GT-031 (NAVIGATE) Statistical Analysis Plan

Study Title: A Seamless, Adaptive, Phase 2b/3, Double-Blind, Randomized, Placebo-controlled, Multicenter, International Study Evaluating the Efficacy and Safety of Belapectin (GR-MD-02) for the Prevention of Esophageal Varices in NASH Cirrhosis

NCT #: NCT04365868

Document Date: 13 November 2024

Statistical Analysis Plan

Galectin Therapeutics Inc.

Protocol number: GT-031 (NAVIGATE)

A Seamless, Adaptive, Phase 2b/3, Double-Blind, Randomized, Placebo-controlled, Multicenter, International Study Evaluating the Efficacy and Safety of Belapectin (GR-MD-02) for the Prevention of Esophageal Varices in NASH Cirrhosis

[redacted]

Document Version: [redacted]

Document Date : [redacted]

[redacted]

[redacted]

Table of Contents

Reviewers

The following reviews of the SAP were conducted:

Name and Title	Role	Version Last Reviewed	Company/ Organization
[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]

Approvals

The undersigned agree that all required reviews of this document are complete, and approve this Statistical Analysis Plan as final. Programming of the tables, figures and listings based upon the specifications within this document can proceed.

Approved by

Galectin Approval:

Signature Date

[redacted] / Chief Medical Officer
Printed Name/Title

Glossary of Abbreviations

Abbreviation	Term
AEs	Adverse Events
ALB	Albumin
ALP	Alkaline Phosphatase
ALT	Alanine Aminotransferase
AR	Autoregressive
AST	Aspartate Aminotransferase
ATC	Anatomical Therapeutic Chemical
AUDIT	Alcohol Use Disorder Identification Test
BMI	Body Mass Index
CEC	Clinical Events Committee
CI	Confidence Interval
CLDQ	Chronic Liver Disease Questionnaire
CMH	Cochran-Mantel-Haenszel
CS	Compound Symmetry
CTCAE	Common Terminology Criteria for Adverse Event
CTP	Child-Turcotte-Pugh
DILI	Drug-Induced Liver Injury
DSMB	Data Safety and Monitoring Board
EAIR	Exposure-Adjusted Incidence Rates
ECG	Electrocardiogram
eCRF	Electronic Case Report Form
eDISH	Evaluation of Drug Induced Serious Hepatotoxicity
EGD	Esophagogastroduodenoscopy
EOPH2b	End of Phase 2b
EOS	End of Study
ET	Early Termination
FAS	Full Analysis Set
FDA	Food and Drug Administration
GLP-1	Glucagon-like Peptide-1
HRQOL	Health-Related Quality of Life
IA	Interim Analysis
ICE	Intercurrent Event
INR	International Normalized Ratio
KM	Kaplan-Meier
LBM	Lean Body Mass
LFT	Liver Function Test
LSMs	Least Square Means
MACE	Major Adverse Cardiac Event
MAR	Missing at Random

MedDRA	Medical Dictionary for Regulatory Activities
MELD	Model for End-stage Liver Disease
mFAS	Modified Full Analysis Set
MMRM	Mixed Effects Model for Repeated Measurements
NASH	Non-Alcoholic Steatohepatitis
NCI	National Cancer Institute
NSBB	Non-selective Beta-blocker
OLE	Open-label Extension
PHG	Portal Hypertensive Gastropathy
PK	Pharmacokinetic
PKAP	PK analysis plan
PPS	Per-Protocol Set
PT	Preferred Term
QTcB	Bazett Corrected QT Interval
QTcF	Fridericia Corrected QT Interval
SAE	Serious Adverse Event
SAF	Safety Set
SAP	Statistical Analysis Plan
SD	Standard Deviation
SOC	System Organ Class (MedDRA classification)
SSR	Sample Size Re-estimation
TB	Total Bilirubin
TEAE	Treatment-Emergent AE
TESAE	Treatment-Emergent SAE
TF	Treatment Failure
TLFB	Timeline Follow-back
WHO	World Health Organization
VCTE	Vibration-Controlled Transient Elastography

1. Source Documents

The Statistical Analysis Plan (SAP) was written based on the following documentation:

Document	Date	Version
Protocol	[redacted]	[redacted]
EGD Independent Review Charter	[redacted]	[redacted]
eCRF	[redacted]	[redacted]

2. Protocol Details

2.1 Study Objectives

Primary Efficacy Objective

To evaluate the efficacy of 2 mg/kg and 4 mg/kg lean body mass (LBM) of belapectin (GR-MD-02) compared to placebo in preventing the development of esophageal varices.

Secondary Efficacy Objectives

- Key Secondary Efficacy Objective
  - To evaluate the effect of belapectin on composite clinical outcomes
- Additional Secondary Efficacy Objectives
  - To evaluate the effect of belapectin on progression of esophageal varices
  - To evaluate the effect of belapectin treatment on liver cirrhosis related clinical events

Exploratory Efficacy Objectives

- To evaluate the effect of belapectin compared to placebo on liver fibrosis
- To assess the effect of belapectin compared to placebo on health-related quality of life (HRQOL)

Safety Objective

- To assess the safety and tolerability of belapectin compared to placebo treatment

Pharmacokinetics (PK) Objective

- To produce a population PK model of belapectin

Note that the PK analysis will be covered in a separate population PK analysis plan (PKAP).

2.2 Overall Study Design

The original study design was for a seamless, adaptive, two-stage, Phase 2b/3, randomized, double-blind, multicenter, parallel group, placebo-controlled study to assess the efficacy, safety, and tolerability of belapectin compared with placebo in subjects with non-alcoholic steatohepatitis (NASH) cirrhosis and clinical signs of portal hypertension but without esophageal varices at baseline.

Based on the feedback provided by the Food and Drug Administration (FDA) on 06SEP2024, the sponsor has decided to designate the pre-planned unblinded interim analysis (IA) as the primary analysis read out for the study. The Data Safety and Monitoring Board (DSMB) will meet as scheduled in December 2024, review the final data for the Phase 2b portion of the study, and make a recommendation with regards to dose selection for a modified Phase 3 portion of the study. The modified Phase 3 portion of the study will be an open-label extension (OLE) for all subjects currently ongoing in the Phase 3 (36-month) rollover until completion of the 36-month infusion or subject discontinuation, whichever comes first. All subjects ongoing in Phase 3 after the final DSMB meeting will switch to the DSMB recommended dose. Further details regarding the changes to the study design and the modified Phase 3 portion can be found in the current study protocol.

2.3 Sample Size and Power

As noted in section 2.2, per the feedback provided by the FDA on 06SEP2024, the pre-planned unblinded IA is being transformed to a primary analysis read out. Although the originally planned sample size for the combined data for Phase 2b and Phase 3 will not be achieved, many of the analysis details will remain the same. Specific changes to the planned analyses to adjust to the changes in the study have been updated throughout the document.

The EAST software (version 6.5, Cytel Inc.) was used for the initial sample size estimations.

Based on a prior belapectin Phase 2 study (GT-026, also known as NASH-CX), we assume the annual incidence rate is [redacted]% for subjects to develop new esophageal varices in the placebo treatment group and is [redacted]% for the belapectin 2 mg/kg LBM or 4 mg/kg LBM group (i.e., [redacted]% reduction in the incidence rate compared to placebo treatment). Derived from the annual incidence rates, the anticipated proportion of subjects developing new esophageal varices at 78 weeks [18 months] is [redacted]% for placebo and [redacted]% for active drug, either the belapectin 2 mg/kg LBM or 4 mg/kg LBM groups. The relative risk reduction from placebo treatment at 78 weeks [18 months] is anticipated to be at least [redacted]%.

For estimating the subject sample size for the primary efficacy endpoint (the incidence of new varices at 78 weeks [18 months]), the following assumptions were made:

1. The anticipated proportion of subjects developing esophageal varices in the placebo group at 78 weeks [18 months] is [redacted]%
2. The anticipated proportion of subjects developing esophageal varices on belapectin (either 2 mg/kg LBM or 4 mg/kg LBM) at 78 weeks [18 months] is [redacted]%
3. Significance level (Type I error) [redacted], 1-sided
4. The randomization ratio is 1:1:1 for 2 mg/kg LBM, 4 mg/kg LBM, and placebo groups, respectively.

Based on the assumptions above, a total subject sample size of 465 would provide at least [redacted]% power for demonstrating that either the belapectin 2 mg/kg LBM or 4 mg/kg LBM treatment is superior to placebo with respect to the primary efficacy endpoint for this study.

As a result of the conversion of the planned IA into a primary analysis read out, the sample size for the primary analysis read out is 357 subjects. This provides approximately [redacted]% power for demonstrating that either the belapectin 2 mg/kg LBM or 4 mg/kg LBM treatment is superior to placebo with respect to the primary efficacy endpoint for this study.

The assumed yearly incidence rate of new varices of [redacted]% is based on data from GT-026 in a subset of subjects without esophageal varices at baseline who subsequently developed varices over 1 year of follow-up (mean rate = [redacted]%; [redacted]% confidence interval [CI]: [redacted]%]).

With the knowledge we have gained about the subject population being studied, it is possible that the overall event rate is more than the estimated [redacted]% we used for the initial sample size calculations. Based on recently published studies in similar patient populations the overall event rate for this subject population could exceed more than [redacted]%.[1]

Given that the actual incidence rate of new varices in this Phase 2b study may be different, a blinded adaptive modification of the subject sample size was planned:

- When at least [redacted]% of the subjects planned for Phase 2b completed 78 weeks (18 months) of treatment, a non-comparative (blinded) sample size re-estimation (SSR) procedure was scheduled to occur. If the observed incidence rate of new varices was less than initially assumed, up to a maximum of 65 additional subjects in each treatment arm could have been enrolled without requiring a protocol amendment, that is, a maximum of 510 ([105+65]×3) subjects could have been enrolled in Phase 2b.

3. Efficacy and Safety Variables

3.1 Primary Efficacy Endpoint

The primary efficacy endpoint is the proportion of subjects in the belapectin treatment groups who develop esophageal varices at 78 weeks (18 months) of treatment, compared to placebo. The primary efficacy endpoint will be adjudicated by a blinded clinical events committee (CEC) to confirm the development of esophageal varices based on the established objective criteria in the CEC Charter.

3.2 Key Secondary Efficacy Endpoint

The key secondary efficacy endpoint is the 78-week cumulative incidence rate of subjects (i.e., proportion of subjects with event over 78 weeks of observation) in the belapectin treatment groups, compared to placebo, who develop any of the following liver-related clinical outcome events:

- varices (esophageal or gastric) requiring treatment
- variceal bleed requiring hospitalization
- clinically significant ascites requiring hospitalization
- spontaneous bacterial peritonitis
- overt hepatic encephalopathy (West Haven score ≥2 and requiring hospitalization)
- mortality (all-cause)
- liver transplant
- model for end stage liver disease (MELD) score ≥15 on 2 consecutive occasions

3.3 Additional Secondary Efficacy Endpoints

The additional secondary efficacy endpoints include:

- The 78-week cumulative incidence rate of subjects in the belapectin treatment group who progress to large esophageal varices or develop red wales compared to placebo
- Event-free survival by time to first cirrhosis related clinical event from the first infusion of study drug in the belapectin treatment group compared to placebo, including the following:
    - progression to large varices or red wales
    - esophageal variceal hemorrhage requiring hospitalization
    - clinically significant ascites requiring hospitalization
    - spontaneous bacterial peritonitis

- overt hepatic encephalopathy (West Haven score ≥2 and requiring hospitalization)
- Child-Turcotte-Pugh (CTP) score increase of ≥2 points (from baseline) on 2 consecutive occasions
- MELD score increases to ≥15 as measured on 2 consecutive occasions
- liver transplant
- liver-related death

3.4 Exploratory Efficacy Endpoints

The exploratory efficacy endpoints include:

- Change in liver stiffness measurement, baseline-adjusted, as determined by vibration-controlled transient elastography (VCTE) (FibroScan) exams during Phase 2b and Phase 3
- Difference in Chronic Liver Disease Questionnaire (CLDQ) scores between belapectin and placebo treatment during Phase 2b and Phase 3

3.5 Safety Variables

The safety endpoints include:

- Incidence and nature of adverse events (AEs), including adjudicated drug-induced liver-injury (DILI) and adjudicated cardiovascular events (Major Adverse Cardiac Event [MACE])
- Changes from baseline in physical examinations, vital signs, and electrocardiogram (ECG) parameters
- Changes from baseline in routine central laboratory tests (clinical chemistry, hematology), including liver enzymes (alanine aminotransferase [ALT], aspartate aminotransferase [AST], gamma glutamyl transferase [GGT], alkaline phosphatase [ALP)], and liver function markers (bilirubin -total and fractions-, albumin, international normalized ratio [INR])

4. Analysis Populations

Per the initial study design, subjects who were enrolled in Stage 1 (Phase 2b) would roll over to Stage 2 (Phase 3/36-month rollover) once they completed their 78 weeks of treatment and would continue for another 78 weeks of follow-up during Stage 2 for a total follow-up period of up to 156 weeks (i.e., 36 months).

Among subjects who rolled over to Phase 3, those who received the selected optimal belapectin treatment or placebo during Stage 1 would continue the same treatment during Stage 2, and those who receive the non-selected belapectin treatment during Stage 1 would switch to the selected optimal belapectin treatment after the IA. Subjects who were newly enrolled in Stage 2 (Phase 3) would have a maximum follow-up of 78 weeks (i.e., 18 months).

Given the change in study design, subjects that are ongoing in Stage 2 (Phase 3/36-month rollover)) will switch to the final DSMB recommended dose and will continue the study as part of an OLE until completion of 36 months. No new subjects will be enrolled into Stage 2 (Phase 3/36-month rollover) after the final DSMB meeting.

The following main analysis populations will be included for this study.

4.1 Full Analysis Set

The Full Analysis Set (FAS) consists [redacted].

4.2 Modified Full Analysis Set

The Modified Full Analysis Set (mFAS) [redacted].

4.3 Per Protocol Set

[redacted]

4.3.1 Important Protocol Deviations Leading to Exclusion from the PPS Analysis

Important protocol deviations considered to have a major effect on the efficacy analysis will lead to complete exclusion of the subjects from the PPS, and may include the following:

- Key inclusion/exclusion criteria deviations
- Treatment compliance does not fall within the range of 80% to 120% (treatment compliance specified in Section 6.6)
- Use of prohibited medications as defined by the protocol and in discussion with the Sponsor and/or Medical Monitor
- Errors in treatment allocation, including allocation of an incorrect treatment from the randomization schedule, and receipt of the wrong treatment at one or more study visits due to packaging or dispensing errors

All important protocol deviations leading to exclusion from the PPS during the study will be reviewed and approved by the Sponsor prior to database lock and unblinding. Should additional important protocol deviations leading to exclusion from the PPS, not anticipated at the time of preparing this SAP, be identified during the study and prior to unblinding, they will be documented in a SAP amendment or other document and included in all relevant protocol deviation reviews and approvals.

4.4 Safety Set

The Safety Set (SAF) includes all randomized subjects who receive at least one infusion of study drug in Phase 2b. The SAF will be the primary analysis population for safety analyses.

4.5 PK Set

The PK Set includes all randomized subjects who receive at least one dose of study drug and contribute PK measurements to allow for calculation of PK parameters.

5. DATA Handling

5.1 Visit Windows

Day 1 is defined as the day of first infusion of study drug. Relative day on or after Day 1 is calculated as (assessment date – Day 1 date) + 1. Relative day prior to Day 1 is calculated as (assessment date – Day 1 date). The day prior to Day 1 is Day -1.

To allow for variations in scheduling and use of most data collected, the visit windows defined for the following endpoints shown in Table 5.1-1 will be used to assign evaluations to a most appropriate nominal visit for by-visit analysis of these endpoints for subjects enrolled in Phase 2b, respectively. By-visit analysis of all other assessments that are not included in Table 5.1-1, such as vital signs, will use the nominal study visit as defined in the Schedule of Assessments from the study protocol and the electronic case report form (eCRF).

Table 5.1-1 Definition of visit windows for subjects enrolled in Phase 2b

				Visit Window		
Visit	Visit Number	Target Day of Visit	Safety laboratory parameters	12-lead ECG	FibroScan	EGD
Baseline		--	≤1	≤1	≤1	≤1
Week 0	Ph2b - 1	Day 1	1-42	1-84	1-91	
Week 12	Ph2b - 7	Day 85	43-126	--	--	
Week 24	Ph2b - 13	Day 169	127-210	85-252	--	
Week 26	Ph2b - 14	Day 183	--	--	92-273	
Week 36	Ph2b - 19	Day 253	211-294	--	--	
Week 48	Ph2b - 25	Day 337	295-378	253-441	--	
Week 52	Ph2b - 27	Day 365	--	--	274-455	
Week 60	Ph2b - 31	Day 421	379-462	--	--	
Week 72	Ph2b - 37	Day 505	463-525	--	--	
Week 78	Ph2b - 40	Day 547	526 to EOPH2B[a]	442 to EOPH2B[a]	456 to EOPH2B[a]	>1 to EOPH2B[a]
Week 90	Ph3 - 7	Day 631	SOPH3 to 672[b]	--	--	
Week 102	Ph3 - 13	Day 715	673-763	SOPH3 to 798[b]	--	
Week 104	Ph3 - 14	Day 729	--	--	SOPH3 to 819[b]	
Week 114	Ph3 - 19	Day 799	764-840	--	--	
Week 126	Ph3 - 25	Day 883	841-938	799-987	--	
Week 130	Ph3 - 27	Day 911	--	--	820-1001	
Week 138	Ph3 - 31	Day 967	939-1008	--	--	
Week 150	Ph3 - 37	Day 1051	1009-1071	--	--	
Week 156	Ph3 - 40	Day 1093	1072 to EOPH3[a]	988 to EOPH3[a]	1002 to EOPH3[a]	SOPH3 to EOPH3[a]

[a] For the last nominal visit in each phase (Phase 2b -Visit 40 or Phase 3 -Visit 40) the visit will be included in the analysis irrespective of the relative day of the visit.
[b] For the first nominal visit in Phase 3, the visit will be included in the analysis for Phase 3 irrespective of the relative day of the visit.
EOPH2B = End of Phase 2b, EOPH3= End of Phase 3, SOPH3 = Start of Phase 3

If there are multiple visits (including both scheduled and unscheduled visits) within a visit window, then the visit closest to the target day of the visit window will be used in the analysis. If there is a tie, the later visit will be used in the analysis.

5.2.3 Incomplete dates for AEs and concomitant medications

Incomplete dates for AEs and concomitant medications will be imputed to determine the treatment-emergent adverse events (TEAEs) or prior/concomitant medication flags.

Imputation of incomplete AE dates

Imputation rules for missing or partial AE start date are defined below:

1. If only Day of AE start date is missing:
    - If the AE start year and month are the same as that for the first infusion date, then:
        - If the full (or partial) AE end date is NOT before the first infusion date or AE end date is missing, then impute the AE start day as the day of first infusion date
        - otherwise, impute the AE start Day as 1
    - Otherwise, impute the AE start Day as 1
2. If Day and Month of AE start date are missing:
    - If AE start year = first infusion year, then:
        - If the full (or partial) AE end date is NOT before the first infusion date or AE end date is missing, then impute the AE start Month and Day as the Month and Day of first infusion date
        - otherwise, impute the AE start Month as January and Day as 1
    - Otherwise, impute the AE start Month as January and Day as 1
3. If Day, Month, and Year of AE start date are all missing:
    - If the full (or partial) AE end date is NOT before the first infusion date or AE end date is missing, then impute the AE start date as the first infusion date
    - Otherwise, impute the AE start Year as year of the first infusion, Month as January, and Day as 1.

Imputation of incomplete prior and concomitant medication dates

1. For medications with partial start dates, missing month will be imputed with January and missing day will be imputed with 1
2. For medications with partial stop dates, missing month will be imputed with December and missing day will be imputed with the last day of the month

6. Statistical Methods

6.1 General Principles

Unless specified otherwise, data will be displayed using the following treatment group labels, in the order presented:

- placebo
- belapectin 2 mg/kg LBM
- belapectin 4 mg/kg LBM
- overall (all subjects) (if applicable)

All data processing, summarization and analyses will be performed using SAS® Version 9.4 (or later) statistical software package.

For continuous data, summary statistics will include number of subjects (n), mean, standard deviation (SD), median, minimum, and maximum. For categorical data, frequency counts and percentages will be presented. The category 'Missing' will be presented if the number of missing is greater than zero for at least one treatment group. Time-to-event endpoints will be summarized and displayed graphically using the Kaplan-Meier (KM) method to estimate the median, and the 25th and 75th percentiles. Continuous endpoints with more than 1 post-baseline measurement will be analyzed using a MMRM model.

Two-sided statistical tests will be performed at the significance level of 0.05, and one-sided tests at the level of 0.025. Two-sided 95% CIs will be produced.

Baseline is defined as the last assessment prior to first infusion of the study drug (i.e., the pre-dosing assessment present in the database that is most proximate to the time of first study drug administration), unless otherwise specified. Measurements that are obtained after the first infusion of the study drug will be considered post-baseline values. Change from baseline is defined as (post-baseline value – baseline value).

All data recorded on the eCRF will be provided in data listings showing individual subject values by treatment group and time of assessment, if applicable. Screen failure data that are recorded on the eCRF will be listed.

6.2 Subject Disposition and Data Sets Analyzed

Subject disposition will be listed and summarized by treatment group and overall and will include the number and percentage of subjects:

- screened
- screen failures and reasons for screen failure
- randomized

- included in each study population (FAS, mFAS, PPS, SAF, PK)
- completed the study treatment
- discontinued the study treatment and reasons for discontinuation
- completed the study
- discontinued the study participation and reasons for discontinuation

A summary of subject enrollment by country and site will also be provided by treatment group and overall for the FAS.

6.3 Protocol Deviations

The protocol deviations will be identified before data are unblinded.

All important protocol deviations leading to exclusion from the PPS (see section 4.3.1) will be listed and summarized by treatment group for the FAS population.

6.4 Demographics and Other Baseline Characteristics

Demographic and baseline characteristics will be listed and summarized by treatment group and overall for subjects in the FAS and SAF (if SAF is different from FAS). Standard descriptive statistics will be presented for the continuous variables of:

- age (years)
- weight (kg)
- height (cm)
- body mass index (BMI) (kg/m^2) [calculated as ($weight/height^2$) where weight is in kg and height is in m]
- liver stiffness (kPa)
- MELD score
- CTP score
- alcohol timeline follow-back (TLFB) – estimated daily number of drinks
- alcohol use disorder identification test (AUDIT) total score
- spleen size (cm)

The total counts and percentages of subjects will be presented for the categorical variables of:

- age group (years) (grouped as ≥18 to <65, and ≥65 to ≤75 years)
- sex (male, female)
    - if female, childbearing potential (yes, no)
- race
- ethnicity
- NASH cirrhosis diagnosis (1a, 1b, 1c, 2a, other)
- abdominal collaterals (yes, no)
- Type 2 diabetes mellitus status (yes, no)
- History of hypertension (yes, no)
- Subjects with prior statin use (yes, no)
- Subjects with prior glucagon-like peptide-1 (GLP-1) agonist use - Status at Baseline

In addition, the following laboratory test results at baseline will be summarized by treatment group and overall using standard descriptive statistics:

- HbA1C
- platelets
- AST
- ALT
- GGT
- ALP
- total bilirubin (TB)
- albumin (ALB)
- INR

No formal tests of statistical significance will be performed on the demographic and baseline data. Other baseline measurements, such as vital signs and laboratory test results, will be summarized by treatment group with the post-baseline measurements.

6.4.1 Medical History

Medical history will be coded using the Medical Dictionary for Regulatory Activities (MedDRA) [Version 23.0 (or a later version if updated during the study)]. All medical history will be listed, and the number and percentage of subjects with any medical history will be summarized for the FAS, and SAF (if SAF is different from FAS), by

system organ class (SOC) and preferred term (PT) for each treatment group and overall. Subjects with more than one medical history within a particular SOC are counted only once for that SOC. Similarly, subjects with more than one medical history within a particular PT are counted only once for that PT. Tables will be sorted in descending overall frequency by SOC and PT, and then alphabetically.

6.4.2 Previous and Concomitant Medications

All prior treatments for liver cirrhosis and NASH (up to 5 years prior to screening) will be recorded in the eCRF. All other prior and concomitant medications taken from 30 days prior to screening and throughout the entire duration of the subject's participation in the study will be documented in the eCRF. Medications will be coded using the World Health Organization (WHO) Drug Dictionary [Version: Global B3-March 2020 (or a later version if updated during the study)], Anatomical Therapeutic Chemical (ATC) Classification codes.

Prior medications and concomitant medications are defined as follows:

- Prior medications are those taken and discontinued prior to the start of first infusion date and time of study drug.
- Concomitant medications are those with a start date and time on or after the first infusion date and time of study drug, or those with a start date and time before the first infusion date and time of study drug and a stop date and time on or after the first infusion date and time of study drug or ongoing at end of study.

If a medication cannot be classified as "prior" or "concomitant" after applying imputation rules for missing/incomplete dates (see section 5.2.3), it will be classified as concomitant.

Prior medications and concomitant medications will be listed together and summarized separately for FAS, and SAF (if SAF is different from FAS), for each treatment group and overall.

The number and percentage of subjects using each medication will be displayed together with the number and percentage of subjects using at least one medication within each therapeutic class (ATC-Level 2), chemical subgroup (ATC-Level 4), and generic term for each treatment group and overall. Tables will be sorted in descending overall frequency by ATC-Level 2, ATC-Level 4, and generic term, and then alphabetically.

6.5 Efficacy

All efficacy data collected will be listed.

6.5.1 Level of Significance

The overall Type I error will be controlled across the primary and key secondary efficacy endpoint at the one-sided [redacted] significance level. Regardless of the significance of the tests for the primary and key secondary efficacy endpoints, the testing of the other endpoints will proceed. These other endpoints will not be adjusted for multiplicity; only nominal p-values will be presented for these analyses.

6.5.2 Multiplicity Adjustment and Type I Error Control

6.5.2.1 Primary Efficacy Endpoint

This study has one primary endpoint and one key secondary endpoint. For the primary endpoint, to account for the multiple comparisons of the two doses levels of belapectin versus placebo, the Step-Down Bonferroni Holm procedure will be used to control the overall Type I error at the [redacted]

For example, if the nominal [redacted]

The nominal p-values will be sorted in ascending order [redacted]

If statistical significance is achieved with the first comparison, testing will proceed for the comparison with the [redacted]

For this example, using the Step-Down Bonferroni Holm multiplicity adjustment method, statistical significance is achieved for both treatment group comparisons versus placebo and the overall Type I error for the primary endpoint has been controlled at [redacted].

In addition, a gatekeeping procedure will be used to adjust for multiplicity with the primary and key secondary endpoints with the following rules.[3]

- If at least one of the treatment group comparisons for the primary endpoint achieves statistical significance according to the Step-Down Bonferroni Holm procedure described above, the study will be considered positive. In that case the key secondary endpoint will be formally tested as described below:
  - If both treatment group comparisons for the primary endpoint achieve statistical significance, then the treatment group comparisons for the key secondary endpoint will use the same Step-Down Bonferroni Holm procedure as described above for the primary endpoint, using the full [redacted].
  - If only one of the belapectin arms achieves statistical significance, then the corresponding belapectin arm comparison against placebo for the key secondary endpoint will use a [redacted].
- If none of the treatment group comparisons for the primary endpoint achieve statistical significance, the treatment group comparison for the key secondary endpoint will not be formally tested and the [redacted].

6.5.2.2 Key Secondary Efficacy Endpoint

The methods of adjustment for the primary efficacy endpoint discussed in section 6.5.2.1 will be applied to the key secondary efficacy endpoint. As discussed in section 6.5.2.1, a hierarchical testing strategy will be applied for the testing of the key secondary endpoint (i.e., the test results of the key secondary endpoint will only be interpreted inferentially if at least one of the treatment group comparisons for the primary endpoint achieves statistical significance).

6.5.3 Primary Efficacy Analyses

6.5.3.1 Primary Estimand

The primary estimand, the main clinical quantity of interest to be estimated in the study, is defined by the following 4 components:

Population: Participants with NASH cirrhosis and clinical signs of portal hypertension, but without esophageal varices at baseline.

Variable: Development of new esophageal varices at Week 78

Intercurrent Events:

- Treatment discontinuation with indication that esophageal varices have developed
- Treatment discontinuation due to non-administrative reasons (see list below)
- Initiation of potential rescue therapy (e.g., variceal band ligation, Transjugular intrahepatic portosystemic shunt [TIPS])
- Death
- Use of Non-selective beta-blockers (NSBBs) for more than 12 months during the Phase 2b portion of the study
- Use of GLP-1 Agonists for more than 12 months during the Phase 2b portion of the study
- Other liver cirrhosis related clinical events. The list of other liver cirrhosis related clinical events are:
    - Varices (esophageal or gastric) requiring treatment
    - Variceal bleed requiring hospitalization
    - Clinically significant ascites requiring hospitalization
    - Spontaneous bacterial peritonitis
    - Overt hepatic encephalopathy (West Haven score ≥2 and requiring hospitalization)
    - Liver transplant
    - MELD score ≥15 on 2 consecutive occasions

The presence of intercurrent events will be handled using a composite strategy: Subjects with any of the intercurrent events listed above before week 78 will be seen as a "Treatment Failure" (i.e., non-responder).

The composite strategy assesses the treatment effects not only based on the variable measurements (i.e., the EGD assessment), but also based on intercurrent

events. If a participant met any of the "Treatment Failure" rules, the participant will be considered as a non-responder. This estimand acknowledges that meeting the TF rule is an unfavorable outcome. This strategy aims to achieve a conservative treatment effect.

If a subject discontinues treatment, did not have an ICE (i.e., discontinued treatment due to an administrative reason) and an end of Phase 2b (EOPH2b) EGD is available, the results of this EGD will be used to determine the response status of this subject. If the EOPH2b EGD is not available, the subject would be counted as a "Treatment Failure" as described in section 5.2.1.

The following reasons for treatment discontinuation would be considered "**Non-Administrative**":

- Adverse Event
- Lack of Efficacy
- Pregnancy
- Large Varices
- Varices with Red Wales
- Variceal Hemorrhage
- Portal Gastropathy Hemorrhage
- Other Complications of Cirrhosis (e.g., ascites, hepatic encephalopathy, hepatic cellular carcinoma)
- Laboratory Safety Assessments
- Drug-Induced Liver Injury
- Death

The following reasons for treatment discontinuation would be considered "**Administrative**":

- Subject Lost to Follow Up
- Investigator decision
- Protocol Non-Compliance
- Study Terminated
- Withdrawal by Subject/Parent/Legal Guardian
- Other

Additional reasons may be added to the "Non-Administrative/Administrative" lists based on a blinded review of the end of treatment data prior to database lock.

Population-level summary: Difference in proportion of non-responders between belapectin groups and placebo group.

6.5.3.2 Supplementary Estimand

A supplementary estimand will be studied using the same estimand attributes as for the primary estimand but with a "Treatment Policy" strategy for intercurrent events instead. The occurrence of intercurrent events is irrelevant: the value for the variable of interest (i.e., the EGD assessment) is used regardless of whether any intercurrent events occurred. Missing EGD assessments will be handled as described in section 5.2.1. This supplementary strategy aims to achieve a robust treatment effect for regulatory decision making for primary and major secondary endpoints.

6.5.3.3 Hypothesis Testing

The global null hypothesis (H0) and the corresponding alternative hypothesis (Ha) are:

[redacted]

[redacted]

where p0 denotes the proportion of subjects with development of esophageal varices in the placebo group at Week 78, and p1 and p2 denote the proportion in the two dose groups of belapectin. The unadjusted p-values comparing the proportions between belapectin groups and the placebo group will be based on the Cochran-Mantel-Haenszel (CMH) test controlling for the stratification variable at randomization (i.e., type 2 diabetes status). The p-values will be adjusted for multiple comparisons using the Step-Down Bonferroni Holm procedure (as discussed in section 6.5.2.1) which will control the overall Type I error at the one-sided 0.025 significance level.

The number and percentage (i.e., proportion) of subjects in the FAS who have a composite endpoint (i.e., develop esophageal varices or have intercurrent events) at Week 78 will be calculated by treatment group. The 2-sided [redacted]% exact binomial CIs of the proportions will be calculated using the Pearson-Clopper method. The difference in proportions between treatment groups (i.e., each belapectin dose groups versus the placebo group) and the associated [redacted]% CIs will also be estimated using the CMH method controlling for the stratification variable at randomization (i.e., type 2 diabetes status).

The following supplementary analyses will be conducted for the primary efficacy endpoint:

1. The proportion of subjects who develop new esophageal varices at Week 78 in the FAS using the combination of a composite strategy and a treatment-policy strategy to account for the intercurrent events. Specifically, death will be handled using a composite strategy (i.e., subjects who died will be considered as having the worst outcome) and all other intercurrent events will be handled

using a treatment-policy strategy (i.e., the EGD assessment will be used regardless of whether or not these intercurrent events occur)
2. the proportion of subjects who develop new esophageal varices at Week 78 in the mFAS using a composite strategy for the intercurrent events as described for the primary estimand in section 6.5.3.1.
3. the proportion of subjects who develop new esophageal varices at Week 78 in the PPS using a composite strategy for the intercurrent events as described for the primary estimand in section 6.5.3.1.

6.5.4 Secondary Efficacy Analyses

6.5.4.1 Key Secondary Estimand

The estimand for the key secondary endpoint will be based on the composite strategy as described in section 6.5.3.1. It is defined by the following 4 components:

Population: Participants with NASH cirrhosis and clinical signs of portal hypertension, but without esophageal varices at baseline.

Variable: Development of other liver cirrhosis related clinical events as specified in Section 3.2.

Intercurrent Events: The occurrence of ICEs (except for other liver cirrhosis related clinical events which define the endpoint) will be handled as described in section 6.5.3.1 for the primary estimand.

Population-level summary: Difference in 78-week cumulative incidence rates between belapectin groups and placebo group.

The 78-week cumulative incidence rate for a treatment group in the FAS will be estimated using all data collected until the end of follow-up and calculated as (number of subjects with the event / total time at risk in weeks over all subjects) × 78 weeks. The total time at risk is calculated from the first dosing date to the date of the first event for subjects with an event, and from the first dosing date to the end of follow-up for subjects without an event. The crude incidence rates, the exposure-adjusted 78-week cumulative incidence rates, the difference of cumulative incidence rates between two groups, the associated 2-sided [redacted]% CI, and p-value will be calculated. The exposure-adjusted incidence rates (EAIR) will be derived based on the methods published by Liu et al in 2006.[4]

The supplementary analyses include:

1. the 78-week cumulative incidence rate of subjects with the composite clinical outcomes in the FAS using a treatment-policy strategy to account for the intercurrent events (i.e., the composite clinical outcomes assessment collected will be used regardless of whether other intercurrent events occur)
2. the 78-week cumulative incidence rate of subjects with the composite clinical outcomes in the mFAS using the composite strategy to account for the intercurrent events as described for the primary estimand
3. the 78-week cumulative incidence rate of subjects with the composite clinical outcomes in the PPS using the composite strategy to account for the intercurrent events as described for the primary estimand

6.5.4.2 Additional Secondary Endpoints

Two additional secondary endpoints will be assessed.

- **Progression to large varices or varices with red wales**

This additional secondary efficacy endpoint is the 78-week cumulative incidence rate (i.e., proportion of subjects with event over 78 weeks of observation) of subjects in the belapectin treatment groups, compared to placebo, who progress to large varices or the development of red wales.

The primary estimand for this secondary efficacy endpoint will be based on the composite strategy as described in section 6.5.3.1.

The crude incidence rates, exposure-adjusted 78-week cumulative incidence rates in the FAS, the difference of cumulative incidence rates between two groups, the associated 95% CI, and p-value will be calculated. The EAIR will be derived based on the methods published by Liu et al in 2006.[4]

The supplementary analyses include:

1. the 78-week cumulative incidence rate of subjects who progress to large varices or develop red wales in the FAS using combination of a composite strategy and a treatment-policy strategy to account for the intercurrent events. Specifically, death will be handled using a composite strategy (i.e., subjects who died will be considered as having the worst outcome) and all other ICEs will be handled using a treatment-policy strategy, in which, the composite clinical outcomes assessment collected [EGD] will be used regardless of whether or not these intercurrent events occur).

2. the 78-week cumulative incidence rate of subjects who progress to large varices or develop red wales in the mFAS with the intercurrent events handled by the composite strategy as described for the primary estimand
3. the 78-week cumulative incidence rate of subjects who progress to large varices or develop red wales in the PPS with the intercurrent events handled by the composite strategy as described for the primary estimand

- **Time to first cirrhosis associated clinical event**

This additional secondary efficacy endpoint is the event-free survival by time to first cirrhosis-associated clinical event (see the list of events in Section 3.2) from the first infusion of study drug in the belapectin treatment groups, compared to placebo.

The primary estimand for this secondary endpoint will be based on the treatment policy strategy as described in section 6.5.3.2 (i.e., the cirrhosis-associated clinical event assessment collected will be used regardless of whether other ICEs occur).

Subjects who do not have any cirrhosis-associated clinical event during the follow-up will be censored at the end of follow-up. The median, 25th and 75th percentiles of time to first cirrhosis-associated clinical event with 95% CIs will be estimated using the KM method for the belapectin dose groups and the placebo group. The KM event-free survival curve will be displayed. The supplementary analyses include:

1. The event-free survival by time to first cirrhosis-associated clinical event in the mFAS using the treatment-policy strategy accounting for the intercurrent events as described for the primary estimand
2. The event-free survival by time to first cirrhosis-associated clinical event in the PPS using the treatment-policy strategy accounting for the intercurrent events as described for the primary estimand

6.5.5 Exploratory Efficacy Analyses

6.5.5.1 Change in Liver Stiffness Measurement

Change in liver stiffness measurement from baseline in the belapectin treatment groups, compared to placebo, will be analyzed using a MMRM analysis in the FAS. The MMRM model will include fixed terms for treatment, time points, baseline value, and an interaction term between treatment and time points. An unstructured covariance matrix will be used to model the within-subject covariance. If a convergence issue occurs, the following covariance structures will be tried in this order: Toeplitz, the first-order autoregressive (AR), and compound symmetry (CS).

The Kenward-Roger approximation will be used to estimate the denominator degrees of freedom. The least square means (LSMs) of change from baseline values at Week 78 for each treatment group will be provided along with the 95% CIs. Similarly, the LSMs of treatment difference between treatment groups, the associated 95% CIs, and p-values will be presented.

The raw values of liver stiffness measurement and changes from baseline will be also summarized using descriptive statistics by treatment group and by scheduled visit.

The average values of the liver stiffness measurements at each timepoint will be used in the MMRM analysis and the descriptive summary.

6.5.5.2 Chronic liver disease questionnaire scores

The CLDQ includes 29 questions with 6 domains: abdominal symptoms (items 1, 5, 7), fatigue (items 2, 4, 8, 11, 13), systemic symptoms (items 3, 6, 21, 23, 27), activity (items 9, 14, 17), emotional function (items 10, 12, 15, 16, 19, 20, 24, 26) and worry (items 18, 22, 25, 28, 29).

The scores of 6 domains and overall CLDQ score will be calculated for each timepoint. Each domain score will be calculated by dividing the sum of item scores by the number of items per domain among items with non-missing scores. Overall CLDQ score will be obtained by adding scores for each item and dividing by the total number of items among items with non-missing scores. Domain scores and overall scores will not be calculated for records with more than half of the items missing.

Domain scores and overall scores will be summarized as continuous variables using descriptive statistics in the FAS by treatment group and by visit. The change from baseline in domain scores and overall scores will also be summarized descriptively by treatment group and by visit.

In addition, the individual 29 item scores of NASH-CLDQ will be summarized by treatment group and by visit as categorical variables using number and percentage of subjects in each score category (ranging from 1 [most impairment] to 7 [least impairment]).

6.5.6 Efficacy Analysis for Phase 2b Subjects that continue into Open Label Extension

At the conclusion of Stage 2 of the study, an 18-36-month blinded data read out and OLE analysis will occur. All efficacy endpoints described above will be analyzed in the Phase 2b FAS using the 156 weeks data collected over Stage 1 and Stage 2 and the same strategy for handling intercurrent events as described for the primary estimand for each endpoint will be used (if applicable).

In addition, sensitivity analyses will be performed by excluding data collected after treatment switching for subjects who switched to the optimal belapectin dose group in the OLE portion of the study.

All analyses for the cumulative 156 Week data for the subjects that continued into the OLE will be considered exploratory. Descriptive Statistics along with nominal p-values will be provided without any adjustments.

6.5.7 Efficacy Subgroup Analyses

The primary efficacy endpoint described in section 6.5.3 will be analyzed for the FAS in the following subgroups:

- Site (Rules for combining low enrolling sites will be determined based on final enrollment data)
- Region (Geographic region definitions will be determined before database lock)
- Gender
- Race
- BMI (median split or clinically agreed upon categories)
- GLP-1 Agonist Use - Status at Baseline
- Type 2 Diabetes Mellitus Status
- Statin Use – Status at Baseline
- Baseline Platelet count (clinically agreed upon categories will be used)
- Baseline Liver Stiffness Measurement as determined by VCTE [FibroScan] – (clinically agreed upon categories will be used)

The subgroup analyses will adopt the same strategy for handling intercurrent events as described for the primary estimand.

The number and percentage of subjects will be presented for the following categorical variables:

- percentage compliance in the following categories:
  - <80.0%
  - 80.0% - 120.0%
  - >120.0%
- duration of exposure in the following categories:
  - <26 weeks
  - ≥26 and <52 weeks
  - ≥52 and <78 weeks
  - ≥78 and <104 weeks
  - ≥104 and <130 weeks
  - ≥130 and <156 weeks
  - ≥156 weeks
- with any dose modification during the study
- with any dose delayed during the study
- with any infusion interruption during the study

6.6.2 Adverse Events

All AEs recorded on the eCRF will be coded using the MedDRA dictionary [Version 23.0 (or a later version if updated during the study)] and classified as either pre-treatment AEs or TEAEs as follows:

- Pre-treatment AEs are events that are reported or worsened after signing the informed consent form and before the date of first infusion of study drug
- TEAEs are events with start date and time on or after the date and time of first infusion of study drug, or events with start date and time prior to the date and time of first infusion of study drug whose severity worsens on or after the first infusion of study drug, through the end of the study

AE severity will be graded using the National Cancer Institute-Common Terminology Criteria for Adverse Events (NCI-CTCAE) version 5.0. TEAEs with missing CTCAE will be reported as a missing category in the summary of TEAEs by CTCAE grade.

The relationship between an AE and study drug is assessed as unrelated or related. A study drug-related AE is an AE considered by the investigator as related to study drug.

All AE data will be listed by treatment group with pre-treatment AEs and TEAEs presented together. Treatment-emergence status will be flagged in the listing. In addition, corresponding listings of serious AEs (SAEs), AEs leading to discontinuation of study drug, AEs leading to discontinuation of study, and AEs resulting in death will be produced.

The number and percentage of subjects reporting each pre-treatment AE will be summarized for each treatment group and overall, by SOC (sorted alphabetically) and PT (sorted by descending overall total) in the SAF.

An overview table will summarize the number and percentage of subjects with at least one of the following TEAEs by treatment group and overall, in the SAF, where subjects with more than one TEAE in a particular category are counted only once in that category:

- any TEAE
- any TEAE by maximum CTCAE grade
- any TEAE of CTCAE grade ≥3
- study drug-related TEAE
- infusion procedure-related TEAE
- TEAE leading to study drug discontinuation
- TEAE leading to discontinuation from the study
- Treatment-emergent SAE (TESAE)
- TESAE by maximum CTCAE grade
- TESAE of CTCAE grade ≥3
- study drug-related TESAE
- infusion procedure-related TESAE
- TESAE leading to study drug discontinuation
- TESAE leading to discontinuation from the study
- TESAE leading to death

The overview table will also summarize the number of the following TEAEs:

- all TEAEs
- all TESAEs

The number and percentage of subjects reporting each TEAE will be summarized by SOC and PT in the SAF. Tables will be sorted alphabetically by SOC. PTs will be sorted by descending overall total. The following summaries will be produced:

- TEAEs, by SOC and PT
- TEAEs, by PT
- TEAEs by maximum CTCAE grade, by SOC and PT
- TEAEs of CTCAE grade ≥3, by SOC and PT
- study drug-related TEAEs, by SOC and PT
- study drug-related TEAEs, by PT
- infusion procedure-related TEAEs, by SOC and PT
- TEAEs leading to study drug discontinuation, by SOC and PT
- TEAEs leading to discontinuation from the study, by SOC and PT
- TESAEs, by SOC and PT
- TESAEs by maximum CTCAE grade, by SOC and PT
- TESAEs of CTCAE grade ≥3, by SOC and PT
- study drug-related TESAEs, by SOC and PT
- infusion procedure-related TESAEs, by SOC and PT
- TESAE leading to study drug discontinuation, by SOC and PT
- TESAE leading to discontinuation from the study, by SOC and PT
- TESAEs leading to death, by SOC and PT

In the above summaries, subjects with more than one AE within a particular SOC are counted only once for that SOC. Similarly, subjects with more than one AE within a particular PT are counted only once for that PT. For summaries by maximum severity, subjects with multiple AEs within a particular SOC or PT will be counted under the category of their most severe AE within that SOC or PT.

No statistical comparisons of AEs between treatment groups will be performed.

6.6.3 Laboratory Evaluations

All laboratory data received from central laboratory will be listed. Data for the following hematology, serum chemistry, serology, and coagulation analytes will be summarized by treatment group and visit in the SAF. If data for any additional analytes are also received, then these will be listed only.

Hematology	Serum Chemistry	Urinalysis
Hematocrit Hemoglobin Platelet count Red blood cell (RBC) count Red cell indices White blood cell (WBC) count WBC differential (% and absolute): Basophils Eosinophils Lymphocytes Monocytes Neutrophils	Albumin ALP ALT AST BUN Calcium Chloride Bicarbonate Creatinine Fasting insulin level GGT Glucose LDH Magnesium Phosphorus Potassium Sodium Total and direct bilirubin Total protein Uric acid	Color Blood Clarity Bilirubin pH and specific gravity Leukocyte esterase Glucose Hyaline and other casts Ketones Bacteria Nitrite Epithelial cells Protein Crystals Microscopic (including RBCs and WBCs) Yeast
Other Tests (Serology)	**Drug Screen (Urine)**	**Coagulation**
Hepatitis A antibody (IgM) Hepatitis B core antibody Hepatitis B surface antibody Hepatitis B surface antigen Hepatitis C antibody HIV antibody	Amphetamines Cocaine (metabolite) Non-prescription opiates Barbiturates Benzodiazepines Marijuana/Cannabinoids (THC) Methadone Methamphetamine Methylenedioxymethamphetamine Opiates Phencyclidine	Prothrombin time Partial thromboplastin time INR **For Women Only** Urine and serum pregnancy test

Abbreviations: ALP = alkaline phosphatase; ALT = alanine aminotransferase; AST = aspartate aminotransferase; BUN = blood urea nitrogen; GGT = gamma glutamyl transferase; IgM = immunoglobulin M; INR = International normalized ratio; LDH = lactate dehydrogenase; RBC = red blood cell; THC = tetrahydrocannabinol, WBC = white blood cell.

All laboratory data will be reported in International System of Units (SI) units. Out-of-reference-range values will be flagged as high (H) or low (L) in the listings.

Laboratory data for hematology, serum chemistry, serology, and coagulation will be summarized by treatment group and visit using standard descriptive statistics for the SAF. Changes from baseline will also be summarized by treatment group and visit.

For hematology and serum chemistry, shift tables presenting movement in and out of reference range from baseline to each scheduled post-baseline visit will be provided for each treatment group. The number and percentage of subjects reporting markedly abnormal clinical laboratory values will also be summarized by treatment group.

For each laboratory analyte, the baseline value will be defined as last scheduled or unscheduled value collected prior to the first infusion of study drug. Assessments carried out on day of first study drug administration are considered to have taken place before the study drug administration if the corresponding times have not been recorded. For post-baseline, data will be included in the summary tables per visit window defined in section 5.1.

6.6.4 Drug-induced Liver Injury

The number and percentage (i.e., proportion) of subjects who had an adjudicated DILI event will be summarized by treatment group and overall in the SAF. The 2-sided [redacted]% exact binomial CIs of the proportions will be reported.

Evaluation of potential DILI will be performed and displayed in an evaluation of drug induced serious hepatotoxicity (eDISH) figure with different symbols representing treatment group, which is a log/log scatter plot of peak on-treatment total bilirubin vs. ALT, both in multiples of ULN, with horizontal and vertical lines indicating Hy's law thresholds, i.e. ALT = 3 × ULN and total bilirubin = 2 × ULN.

The scatter plot of maximal on-treatment values of AST, ALT, ALP, and TB expressed as the multiples of upper normal range with the baseline values on the x-axis will be provided to assess the shift from baseline.

For these analyses, baseline liver function test (LFT) measurements are defined as:

- for ALT, AST, and ALP: the mean value of LFT values from Screening Visit 1, 2, 3 (if applicable), and Infusion Visit 1 (pre-infusion)
- for TB: the pre-dose value at Infusion Visit 1

6.6.5 Major Adverse Cardiovascular Events

The number and percentage of subjects who had an adjudicated MACE, including a breakdown of MACE events (i.e., myocardial infarction or hospitalization for unstable angina, stroke or transient ischemic attack [TIA], and heart failure), will be summarized by treatment group and overall in the SAF. The 2-sided [redacted]% exact binomial CIs of the proportions of subjects who had a MACE will be reported.

6.6.6 Vital Signs

The following vital signs will be listed and summarized by treatment group and visit.

- systolic and diastolic blood pressure (mmHg)
- heart rate (beats/minute)
- respiratory rate (breaths/min)
- body temperature (°C)

Vital signs data and changes from baseline in vital signs will be summarized by treatment group and visit using standard descriptive statistics in the SAF.

The baseline value will be defined as last scheduled or unscheduled value collected prior to the first infusion of study drug. Assessments carried out on day of first study drug administration are considered to have taken place before the study drug administration if the corresponding times have not been recorded. For post-baseline, only data from scheduled visits will be included in the summary tables.

6.6.7 Electrocardiograms

The following quantitative ECG measurements will be taken during the study:

- heart rate (bpm)
- RR interval (msec)
- PR interval (msec)
- QRS interval (msec)
- QT interval (msec)
- Bazett corrected QT interval (QTcB) (msec)
- Fridericia corrected QT interval (QTcF) (msec)

An overall Investigator assessment of ECG will be provided (categories "normal", "abnormal, not clinically significant", and "abnormal, clinically significant").

The ECG measurements and changes from baseline in ECG will be listed and summarized by treatment group and visit using standard descriptive statistics in the SAF.

The baseline value will be defined as last scheduled or unscheduled value collected prior to the first infusion of study drug. Assessments carried out on day of first study drug administration are considered to have taken place before the study drug administration if the corresponding times have not been recorded. For post-baseline, data will be included in the summary tables per visit window defined in section 6.1.

6.6.8 Physical Examination

Physical examination data will be listed. For each physical examination body system, the number and percentage of subjects with abnormalities at baseline and post-baseline will be summarized by treatment group in the SAF.

6.6.9 Safety Analysis for Phase 2b Subjects Only

All safety endpoints described above will be analyzed in the Phase 2b SAF using the 156 weeks data collected during both Stages (i.e., Phase 2 and Phase 3) for subjects enrolled in Phase 2b.

7. Interim Analyses

7.1 Non-comparative Sample Size Re-estimation

A non-comparative SSR was planned to be performed when at least [redacted] of subjects completed 78 weeks of Phase 2b treatment. The proportion of subjects who develop esophageal varices at Week 78 was to be calculated among all subjects who had post-baseline EGD assessment. Based on the observed overall proportion, the sample size for Phase 2b was to be re-estimated using the nQuery 8.5.2.0 software for comparing two proportions using a Chi-square test at a one-sided significance level of 0.025.

Based on the findings of the non-comparative SSR, the sponsor had the option to decide to increase the total sample size to ensure that an adequate number of Phase 2b subjects were to be included in the study. Additional subjects could have been enrolled in Phase 2b up to a maximum of 65 additional subjects in each Phase 2b treatment group (i.e., a maximum of 510 [315 +195] subjects enrolled in Phase 2b), without requiring a protocol amendment. Since the review was to be blinded and the efficacy analyses were not to be reviewed by treatment group, no adjustment of p-values was to be required. Nevertheless, the sponsor was to take steps to minimize the distribution of results to avoid introducing any bias.

7.2 Unblinded Interim Analysis

Not Applicable.

8. Changes in Planned Analysis

The planned IA, which was to inform various adaptive modifications for Stage 2 of the study, was converted into a primary analysis read out which will inform selection of the optimal belapectin dose for the Stage 2 OLE portion of the study. Additionally, at the conclusion of the study (at the end of Stage 2), there will be an 18-36-month blinded data read out (which will cover the blinded/rollover Stage 2 portion of the study) and OLE analysis (which will cover long term safety). Both will assess the same endpoints and outcomes but will be analyzed differently. The details for these

Stage 2 analyses will be documented via an amendment to the SAP or provided in a supplementary analysis plan.

9. References

1. Harrison, S.A., Ruane, P.J., Freilich, B.L. et al. Efruxifermin in non-alcoholic steatohepatitis: a randomized, double-blind, placebo-controlled, phase 2a trial. *Nat Med.* 2021; 27:1262–1271.
2. Holm S. A simple sequentially rejective multiple test procedure. *Scandinavian Journal of Statistics*. 1979;6:65–70.
3. Dmitrienko A, Tamhane AC, Bretz F (2010). Multiple Testing Problems in Pharmaceutical Statistics. Chapman & Hall.
4. G. Frank Liu, Junyuan Wang, Kenneth Liu and Duane B. Snavely. Confidence intervals for an exposure adjusted incidence rate difference with applications to clinical trials. *Statist. Med.* 2006; 25:1275–1286.

10. Appendices

10.1 Appendix I – Document History

Document History

Document Version, Status, Date	Summary/Reason for Changes

[redacted]

	• [redacted]
[redacted]	[redacted]

[redacted]	[redacted]