STATISTICAL ANALYSIS PLAN

Protocol SB2640-CLIN-007

FASCINATE-2: A Phase 2b, Multi-Center, Double-Blind, Randomized, Placebo -Controlled Study of the Safety and Efficacy of TVB-2640 in Subjects with Nonalcoholic Steatohepatitis

Final Version
Date of Plan:

04December2023

Study Drug:
TVB-2640 (denifanstat)

Prepared for:

Sagimet Biosciences

[redacted]

Approval Signature Page

Summit Analytical:

[redacted]

Date

Sagimet Biosciences:

[redacted]
Eduardo Bruno Martins, MD, DPhil

[redacted]

Table of Contents

1. ABBREVIATIONS

Abbreviation	Explanation
AE	Adverse Event
ALT	Alanine aminotransferase
ANCOVA	Analysis of covariance
AST	Aspartate aminotransferase
ATC	Anatomic therapeutic class
CI	Confidence interval
CMH	Cochran-Mantel-Haenszel
CRF	Case Report Form
CRN	Clinical Research Network
CRO	Clinical Research Organization
CSR	Clinical Study Report
CTCAE	Common Terminology Criteria for Adverse Events
ECG	Electrocardiogram
eCRF	Electronic case report form
ELF	Enhanced liver fibrosis test
FAST	FibroScan-AST score
FGF-21	Fibroblast Growth Factor 21
FIB-4	Fibrosis-4 index
GGT	Gamma-glutamyl transferase
HbA1c	Glycated hemoglobin
ICH	International Council for Harmonization
IDMC	Independent Data Monitoring Committee
INR	International normalized ratio
ITT	Intent-to-treat population
IWRS	Interactive Web Response System
kg	Kilogram
LDL-C	Low-density lipoprotein cholesterol
LLN	Lower limit of normal
max	Maximum
MedDRA	Medical Dictionary for Regulatory Activities
min	Minimum
mITT	Modified intent-to-treat population
mmHg	Millimeters of mercury
MMRM	Mixed model for repeated measures
MRI-PDFF	Magnetic Resonance Imaging-Proton Density Fat Fraction
ms	Millisecond
N or n	Number of subjects
NAFLD	Nonalcoholic fatty liver disease
NAS	NAFLD activity score
NASH	Nonalcoholic steatohepatitis

Abbreviation	Explanation
NCI	National Cancer Institute
PD	Pharmacodynamic
PK	Pharmacokinetic
PO	Orally
PRO-C3	N-Terminal Type III Collagen Propeptide
PT	Preferred term
Q1	First Quartile
Q3	Third Quartile
QD	Once daily
SAE	Serious adverse event
SAF	Safety Analysis Population
SAP	Statistical analysis plan
SD	Standard Deviation
SI	International System of Units
SNPs	Single-Nucleotide Polymorphisms
SOC	System organ class
T2DM	Type 2 Diabetes Mellitus
TEAE	Treatment-Emergent Adverse Event
TIMP-1	Tissue Inhibitor of Metalloproteinase-1
ULN	Upper limit of normal

2. INTRODUCTION

The statistical analysis plan (SAP) corresponds to the study number SB2640-CLIN-007: "A Phase 2b, Multi-Center, Double-Blind, Randomized, Placebo Controlled Study of the Safety and Efficacy of TVB-2640 in Subjects with Nonalcoholic Steatohepatitis (FASCINATE-2)". The SAP comprises the protocol pre-planned interim and final analyses and will be the basis for the analyses in support of the Clinical Study Report (CSR). The SAP was based on the study protocol version 4.0 dated 06 September 2023 and CRF dated 24 February 2023. The final sponsor-approved version of the SAP must occur prior to unblinding of the study database and the final analyses.

3. STUDY OBJECTIVES

3.1. Co-primary Objectives

- To evaluate the effect of TVB-2640 50 mg PO QD compared with matching placebo in noncirrhotic subjects with nonalcoholic steatohepatitis (NASH) and F2-F3 fibrosis
    - Histological improvement at Week 52 in nonalcoholic fatty liver disease (NAFLD) activity score (NAS) (i.e., ≥2 points improvement in NAS with ≥1 point improvement in ballooning or inflammation) and without worsening of fibrosis (by NASH Clinical Research Network [CRN] fibrosis score)
    - Resolution of steatohepatitis and no worsening of liver fibrosis (by NASH CRN fibrosis score). Resolution of steatohepatitis is defined as absence of fatty liver disease or isolated or simple steatosis without steatohepatitis and a NAS of 0 or 1 for inflammation, 0 for ballooning, and any value for steatosis

      AND

      Histological improvement in NAS (≥2 points improvement in NAS) at Week 52

- To evaluate the safety and tolerability of TVB-2640 50 mg PO QD in subjects with confirmed NASH and liver fibrosis.

3.2. Secondary Objectives

- To evaluate the effect of TVB-2640 50 mg QD compared with matching placebo on the proportion of subjects with ≥8% liver fat content at Baseline who achieve a ≥30% relative reduction in liver fat content as assessed by magnetic resonance imaging-proton density fat fraction (MRI-PDFF) at Week 26.
- To evaluate the effect of TVB-2640 50 mg QD compared with matching placebo on the proportion of subjects with ≥8% liver fat content at Baseline who achieve a ≥30% relative reduction in liver fat content as assessed by MRI-PDFF at Week 52.
- To evaluate the effect of TVB-2640 50 mg QD compared with matching placebo on fibrosis as assessed by ≥1 stage of fibrosis improvement by NASH CRN score) without worsening of steatohepatitis (no increase in NAS for ballooning, inflammation, or steatosis) at Week 52.

- To evaluate the effect of TVB-2640 50 mg QD compared with matching placebo on fibrosis as assessed by ≥1 stage of fibrosis improvement by NASH CRN score without increasing of total NAS scores at Week 52.
- To evaluate the effect of TVB-2640 50 mg QD compared with matching placebo on fibrosis as assessed by ≥1 stage of fibrosis improvement by NASH CRN score without any worsening of NASH (no worsening of ballooning and lobular inflammation, a 1 grade change in steatosis may be acceptable) at Week 52.
- To evaluate the effect of TVB-2640 50 mg QD compared with matching placebo on the proportion of subjects experiencing **both** of the following at Week 52:
    - Resolution of steatohepatitis on overall histopathological reading and no worsening of liver fibrosis on NASH CRN fibrosis score. Resolution of steatohepatitis is defined as absent fatty liver disease or isolated or simple steatosis without steatohepatitis and a NAS score of 0–1 for inflammation, 0 for ballooning, and any value for steatosis.

  AND

    - Improvement in liver fibrosis greater than or equal to one stage (NASH CRN fibrosis score) and no worsening of steatohepatitis (defined as no increase in NAS for ballooning, inflammation, or steatosis).
- To evaluate the effect of TVB-2640 50 mg QD compared with matching placebo on the proportion of subjects experiencing **both** of the following at Week 52:
    - Histological improvement at Week 52 in NAS (i.e., ≥2 points improvement in NAS with ≥1 point improvement in ballooning or inflammation) and without worsening of fibrosis (by NASH CRN fibrosis score)

  AND

    - Improvement in liver fibrosis greater than or equal to one stage (NASH CRN fibrosis score) and no worsening of steatohepatitis (defined as no increase in NAS for ballooning, inflammation, or steatosis).
- To evaluate the effect of TVB-2640 50 mg QD compared with matching placebo on resolution of steatohepatitis and no worsening of liver fibrosis (by NASH CRN fibrosis score). Resolution of steatohepatitis is defined as absence of fatty liver disease or isolated or simple steatosis without steatohepatitis and a NAS of 0 or 1 for inflammation, 0 for ballooning, and any value for steatosis at Week 52.
- To evaluate the effect of TVB-2640 50 mg QD compared with matching placebo on histological improvement (decrease) at Week 52 in NAS by ≥1 point.
- To evaluate the effect of TVB-2640 50 mg QD compared with matching placebo on histological improvement (decrease) at Week 52 in NAS by ≥2 points.
- To evaluate the effect of TVB-2640 50 mg QD compared with matching placebo on liver fat content as assessed by MRI-PDFF at Week 26.
- To evaluate the effect of TVB-2640 50 mg QD compared with matching placebo on liver fat content as assessed by MRI-PDFF at Week 52.

- To evaluate the effect of TVB-2640 50 mg QD compared with matching placebo on NASH resolution (defined as absence of fatty liver disease or isolated or simple steatosis without steatohepatitis and a NAS of 0 or 1 for inflammation, 0 for ballooning, and any value for steatosis) at Week 52.
- To determine the effect of TVB-2640 50 mg QD compared with matching placebo on change from Baseline at 26 and 52 weeks in alanine aminotransferase (ALT), aspartate aminotransferase (AST), and gamma-glutamyl transferase (GGT).
- To determine the effect of TVB-2640 50 mg QD compared with matching placebo on the change from Baseline at 26 and 52 weeks in low-density lipoprotein cholesterol (LDLC) and other lipid levels.
- To determine the effect of TVB-2640 50 mg QD compared with matching placebo on the change from Baseline at 52 weeks in the amount of collagen/fibrous area and fibrosis score, assessed by digital pathology.
- To determine the effect of TVB-2640 50 mg QD compared with matching placebo on the change from Baseline at 26 and 52 weeks in fasting insulin, fasting glucose, homeostatic model assessment of insulin resistance (HOMA-IR), adipose tissue insulin resistance (adipo-IR), and glycosylated hemoglobin (HbA1c) levels.
- To determine the effect of TVB-2640 50 mg QD compared with matching placebo on the change from Baseline at 26 and 52 weeks in fibroblast growth factor 21 (FGF-21), adiponectin, and other NASH biomarker levels.
- To determine the effect of TVB-2640 50 mg QD compared with matching placebo on change from Baseline at 26 and 52 weeks in FibroScan® and controlled attenuation parameter (CAP) score.
- To determine the effect of TVB-2640 50 mg QD compared with matching placebo on N-terminal type III collagen propeptide (PRO-C3) levels and other fibrosis biomarkers on change from Baseline at Weeks 4, 13, 26, and 52.
- To determine the effect of TVB-2640 50 mg QD compared with matching placebo on change from Baseline at 26 and 52 weeks on enhanced liver function (ELF) score.

[redacted]

3.3. Exploratory Objectives

- To evaluate the effect of TVB-2640 50 mg QD compared with matching placebo on the proportion of subjects with ≥8% liver fat content at Baseline who achieve a ≥30% relative reduction in liver fat content as assessed by MRI PDFF AND achieve ≥17 U/L reduction in ALT at Week 52.
- To determine the effect of TVB-2640 50 mg QD compared with matching placebo on FibroScan-AST (FAST) score at Weeks 26 and 52.
- To determine the effect of TVB-2640 50 mg QD compared with matching placebo on

select fibrosis biomarkers (including at least tissue inhibitor of metalloproteinase-1 [TIMP1]) at Weeks 4, 13, 26, and 52.

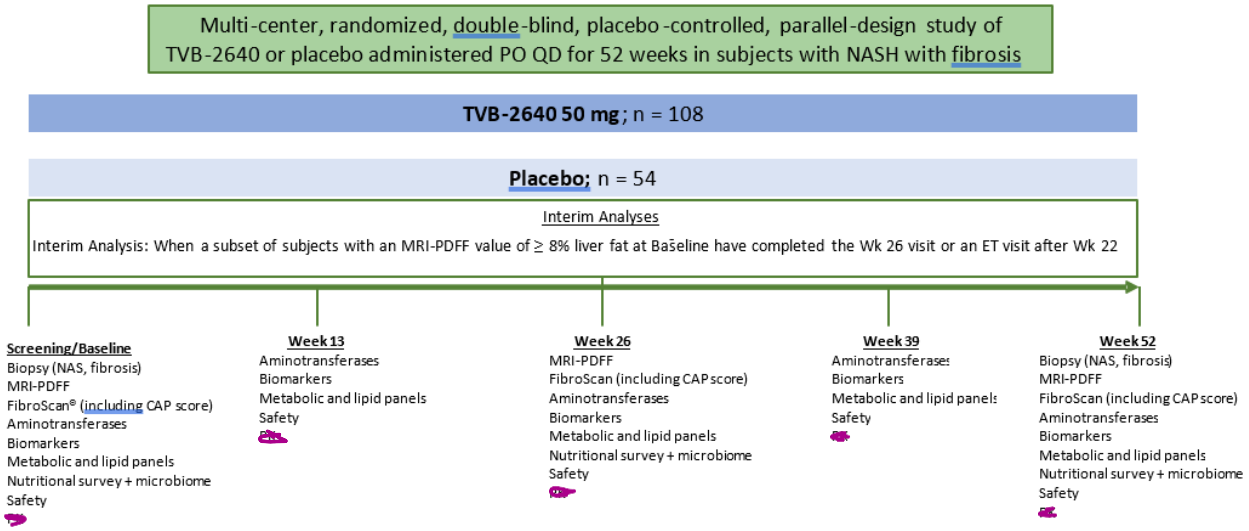
- To determine the effect of TVB-2640 50 mg QD compared with matching placebo on lipidomic biomarkers (including tripalmitin, ceramides, bile acids, diacylglycerols, and other classes of metabolites) at Weeks 4, 13, 26, and 52.
- To determine the effect of TVB-2640 50 mg QD compared with matching placebo on glycosylated hemoglobin (HbA1c) levels among subjects with type 2 diabetes mellitus (T2DM) at Weeks 13, 26, 39, and 52.
- To determine the effect of TVB-2640 50 mg QD compared with matching placebo on Fibrosis 4 score (FIB-4) at Weeks 4, 13, 26, and 52.
- To use digital pathology to perform zonal analysis of fibrosis at Week 52.
    - To use digital pathology to evaluate the concurrence of fibrosis with components of the NAS score.
    - To use digital pathology to evaluate the NAS components.
- To correlate changes in MRI-PDFF with changes in, or baseline levels of, serum biomarkers and explore predictive markers of response.
- To explore the relationship between changes in, or Baseline levels of, serum biomarkers and histology results, and explore predictive markers of response.
- To explore the relationship between MRI-PDFF results and histology results.
- To correlate the presence of relevant single-nucleotide polymorphisms (SNPs) or other genes relevant to NAFLD or NASH with clinical and histopathologic response.
- To determine the effect of TVB-2640 50 mg QD compared with matching placebo on gut microbiome change from Baseline to Weeks 26 and 52 in a subset of subjects.
- To explore the relationship between nutrition, as measured by the HepVita Nutritional Survey, and clinical, histopathological, and biomarker changes in NASH at Weeks 26 and 52 in a subset of subjects.

4. STUDY DESIGN

4.1. Study Design and Population

This is a multi-center, double-blind, randomized, placebo-controlled, parallel-design study (see Figure 1, below) to evaluate the safety and efficacy of TVB-2640 compared with matching placebo in subjects with NASH who are aged ≥18 years at Screening.

Figure 1: FASCINATE-2 Study Design



Abbreviations: CAP = controlled attenuation parameter; MRI = magnetic resonance imaging: NAS = nonalcoholic fatty liver disease activity score; NASH = nonalcoholic steatohepatitis; PDFF = protein density fat fraction; PK = pharmacokinetic; PO = orally; QD = once daily.

Histological eligibility will be determined during the Screening period or prior to Screening but within 180 days of the Screening visit in 1 of 2 ways:

- If an 'eligible' and adequate (this refers to the quality and quantity of remaining tissue) liver histology specimen, in the judgement of the central pathologist, exists from the time period within 180 days before the subject's Screening visit, the subject may participate so long as other eligibility criteria are satisfied.
- If no liver histology specimen exists from the period within 180 days of the Screening visit, the subject must undergo liver biopsy during the Screening period.

Approximately 162 unique subjects with liver fibrosis stage F2-F3 will be enrolled and randomized (2:1 ratio) with approximately 108 subjects on TVB-2640 50 mg and 54 subjects on placebo. Dynamic allocation is used for the randomization, stratified by 3 factors: T2DM status (yes or no), region (North America or not North America), and amount of fibrosis (F2 or F3).

Subjects will be screened for study eligibility within 90 days before randomization (randomization is within 3 days of Baseline [Day 1]). Subjects who are determined to be eligible for the study based on Screening assessments are to be randomized into the study within 3 days before Baseline (i.e., Day 1, the first day of study drug administration). Subjects who are randomized are considered to be enrolled in the study.

Subjects will receive TVB-2640 or matching placebo PO QD for 52 weeks, with the first dose administered on Day 1. During the 52-week Treatment period, subjects are to attend study center visits at Weeks 4, 8, 13, 26, 39, and 52. After completion of the 52-week Treatment period,

subjects are to attend a Follow-up visit at Week 56 for posttreatment safety and efficacy assessments.

In the event a subject is unable to complete the end of treatment MRI-PDFF and/or liver biopsy within the protocol-specified window (Week 52 +/- one week), extended dosing is permitted to maintain continuous dosing up to the time both assessments are completed, to a maximum total of 56 weeks of dosing. In this instance, the Follow-up visit at Week 56 for post treatment safety and efficacy assessments is to be adjusted to occur four weeks after the last dose (+/- one week).

In the instance of a dose hold due to an AE or SAE, the decision to resume study drug will require consultation with the Investigator, the CRO's Medical Monitor, and the Sponsor's Chief Medical Officer.

Improvement of NAS and improvement of fibrosis score will be evaluated by comparison of Baseline and Week 52 liver histology specimens. Subjects will also undergo MRI-PDFF prior to the start of treatment, at Week 26, and at Week 52 or Early Termination unless it presents a logistical or physical hardship to the subject. MRI-PDFF is not required at Early Termination if a subject has completed less than 22 weeks of study treatment. Other efficacy and pharmacodynamic (PD) assessments include FibroScan with CAP, and measurement of liver aminotransferases, lipid parameters, and other noninvasive biomarkers of NASH and liver fibrosis throughout the study.

During the study, safety will be assessed by vital signs, 12-lead electrocardiograms (ECGs), physical examinations including eye examinations, and clinical laboratory testing. Subjects will be evaluated for AEs and concomitant medication use throughout the study.

Blood samples will be collected at Baseline (Day 1) and at Weeks 4, 13, 26, and 52 for the assessment of sparse PK. In addition, blood samples will be collected on Day 1 and at Week 4 from a subset of subjects for serial PK.

An independent data monitoring committee (IDMC) will be established by charter to provide independent review and assessment of study data and to monitor the overall study conduct in a systematic manner to safeguard the safety of study subjects. The IDMC will be tasked with making recommendations to the Sponsor to continue, amend, or stop the study based on these assessments.

4.2. Randomization and Blinding

This will be a double-blind, randomized study. Subjects will be centrally assigned to randomized study treatment using an Interactive Web-based Response System (IWRS).

Approximately 162 unique subjects with liver fibrosis stage F2-F3 will be enrolled and randomized (2:1 ratio) to receive TVB-2640 50 mg or placebo (approximately 108 for TVB-2640 50 mg and approximately 54 subjects for placebo) PO QD for 52 weeks. Dynamic allocation is used for the randomization, stratified by T2DM status (Yes, No), region (North America, not North America), and amount of fibrosis (F2 or F3).

All subjects, Investigators, the Sponsor, and the clinical research organization (CRO) will be blinded to the treatment assignment of each individual subject. Members of the IDMC and the IDMC's independent statistician will be unblinded to treatment assignment as needed. In addition, specific vendors whose role in study conduct requires them to be unblinded (e.g., personnel operationally associated with the IWRS) may also be unblinded.

4.3. Sample Size Considerations

The study is powered to test the one-sided (statistical) superiority hypothesis stipulating that a TVB-2640 50-mg dose provides better benefit than placebo based on the primary efficacy endpoint. The study estimated total sample size is approximately 162 randomized subjects (108 on TVB-2640 and 54 on placebo). The sample size was derived using the normal approximation to the binomial distribution, and the estimated power to detect a difference in the primary efficacy response of 20% (TVB-2640 40% and placebo 20% at Week 52) was at least 80% (one-sided test at the 0.05 significance level). A 2-to-1 randomization allocation was assumed (TVB-2640 to placebo), and an anticipated dropout rate of 18.5% was also applied to estimate the final total sample size (N=162).

4.4. Independent Data Monitoring Committee

An IDMC will be established by charter to provide independent review and assessment of study data and to monitor the overall study conduct in a systematic manner to safeguard the safety of study subjects. Periodic meetings will occur as outlined in the IDMC charter. An interim analysis, which will be reviewed by the IDMC, will be performed when approximately 60 subjects with an MRI-PDFF value of ≥8% liver fat at Baseline have completed the Week 26 visit or an Early Termination Visit after Week 22. The IDMC will be provided with selected biomarker and safety interim results to allow for a more complete examination of the totality of the data. Upon review of the interim analysis findings, the IDMC may communicate with the Sponsor by making recommendations to the Sponsor to continue, amend, or stop the study based on these interim analysis results. The Sponsor will maintain its blind to any individual treatment assignments.

4.5. Interim Analyses

A single (secondary endpoint) efficacy interim analysis is planned when approximately 60 subjects with an MRI-PDFF value of ≥8% liver fat at Baseline have completed Week 26 or an Early Termination visit after Week 22. The purpose of the proposed interim analysis is to examine the TVB-2640 benefit over placebo based on a Week 26 secondary efficacy endpoint consisting of "proportion of MRI-PDFF ≥30% responders at Week 26, where an MRI-PDFF ≥30% responder is defined as a subject with ≥8% liver fat content at Baseline who achieves a percent change (reduction) from Baseline in MRI-PDFF ≥30%".

Interim safety summaries will accompany the interim analyses of the secondary efficacy endpoint.

An interim analysis of the primary efficacy endpoint is not feasible since the primary endpoint is only evaluated at Week 52, and not at Week 26. Therefore, a statistical penalty is not envisaged in conjunction with the Week 26 interim analysis of the above-stated secondary endpoint.

5. STUDY ENDPOINTS

5.1. Efficacy

The co-primary efficacy endpoints are the proportion of subjects experiencing the following:

- Histological improvement at Week 52 in NAS (i.e., ≥2-point improvement in NAS with ≥1-point improvement in ballooning or inflammation) and without worsening of fibrosis score (by NASH CRN fibrosis score)
- Resolution of steatohepatitis and no worsening of liver fibrosis (by NASH CRN fibrosis score). Resolution of steatohepatitis is defined as absence of fatty liver disease or isolated or simple steatosis without steatohepatitis and a NAS of 0 or 1 for inflammation, 0 for ballooning, and any value for steatosis

  AND

  Histological improvement in NAS (≥2 points improvement in NAS) at Week 52.

The secondary efficacy and pharmacodynamic endpoints are:

- Proportion of MRI-PDFF ≥30% responders at Week 26, where an MRI-PDFF ≥30% responder is defined as a subject with ≥8% liver fat content at Baseline who achieves a relative reduction from Baseline in MRI-PDFF ≥30%.
- Proportion of MRI-PDFF ≥30% responders at Week 52 where an MRI-PDFF ≥30% responder is defined as a subject with ≥8% liver fat content at Baseline who achieves a relative reduction from Baseline in MRI-PDFF ≥30%.
- Proportion of MRI-PDFF ≥50% responders at Week 52 where an MRI-PDFF ≥50% responder is defined as a subject with ≥8% liver fat content at Baseline who achieves a relative reduction from Baseline in MRI-PDFF ≥50%.
- Proportion of subjects with improvement in liver fibrosis ≥1 stage by NASH CRN fibrosis score) without worsening of steatohepatitis (defined as no increase in NAS for ballooning, inflammation, or steatosis) at 52 weeks.
- Proportion of subjects with improvement in liver fibrosis ≥1 stage by NASH CRN score without increase of total NAS score at Week 52.
- Proportion of subjects with improvement in liver fibrosis ≥1 stage by NASH CRN score without any worsening of NASH (defined as no worsening of ballooning and lobular inflammation; a 1 grade change in steatosis may be acceptable) at Week 52.
- Proportion of subjects experiencing **both** of the following at Week 52:
    - Resolution of steatohepatitis on overall histopathological reading and no worsening of liver fibrosis on NASH CRN fibrosis score. Resolution of

steatohepatitis is defined as absent fatty liver disease or isolated or simple steatosis without steatohepatitis and a NAS score of 0-1 for inflammation, 0 for ballooning, and any value for steatosis.

AND

- Improvement in liver fibrosis greater than or equal to one state (NASH CRN fibrosis score) and no worsening of steatohepatitis (defined as no increase in NAS for ballooning, inflammation, or steatosis).

- Proportion of subjects experiencing **both** of the following at Week 52:
  - Histological improvement at Week 52 in nonalcoholic fatty liver disease (NAFLD) activity score (NAS) (i.e., ≥2 points improvement in NAS with ≥1 point improvement in ballooning or inflammation) and without worsening of fibrosis (by NASH Clinical Research Network [CRN] fibrosis score)

  AND

  - Improvement in liver fibrosis greater than or equal to one stage (NASH CRN fibrosis score) and no worsening of steatohepatitis (defined as no increase in NAS for ballooning, inflammation, or steatosis).
- Proportion of subjects with resolution of steatohepatitis and no worsening of liver fibrosis (by NASH CRN fibrosis score). Resolution of steatohepatitis is defined as absence of fatty liver disease or isolated or simple steatosis without steatohepatitis and a NAS of 0 or 1 for inflammation, 0 for ballooning, and any value for steatosis at Week 52.
- Proportion of subjects with histological improvement (decrease) at Week 52 in NAS by ≥1 point.
- Proportion of subjects with histological improvement (decrease) at Week 52 in NAS by ≥2 points.
- Change and percent change from Baseline in MRI-PDFF liver fat (%) at Week 26.
- Change and percent change from Baseline in MRI-PDFF liver fat (%) at Week 52.
- Proportion of subjects with NASH resolution (defined as absence of fatty liver disease or isolated or simple steatosis without steatohepatitis and a NAS of 0 or 1 for inflammation, 0 for ballooning, and any value for steatosis) at 52 weeks.
- Change from Baseline in ALT, AST, alkaline phosphatase (ALP), and GGT at Weeks 26 and 52 and at each study visit.
- Change from Baseline in LDL-C and other lipid levels (including LDL/HDL ratio) at Weeks 26 and 52 and at each study visit.
- Change from Baseline in collagen/fibrous area and fibrosis score, assessed by digital pathology, at Week 52.
  - The primary digital pathology variables that will be analyzed include the following:
    - qFibrosis – continuous value

  - qFibrosis – stage
  - qSteatosis – continuous value
  - qSteatosis – grade
  - Steatosis corrected qFibrosis – continuous value
  - Steatosis corrected qFibrosis – stage

- Change from Baseline in metabolic parameters, including fasting insulin, fasting glucose, HOMA-IR, adipo-IR, and HbA1c levels at Weeks 26 and 52 and at each study visit.
- Change from Baseline in FGF-21, adiponectin, and other NASH biomarker levels at Weeks 26 and 52 and at each study visit.
- Change from Baseline in FibroScan Liver Stiffness Measurement (kPa)) and CAP score results at Weeks 26 and 52.
- Change from Baseline in PRO-C3 and other fibrosis biomarkers at Weeks 4, 13, 26, and 52.
- Change from Baseline in ELF score at Weeks 26 and 52.
- Plasma concentrations of TVB-2640 at Weeks 1, 4, 13, 26, and 52.
- Proportion of subjects with histological improvement at Week 52 in NAS (i.e., ≥2 points improvement in NAS with ≥1 point improvement in ballooning or inflammation) and without worsening of fibrosis (by NASH CRN fibrosis score)
- Proportion of subjects with resolution of steatohepatitis and no worsening of liver fibrosis (by NASH CRN fibrosis score). Resolution of steatohepatitis is defined as absence of fatty liver disease or isolated or simple steatosis without steatohepatitis and a NAS of 0 or 1 for inflammation, 0 for ballooning, and any value for steatosis.

The exploratory efficacy and pharmacodynamics endpoints are:

- Proportion of MRI-PDFF ≥30% responders at Week 52 who have ≥17 U/L reduction in ALT.
- Change from Baseline in FAST score at Weeks 26 and 52.
- Change from Baseline in Agile 3+ score at Weeks 26 and 52.

  The formula from Sanyal et al (2023) for the Agile 3+ score is:

  $$Agile\ 3+ = e^{logit(p_{F\geq 3})}/(1 + e^{logit(p_{F\geq 3})})$$

  with

  $logit(p_{F\geq 3})$ = −3.92368 + 2.29714×ln(LSM) − 0.00902×PLT − 0.98633×AAR^{-1} + 1.08636×Diabetes status − 0.38581×Sex + 0.03018×Age

  Where ln is natural log; LSM=Liver Stiffness Measurement (kPa) by vibration-controlled transient elastography (VCTE); AAR=AST/ALT ratio; PLT=platelets (G/L); Sex: male=1, female=0; and Diabetes status: yes=1, no=0.
- Change from Baseline in levels of select fibrosis biomarkers (e.g., TIMP-1) at Weeks 4,

13, 26, and 52.

- Change from Baseline in levels of lipidomic/metabolomic biomarkers (e.g., tripalmitin, ceramides, bile acids, diacylglycerols, and other classes of metabolites) at Weeks 4, 13, 26, and 52.
- Change from Baseline in HbA1c levels among subjects with T2DM at Weeks 13, 26, 39, and 52.
- Change from Baseline in FIB-4 at Weeks 4, 13, 26, and 52.
- Change from Baseline in additional fibrosis parameters such as zonal analyses of fibrosis, NAS, and individual NAS components as determined by digital pathology at Week 52.
- Correlation of changes in MRI-PDFF with changes in other endpoints.
- Presence or absence of SNPs relevant for NASH.
- Correlation of predictive signature with other endpoints.
- Correlation of lipids (e.g. tripalmitin) with other endpoints.
- Change from Baseline in gut microbiome analyses at Weeks 26 and 52 in a substudy of subjects.
- Change from Baseline in HepVita Nutritional Survey at Weeks 26 and 52 in a substudy of subjects.
- Change in MELD score from baseline to Week 52

  The formula from Kamath et al (2001) for the MELD score is 3.8×ln(bilirubin [mg/dL]) + 11.2×ln(INR) + 9.6×ln(creatinine [mg/dL]) + 6.4×(etiology: 0 if cholestatic or alcoholic, 1 otherwise).

  Where ln is natural log.

5.2. Safety

The safety endpoints are:

- Adverse events (AEs) throughout the study.
- Electrocardiograms (ECGs) at Weeks 26 and 52.
- Vital signs at each study visit.
- Physical examination findings at each study visit.
- Eye examination findings at each study visit.
- Hematology and coagulation results at each study visit.
- Liver aminotransferase levels at each study visit.
- Clinical chemistry results at each study visit.
- Urinalysis results at each study visit.

- Pregnancy test results throughout the study.
- Prior and concomitant medications throughout the study.
- Results of anthropometric measurements at Week 52.

6. ANALYSIS POPULATIONS/SETS

Intention-to-Treat (ITT) Population: The principal analysis of the primary and secondary efficacy endpoints will be based on the Intention-to-Treat (ITT) population, which comprises all randomized subjects; this population will serve as the basis for all efficacy analyses. Subjects' data will be analyzed according to their randomized treatment assignment.

Modified Intention-to-Treat (mITT) Population: Comprises all subjects in the ITT population who have completed at least 42 weeks of treatment and have an evaluable post treatment histological assessment. Subjects will be analyzed according to the randomized treatment assignment.

Interim Analyses Modified Intention-to-treat (IAmITT) Population: Comprises a subset of approximately 60 subjects in the ITT population with 8% >/= MRI-PDFF at screening or at baseline with at least 1 evaluable post-Baseline MRI-PDFF value, defined as an MRI-PDFF assessment obtained at least 22 weeks after the first dose of study drug.

Safety (SAF) Population: Consists of all randomized subjects who received at least 1 dose (partial or complete) of study drug. This population will be used for all summaries of subject accountability, demographic and baseline data, and safety information, including AE incidence, clinical laboratory data, ECG data, and vital signs data. Subjects' data will be summarized according to the treatment they received. If all subjects were dosed according to randomized treatment assignment, then the safety and ITT analysis sets are identical.

PK Population: Comprises all ITT subjects with available PK measurements.

Serial PK Population: Comprises the subset of subjects who provided blood samples for evaluation of serial PK at timepoints on Day 1 and at Week 4.

7. GENERAL DATA HANDLING CONSIDERATIONS

All analyses will be conducted using SAS software Version 9.4 or higher.

All data in the database will be presented in by-subject data listings. Unless otherwise stated, all listings will be sorted by treatment group, center ID, subject number, and assessment date and time, if available.

Unless stated otherwise, continuous data will be summarized by treatment group using descriptive statistics including number of subjects (n), mean, median, standard deviation (SD), first quartile (Q1), third quartile (Q3), minimum (min) value, maximum (max) value.

Unless stated otherwise, categorical data will be summarized by treatment group using counts and percentage based on the number of non-missing values. For selected summaries, the number of missing values will be presented as a separate category with no percentage, if one or more subjects have missing data. Counts of zero will be presented without percentages.

Descriptive statistics will be presented with the following numerical precision:

- Minimum and Maximum: same as the raw data collected
- Mean, Median, Q1, and Q3: one additional decimal place to that reported for Minimum and Maximum
- SD: two additional decimal places than the Minimum and Maximum
- Percentages: reported to one decimal place if percent is <100. If the value is 100, no decimal place will be reported.

Unless otherwise noted, statistical inference will be based on two-sided testing with a 5% significance level. Confidence intervals with be at the 90% and 95% levels.

P-values will be reported to four decimal places. If the value is below 0.0001 it will be noted as < 0.0001; if the value above 0.9999 it will be noted as > 0.9999.

All data up to the time of study completion/withdrawal from the study will be included in the analysis, regardless of duration of treatment.

Numbering for data displays will be based on International Council for Harmonization (ICH) E3 whenever possible.

7.1. Stratification and Covariates

Subjects will be randomized to achieve a 2:1 ratio of TVB-2640 50 mg to placebo, with the randomization stratified three factors: T2DM status (Yes, No), and amount of fibrosis (F2 or F3). Selected statistical models and tests will be stratified based on the randomization strata. Details on the models and tests are provided in the relevant sections below.

7.2. Evaluation of Subgroups

Select analyses for efficacy and safety endpoints will be repeated for a variety of subgroups. The subgroups are summarized in the table below. If a subject does not have a value for one of the subgroup variables they will be excluded from the subgroup. Efficacy endpoints will be analyzed by subgroup for both the ITT and mITT populations. Safety endpoints will be summarized by subgroup for the Safety population.

Endpoint	**Subgroups**
Co-Primary Efficacy Endpoints	

Endpoint	**Subgroups**
Proportion of subjects experiencing histological improvement at Week 52 in NAS (i.e., ≥2 points improvement in NAS with ≥1 point improvement in ballooning or inflammation) and without worsening of fibrosis (by NASH CRN fibrosis score)	• ≥30% reduction in MRI-PDFF from baseline to Week 52 (yes or no) – includes subjects with ≥8% liver fat content at Baseline • ≥50% reduction in MRI-PDFF from baseline to Week 52 (Yes or No) – includes subjects with ≥8% liver fat content at Baseline • Stage of fibrosis at baseline (F2 or F3) • T2DM at baseline (Yes or No) • GLP-1 at baseline (Yes or No) [1] • SGLT2 at baseline (Yes or No) [2] • Statin use at baseline (Yes or No) [3] • Diabetes medication use at baseline (Yes or No) [4] • <5% reduction in weight from baseline to Week 52 (Yes or No) • Liver fat ≥8% at baseline (Yes or No) • ≥17 U/L reduction in ALT from baseline to Week 52 (Yes or No) • NAS at baseline (4/5 or 6/7/8) • Race (White, Black, Asian, and other) • Hispanic (Yes or No) • Liver biopsy length ≥ 15mm (both baseline and Week 52 biopsies) vs < 15mm • Age categories (<40 years, 40 to <65 years, and ≥65 years)
Proportion of subjects experiencing **both** of the following at Week 52: o Resolution of steatohepatitis and no worsening of liver fibrosis (by NASH CRN fibrosis score). Resolution of steatohepatitis is defined as absence of fatty liver disease or isolated or simple steatosis without steatohepatitis and a NAS of 0 or 1 for inflammation, 0 for ballooning, and any value for steatosis. AND	• ≥30% reduction in MRI-PDFF from baseline to Week 52 (yes or no) – includes subjects with ≥8% liver fat content at Baseline • ≥50% reduction in MRI-PDFF from baseline to Week 52 (Yes or No) – includes subjects with ≥8% liver fat content at Baseline • Stage of fibrosis at baseline (F2 or F3) • T2DM at baseline (Yes or No) • GLP-1 at baseline (Yes or No) [1] • SGLT2 at baseline (Yes or No) [2] • Statin use at baseline (Yes or No) [3] • Diabetes medication use at baseline (Yes or No) [4] • <5% reduction in weight from baseline to Week 52 (Yes or No) • Liver fat ≥8% at baseline (Yes or No) • ≥17 U/L reduction in ALT from baseline to Week 52 (Yes or No) • NAS at baseline (4/5 or 6/7/8)

Endpoint	**Subgroups**
○ Histological improvement in NAS (≥2 points improvement in NAS).	• Race (White, Black, Asian, and other) • Hispanic (Yes or No) • Liver biopsy length ≥ 15mm (both baseline and Week 52 biopsies) vs < 15mm • Age categories (<40 years, 40 to <65 years, and ≥65 years)
Selected Secondary Efficacy and Pharmacodynamic Endpoints	
Proportion of subjects with improvement in liver fibrosis ≥1 stage by NASH CRN fibrosis score without worsening of steatohepatitis (defined as no increase in NAS for ballooning, inflammation, or steatosis) at Week 52.	• Stage of fibrosis at baseline (F2 or F3) • GLP-1 at baseline (Yes or No) [1] • SGLT2 at baseline (Yes or No) [2]
Proportion of subjects with improvement in liver fibrosis ≥1 stage by NASH CRN score without increase of total NAS score at Week 52.	• Stage of fibrosis at baseline (F2 or F3) • GLP-1 at baseline (Yes or No) [1] • SGLT2 at baseline (Yes or No) [2]
Proportion of subjects with improvement in liver fibrosis ≥1 stage by NASH CRN score without any worsening of NASH (defined as no worsening of ballooning and lobular inflammation; a 1 grade change in steatosis may be acceptable) at Week 52.	• Stage of fibrosis at baseline (F2 or F3) • GLP-1 at baseline (Yes or No) [1] • SGLT2 at baseline (Yes or No) [2]
Proportion of subjects experiencing **both** of the following at Week 52: ○ Resolution of steatohepatitis on overall histopathological reading and no worsening of liver fibrosis on NASH CRN fibrosis score. Resolution of steatohepatitis is defined as absent fatty liver disease or isolated or simple steatosis without steatohepatitis and a	• Stage of fibrosis at baseline (F2 or F3) • GLP-1 at baseline (Yes or No) [1] • SGLT2 at baseline (Yes or No) [2]

Endpoint	Subgroups
NAS score of 0-1 for inflammation, 0 for ballooning, and any value for steatosis. AND o Improvement in liver fibrosis greater than or equal to one state (NASH CRN fibrosis score) and no worsening of steatohepatitis (defined as no increase in NAS for ballooning, inflammation, or steatosis).	
Proportion of subjects experiencing **both** of the following at Week 52: o Histological improvement at Week 52 in nonalcoholic fatty liver disease (NAFLD) activity score (NAS) (i.e., ≥2 points improvement in NAS with ≥1 point improvement in ballooning or inflammation) and without worsening of fibrosis (by NASH Clinical Research Network [CRN] fibrosis score) AND o Improvement in liver fibrosis greater than or equal to one stage (NASH CRN fibrosis score) and no worsening of steatohepatitis (defined as no increase in NAS for ballooning,	• Stage of fibrosis at baseline (F2 or F3) • GLP-1 at baseline (Yes or No) [1] • SGLT2 at baseline (Yes or No) [2]

Endpoint	Subgroups
inflammation, or steatosis).	
Proportion of subjects experiencing resolution of steatohepatitis and no worsening of liver fibrosis (by NASH CRN fibrosis score). Resolution of steatohepatitis is defined as absence of fatty liver disease or isolated or simple steatosis without steatohepatitis and a NAS of 0 or 1 for inflammation, 0 for ballooning, and any value for steatosis at Week 52.	• Stage of fibrosis at baseline (F2 or F3) • GLP-1 at baseline (Yes or No) [1] • SGLT2 at baseline (Yes or No) [2]
Proportion of subjects with histological improvement (decrease) at Week 52 in NAS by ≥1 point.	• Stage of fibrosis at baseline (F2 or F3) • GLP-1 at baseline (Yes or No) [1] • SGLT2 at baseline (Yes or No) [2]
Proportion of subjects with histological improvement (decrease) at Week 52 in NAS by ≥2 points.	• Stage of fibrosis at baseline (F2 or F3) • GLP-1 at baseline (Yes or No) [1] • SGLT2 at baseline (Yes or No) [2]
Change from Baseline in ALT, AST, and GGT at Weeks 26 and 52 and at each study visit.	• Stage of fibrosis at baseline (F2 or F3) • T2DM at baseline (Yes or No)
Change from Baseline in LDL-C and other lipid levels (including LDL/HDL ratio) at Weeks 26 and 52 and at each study visit.	• T2DM at baseline (Yes or No) • Statin use at baseline (Yes or No) [3]
Changes from Baseline in metabolic parameters, including fasting insulin, fasting glucose, HOMA-IR, adipo-IR, and HbA1c levels at Weeks 26 and 52 and at each study visit.	• T2DM at baseline (Yes or No)
Change from Baseline in FibroScan (Liver Stiffness Measure (kPA)) and CAP score results at Weeks 26 and 52.	• ≥30% reduction in MRI-PDFF from baseline to Week 52 (yes or no) – includes subjects with ≥8% liver fat content at Baseline • ≥50% reduction in MRI-PDFF from baseline to Week 52 (Yes or No) – includes subjects with ≥8% liver fat content at Baseline

Endpoint	Subgroups
Selected Exploratory Endpoints	
Change from Baseline in FAST score at Weeks 26 and 52.	• 30% reduction in MRI-PDFF from baseline to Week 52 (yes or no) – includes subjects with ≥8% liver fat content at Baseline • ≥50% reduction in MRI-PDFF from baseline to Week 52 (Yes or No) – includes subjects with ≥8% liver fat content at Baseline
Change from Baseline in Agile 3+ score at Weeks 26 and 52.	• 30% reduction in MRI-PDFF from baseline to Week 52 (yes or no) – includes subjects with ≥8% liver fat content at Baseline • ≥50% reduction in MRI-PDFF from baseline to Week 52 (Yes or No) – includes subjects with ≥8% liver fat content at Baseline
Change from Baseline in levels of lipidomic/metabolomic biomarkers (e.g., tripalmitin, ceramides, bile acids, diacylglycerols, and other classes of metabolites) at Weeks 4, 13, 26, and 52.	• On a GLP-1 at baseline (Yes or No) [1]
Selected Safety Endpoints	
Skin and subcutaneous tissue disorders TEAEs	• Diabetes medication use at baseline (Yes or No) [4] • COVID-19 diagnosis or COVID-19 vaccination within 30 days of skin and subcutaneous tissue disorders TEAE onset (Yes or No) • Statin use at baseline (Yes or No) [3] • GLP-1 or SGLT2 at baseline (Yes or No) [1] [2] • Female age (>50 or ≤50 years)
Eye disorder TEAEs	• Diabetes medication use at baseline (Yes or No) [4] • COVID-19 diagnosis or COVID-19 vaccination within 30 days of eye disorder TEAE onset (Yes or No) • Statin use at baseline (Yes or No) [3] • GLP-1 or SGLT2 at baseline (Yes or No) [1] [2]

[1] GLP-1 medications are identified using ATC code = 'A10BJ'

[2] SGLT2 medications are identified using ATC code = 'A10BK'
[3] Statin medications are identified using ATC code = 'C10AA'
[4] Diabetes medications are identified using ATC code = 'A10' at baseline.

7.3. Multiple Comparisons and Multiplicity

The study co-primary hypotheses are the (statistical) one-sided superiority of a TVB-2640 50-mg dose over placebo based on each of the two co-primary efficacy endpoints at the 0.05 significance level, where a higher responder proportion is a better outcome. The statistical analysis (one-sided) test (CMH) of the primary hypothesis is declared if $p<0.05$. No adjustment for the two co-primary efficacy tests will be performed, both will be tested at the 0.05 level using one-sided tests.

Both one-sided and two-sided 0.05 significance level tests will be used for the efficacy endpoints. The table below specifies whether the test of significance is one- or two-sided along with the expected direction of effect. No other control for multiplicity will be provided.

Note that endpoints that are being tested as one-sided will also have the two-sided p-values provided. Both 90% and 95% two-sided confidence intervals will be shown.

7.4. Efficacy Endpoints: 1-sided and 2-sided Testing Allocation

Endpoint Number	**Endpoint Description**	**1-sided or 2-sided**	**Directionality for better outcome for 1-sided tests**
	Co-Primary Efficacy Endpoints		
1A	Proportion of subjects experiencing histological improvement at Week 52 in NAS (i.e., ≥2 points improvement in NAS with ≥1 point improvement in ballooning or inflammation) and without worsening of fibrosis (by NASH CRN fibrosis score)	1-sided	Higher
1B	Proportion of subjects experiencing **both** of the following at Week 52: o Resolution of steatohepatitis and no worsening of liver fibrosis (by NASH CRN fibrosis score). Resolution of steatohepatitis is defined as absence of fatty liver disease or isolated or simple steatosis without steatohepatitis and a NAS of 0 or 1 for inflammation, 0 for ballooning, and any value for steatosis.	1-sided	Higher

Endpoint Number	Endpoint Description	1-sided or 2-sided	Directionality for better outcome for 1-sided tests
	AND o Histological improvement in NAS (≥2 points improvement in NAS).		
	Secondary Efficacy and Pharmacodynamic Endpoints		
2	Proportion of MRI-PDFF ≥30% responders at Week 26, where an MRI-PDFF ≥30% responder is defined as a subject with ≥8% liver fat content at Baseline who achieves a relative reduction from Baseline in MRI-PDFF ≥30%.	2-sided	
3	Proportion of MRI-PDFF ≥30% responders at Week 52, where an MRI-PDFF ≥30% responder is defined as a subject with ≥8% liver fat content at Baseline who achieves a relative reduction from Baseline in MRI-PDFF ≥30%.	2-sided	
4	Proportion of MRI-PDFF ≥50% responders at Week 52, where an MRI-PDFF ≥50% responder is defined as a subject with ≥8% liver fat content at Baseline who achieves a relative reduction from Baseline in MRI-PDFF ≥50%.	2-sided	
5	Proportion of subjects with improvement in liver fibrosis ≥1 stage by NASH CRN fibrosis score without worsening of steatohepatitis (defined as no increase in NAS for ballooning, inflammation, or steatosis) at Week 52.	1-sided	Higher
6	Proportion of subjects with improvement in liver fibrosis ≥1 stage by NASH CRN score without increase of total NAS score at Week 52.	1-sided	Higher
7	Proportion of subjects with improvement in liver fibrosis ≥1 stage by NASH CRN score without any worsening of NASH (defined as no worsening of ballooning and lobular	1-sided	Higher

Endpoint Number	Endpoint Description	1-sided or 2-sided	Directionality for better outcome for 1-sided tests
	inflammation; a 1 grade change in steatosis may be acceptable) at Week 52.		
8	Proportion of subjects experiencing **both** of the following at Week 52: o Resolution of steatohepatitis on overall histopathological reading and no worsening of liver fibrosis on NASH CRN fibrosis score. Resolution of steatohepatitis is defined as absent fatty liver disease or isolated or simple steatosis without steatohepatitis and a NAS score of 0-1 for inflammation, 0 for ballooning, and any value for steatosis AND o Improvement in liver fibrosis greater than or equal to one state (NASH CRN fibrosis score) and no worsening of steatohepatitis (defined as no increase in NAS for ballooning, inflammation, or steatosis).	1-sided	Higher
9	Proportion of subjects experiencing **both** of the following at Week 52: o Histological improvement at Week 52 in nonalcoholic fatty liver disease (NAFLD) activity score (NAS) (i.e., ≥2 points improvement in NAS with ≥1 point improvement in ballooning or inflammation) and without worsening of fibrosis (by NASH Clinical Research Network [CRN] fibrosis score) AND o Improvement in liver fibrosis greater than or equal to one stage	1-sided	Higher

Endpoint Number	Endpoint Description	1-sided or 2-sided	Directionality for better outcome for 1-sided tests
	(NASH CRN fibrosis score) and no worsening of steatohepatitis (defined as no increase in NAS for ballooning, inflammation, or steatosis).		
10	Proportion of subjects experiencing resolution of steatohepatitis and no worsening of liver fibrosis (by NASH CRN fibrosis score). Resolution of steatohepatitis is defined as absence of fatty liver disease or isolated or simple steatosis without steatohepatitis and a NAS of 0 or 1 for inflammation, 0 for ballooning, and any value for steatosis at Week 52.	1-sided	Higher
11	Proportion of subjects with histological improvement (decrease) at Week 52 in NAS by ≥1 point.	1-sided	Higher
12	Proportion of subjects with histological improvement (decrease) at Week 52 in NAS by ≥2 points.	1-sided	Higher
13	Change and percent change from Baseline in MRI-PDFF liver fat (%) at Week 26.	2-sided	
14	Change and percent change from Baseline in MRI-PDFF liver fat (%) at Week 52.	2-sided	
15	Proportion of subjects with NASH resolution (defined as absence of fatty liver disease or isolated or simple steatosis without steatohepatitis and a NAS of 0 or 1 for inflammation, 0 for ballooning, and any value for steatosis) at 52 weeks.	1-sided	Higher
16	Change from Baseline in ALT, AST, ALP, and GGT at Weeks 26 and 52 and at each study visit.	2-sided	
17	Change from Baseline in LDL-C and other lipid levels (including LDL/HDL ratio) at Weeks 26 and 52 and at each study visit.	2-sided	

Endpoint Number	Endpoint Description	1-sided or 2-sided	Directionality for better outcome for 1-sided tests
18	Change from Baseline in collagen/fibrous area and fibrosis score, assessed by digital pathology, at Week 52.	1-sided	Lower
19	Changes from Baseline in metabolic parameters, including fasting insulin, fasting glucose, HOMA-IR, adipo-IR, and HbA1c levels at Weeks 26 and 52 and at each study visit.	2-sided	
20	Change from Baseline in FGF-21, adiponectin, and other NASH biomarker levels at Weeks 26 and 52 and at each study visit.	2-sided	
21	Change from Baseline in FibroScan (Liver Stiffness Measure (kPa)) and CAP score results at Weeks 26 and 52.	1-sided	Lower
22	Change from Baseline in PRO-C3 and other fibrosis biomarkers at Weeks 4, 13, 26, and 52.	2-sided	
23	Change from Baseline in ELF score at Weeks 26 and 52.	2-sided	
24	Plasma concentrations of TVB-2640 at Weeks 1, 4, 13, 26, and 52.	N/A	
	Exploratory Efficacy and Pharmacodynamic Endpoints		
25	Proportion of MRI-PDFF ≥30% responders at Week 52 who have ≥17 U/L reduction in ALT.	2-sided	
26	Change from Baseline in FAST score at Weeks 26 and 52.	1-sided	Lower
27	Change from Baseline in Agile 3+ score at Weeks 26 and 52.	1-sided	Lower
28	Change from Baseline in levels of select fibrosis biomarkers (e.g., TIMP-1) at Weeks 4, 13, 26, and 52.	2-sided	

Endpoint Number	Endpoint Description	1-sided or 2-sided	Directionality for better outcome for 1-sided tests
29	Change from Baseline in levels of lipidomic/metabolomic biomarkers (e.g., tripalmitin, ceramides, bile acids, diacylglycerols, and other classes of metabolites) at Weeks 4, 13, 26, and 52.	2-sided	
30	Change from Baseline in HbA1c levels among subjects with T2DM at Weeks 13, 26, 39, and 52.	2-sided	
31	Change from Baseline in FIB-4 at Weeks 4, 13, 26, and 52.	2-sided	
32	Change from Baseline in additional fibrosis parameters such as zonal analyses of fibrosis, NAS, and individual NAS components as determined by digital pathology at Week 52.	N/A*	
33	Correlation of changes in MRI-PDFF with changes in other endpoints.	N/A*	
34	Presence or absence of SNPs relevant for NASH.	N/A	
35	Correlation of predictive signature with other endpoints.	N/A*	
36	Correlation of lipids (e.g. tripalmitin) with other endpoints.	N/A*	
37	Change from Baseline in gut microbiome analyses at Weeks 26 and 52 in a substudy of subjects.	N/A*	
38	Change from Baseline in HepVita Nutritional Survey at Weeks 26 and 52 in a substudy of subjects.	N/A*	
39	Change in MELD score from baseline to Week 52	1-sided	Lower

* These endpoints may be analyzed at a later date.

7.5. Reference Dates

- Screening date is defined as the electronic case report form (eCRF) provided date on which a subject was screened for study entry.
- Randomization date is defined as the date on which the subject is randomized to study treatment.
- Treatment start date is defined as the date of first dose of study drug (also referred to as Day 1).
- Treatment end date is defined as the date of last dose of study drug.
- The calculation of age will use the informed consent date as its reference date.
- Safety data, such as AEs and laboratory assessments will use the treatment start date as a reference date to assign treatment emergence.
- Study day will be based on treatment start date as a reference date.

Reference day calculations will be defined as:

- date of interest – reference date + 1 when the date of interest ≥ reference date.
- otherwise, date of interest – reference date.

For example, duration of treatment is defined as treatment end date – treatment start date + 1.

If either date is missing, reference date calculations will not be performed.

7.6. Baseline and Post-Baseline Changes

Unless stated otherwise, baseline will be based on the last non-missing value collected prior to the start of study treatment based on the treatment start date and time, if applicable. Post-baseline values will be those collected after the treatment start date and time, if applicable.

Change from baseline is defined as: post-baseline value – baseline value.

Percentage change from baseline is defined as: (post-baseline value – baseline value)/baseline value x 100%.

7.7. Handling Below Limit of Quantification (BLQ) or Lower Limit of Quantification (LLOQ) Values

Any parameters with values identified as <x, ≤x, >x, or ≥x will be analyzed using the ‘x’ value (the analysis will ignore the <, ≤, >, or ≥ symbol).

7.8. Imputation of Partial Dates

Imputed dates will be used to classify treatment emergent AEs (TEAEs) and prior and concomitant medications.

Adverse Events

- If the AE start date is completely missing the event will be assumed to have started on the first day of treatment and will be considered a TEAE. If a subject was not treated, then no imputation will be conducted for missing start dates.
- If the AE start date is missing day and month, do the following:
    - If the AE start year is the same year as that of the first treatment and the AE contains information to indicate that the event ended before the treatment start date (e.g., AE end date month and year are earlier than the treatment start date or the full AE end date is known and occurs earlier than the treatment start date), then set the AE day/month to '1 January'.
    - If the AE start year is the same as that of the first treatment and the AE contains information to indicate that the event ended after the treatment start date, then set the AE start day/month to the treatment start day/month
    - If the AE start year is later than that of the first treatment, then set the AE start day/month to '1 January.'
    - If the AE start year falls before that of the first treatment, then the AE start day/month will not be imputed.
- If only the AE start day is missing, do the following:
    - If the AE start year is the same year as that of the first treatment and the AE month is prior to the first treatment date, then set the AE day to '1'.
    - If the AE start month/year are the same as that of the first treatment and the AE contains information to indicate that the event ended before the treatment start date (e.g., the full AE end date is known and occurs earlier than the treatment start date), then set the AE day to '1'.
    - If the AE start month/year are the same as the first treatment and the AE end date is after the treatment start date, then set the AE day to the same day as treatment start day.
    - If the AE start year is the same as that of the first treatment and the AE start month is after the first treatment, or the AE start year is later than that of the first treatment, then set AE day to '1'.
- AE end dates will not be imputed.

Prior and Concomitant Medications

- The imputation rules for AE start dates will be used for prior and concomitant medication start dates.
- Prior and concomitant medication stop dates will be imputed as follows:
    - If the stop date is only missing the day, then the stop day is the last day of the month
    - If the stop date is missing both the day and month, then the stop month and day is December 31

- If the stop date is completely missing, no imputation is performed, and the medication will be classified as a concomitant medication for subjects who were treated.

7.9. Multiple Assessments and Visit Windows

Nominal visits (e.g., those identified by the study eCRF) will be the basis of summarization and statistical analysis; no visit date windowing will be conducted. Unless specified otherwise, data from unscheduled visits will be included in: (1) summaries of most extreme and baseline data, (2) summaries of specific abnormalities at any time post-baseline, (3) adverse events and any other summaries not based on study visit, (4) subject data listings, and (5) analyses that utilize MRI-PDFF, PK, lipidomic/metabolic biomarkers, collagen/fibrous area and fibrosis score assessed by digital pathology, SNPs relevant for NASH, and PRO-C3 data where unscheduled visit results may be used in analyses at the closest scheduled visit; if the closest scheduled visit already has valid MRI-PDFF data then the unscheduled visit results will not be used in the analysis.

7.10. Handling of Early Termination Visit Information

If a subject is terminated early from this study the early termination visit data will be analyzed at the closest scheduled visit. If the closest visit has valid data, the early termination data will be assigned to the next available visit.

7.11. Missing Data

Imputation will only be performed on the primary and selected secondary efficacy endpoints based on histological and MRI-PDFF data. Imputed data will be analyzed using both the ITT and mITT populations.

For the primary efficacy endpoint and secondary efficacy endpoints based on histology, a subject who discontinued treatment prior to Week 42, does not provide a histological assessment on or after Week 42, or who otherwise has missing data will be considered a non-responder. A pre-specified sensitivity analysis of the primary efficacy endpoint will employ the multiple imputation technique.

For the secondary endpoint MRI-PDFF ≥ 30% responder analysis at Week 26, a subject who discontinued treatment prior to Week 20, does not provide an MRI-PDFF assessment between Week 20 through Week 27, or who otherwise has missing data will be considered a non-responder. For analysis of MRI-PDFF, missing data at Week 26 will be imputed using the last post-baseline observation obtained at least 20 weeks to 27 weeks after the first dose. For the MRI-PDFF ≥ 30% responder analysis at Week 52, a subject who discontinued treatment prior to Week 42, does not provide an MRI-PDFF assessment after Week 42 or who otherwise has missing data will be considered a non-responder. Missing data at Week 52 will be imputed using the last post-baseline observation obtained at least 42 weeks after the first dose. These missing data techniques will also be applied to MRI-PDFF ≥ 50% responder analyses.

Additional details are provided in the relevant analysis sections below.

8. STUDY SUBJECT DATA

8.1. Subject Disposition

Summaries of analysis population membership, final subject status (completed or withdrawn), including reasons for withdrawal, sub study participation, and study duration in days will be produced based on the number of randomized subjects. The number of subjects screened will be provided. Data will be presented by treatment group as well as overall for the study, with exception of the number of screen failures (to be displayed without respect to treatment).

Screen failures, analysis populations, and final subject disposition status will be listed.

8.2. Protocol Deviations

Protocol deviations will be identified and classified as important or not important. Important protocol deviations may include but are not limited to:

- Violation of Inclusion/Exclusion Criteria
- Study drug compliance ≤ 80% or ≥ 120%
- Use of prohibited therapies
- Incorrect treatment

Protocol deviations including category will be summarized by treatment group and overall for the study. Protocol deviations due to COVID impact will be identified as such.

A listing of all protocol deviations will be provided.

8.3. Demographic and Baseline Characteristics

Subject demographics and baseline characteristics will be summarized and listed. Data will be presented by treatment group as well as overall for the study.

Demographic and baseline characteristics summaries will be provided for the ITT population, mITT population, safety population, and Serial PK population.

8.4. NASH History

Subject NASH history will be summarized and listed. Days since biopsy, fibrosis stage, NAS, Steatosis score, ballooning degeneration score, lobular inflammation score, Fibrosis-4 (FIB-4) Score, and MRI-PDFF.

NASH history summaries will be provided for the ITT population, mITT population, safety population, Serial PK population. NASH history will also be summarized for T2DM vs not T2DM.

NASH confirmation data collected at screening will be presented in a data listing.

8.5. Medical History

Medical history will be coded using the Medical Dictionary for Regulatory Activities (MedDRA) version 24.0 or higher for reporting by system organ class (SOC) and preferred term (PT). Medical history data will be presented in a listing.

8.6. Prior and Concomitant Medication

The incidence of medication use will be summarized by WHO Drug Dictionary, Global B3 Mar2021 or higher, anatomic therapeutic class (ATC) Level 2 classification (i.e., therapeutic main group) and preferred name. A subject will be counted only once at each level of reporting. Prior medications are medications that have been discontinued prior to the treatment start date (e.g., taken exclusively during the pre-therapy period). Concomitant medications are medications taken at any point during the on-therapy or post-therapy periods. Subjects could start a concomitant medication prior to study start which is ongoing at time of first treatment or subjects could start a new medication during the on-therapy or post-therapy periods.

Concomitant medication use will be summarized separately and presented by treatment group for the ITT population.

All prior and concomitant medication data will be listed including the verbatim and preferred drug name and ATC Level 2.

8.7. Study Drug Exposure and Compliance

Duration of drug exposure will be summarized both as a continuous and a categorical variable for the safety population. Exposure will be calculated as:

study drug stop date – study drug start date + 1 [converted to weeks (days divided by 7)]

Categories for reporting exposure will include < 13 weeks; 13 weeks - < 26 weeks; 26 weeks - <39 weeks; 39 weeks - < 52 weeks; ≥ 52 weeks.

Drug accountability is captured in the eCRF at each visit. The eCRF captures the number of pills returned and the number missed during the dosing period. Overall compliance will be determined based on pill counts over each visit period and taking into account the number missed. Compliance will be derived as:

$$\frac{[(\#\text{ of pills dispensed} - \#\text{ of pills returned}) - \#\text{ of missed pills}] \times 100}{2 \times (\text{number of days within the visit period})}$$

Where number of days within the visit period is derived as:

Date of Visit (X+1) – Date of Visit (X) + 1

Compliance will also be summarized for the following categories < 80%, 80% - 120%, and > 120%.

Mean daily dose (mg) and total cumulative dose (mg) received will also be summarized. Total cumulative dose will be summarized overall and separately for those that had extended dosing.

Listings of planned and actual treatments, overall compliance, and drug exposure will be produced.

9. EFFICACY

9.1. Primary Efficacy Endpoint Analyses

9.1.1. Principal Analysis of the Primary Efficacy Endpoints

The principal efficacy analyses will be conducted on the ITT population. Additional sensitivity and supplementary analyses will be performed for the primary efficacy endpoints, as described in the next section.

The treatment groups will be compared based on the primary efficacy endpoints using one-sided Cochran-Mantel-Haenszel (CMH) tests stratified by the randomization stratifications T2DM status at baseline (yes or no) and amount of fibrosis at baseline (F2 or F3).

Mathematically, the null and alternative hypotheses for the primary efficacy endpoints are stated as:

Ho: $p_T \leq p_p$

Ha: $p_T > p_p$

Where, p_T is the proportion of responders in the TVB-2640 50-mg treatment group and p_p is the proportion of responders in the placebo treatment group. The higher responder proportion is a better outcome.

A Clopper-Pearson 90% and 95% CI will be derived for each treatment group, and the stratified CMH odds ratio (with its 90% and 95% CI, using the control group as a reference group) will be produced to assess the treatment effect magnitude. The difference in treatment groups response rates will be provided and also be characterized by the 90% and 95% CIs using Miettinen and Nurminen (1985).

Subgroup analyses of the primary efficacy endpoints will be performed to assess whether the treatment effect is concordant among subgroups. The planned subgroup analyses are described in Section 7.2. Analyses will be performed if there is an adequate number of subjects in each subgroup.

A subject who discontinues treatment prior to Week 42, does not provide a histological assessment on or after Week 42, or who otherwise has missing data will be considered a non-responder.

9.1.2. Sensitivity Analyses of the Primary Efficacy Endpoints

Evaluation of the robustness of the principal analysis of the primary efficacy endpoints will be carried out using the following (4) four approaches:

- A logistic regression with response as a dependent variable; treatment group, randomization stratification factors (T2DM status at baseline [yes or no] and amount of fibrosis at baseline [F2 or F3]) as independent factors; and age as a baseline covariate.
- The analysis of the primary efficacy endpoints will be repeated for the modified Intention-to-Treat (mITT) population to ascertain any changes in the conclusion of the principal analysis findings.
- A repeat of the primary efficacy analysis principal analysis (CMH test) for the following three ITT subsets:
  - ITT Subjects who achieve <u>at least 50% of total cumulative study drug dose</u> (i.e., at least **9100 mg** 52-week total dose (based on a total of 50 mg.kg x 52 weeks x 7 days = 18200 mg x weeks x days).
  - ITT Subjects who achieve <u>at least 70% of total cumulative study drug dose</u> (i.e., at least **12740 mg** 52-week total dose (based on a total of 50 mg x 52 weeks x 7 days = 18200 mg x weeks x days).
  - ITT Subjects who achieve <u>at least 80% of total cumulative study drug dose</u> (i.e., at least **14560 mg** 52-week total dose (based on a total of 50 mg.kg x 52 weeks x 7 days = 18200 mg x weeks x days).
- To assess the impact of missing data in the primary endpoint analyses, multiple imputation procedure will be carried out using MAR assumption with Fully Conditional Specification method for binary endpoint in SAS PROC MI. 100 imputations will be performed. The CMH test statistic from each imputed dataset will be based on Wilson-Hilferty transformation (Wilson & Hilferty, 1931) and will be combined using PROC MIANALYZE. Combined CMH test statistic and p-value will be computed. The method described in Ratitch, Lipkovich, and O'Kelly (2013) will be used to combine the results.

9.2. Secondary Efficacy Endpoints and Analyses

Secondary efficacy analyses for the other histological responder endpoints will be analyzed in the ITT and mITT populations using the methods described for the primary efficacy endpoint.

Secondary efficacy analyses of endpoints expressed as percent change or change from Baseline only at Week 26 or only at Week 52 will be analyzed in the ITT analysis set using an analysis of covariance (ANCOVA) model with fixed effects for T2DM status at baseline (Yes, No) and treatment group, and a covariate term for baseline analysis value of response. Least squares (LS) means for the percent change from Baseline by treatment and the associated standard errors, the

LS means for the difference between treatment groups, and the associated two-sided 95% CIs and two-sided p-values, based on t-tests, will be derived from the ANCOVA model.

For secondary endpoints assessed at multiple visits, a linear mixed-effects model for repeated measures (MMRM) may be used. The model will include T2DM status at baseline (Yes, No), treatment group, visit, and treatment-by-visit interaction as the fixed effects, and baseline value as a covariate, and the model will use an unstructured variance-covariance matrix. The point estimates for the least-squares mean of the difference treatment groups at each visit and the corresponding 95% confidence interval and two-sided p-value will be summarized. If convergence problems arise when fitting the model, a spatial power covariance structure will be used as a first step. If the model still has convergence issues, compound symmetry covariance structure will be employed. The Kenward-Roger correction for the denominator degrees of freedom will be applied.

ALT, AST. ALP, and GGT will also be summarized in shift tables of baseline to each visit based on range categories of low (below lower limit of normal [LLN], normal, and high [above upper limit of normal [ULN]).

The continuous digital pathology variables (qFibrosis, qSteatosis, and Steatosis corrected qFibrosis) will be summarized descriptively with change from baseline being analyzed using MMRM methods. Digital pathology stage/grade variables (qFibrosis – stage, qSteatosis – grade, and Steatosis corrected qFibrosis – stage) will be summarized descriptively with change from baseline being analyzed using ANCOVA methods. The ANCOVA models will include fixed effects for treatment group, T2DM status at baseline (Yes, No), and amount of fibrosis at baseline (F2 or F3), and baseline value as a covariate. Shifts in stage from baseline to each visit will also be presented. The shifts will include decrease in stage, no change in stage, and increase in stage. The shifts will be compared between treatment groups using a CMH-row mean scores test.

Graphical display for selected secondary endpoints will be provided.

9.3. Exploratory Efficacy Endpoints Analyses

Analysis of exploratory endpoints will be based on the ITT and mITT populations. Exploratory efficacy endpoints expressed as a proportion will be analyzed using the methods described for the primary efficacy endpoint. While the endpoints expressed as change from baseline and percent change from baseline will be analyzed using descriptive summary statistics and with ANCOVA or MMRM at a later date, if warranted.

9.4. Pharmacokinetics (PK)

All PK analyses will be described in a separate PK analysis plan. PK concentration data will be listed.

10. SAFETY

Safety will be evaluated based on data collected for adverse events, vital signs, physical examinations, electrocardiogram (ECG), ophthalmologic examinations, and laboratory data. No formal statistical testing will be conducted for the safety analyses. Descriptive statistics will be used to evaluate safety data in this study. Summaries will be presented by treatment group.

All safety analysis reporting will be based on the Safety population. All safety data will be listed by treatment group and subject.

10.1. Adverse Events

An AE is any untoward medical occurrence in a subject or clinical investigation subject administered a pharmaceutical product, and which does not necessarily have to have a causal relationship with this treatment. An AE can therefore be any unfavorable and unintended sign (including abnormal laboratory findings), symptom, or disease temporally associated with the use of an investigational product, whether related to the investigational product or not.

An unexpected AE is any event for which the nature or severity is not consistent with the information in the current Investigator's Brochure.

Any abnormal clinical or laboratory finding considered by the Investigator to be clinically significant is to be recorded in the eCRF as part of the subject's medical history if occurring prior to the start of study drug administration and as an AE if occurring after the start of study drug administration at baseline, where the finding represents a change from baseline.

A treatment emergent AE (TEAE) is defined as any AE occurring after the first dose of study treatment. All summaries of AEs will be based on TEAEs. All AEs recorded on the eCRF, whether treatment emergent or not, will be presented in the data listing. Non-TEAEs will be flagged in the listings.

Adverse Events will be coded using MedDRA version 24.0 or higher for reporting by SOC and PT. Summaries of TEAEs by SOC and PT will be presented in descending order of overall incidence. TEAEs will also be summarized by maximum intensity as classified according to the National Cancer Institute (NCI) Common Terminology Criteria for Adverse Events (CTCAE), Version 5.0 (or higher). If the AE is not included in the NCI CTCAE, then the Investigator is to determine the severity of the AE according to the criteria provided in the study protocol.

The causal relationship of each AE to study drug will be determined by the Investigator according to best medical judgment as definitely related, probably related, possibly related, unlikely related, or unrelated. If the relationship between the AE and study drug is determined to

be “possible”, “probable”, or “definite”, the event will be considered treatment-related for the safety analyses.

Missing severity assessments will be listed as missing information. Missing relationship assessments will also be listed as missing information. All AE summaries will be produced using the safety population and displayed by treatment group and overall.

An overview of TEAEs will be produced, including counts and percentages of subjects with any incidences of: TEAEs, TEAEs related to study treatment, serious adverse events (SAEs), TEAEs leading to study drug discontinuation, TEAEs with CTCAE grade 3 or higher, and fatal SAEs. The number of AEs will also be included in the summary.

In addition to the overview of TEAEs, the following summaries of AEs will be provided for the safety population by treatment group and overall:

- TEAEs by SOC and PT
- TEAEs related to study treatment by SOC and PT
- CTCAE Grade 3 or higher TEAEs by SOC and PT
- CTCAE Grade 3 or higher TEAEs related to study treatment by SOC and PT
- SAEs by SOC and PT
- TEAEs leading to treatment discontinuation by SOC and PT
- Skin and subcutaneous tissue disorders TEAEs by PT
- Eye disorder TEAEs by PT

Summaries of TEAEs by SOC, PT, and maximum severity as well as TEAEs by SOC, PT, and maximum relationship to study treatment will also be prepared.

Skin and subcutaneous tissue disorders and eye disorder TEAE summaries will also be performed for the mITT population.

A comprehensive listing of all AEs will be provided in a by-subject data listing. In addition, the following listings will be provided:

- TEAEs related to study treatment
- SAEs
- TEAEs leading to treatment discontinuation; and
- Fatal AEs.

10.2. Clinical Laboratory Evaluations

Clinical chemistry, hematology, coagulation, and urinalysis parameters will be reported based on conventional units as applicable. The following laboratory evaluations will be reported in data summaries:

- Chemistry: Chloride, Sodium, Blood urea nitrogen, Creatinine, Albumin, Total protein, AST, GGT, Total bilirubin, Indirect and direct bilirubin, Carbon dioxide, Potassium, Calcium, Magnesium, Glucose, Alkaline phosphatase, ALT, Creatine phosphokinase (total and fractionated) (screening and baseline only). Note that AST and ALT collected as part of the clinical chemistry panel are also applicable for efficacy assessments.
- Hematology: Hematocrit, Hemoglobin, Red blood cell count, MCV, Platelet count, White blood cell count with differential, Absolute neutrophil count, MCHC.
- Coagulation Studies: Prothrombin time, Fibrinogen, Activated partial thromboplastin time, international normalized ratio (INR).
- Urinalysis: Specific gravity, Protein, pH, Ketones, Blood, Microscopic examination of sediment, Glucose.

Observed values and changes from baseline for laboratory evaluations will be summarized at each visit and will also include the most extreme change from baseline. The percent change from baseline at each visit and the most extreme percent change from baseline will also be summarized. Select laboratories mean change from baseline by study visit will be depicted graphically in a longitudinal line plot with standard error of the mean.

Laboratory data will also be summarized in shift tables of baseline to most extreme change based on range categories of low (below lower limit of normal [LLN], normal, and high [above upper limit of normal [ULN]). Urinalysis results and shifts will be based on the categories present in the data.

10.3. Vital Signs

Vital signs include oral temperature, pulse (bpm), systolic and diastolic blood pressure (mmHg), and respiration rate (breaths/minute). Note that systolic and diastolic blood pressure will be summarized with the other anthropometric parameters. Observed values and changes from baseline for vital signs will be summarized at each visit and time point, as well as for most extreme change from baseline.

10.4. Electrocardiogram (ECG)

Electrocardiogram (ECG) parameters include HR (bpm), PR (ms), QRS (ms), QT (ms), QTcF (ms), and RR interval (ms). Observed values and changes from baseline for ECG parameters will be summarized by treatment group at each visit and time point.

Investigator reported ECG result will be tabulated by visit and worst-case post-baseline. In addition, the shifts from baseline to each visit and worst-case post-baseline will be summarized. Worst case post-baseline will be based on the worst observed value on or after the treatment start date. Categories will include within normal limits, abnormal without clinical significance, and abnormal with clinical significance.

10.5. Ophthalmologic Examinations

The proportion of subjects experiencing ophthalmologic abnormalities will be tabulated by visit and by treatment group. The clinical interpretation of the ophthalmologic examination will be tabulated as normal or abnormal. The worst-case post-baseline will also be summarized.

10.6. Physical Examinations

Physical examination results will be presented in subject data listings.

11. CHANGES TO THE PLANNED ANALYSES

The following changes were made to analyses specified in the protocol. Any additional changes to the planned analyses as detailed in this SAP will be described in the CSR.

- Co-primary objectives and endpoints were added.
- Added 'To evaluate the effect of TVB-2640 50 mg QD compared with matching placebo on resolution of steatohepatitis and no worsening of liver fibrosis (by NASH CRN fibrosis score). Resolution of steatohepatitis is defined as absence of fatty liver disease or isolated or simple steatosis without steatohepatitis and a NAS of 0 or 1 for inflammation, 0 for ballooning, and any value for steatosis at Week 52.' as a secondary objective and endpoint.
- Added 'To evaluate the effect of TVB-2640 50 mg QD compared with matching placebo on histological improvement (decrease) at Week 52 in NAS by ≥2 points.' as a secondary objective and endpoint.
- Removed 'To evaluate the effect of TVB-2640 50 mg QD compared with matching placebo on fibrosis as assessed by ≥1 stage of fibrosis improvement by NASH CRN score without worsening of NAS (defined as no increase in any component of NAS) at Week 52.' as a secondary objective and endpoint.
- Added proportion of MRI-PDFF ≥50% responders, MELD, and Agile3+ endpoints.
- Region (North America or not North America) was removed as an effect in all analyses.

12. REFERENCES

International Conference on Harmonization of Technical Requirements for Registration of Pharmaceuticals for Human Use. *ICH Harmonized Tripartite Guideline: Structure and content of clinical study reports* E3. Step 4. 1995.
https://database.ich.org/sites/default/files/E3_Guideline.pdf

Kamath PS, Wiesner RH, Malinchoc M, Kremers W, Therneau TM, Kosberg CL, D'Amico G, Dickson ER, Kim WR. (2001); A model to predict survival in patients with end-stage liver disease. Hepatology; 33:464-470.
https://doi.org/10.1053/jhep.2001.22172

Miettinen OS, Nurminen M. (1985); Comparative analysis of two rates. Statistics in Medicine 4, 213-226.

Ratitch B, Lipkovich I, O'Kelly M. Combining analysis results from multiply imputed categorical data. PharmaSUG Conference Proceedings 2013; Paper SP03.

Sanyal AJ, Foucquier J, Younossi ZM, Harrison SA, Newsome PN, Chan W-K, Yilmaz Y, De Ledinghen V, Costentin C, Zheng M-H, Wong VW, Elkhashab M, Huss RS, Myers RP, Roux M, Labourdette A, Destro M, Fournier-Poizat C, Miette V, Sandrin L, Boursier J. (2023); Enhanced diagnosis of advanced fibrosis and cirrhosis in individuals with NAFLD using FibroScan-based Agile scores. J Hepatol. 2023 February; 78(2): 247–259.
https://doi.org/10.1016/j.jhep.2022.10.034