlled Attenuation Parameter (CAP; dB/m)
- Baseline LFC (%) by MRI-PDFF
- Baseline cTI (msec) by MRI including the following subgroup: Low (<800 msec), Elevated (800 – 875 msec), and High (>875 msec).
- Liver Biopsy within 2 years of randomization
  - Baseline fibrosis stage (F1, F2, F3)
  - Baseline NAFLD Activity Score (NAS)

- Baseline Comorbidities will be based on Medical History case report form, Medical Dictionary of Regulatory Activities (MedDRA) preferred term (PT)
  - Diabetes status (yes, no) will be defined as PT preferred term (PT) includes "diabetes mellitus" or "Diabetic neuropathy" will be assigned yes, otherwise no.
  - Hypertension (yes, no) will be defined as PT of "Hypertension" will be assigned yes, otherwise no.
  - Dyslipidemia (yes, no) will be defined as Standardized MedDRA query (SMQ) of "Dyslipidemia (SMQ)" will be assigned yes, otherwise no.
- Baseline medication use, defined as starting prior to first administration of study drug, regardless of end date:
  - Statins defined by the WHODD Standardized Drug Grouping (SDG) of "Statins"
  - Antihypertension medications defined by the WHODD SDG of "Antihypertensives"
  - Antithrombotic medications defined by WHODD SDG of "Antithrombotic drugs"
  - Antidiabetic medications defined by the WHODD SDG of "Drugs used in diabetes"
  - NASH Medications, defined as the following medications with potential beneficial effect on NASH:
    - Vitamin E defined by the WHODD drug codes beginning in "001105" or if "VITAMIN E" is found in the preferred drug name
    - Obeticholic acid defined by WHODD drug codes of 064968 01 001 through 064968 01 003
    - Pioglitazone or other PPARγ agonists defined by WHODD ATC level 4 A10BD and A10BG
    - GLP-1 analogues defined by WHODD ATC level 4 code (term) A10BJ (Glucagon-like peptide-1 [GLP-1] analogues)
    - DPP-4 inhibitors defined by WHODD ATC level 4 code (term) A10BH (Dipeptidyl peptidase 4 [DPP-4] inhibitors)
    - SGLT2 inhibitors defined by WHODD ATC level 4 code (term) A10BK (Sodium-glucose co-transporter 2 [SGLT2] inhibitors)
  - Steatohepatitis Medications, defined as the following medications that may potentially worsen steatohepatitis:
    - Corticosteroids (of > 10 mg/day for > 2 weeks [prednisone or prednisone equivalent dose (Appendix 1)]) defined by WHODD ATC level 2 code (term) H02 (Corticosteroids for Systemic Use)
    - Amiodarone defined by WHODD drug codes 001331 01 001 through 001331 01 029 and 001331 02 001 through 001331 02 250
    - Tamoxifen defined by WHODD drug codes of 003887 01 001 through 003887 01 020 and 003887 02 001 through 003887 02 273

- Methotrexate defined by WHODD drug codes of 001138 01 001 through 001138 01 172 and 001138 02 001 through 001138 02 189

The following baseline laboratory tests will be summarized as continuous variables and as frequencies for categorizations described:

- Baseline ALT level (IU/L):
  - ≤60 U/L versus >60 U/L (overall and by sex)
  - <2x ULN versus ≥2x ULN to < 5x ULN
- Baseline AST level (IU/L): The same subgroups as specified for ALT above will be used.
- Baseline GGT level (IU/L): ≤ULN, >ULN
- Baseline ALP level (IU/L): ≤ULN, >ULN
- Baseline total bilirubin level (mg/dL): ≤ULN, >ULN
- Baseline fasting HDL cholesterol (mg/dL)
- Baseline fasting LDL cholesterol (mg/dL)
- Baseline fasting total cholesterol (mg/dL)
- Baseline fasting triglycerides (mg/dL)
- Baseline fasting glucose level (mg/dL)
- HbA1c (%): < 6.5%, ≥ 6.5%
- Baseline platelet count ($10^9$/L):
  - <75, ≥75 × $10^9$/L
  - <LLN, ≥LLN to ≤ULN, >ULN

Demographics and baseline characteristics will be summarized by treatment group and overall, for Efficacy, Safety, PP, and Treatment Completer Analysis Sets. The PP and Treatment Completer Analysis Sets will only be presented if the analysis set differs from the Efficacy Analysis Set by more than 10 patients total.

All demographic and baseline characteristics data will be presented in by-patient data listings.

No inferential statistical comparisons will be performed.

### 7.3. Protocol Deviations

[redacted]

### 7.4. General Medical History

All medical history conditions will be coded using the Medical Dictionary of Regulatory Activities (MedDRA). Medical history will be summarized by treatment group and overall total by system organ class and preferred term, using the Safety Analysis Set. Summaries will be ordered by descending order of the overall incidence of system organ class and preferred term within each system organ class.

### 7.5. Prior and Concomitant Medications

Prior medications are defined as medications that started before first study drug administration and either stopped before or continued after first study drug administration. Concomitant medications are defined as medications that are being taken while on study drug through 30 days after last dose or through the Follow-Up Period (Week 16). Medications that are ongoing on the date of the first administration of study drug will be classified as both prior and concomitant. Any medication that cannot be confirmed as stopping before the start of study drug will be classified as both a prior and a concomitant medication. Imputation of partial dates is defined in Appendix 2.

Prior and concomitant medications will be summarized separately and the number and percentage of patients in each treatment group who took at least one prior (concomitant) medication as well as the number and percentage of patients who took each type of medication will be summarized by Anatomic Therapeutic Class (ATC) Level 2, ATC Level 4, and preferred name for the Safety Analysis Set. If a patient has more than one occurrence of the same preferred name, then the preferred name will be counted only once for that patient. Similarly, if a patient has more than one preferred name within ATC Level 4 or ATC Level 4 within ATC Level 2, then the patient will be counted only once in that ATC Level 4 or ATC Level 2.

Any medications with potential impact on NASH outcome and steatohepatitis will be presented in a separate by-patient listings.

## 8. EFFICACY ANALYSES

### 8.1. Adjustments for Covariates

For comparison of treatment groups with respect to change and percent change from baseline, analysis of covariance (ANCOVA) [redacted] will be used. The corresponding baseline value will be used as a covariate in the model.

## 8.2. Handling of Dropouts or Missing Data

Primary analyses will be based on observed data.

## 8.3. Interim Analyses and Data Monitoring

## 8.4. Multicenter Studies

The randomization is not stratified by site. Likewise, analyses of data will not be stratified by study site.

## 8.5. Use of an "Efficacy Subset" of Patients

Not Applicable.

## 8.6. Multiple Comparisons/Multiplicity

There will be no adjustment for multiplicity. Statistical testing will be considered nominal, descriptive, and exploratory.

## 8.7. Secondary Efficacy Analyses: Percent Change from Baseline in ALT Levels at Week 12

Descriptive statistics of ALT values, change from baseline, and percent change from baseline values will be presented by treatment group for each visit using the Efficacy Analysis Set. In addition, figures of mean (±SD) ALT values, change from baseline, and percent change from baseline will be presented over time by treatment group.

Analyses of percent change (and separately change) from baseline will be carried out using an ANCOVA model at Week 12 with percent change (and separately change) from baseline as the dependent variable including treatment group as a fixed effect and baseline as a covariate. The primary analysis of percent change (and separately change) in ALT at Week 12 will be carried out using available observed data only, without imputation for missing data.

The estimates of least-square (LS) means, standard errors (SE), and 90% and 95% CIs will be presented by treatment group. Estimates of the LS mean difference between pooled TERN-101 (All TERN-101) and placebo, and each pairwise comparison between placebo and each active dose of TERN-101, will be presented with the associated standard error of the difference, and 90% and 95% CIs of the difference.

The same ANCOVA model will be used to analyze the last on-treatment value for Week 12, to be labelled as "End of Treatment". For patients who complete Week 12, the Week 12 value will be used. For patients who do not complete treatment, the last value on treatment will be used.

The same ANCOVA model will be presented for each post-baseline visit using observed data only.

The distribution of percent change values will be evaluated visually for normality assumptions. If normality assumptions are violated, i.e. data are not normally distributed, a Wilcoxon Rank Sum Test will be used to compare each TERN-101 dose to placebo, and pooled TERN-101 (All TERN-101) to placebo. The median differences and 90% and 95% CIs of the median differences between treatment groups will be constructed using Hodges-Lehmann estimate.

### 8.7.1. Additional Exploratory Analyses of ALT

#### 8.7.1.1. Change from Baseline in ALT

Analyses will also be carried out using a REML based MMRM to evaluate the effect over time on treatment using the Efficacy Analysis Set. The dependent variable will be the percent change

[redacted]

### 8.7.1.3. ALT Responder Analyses

The number and percentage of patients meeting the following definitions of responder and non-responder based on ALT will be summarized by treatment group at Week 12 [redacted]

- Change from baseline
  - Responder: Decrease in ALT of at least 17 U/L
  - Non-responder: all other non-missing values
- Percent change from baseline
  - Responder: Decrease in ALT of at least 30%
  - Non-responder: all other non-missing values
- ALT result
  - Responder: Values 0 to ≤ 30 U/L
  - Non-responder: Values >30 U/L

[redacted]

In addition, for each definition of response, the number and percentage of patients who ever met the criteria over the course of the study will be summarized.

A Chi-square test at each post-baseline visit will be used to compare pooled TERN-101 (All TERN-101) and placebo, and each pairwise comparison between placebo and each active dose of TERN-101.

## 8.8. Exploratory Efficacy Analyses

### 8.8.1. Magnetic Resonance Imaging

MRIs will be evaluated by a central reader. The following analyses of MRI will use the Efficacy Analysis Set.

#### 8.8.1.1. Liver Fat Content by MRI Proton Density Fat Fraction (PDFF)

The PDFF metric is displayed as a median and interquartile range (IQR), the median will be used in all analyses.

Analyses of liver fat content (LFC) by MRI by MRI-PDFF will be based on the Efficacy Analysis Set. For the purposes of LFC analyses, absolute change is equivalent to change and relative change is equivalent to percent change (see Section 4.5). Descriptive statistics of the LFC by MRI-PDFF results, absolute change from baseline, and relative change will be summarized by treatment group and visit.

Absolute change (and separately for relative change) from baseline in LFC will be analyzed using the same ANCOVA model presented in Section 8.7.

Assessments of normality will be performed as described in Section 8.7.

LFC by MRI-PDFF responders will be defined as any patient with a relative decrease of at least 30% (i.e., ≤-30%). Non-responders will be defined as any patient who did not have at least a 30% relative decrease (i.e., >-30%). The number and percentage, with associated two-sided exact (Clopper-Pearson) 90% and 95% CIs, of patients in each category (response, nonresponse) will be presented by treatment group at Week 12. A Chi-square test Week 12 will be used to compare pooled TERN-101 (All TERN-101) and placebo, and each pairwise comparison between placebo and each active dose of TERN-101.

[redacted]

#### 8.8.1.2. cTI

Corrected T1 (cT1) will be collected along with MRI-PDFF at sites with this capability.

cT1 related to the amount of extracellular fluid present in the liver parenchyma. cT1 is derived [redacted] The cT1 metric is displayed as a median and IQR, the median will be used in all analyses.

cT1will be analyzed in the same manner as LFC.

cT1 will be categorized as low (<800 msec), elevated (800 – 875 msec), and high (>875 msec). In addition, cT1 responders will be defined any patient with a decrease of at least 80 msec at Week 6 and 12. Non-responders will be defined any patient who did not have at least an 80 msec decrease at Week 6 and 12. For both endpoints, the number and percentage, with associated two-sided exact (Clopper-Pearson) 90 and 95% CIs, of patients in each category will be presented by treatment group at each post-baseline visit. A Chi-square test at each post-baseline visit will be used to compare pooled TERN-101 (All TERN-101) and placebo, and each pairwise comparison between placebo and each active dose of TERN-101.

[redacted]

### 8.8.2. Liver Function Tests

LFTs for efficacy analyses include AST, ALT, and GGT. ALT analyses are described in Section 8.7). Descriptive statistics of LFT values, change from baseline, change from end of treatment, percent change from baseline, and percent change from end of treatment will be presented by treatment group for each visit using the Efficacy Analysis Set.

Change (and separately percent change) from baseline will be analyzed using the same ANCOVA model presented in Section 8.7 and the MMRM model in Section 8.7.1 for each post-

baseline and follow up visit using observed data only. Change (and separately percent change) from end of treatment in LFTs will be analyzed using the same ANCOVA model presented in Section 8.7. [redacted]

Figures of LS mean (±SE) AST and GGT change from baseline and percent change from the MMRM model will be presented over time by treatment group.

Assessments of normality will be performed as described in Section 8.7.

### 8.8.3. NASH and Fibrosis Biomarkers

Markers of hepatic fibrosis and apoptosis include CK-18 (M30 and M65), PIIINP, TIMP-1, HA, and Pro-C3. Descriptive statistics of biomarker values, change from baseline, and percent change from baseline will be presented by treatment group for each visit using the Efficacy Analysis Set.

Change (and separately percent change) from baseline will be analyzed using the same ANCOVA model presented in Section 8.7. Sensitivity analyses using the PP and Treatment Completer Analysis Sets will be presented if the analysis set differs from the Efficacy Analysis Set by more than 10 patients total.

Assessments of normality will be performed as described in Section 8.7.

[redacted]

#### 8.8.3.1. Fibrosis-4 (FIB-4) Score

The Fibrosis-4 (FIB-4) score helps to estimate the amount of scarring in the liver. The FIB-4 score is calculated as: [Age (years) × AST (U/L)] ÷ [Platelet Count ($10^9$/L) × ALT (U/L)$^{0.5}$]. Age at informed consent will be used. Lab values at each visit, from the same visit, will be used. The central lab calculation of FIB-4 will be used in analyses. Programming will also calculate the score programmatically for validation. Baseline AST, Platelets, and ALT will be the mean of all available evaluations, including any unscheduled or repeat assessments, prior to the first administration of study drug, per Section 4.7.2 The FIB-4 score will be calculated at baseline and at Week 6, Week 12, and Week 16.

The FIB-4 score will be analyzed in the same manner as specified in Section 8.8.3.

#### 8.8.3.2. Enhanced Liver Fibrosis (ELF) Score

The enhanced liver fibrosis (ELF) score will be calculated as: 2.278+0.851 ln (HA) +0.751 ln (PIIINP) +0.394 ln (TIMP-1). Lab values from the same visit will be used to calculate the ELF score. The central lab calculation of ELF will be used in analyses. Programming will also

calculate the score programmatically for validation. The ELF score will be calculated at baseline and post-baseline visits Week 6 and Week 12.

The ELF score will be analyzed in the same manner as specified in Section 8.8.3.

[redacted]

#### 8.8.3.3. Nonalcoholic Fatty Liver Disease (NAFLD) Fibrosis Score

The Nonalcoholic Fatty Liver Disease (NAFLD) fibrosis score will be calculated as: -1.675 + 0.037 × Age (years) + 0.094 × BMI ($kg/m^2$) + 1.13 × baseline diabetes status (yes = 1, no = 0) + 0.99 × AST/ALT ratio - 0.013 × Platelet Count ($10^9/L$) - 0.66 × Albumin (g/dL). Age at informed consent and baseline diabetes status will be used. BMI at each visit will be calculated using the weight from that visit and height from screening. The baseline weight will be used for Week 2 since weight is not collected at Week 2. Lab values at each visit, from the same visit, will be used. Programming will calculate the score at each to be used in analyses. The NAFLD fibrosis score will be calculated at baseline and at Week 6, Week 12, and Week 16.

The NAFLD fibrosis score will be analyzed in the same manner as specified in Section 8.8.3.

[redacted]

[redacted]

[redacted]

[redacted]

## 8.10. Pharmacodynamic (PD) Analyses

Samples for FGF19 are collected via intensive sampling according to the schedule of activities in Section 1.3 of the protocol. Samples for 7αC4 are collected via trough and intensive sampling according to the schedule of activities in Section 1.3 of the protocol.

[redacted]

### 8.10.1. Below the Quantification Level (BLQ) Values

Plasma concentrations for 7αC4, FGF19, [redacted] parameters that are BLQ will be set to ½ × BLQ.

### 8.10.2. Concentration-Time Data

Individual plasma concentration data for [redacted] 7αC4 will be listed for all patients (trough samples) and separately for PK/PD substudy patients only (7αC4 and FGF19 only, intensive samples). Concentrations will be summarized by nominal sampling time point and treatment group with descriptive statistics (number of non-missing observations, arithmetic mean, SD, median, minimum, maximum, geometric mean and geometric coefficient of variation [CV%]). The number of values BLQ will be presented for all patients and separately for PK/PD substudy patients (7αC4 and FGF19 only, intensive samples).

The following rules will be applied in the presentation of plasma concentration data if there are values that are BLQ or if there are missing values (e.g., no result [NR]) in a plasma concentration data series to be summarized.

- For the calculation of summary statistics, BLQ values will be set to ½ × BLQ.
- Where there is NR, these will be set to missing.
- If there are less than three values in the data series, only the minimum, maximum and number of observations will be presented. The other summary statistics will be denoted as not calculated (NC). BLQ is considered a value.
- If all the values are BLQ, then the arithmetic mean, arithmetic SD, median, minimum and maximum will be presented as ½ × BLQ, and the geometric mean and geometric CV% will be denoted as NC.

Individual patient and geometric mean concentration-time profiles, for each TERN-101 dose, will be presented graphically on linear and logarithmic concentration scales for PK/PD substudy patients for Day 1 and Week 12.

Concentration data in listings will be presented to the same number of significant figures as provided by the analytical laboratory. Descriptive statistics will be displayed according to the rules outlined in Section 3.2.

### 8.10.3. Change from Baseline

Descriptive statistics of [redacted], 7αC4, and FGF19 values, change from baseline, and percent change from baseline values will be presented by treatment group for each visit using the PD Analysis Set. Analyses of 7αC4 will include both trough and intensive samples.

Analyses of change and percent change from baseline will be carried out using an ANCOVA model at each post-baseline visit with change (percent change) from baseline as the dependent variable including treatment group as a fixed effect and baseline as a covariate. The primary analysis of change (percent change) will be carried out using available observed data only, without imputation for missing data.

The estimates of LS means, SE, and 90% and 95% CIs will be presented by treatment group. Estimates of the LS mean difference between pooled TERN-101 (All TERN-101) and placebo, and each pairwise comparison between placebo and each active dose of TERN-101, will be presented with the associated standard error of the difference, and 90% and 95% CIs of the difference.

The same ANCOVA model will be used to present the last on treatment value for Week 12, to be labelled as "End of Treatment". For patients who complete Week 12, the Week 12 value will be used. For patients who do not complete treatment, the last value on treatment will be used.

Figures of LS mean (±SE) changes from baseline and percent changes from baseline will be presented over time by treatment group.

The distribution of change (percent change) values will be evaluated visually for normality assumptions. If normality assumptions are violated, i.e. data are not normally distributed, a Wilcoxon Rank Sum Test will be used to compare each TERN-101 dose to placebo, and pooled TERN-101 (All TERN-101) to placebo. The median differences and 90% and 95% CI of the median differences between treatment groups will be constructed using Hodges-Lehmann estimate.

[redacted]

[redacted]

[redacted]

## 9. SAFETY ANALYSES

Safety analyses will be conducted using the Safety Analysis Set.

No inferential comparison of safety endpoints will be performed, unless otherwise specified.

[redacted]

[redacted]

Study drug administration including missed doses, delays, and overdoses, and study drug dispensing and accountability, calculated durations of exposure, and calculated compliances will be presented in by-patient data listings. [redacted]

## 9.2. Adverse Events

Adverse event (AE) verbatim terms on eCRFs will be mapped to preferred terms (PT) and system organ classes (SOC) using MedDRA. AE severity will be categorized using the National Cancer Institute Common Terminology Criteria for Adverse Events (NCI CTCAE, Version 5).

All AEs will be listed, but summaries will be restricted to TEAEs, defined as any AE with a start date on or after the date of first administration of study drug through 30 days after the last administration of study drug or through the Follow-Up Period (Week 16). Related SAEs after Week 16 will be considered TEAEs. If it cannot be determined whether an AE is treatment emergent due to a partial onset date, then it will be counted as such. Methodology for imputation of partial AE start and stop dates is described in Appendix 2. Each summary will be displayed by treatment group.

### 9.2.1. Overall Summary of Adverse Events

An overall summary of TEAEs will be presented by treatment group. The number and percentage of patients who experience at least one of the following:

- Any TEAE, All CTCAE grades
- Grade 3 or higher TEAE
- Serious TEAE
- TEAE leading to death
- Treatment-related TEAE, All CTCAE grades
- Treatment-related Grade 3 or higher TEAE
- Treatment-related serious TEAE
- TEAE leading to study drug interruption
- TEAE leading to study drug discontinued
- TEAE leading to discontinuation from the study

### 9.2.2. Patient Incidence of Adverse Events

Summaries will be displayed by SOC and PT, and will be ordered by descending order of all TERN-101 incidence of system organ class and preferred term within each system organ class. Summaries of the following types will be presented:

- Patient incidence of TEAEs by MedDRA SOC and PT. [redacted]
- Patient incidence of TEAEs by MedDRA SOC, PT, and highest severity (CTCAE grade). At each level of patient summarization, a patient is classified according to the highest severity if the patient reported 1 or more events. AEs with missing severity (CTCAE grade) will be considered Grade 3 (severe) for this summary.
- Patient incidence of CTCAE Grade 3 or higher TEAEs by MedDRA SOC and PT. At each level of patient summarization, a patient is classified according to the highest severity if the patient reported 1 or more events. AEs with missing severity (CTCAE grade) will be considered Grade 3 (severe) for this summary.
- Patient incidence of treatment-related TEAEs by MedDRA SOC and PT. Related AEs are those with relationships reported as "Definitely Related" or "Possibly Related". At each level of patient summarization, a patient is classified according to the closest relationship to study drug if the patient reported 1 or more events. AEs with a missing relationship will be considered related for this summary.
- Patient incidence of treatment-related CTCAE Grade 3 or higher TEAEs by MedDRA SOC and PT. At each level of patient summarization, a patient is classified according to the highest severity if the patient reported 1 or more events. AEs with missing severity (CTCAE grade) will be considered Grade 3 (severe) for this summary. AEs with a missing relationship will be considered related for this summary.
- Patient incidence of serious TEAEs by MedDRA SOC and PT.
- Patient incidence of TEAEs leading to study drug discontinued by MedDRA SOC and PT. This is a subset of the AEs where Action Taken with Study Drug is checked as "Study drug discontinued."

The following listings will be presented by treatment group and patient, include study day event started and the duration of event:

- All adverse events.
- Serious adverse events (subset of the AEs where serious is marked as "Yes").
- CTCAE Grade 3 or higher adverse events (subset of AEs where severity is marked as CTCAE Grade 3, 4, or 5).
- Related adverse events (subset of the AEs where relationship marked as "Definitely Related" or "Possibly Related").

- Adverse events leading to study drug interruption (subset of the AEs where Action Taken with Study Drug is checked as "Study drug interrupted").
- Adverse events leading to study drug discontinued (subset of the AEs where Action Taken with Study Drug is checked as "Study drug discontinued").
- Adverse events leading to death (subset of the AEs where outcome is indicated as "Fatal" or the CTCAE grade is 5 or seriousness criteria is Death).

### 9.2.3. Pruritus

For patients who report AEs consistent with pruritus at the visit, a qualified staff member will administer the Pruritus Numerical Rating Scale (see Appendix 8 of the study protocol). The Pruritus Numerical Rating Scale assesses a patient's severity of itching on average and at its worst point, in the past 24 hours, on a scale of 0 (no itching) to 10 (worst imaginable itch).

Treatment-emergent pruritus is defined as any preferred term including "Prur". Pruritus will be summarized overall with

An overall summary of pruritus will be presented by treatment group. The number and percentage of patients who experience at least one of the following:

- Any Pruritus TEAE, All CTCAE grades
- Pruritus TEAEs by Maximum Severity (Grade 1, 2, or 3)
- Serious Pruritus TEAE
- Treatment-related Pruritus TEAE, All CTCAE grades
- Treatment-related Grade 3 or higher Pruritus TEAE
- Treatment-related serious Pruritus TEAE
- Pruritus TEAE leading to study drug interruption
- Pruritus TEAE leading to study drug discontinued
- Pruritus TEAE leading to discontinuation from the study

A separate by-patient listing for pruritus events will be provided with associated Pruritus Numerical Rating Scale results, if applicable.

## 9.3. Clinical Safety Laboratory Evaluations

All laboratory assessments are specified in Appendix 1 of the study protocol. Laboratory parameters will be summarized in conventional units. Coagulation parameters, hemoglobin A1c, direct bilirubin, and indirect bilirubin will not be summarized.

Quantitative serum chemistry and hematology results will be summarized by treatment group using descriptive statistics at baseline and at each post-baseline visit. The change and percent change from baseline will also be summarized.

All clinical safety laboratory data will be presented in by-patient data listings. A separate listing for central laboratory normal ranges, by lab category, lab test, lab parameter, sex, and age, will be presented. Any local laboratory normal ranges for COVID-19 will be presented separately.

### 9.3.1. Treatment Emergent Abnormalities

Quantitative laboratory tests will be assigned grades based on National Cancer Institute Common Terminology Criteria for Adverse Events (NCI CTCAE) Version 5. For each applicable laboratory test, a treatment emergent abnormality will be defined as a post-baseline through Week 16 CTCAE grade increase of ≥2 grades.

The number and percentage of patients overall and by laboratory test, with any treatment emergent laboratory abnormality, CTCAE increase of ≥2 grades, will be presented. Similarly, the number and percentage of patients overall and by laboratory test, with CTCAE increase of ≥1 grade will be presented.

Estimated glomerular filtration rate (eGFR) is calculated by the central lab using the Chronic Kidney Disease Epidemiology Collaboration (CKD-EPI) equation. eGFR does not have a normal range within the central lab data. A normal range will be defined as ≥90 mL/min/1.73m$^2$.

LDL-c cannot be graded via NCI CTCAE version 5. For the purposes of treatment emergent abnormalities, LDL will be graded as follows:

- Optimal: < 100 mg/dL
- Near or Above Optimal: 100 – 129 mg/dL
- Borderline High: 130 – 159 mg/dL
- High: 160 – 189 mg/dL
- Very High: >189 mg/dL

LDL-c treatment emergent abnormalities will be summarized separately.

Listings will be presented for any patient with any treatment emergent laboratory abnormality (any increase in CTCAE grade). Normal ranges provided by the central laboratory will be presented in a listing.

### 9.3.2. Special Assessments of Potential Hepatotoxicity

Number and percentage of patients meeting each of the below criteria in Table 9 will be summarized by treatment group.

Modified Hy's rule by Bjornsson (Bjornsson 2005), the incidence of patients with ALT or AST values >3x ULN and total bilirubin >2x ULN (based on the same blood draw of ALT, AST, and total bilirubin while on study drug, or within 30 days of last received dose) will be used to assess the potential for study drug-induced liver toxicity. The incidence rates of this laboratory abnormality constellation will be presented. Any patient meeting any of the criteria specified in Table 9 will be presented.

**Table 9: Modified Hy's Rule Laboratory Abnormalities**

| Hepatocellular Injury and Liver Function Tests | Range of Test Values |
|---|---|
| AST | 3 to 4.99 × ULN |
| | 5 to 9.99 × ULN |
| | 10 to 19.99 × ULN |
| | ≥20 × ULN |
| ALT | 3 to 4.99 × ULN |
| | 5 to 9.99 × ULN |
| | 10 to 19.99 × ULN |
| | ≥20 × ULN |
| AST or ALT | 3 to 4.99 × ULN |
| | 5 to 9.99 × ULN |
| | 10 to 19.99 ×ULN |
| | ≥20 × ULN |
| Total bilirubin | > ULN |
| | >1.5 × ULN |
| | >2 × ULN |
| AST or ALT plus total bilirubin | >3 × ULN (AST or ALT) plus >1.5 × ULN (total bilirubin) |
| AST or ALT plus total bilirubin | >3 × ULN (AST or ALT) plus >2 × ULN (total bilirubin) |
| Alkaline phosphatase | >1.5 × ULN |
| ALT = alanine aminotransferase; AST = aspartate aminotransferase; ULN = upper limit of normal | |

[redacted]

LFT values in units of ULN will be calculated as the result/upper limit of normal.

### 9.3.3. Fasting Glucose and Lipids

Analyses of fasting glucose and lipid parameters (LDL-c, HDL-c, VLDL-c, total cholesterol, triglycerides) will be carried out using an ANCOVA model at each visit with change (and percent change [except for glucose]) from baseline as the dependent variable including treatment group as a fixed effect and baseline as a covariate. The same analysis will be carried out using change from baseline as the dependent variable. The estimates of least-square (LS) means, standard errors, and 90% and 95% CIs will be presented by treatment group. Estimates of the LS mean difference between pooled TERN-101 (All TERN-101) and pooled placebo, and each pairwise comparison between pooled placebo and each active dose of TERN-101, will be presented with the associated standard error of the difference, and 90% and 95% CI of the difference, the standard error of the difference, and 90% and 95% CI of the difference will be presented.

Figures of mean (±SD) LDL and HDL values will be presented over time by treatment group. In addition, figures of LS mean (±SE) changes from baseline will be presented over time by treatment group.

[redacted]

### 9.3.4. Shifts in CTCAE Grade

Shifts in CTCAE grade of laboratory tests will be presented from baseline to worst post-baseline value and at each post-baseline visit. Summaries will present the number and percentage of patients with shifts in laboratory grade by treatment group. Denominators for percentages will be the number of patients with non-missing data at the specific assessment and baseline.

LDL-c will be summarized separately using the categories defined in Section 9.3.1.

Some lab tests include additional clinical criteria which cannot be determined programmatically. Table 11 below specifies the grade with clinical criteria in italics. The text in italics will not be considered when assigning grades programmatically.

**Table 11: CTCAE Laboratory Tests with Clinical Criteria**

| |
|---|
| Bicarbonate (Blood bicarbonate decreased):<br>• Grade 1: <LLN *and no intervention initiated.* |
| Prothrombin Time (PT)/International Normalized Ratio (INR) (INR increased):<br>• Grade 1: >1.2 - 1.5; *>1 - 1.5 x baseline if on anticoagulation; monitoring only indicated*<br>• Grade 2: >1.5 - 2.5; *>1.5 - 2.5 x baseline if on anticoagulation; dose adjustment indicated*<br>• Grade 3: >2.5; *>2.5 x baseline if on anticoagulation; bleeding* |

| |
|---|
| eGFR (Chronic Kidney Disease):<br>• Grade 4: <15 ml/min/1.73m$^2$; *dialysis or renal transplant indicated* |
| Glucose (Hypoglycemia):<br>• Grade 4: <30 mg/dL; <1.7 mmol/L; *life-threatening consequences; seizures* |
| Potassium (Hypokalemia)<br>• *Grade 2: Symptomatic with <LLN - 3.0 mmol/L; intervention indicated*<br>• Grade 3: <3.0 - 2.5 mmol/L; *hospitalization indicated*<br>• Grade 4: <2.5 mmol/L; *life-threatening consequences* |
| Potassium (Hyperkalemia):<br>• Grade 2: >5.5 - 6.0 mmol/L; *intervention initiated*<br>• Grade 3: >6.0 - 7.0 mmol/L; *hospitalization indicated*<br>• Grade 4: >7.0 mmol/L; *life-threatening consequences* |
| Sodium (Hyponatremia):<br>• Grade 2: 125-129 mmol/L *and asymptomatic*<br>• Grade 3: *125-129 mmol/L symptomatic;* 120-124 mmol/L regardless of symptoms<br>• Grade 4: <120 mmol/L; *life-threatening consequences* |
| Sodium (Hypernatremia):<br>• Grade 2: >150 - 155 mmol/L; *intervention initiated*<br>• Grade 3: >155 - 160 mmol/L; *hospitalization indicated*<br>• Grade 4: >160 mmol/L; *life-threatening consequences* |
| Albumin (Hypoalbuminemia):<br>• *Grade 4: Life-threatening consequences; urgent intervention indicated* |

### 9.3.6. COVID-19

SARS-CoV-2 test for active infection and SARS-CoV-2 antibody test results, central and local labs, will be presented in a by-patient listing. A separate listing for central and local laboratory normal ranges, by lab test, lab parameter, sex, and age, will be presented.

At study initiation, the EDI Epitope assay was used for SARS-CoV-2 antibody testing. During the study, the EDI Epitope assay was discontinued. Where possible, samples were re-assayed using the Abbott assay. Both assay results are included in the central laboratory results. If both results are available, the Abbott assay results will be reported instead of the EDI Epitope assay results.

The number and percent of patients with any positive COVID-19 test result (i.e., either test method) and separately by test method at baseline and at any post-baseline visit will be presented by treatment group. Antibody test results will be further separated by IgG and IgM.

[redacted]

## 11. CHANGES TO PROTOCOL PLANNED ANALYSES

The Enrolled Analysis Set was renamed to Screened Analysis Set.

The Safety Analysis Set was changed to remove 'randomized', as per regulatory guidance, any patient who receives study drug, regardless of whether the patient was randomized or not, will be included in analyses. In addition, as all patients are evaluated for safety post dose (e.g. AEs), the specification of 'and have post dose safety data' was removed.

Study endpoints were further clarified.

[redacted]

# 12. REFERENCES

Björnsson. Drug-induced liver injury: Hy's rule revisited. Clinical Pharmacology Therapeutics. 2006;79(6):521-8.

Bornkamp, et. al., MCPMod: An R Package for the Design and Analysis of Dose-Finding Studies, Journal of Statistical Software, February 2009, Volume 29, Issue 7.

Guidance for Industry *E3: Statistical Principles for Clinical Trials*. U.S. Department of Health and Human Services, Food and Drug Administration, Center for Drug Evaluation and Research (CDER), Center for Biologics Evaluation and Research (CBER), September 1998.

Guidance for Industry *E9: Statistical Principles for Clinical Trials*. U.S. Department of Health and Human Services, Food and Drug Administration, Center for Drug Evaluation and Research (CDER), Center for Biologics Evaluation and Research (CBER), September 1998.

Guidance for Industry *E14: The Clinical Evaluation of QT/QTc Interval Prolongation and Proarrhythmic Potential for Non-antiarrhythmic Drugs*. U.S. Department of Health and Human Services, Food and Drug Administration, Center for Drug Evaluation and Research (CDER), October 2005.

Guidance for Industry: *Drug Induced Liver Injury: Premarketing Clinical Evaluation*, CDER, FDA (2009).

Guidance for Industry: *Providing Regulatory Submissions in Electronic Format — Certain Human Pharmaceutical Product Applications and Related Submissions Using the eCTD Specifications*, Center for Drug Evaluation and Research (CDER), Center for Biologics Evaluation and Research (CBER), September 2016.

# 13. APPENDICES

[redacted]

[redacted]

[redacted]

[redacted]



[redacted]

# APPENDIX 3. IMPUTATION RULES FOR MISSING OR PARTIAL DATES FOR ADVERSE EVENTS AND CONCOMITANT MEDICATIONS

| Date | Situation | Imputation Rule |
|---|---|---|
| Start Date | Only month and year are known and month and year are prior to first dose date | Use the last day of the month |
| | Only month and year are known and month and year are the same as first dose date | Use the first study drug administration date |
| | Only month and year are known and month and year are after first dose date | Use the first day of the month |
| | Only year is known and year is before first dose date | Use Dec 31 of that year |
| | Only year is known and year is same as first dose date | Use the first study drug administration date |
| | Only year is known and year is after first dose date | Use Jan 1 of that year |
| | Entire date is missing | Use the first study drug administration date |
| | The estimated start date is after a complete or imputed AE stop date | Use the first day of the month of the AE stop date |
| Stop/End Date | Only month and year are known and month and year are prior to last dose date | Use the last day of the month |
| | Only month and year are known and month and year are the same as last dose date | Use the last dose date |
| | Only month and year are known and month and year are after last dose date | Use the last day of the month |
| | Only year is known and year is before last dose date | Use Dec 31 of that year |
| | Only year is known and year is same as last dose date | Use the last study drug administration date |
| | Only year is known and year is after last dose date | Use Dec 31 of that year |
| | Entire date is missing | Use the last study drug administration date |
| | The estimated stop date is before a complete or imputed AE start date | Use the last day of the month of the AE start date |

AE = adverse event, Dec = December, Jan = January

# APPENDIX 4. ABBREVIATIONS

| Abbreviation or Special Term | Explanation |
|---|---|
| 7αC4 | 7α-hydroxy-4-cholesten-3-one |
| ADaM | Analysis Data Model |
| AE | adverse event |
| ALP | alkaline phosphatase |
| ALT | alanine transaminase |
| ANCOVA | analysis of covariance |
| AST | aspartate transaminase |
| ATC | Anatomical Therapeutic Class |
| AUC | area under the concentration-time curve |
| BLQ | below limit of quantification |
| BMI | body mass index |
| CAP | Controlled Attenuation Parameter |
| CDISC | Clinical Data Interchange Standards Consortium |
| CHG | change from baseline ADaM basic dataset structure defined variable |
| CI | confidence interval |
| CK-18 | cytokeratin-18 |
| COVID-19 | Coronavirus disease 2019 |
| $C_{max}$ | maximal concentration |
| CRF | case report form |
| CSR | clinical study report |
| CTCAE | Common Terminology Criteria for Adverse Events |
| CTMS | clinical trial management system |
| CV | coefficient of variation |
| DPP-4 | Dipeptidyl peptidase 4 |
| eGFR | estimated glomerular filtration rate |
| eCRF | electronic case report form |
| ECG | electrocardiogram |
| ET | early termination |
| FDA | Food and Drug Administration |
| FGF | fibroblast growth factor |
| FIB-4 | Fibrosis-4 Score |
| FSH | follicle stimulating hormone |
| FXR | farnesoid X receptor |
| GGT | gamma-glutamyl transpeptidase |
| GLP-1 | Glucagon-like peptide-1 |
| HA | hyaluronic acid |
| HDL | high-density lipoprotein |
| hs-CRP | high sensitivity C-reactive protein |
| ICF | informed consent form |
| ICH | International Council for Harmonisation |

| Abbreviation or Special Term | Explanation |
|---|---|
| IG | implementation guide |
| IMP | investigational medicinal product |
| INR | international normalized ratio |
| IWRS | interactive web response system |
| LDL | low-density lipoprotein |
| LFC | liver fat content |
| LFT | liver function test |
| LLN | lower limit of normal |
| ln | natural log |
| LS | least squares |
| MedDRA | Medical Dictionary for Regulatory Activities |
| MMRM | mixed model for repeated measures |
| MRI-PDFF | Magnetic Resonance Imaging Proton Density Fat Fraction |
| NAFLD | non-alcoholic fatty liver disease |
| NAS | non-alcoholic fatty liver disease activity score |
| NASH | non-alcoholic steatohepatitis |
| NC | not calculated |
| NCA | noncompartmental analysis |
| NCI | National Cancer Institute |
| PCHG | percent change from baseline ADaM basic dataset structure defined variable |
| PIIINP | procollagen III n-terminal propeptide |
| PD | pharmacodynamics |
| PK | pharmacokinetics |
| Pro-C3 | pro-peptide of type III collagen |
| Q1 | 25$^{th}$ percentile, first quartile |
| Q3 | 75$^{th}$ percentile, third quartile |
| REML | restricted maximum likelihood |
| SAE | serious adverse event |
| SAP | statistical analysis plan |
| SARS-CoV-2 | severe acute respiratory syndrome coronavirus 2 |
| SD | standard deviation |
| SDG | Standardized Drug Grouping |
| SDTM | Study Data Tabulation Model |
| SE | least squares mean standard error |
| SGLT-2 | Sodium-glucose co-transporter 2 |
| TEAE | treatment emergent adverse event |
| TERN-101 | investigational agonist of Farnesoid X receptor |
| TIMP-1 | tissue inhibitor of metalloproteinases-1 |
| ULN | upper limit of normal |
| WHODD | World Health Organization Drug Dictionary |

# APPENDIX 5. DOCUMENT HISTORY

| Version | Date | Author | Description |
|---|---|---|---|
| 1.0 | 19MAY2021 | [redacted] | Final Version 1 |

# APPENDIX 6. APPROVAL PAGE

I confirm that I have reviewed this document and agree with the content.

| APPROVALS | | |
|---|---|---|
| *Terns, Inc. (Sponsor)* | | |
| [redacted] | | 5/19/2021 |
| [redacted] | | **Date** (dd-mmm-yyyy)<br>5/19/2021 |
| [redacted] | | **Date** (dd-mmm-yyyy)<br>5/19/2021 |
| [redacted] | | **Date** (dd-mmm-yyyy) |
| *ntract Research Organization)* | | |
| [redacted] | | 5/19/2021 |
| [redacted] | | **Date** (dd-mmm-yyyy)<br>5/19/2021 |
| [redacted] | | **Date** (dd-mmm-yyyy) |