STATISTICAL ANALYSIS PLAN

Evaluation of Efficacy, Safety and Tolerability of NGM282 (Aldafermin) in a Phase 2b, Randomized, Double-blind, Placebo-controlled, Multi-center Study in Subjects with Compensated Cirrhosis Due to Nonalcoholic Steatohepatitis (ALPINE 4)

Investigational Product:	Aldafermin (NGM282)
Protocol Number:	282-CC-207 (ALPINE 4)
Development Phase:	2b
Sponsor:	NGM Biopharmaceuticals, Inc. 333 Oyster Point Boulevard South San Francisco, CA 94080
SAP Version	30 Jan 2023

Confidentiality Statement

This document contains proprietary and confidential information of NGM Biopharmaceuticals, Inc. Acceptance of this document constitutes agreement by the recipient that no previously unpublished information contained herein will be published or disclosed without the prior written approval of NGM Biopharmaceuticals, Inc., with the exception that this document may be disclosed to study personnel under your supervision who need to know the contents for conducting the study and appropriate Institutional Review Boards/Human Research Ethics Committees under the condition that the personnel have agreed to keep this information confidential. The foregoing shall not apply to disclosure required by governmental regulations or laws; however, NGM Biopharmaceuticals, Inc. shall be promptly notified of any such disclosure.

SIGNATURE PAGE

STUDY TITLE: Evaluation of Efficacy, Safety and Tolerability of NGM282 (Aldafermin) in a Phase 2b, Randomized, Double-blind, Placebo-controlled, Multi-center Study in Subjects with Compensated Cirrhosis Due to Nonalcoholic Steatohepatitis (ALPINE 4)

Prepared by **Date**

[redacted] 30 January 2023

Sr. Director of Biometrics

NGM Biopharmaceuticals, Inc.

We, the undersigned, have reviewed and approved this statistical analysis plan.

Signature **Date**

{See appended electronic signature page}

[redacted]

Executive Director and head of Biometrics

NGM Biopharmaceuticals, Inc.

{See appended electronic signature page}

[redacted]

Chief Medical Officer, Clinical Development

NGM Biopharmaceuticals, Inc.

GLOSSARY OF ABBREVIATIONS

Abbreviation	Definition
ADA	Antidrug antibody
AE	Adverse event
ALP	Alkaline phosphatase
ALT	Alanine aminotransferase
AR1	Autoregressive(1)
AST	Aspartate transaminase
ATC	Anatomical-Therapeutic-Chemical
BMI	Body mass index
C4	7-alpha-hydroxy-4 cholesten-3-one
CI	Confidence interval
CMH	Cochran-Mantel-Haenszel
CRN	Clinical Research Network
CTCAE	Common Terminology Criteria for Adverse Events
DNA	Deoxyribonucleic acid
ECG	Electrocardiogram
ELF	Enhanced liver fibrosis
eCRF	Electronic case report form
EOS	End of study
EOT	End of treatment
HDL-C	High density lipoprotein cholesterol
ITT	Intent to treat
[redacted]	
LDL-C	Low density lipoprotein cholesterol
LISSA	Local injection-site symptom assessment
MAR	Missing at random
MCP-Mod	Multiple comparison procedure – modelling
MED	Minimum effective dose

Abbreviation	Definition
MedDRA	Medical Dictionary for Regulatory Activities
MI	Multiple imputation
MMRM	Mixed model repeated measures
NAB	Neutralizing antibody
NAFLD	Nonalcoholic fatty liver disease
NAS	NAFLD activity score
NASH	Nonalcoholic steatohepatitis
NCI	National Cancer Institute
NIH	National Institutes of Health
PD	Protocol deviations
PK	Pharmacokinetic
PoC	Proof of concept
PP	Per protocol
Pro-C3	N-terminal Type III collagen
PT	Preferred term
RUQ	Right upper quadrant
SAE	Serious adverse event
SAP	Statistical analysis plan
SAS	Statistical Analysis System
SC	Subcutaneous
SD	Standard deviation
SOC	System organ class
TEAE	Treatment emergent adverse event
WHO	World Health Organization

TABLE OF CONTENTS

SIGNATURE PAGE 2
1. INTRODUCTION 7
1.1. Primary Study Objectives and Design 7
1.2. Secondary and Other Objectives 8
1.3. Sample Size 8
2. ANALYSIS SETS AND DATA CONVENTIONS 9
2.1. Analysis Sets 9
2.2. Data Conventions and General Considerations 10
2.3. Visit Windows 10
3. DISPOSITION, EXIT STATUS AND PROTOCOL DEVIATIONS 13
3.1. Disposition and Exit Status 13
3.2. Protocol Deviations (PD) 13
4. DEMOGRAPHICS AND OTHER BASELINE CHARACTERISTICS 13
4.1. Demographics 13
4.2. Disease Characteristics 13
4.3. Prior and Concomitant Medications 13
4.4. Medical History and Concurrent Procedures 14
5. EFFICACY ANALYSES 14
5.1. Primary Efficacy Analyses 14
5.1.1. Primary Efficacy Endpoint and Analysis 14
5.1.2. Primary Estimand 14
5.1.3. Statistical Hypotheses 15
5.1.4. Primary Analysis 15
5.1.5. Sensitivity Analysis 16
5.2. Secondary Efficacy Analyses 16
5.2.1. Analysis for Fibrosis Responses 17
5.3. Other Efficacy Analyses 18
5.4. Subgroup Analyses for Efficacy Endpoints 19
6. PK ANALYSIS 19
7. SAFETY ANALYSES 19
7.1. Study Treatment Exposure and Compliance 19

7.1.1. Exposure to Study Treatment 19
7.2. Adverse Events 20
7.3. Clinical Laboratory Evaluations 21
7.4. Vital Signs, Body Weight, BMI and Waist Circumference....21
7.5. 12-Lead Electrocardiograms....21
7.6. Injection Site Reactions 21
7.7. Physical Examinations and RUQ Pain Assessment....**Error! Bookmark not defined.**
7.7.1. Study Treatment Compliance 21
7.8. Pregnancy Test....**Error! Bookmark not defined.**
8. BIOMARKER, AND IMMUNOGENICITY DATA....22
9. HEALTH OUTCOMES DATA ANALYSES 22
10. INTERIM ANALYSES AND DATA MONITORING 23
11. DEVIATIONS FROM THE PROTOCOL....23
12. AMENDMENTS 23
13. LITERATURE REFERENCES....24
14. APPENDICES 25
APPENDIX A. NASH CLINICAL RESEARCH NETWORK NAFLD ACTIVITY SCORE AND FIBROSIS SCORE 25
APPENDIX B. SAMPLE SAS CODE 26
ELECTRONIC SIGNATURE PAGE....28

LIST OF TABLES

Table 1: Power for the Primary Efficacy Analysis (two-sided alpha = 5%) 9
Table 2: ELF, PRO-C3, CLDQ Visit Windows 11
Table 3: Fibroscan Visit Windows 11
Table 4: Liver Biopsy and Bile Acids Visit Windows 11
Table 5: ALT and AST Visit Windows 12
Table 6: C4 Visit Windows....12

1. INTRODUCTION

This Statistical Analysis Plan (SAP) provides the details of the statistical analyses to be performed for NGM Biopharmaceuticals clinical study 282-CC-207 (ALPINE 4) entitled "Evaluation of Efficacy, Safety and Tolerability of NGM282 (Aldafermin) in a Phase 2b, Randomized, Double-blind, Placebo-controlled, Multi-center Study in Subjects with Compensated Cirrhosis Due to Nonalcoholic Steatohepatitis (ALPINE 4)" dated 11 November 2021, Version 6.0.

There will be one database lock after all study subjects have exited study and all data clarifications have been resolved.

Exploratory analyses, not identified in this SAP, may be performed to support the clinical development program. Any post-hoc, or unplanned, analyses that are performed but not identified in this SAP will be identified in the clinical study report.

Should the SAP and the protocol be inconsistent with respect to the planned analyses, the language of the SAP prevails.

1.1. Primary Study Objectives and Design

The primary objective of this study is to evaluate the efficacy and safety of aldafermin compared to placebo.

- The primary efficacy endpoint is the change in Enhanced Liver Fibrosis (ELF) score from baseline after 48 weeks of treatment with aldafermin *vs*. matched placebo.
- The primary safety endpoint is frequency and severity of adverse events (AEs) and serious adverse events (SAEs).

This is a multiple center evaluation of aldafermin administered for 48 weeks as a daily subcutaneous (SC) injection in a randomized, double-blind, placebo-controlled study in subjects with histologically confirmed cirrhosis with NASH. Eligible subjects must have liver biopsies consistent with cirrhosis according to the NASH CRN classification (NASH CRN fibrosis score of 4) with a small proportion (~10%) of patients with a clinical diagnosis of NASH cirrhosis despite a NASH CRN fibrosis score of 3, as assessed by the central reader.

Approximately 150 eligible subjects will be randomized into 1 of 4 treatment groups in a 3:2:2:3 ratio to receive either subcutaneous placebo, 0.3 mg aldafermin, 1 mg aldafermin or 3 mg aldafermin, respectively. Under protocol Version 5.0, eligible subjects will no longer be randomized into the 0.3 mg aldafermin treatment group. Instead, subjects will be randomized into 1 of 3 treatment groups in a 4:3:4 ratio to receive either subcutaneous placebo, 1 mg aldafermin or 3 mg aldafermin, respectively. Subjects randomized to the 0.3 mg aldafermin treatment group prior to protocol Version 5.0 will continue to receive the same until treatment completion.

On Day 1, eligible subjects will be randomized into one of the treatment groups as described above. The randomization will be stratified by clinical diagnosis of T2D versus no T2D to ensure planned allocation ratio of treatment groups within each stratum. Treatment assignment will be blinded to the investigational sites, study subjects, sponsor study team, and sponsor Medical

for the 3 mg and placebo groups will provide 92% power to reject the null hypothesis that there is no treatment effect between the aldafermin 3 mg group and placebo group, based on a two sample t-test at the 5% significance level (two-sided).

Based on the modified primary endpoint of the change from baseline in ELF at Week 48 and the analysis results from ALPINE 2/3, the following table presents the powers associated with different change from baseline values.

Table 1: Power for the Primary Efficacy Analysis (two-sided alpha = 5%)

Scenario	Placebo	3 mg	Power
1	0.1	-0.3	77%
2	0.1	-0.4	92%
3	0.1	-0.5	98%

2. ANALYSIS SETS AND DATA CONVENTIONS

2.1. Analysis Sets

Subjects in this study will be analyzed using the following analysis sets:

- **Intent-to-Treat Analysis Set**: All randomized subjects will be included in the Intent-to-Treat (ITT) analysis set. The ITT analysis set will be used for the primary and secondary efficacy analyses and will be based on the randomized treatment if it differs from the actual treatment received.
- **Modified ITT (mITT) Analysis Set**: All ITT subjects have baseline and at least one valid, post-dose, on-treatment efficacy value of ELF score. The mITT analysis set will be used for sensitivity analyses to support the ITT analyses and will be used for the primary and secondary efficacy analyses and will be based on the randomized treatment if it differs from the actual treatment received.

[redacted]

2.2. Data Conventions and General Considerations

The following data conventions and general considerations will be used for all analyses described in this SAP.

- All Tables, Listings and Figures will be produced using SAS version 9.4 or higher except the analysis for the primary efficacy endpoint.
- Key subject-level data will be provided in data listings. All data listings will be sorted by treatment, subject ID, and visit.
- Study day will be calculated as:
  - visit date – first dose date, for screening visits.
  - visit date – first dose date + 1, for post-baseline/treatment visits.
- Baseline value will be defined as the last available, valid, non-missing assessment reported before the start of first study drug administration. The randomization date (and time, if available) will be used if the study drug start date is missing or not available for any reason.
- Change from baseline (change) is calculated as post-baseline value – baseline value. Percent (or relative) change from baseline (percent/relative change) is calculated as 100 × (change from baseline/baseline value).
- Descriptive statistics including the number of non-missing observations (n), arithmetic mean (mean), standard deviation (SD), median, minimum (min), and maximum (max) will be presented for continuous endpoints. Frequency, percentage, odds ratio or proportion difference will be presented for categorical endpoints.
- Statistical hypotheses for the primary efficacy endpoint will be tested for statistical significance with two-sided alternatives at alpha level of 0.05, Significance level of 0.1 will be used for all secondary and other efficacy analyses. All confidence intervals (CIs) will be two-sided, unless stated otherwise.
- For the primary endpoint, secondary endpoints and endpoint of CLDQ, unscheduled and EOT visits will be mapped into protocol scheduled visits using the analysis window if possible. For details regarding the analysis window definition, please refer to section 2.3.
- Missing data for selected histological endpoints (e.g., secondary) will be imputed using the methods as described in Section 5. Missing or incomplete dates for adverse events (AEs) and medications may be imputed as needed (e.g., to determine treatment emergence or distinguish between prior and concomitant medications) based on programming conventions. No imputations will be performed for missing or incomplete data of other types and, in this case, analyses will be performed based on the observed data only.

2.3. Visit Windows

Unscheduled and EOT visits for below endpoints will be assigned to analysis visits according to the following visit windows:

Table 2: ELF, PRO-C3, CLDQ Visit Windows

Analysis Visit	Target Analysis Day	Lower End Analysis Window	Higher End Analysis Window
Day 1	1	NA	NA
Week 24	169	85	252
Week 48	337	253	358
Week 54	379	NA	NA

Table 3: Fibroscan Visit Windows

Analysis Visit	Target Analysis Day	Lower End Analysis Window	Higher End Analysis Window
Day 1	**1**	**NA**	**NA**
Week 24	**169**	**85**	**252**
Week 48	**337**	**253**	**EOS**

Table 4: Liver Biopsy and Bile Acids Visit Windows

Analysis Visit	Target Analysis Day	Lower End Analysis Window	Higher End Analysis Window
Day 1	1	NA	NA
Week 48	337	169	EOS

Table 5: ALT and AST Visit Windows

Analysis Visit	Target Analysis Day	Lower End Analysis Window	Higher End Analysis Window
Day 1	1	NA	NA
Week 2	15	8	29
Week 6	43	30	64
Week 12	85	65	106
Week 18	127	107	148
Week 24	169	149	190
Week 30	211	191	232
Week 36	253	233	274
Week 42	295	275	316
Week 48	337	317	341
Week 49	344	NA	NA
Week 54	379	NA	NA

Table 6: C4 Visit Windows

Analysis Visit	Target Analysis Day	Lower End Analysis Window	Higher End Analysis Window
Day 1	1	NA	NA
Week 2	15	8	29
Week 6	43	30	64
Week 12	85	65	106
Week 18	127	107	148
Week 24	169	149	190
Week 30	211	191	232
Week 36	253	233	274
Week 42	295	275	316
Week 48	337	317	EOS

Within an analysis visit window, the measurement from the scheduled visits will be used if available. If there is no scheduled visit, the on-treatment measurement from the visit closest to the defined target analysis day within the window will be used. If there is more than one

measurement with equal distance to the defined target analysis day, the latter will be used. If no visits occur within a visit window, the measurement for this visit will be treated as missing.

3. DISPOSITION, EXIT STATUS AND PROTOCOL DEVIATIONS

3.1. Disposition and Exit Status

Subject disposition will be summarized by treatment group and overall using the ITT analysis set and mITT analysis set. The number and percent of subjects screened (ITT only), enrolled (ITT only), randomized, completed, and prematurely discontinued, along with the primary reason for the discontinuation (per study eCRF specifications) will be presented.

3.2. Protocol Deviations (PD)

Subject data will be examined prior to database lock or unblinding for evidence of protocol deviations. Classification of possible protocol deviations including both major (CSR-reportable) and minor (non-CSR-reportable) PDs will be reviewed and approved by the NGM study team based on predefined criteria. Numbers and percentages of subjects with major PDs will be summarized using the ITT analysis set. Protocol deviations will be provided in data listing as well.

4. DEMOGRAPHICS AND OTHER BASELINE CHARACTERISTICS

4.1. Demographics

Demographics will be summarized with descriptive statistics by treatment group and total for the ITT and Safety analysis set. These variables include age (years), sex (Male or Female), race (Asian, Black, or African American, White, etc.), ethnicity (Hispanic/Latino, Not Hispanic/Latino, Not Reported, or Unknown), body weight (kg), height (cm), body mass index (BMI, kg/m^2) and waist circumference (cm) at baseline.

4.2. Disease Characteristics

Baseline disease characteristics will be summarized with descriptive statistics by treatment group and total for the ITT and Safety analysis set. These variables include but not limited to ELF, fibrosis score, NAS score, type 2 diabetes status (Yes/No), Pro-C3, C4, LSM and statin status (naïve/experienced).

4.3. Prior and Concomitant Medications

Medications taken at least once within 90 days prior to Screening visit and during the study period as well as the reason for use will be recorded in the source documents and the eCRFs.

- **Prior medications**: are defined as any medication which is administered any time prior to the study treatment starting time, regardless of when the medication stops.

- **Concomitant medications**: are defined as any medication which is administered any time after the study treatment starting time, regardless of when the medication starts or stops.

A given medication may be classified both as a prior medication and as a concomitant medication.

All medications will be coded using the World Health Organization Drug Dictionary Global (WHO Drug Dictionary Enhanced B3, September 2019). The number and percentage of subjects taking prior/concomitant medications will be summarized by the drug class from the Anatomical-Therapeutic-Chemical (ATC) classification level 4 and WHODrug Global drug name (preferred name) for each treatment group and overall using the safety analysis set. For these summaries, subjects who take the same medication (i.e., the same preferred name) more than once will be counted only once for that medication.

4.4. Medical History and Concurrent Procedures

Medical history will be coded using MedDRA version 23.0. The number and percentage of subjects with medical history will be summarized by primary SOC and preferred term by treatment group and overall using the safety analysis set.

5. EFFICACY ANALYSES

All efficacy endpoints will be summarized with descriptive statistics by treatment group (0.3 mg aldafermin, 1 mg aldafermin, 3 mg aldafermin, or placebo) and visit using the ITT and mITT analysis sets. Summaries for continuous endpoints will be provided for the actual value, change and percent change from baseline (if applicable), respectively. Figures displaying LS means over time will be provided for selected endpoints.

5.1. Primary Efficacy Analyses

5.1.1. Primary Efficacy Endpoint and Analysis

The primary efficacy endpoint is the change from baseline in ELF at Week 48.

5.1.2. Primary Estimand

The estimand for the primary interest of this study is defined as follows:

1. Treatment: aldafermin 0.3 mg, aldafermin 1 mg, aldafermin 3 mg, and placebo.
2. Population: subjects with histologically confirmed cirrhosis (NASH CRN fibrosis Stage 4) due to NASH, as defined by the inclusion-exclusion criteria of this study.
3. Endpoint: the change from baseline in ELF at Week 48 as defined in Section 5.1.1.

4. Inter-current events and their corresponding strategies:

Intercurrent Events	Estimand Strategies
study discontinuation	Hypothetical "as own treatment group" strategy: clinical course is the same as other subjects in the same treatment group, who have not discontinued study treatment.
treatment non-compliance or discontinuation	Treatment Policy strategy: all observed values of the variable are used, regardless of whether or not the subject has experienced the intercurrent event.
treatment interruption or dose reduction	Treatment Policy strategy
initiation or adjustment of concomitant medication	Treatment Policy strategy
any other major protocol deviations	Treatment Policy strategy

5. Population level summary: the treatment group difference of change from baseline in ELF at Week 48 between each aldafermin dose group and the placebo group.

5.1.3. Statistical Hypotheses

The statistical hypotheses (null and alternative) to be tested with respect to the primary efficacy endpoint:

- **H_0 (Null Hypothesis)**: there is no difference between the aldafermin and placebo groups in the change from baseline to week 48 in ELF.
- **H_1 (Alternative Hypothesis)**: there is a difference between the aldafermin and placebo groups in the change from baseline to week 48 in ELF.

Since the aldafermin 0.3 mg group was terminated early, there will be no statistical comparison being done between the aldafermin 0.3 mg group and the placebo group. Aldafermin 0.3mg data will be presented descriptively by visits.

In order to account for the additional variation and avoid potential bias introduced by randomization change during the study, a covariate 'randomization ratio' representing randomization schema under which the subject was randomized will be included in the efficacy analyses when applicable.

5.1.4. Primary Analysis

The mixed model for repeated measures (MMRM) will be used to analyze the primary endpoint. The model will use the change from baseline at Week 24 and Week 48 as the outcome variable, include fixed effects for treatment, visit (categorical), treatment-by-visit interaction, and the baseline ELF value, baseline T2D status (Yes/No), and randomization ratio as covariates.

The within-subject covariance structure for this model will be assumed to be unstructured. SAS PROC MIXED will be used with denominator degrees of freedom estimated by the KENWARDROGER2 approach. Within the framework of this model, a point estimate of the treatment group difference between each aldafermin dose and the placebo group at each visit will be provided. The corresponding two-sided 95% confidence interval and p-value for the point estimate at Week 48 will be presented.

A hierarchical testing procedure will be used to control the family-wise type I error rate associated with the tests of the two aldafermin dose groups (aldafermin 3 mg vs placebo and aldafermin 1 mg vs placebo). Specifically, the comparison between aldafermin 3 mg and placebo will be performed first. If this test is statistically significant at the two-sided 5% level, then the comparison between aldafermin 1 mg and placebo will be performed. If the first test is not statistically significant at the two-sided 5% level, then the comparison between aldafermin 1mg and placebo will be considered exploratory.

5.1.5. Sensitivity Analysis

The analysis and estimation described in Section 5.1.4 will be performed using the mITT analysis set.

In addition, the following multiple imputation method will be performed to assess the impact of missing data on the primary analysis using ITT analysis set. In this method, missing data will be imputed 100 times via SAS PROC MI. Once data is imputed, the same model specified as the primary analysis in Section 5.1.4. will be run for each imputation, and the pooled estimate of the change from baseline in ELF at Week 48 along with its standard error and 95% CI for each dose level will be obtained using the method developed by Rubin (1987) using PROC MIANALYZE.

- Data will be imputed using a placebo-based pattern mixture model to evaluate the MAR assumption. For this imputation, intermittent missing data (non-monotone) pattern will first be multiply imputed under MAR assumptions using MCMC to generate a monotone pattern. The imputation to generate monotone data will be performed overall (e.g., not by treatment). The remaining missing data (e.g., after study discontinuation) will be imputed under MNAR via a sequential placebo-based pattern imputation using regression methods in SAS PROC MI using the MNAR option.

5.2. Secondary Efficacy Analyses

The secondary objectives are to evaluate the efficacy of aldafermin and the effect of aldafermin on pharmacokinetics and on biomarkers of target engagement, fibrogenesis, and imaging.

The following endpoints will be measured:

1. 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), after 48 weeks of treatment with aldafermin or matched placebo, in subjects who had a minimum 1 point in each category (hepatocellular ballooning, steatosis, and lobular inflammation) of NAFLD Activity Score (NAS) at baseline as determined by central pathologist evaluation.

2. Improvement in liver fibrosis greater than or equal to one stage (NASH CRN fibrosis score) after 48 weeks of treatment with aldafermin or placebo.
3. Changes from baseline in fibrosis as measured by Collagen Proportional Area (CPA) and/or Second Harmonic Generation (SHG) methodologies at Week 48 (if data is available at final database lock)
4. Aldafermin concentrations pre-dose at all treatment visits and 4-hours post-dose at Day 1, Week 24, and Week 48 in all subjects. In a subset of subjects participating in an optional PK study, full PK profiles up to 24 hours will be assessed at Day 1 and steady state at end of treatment (EOT), i.e., 48 weeks. PK analysis will be described in Section 6.

5.2.1. Analysis for Fibrosis Responses

Between-group comparisons (aldafermin doses vs placebo) of the fibrosis response (first two endpoints listed above) at Week 48 (by NASH CRN criteria) will be performed using the Cochran-Mantel-Haenszel (CMH) test stratified by baseline T2D status (Yes/No) and randomization ratio. The treatment group difference of the response rates and odds ratio along with the 90% CI and 95% CI will be provided. The CIs will be constructed using the normal approximation method.

Missing fibrosis responses will be imputed using a multiple imputation (MI) method under the assumption of missing at random (MAR). For fibrosis response. The MI inference involves the following three distinct steps: imputation, analysis and pooling. The details of these steps are described as follows:

1. Imputation: Missing fibrosis responses are imputed 100 times to generate 100 complete data sets. The imputed values will be obtained using a logistic model with fixed effects for treatment group, randomization ratio and baseline T2D status (Yes/No). At the end of this step, all the 100 complete data sets contain the same non-missing values while the missing values are replaced by imputed values.
2. Analysis: Each of the 100 competed data sets will be analyzed using the Cochran-Mantel-Haenszel (CMH) test described above.
3. Transformation: the Wilson-Hilferty transformation (Wilson and Hilferty 1931; Goria 1992) will be applied to each CMH statistic from the analysis step above in order to normalize the CMH statistic.
4. Pooling: The results from the 100 transformed CMH statistics are combined to generate the testing statistic and p-value. The pooled estimate of the response rate along with its standard error and 90%/95% CI for each dose level will be obtained using the method developed by Rubin (1987).

For fibrosis response, the following sensitivity analysis will be performed to evaluate the impact of missing data.

- Subjects with missing baseline or postbaseline data will be dropped out for analysis, only completers with both baseline and postbaseline results will be included (Completer Analysis)

- Missing data will be imputed using the multiple imputation (MAR) as described above using mITT analysis set
- Missing fibrosis responses are imputed using a placebo-based MI method (i.e., only data from placebo group will be used in the MI procedure)

In addition, analyses will be performed for the following secondary efficacy variables:

- Changes from baseline in C4 and serum bile acids
- Changes from baseline in Pro-C3
- Changes from baseline in ALT, AST
- Changes from baseline in liver stiffness measure (LSM by FibroScan®) and steatosis (CAP by FibroScan®)

For the continuous secondary efficacy endpoints with more than one post baseline measurements, change and percent change from baseline will be analyzed using MMRM model, include fixed effects for treatment, visit (or visit-timepoint) (categorical), treatment-by-visit interaction (or treatment-by-visit-timepoint) interaction, baseline T2D status (Yes/No), randomization ratio and the baseline outcome value as a covariate. For subjects participating in the optional PK sub-study, the C4 values collected at additional timepoints will only be summarized descriptively, not included in the MMRM model. The within-subject covariance structure for this model will be assumed to be unstructured. SAS PROC MIXED will be used with denominator degrees of freedom estimated by the KENWARDROGER2 approach. Within the framework of the MMRM model, a point estimate of the treatment group difference between each aldafermin dose and placebo will be provide along with the corresponding two-sided p-value and 90%/95% CI. Before performing analysis on the continuous variables, normality of the data distribution will be tested. If the data are not normally distributed, Wilcoxon rank-sum test will be used.

For the secondary endpoints with one post baseline measurement, change from baseline will be analyzed by analysis of covariance (ANCOVA) model with effects for treatment, baseline T2D status (Yes/No), randomization ratio and the baseline outcome value.

All secondary efficacy analyses will be performed using the ITT population. mITT population will be used as sensitivity analyses.

All statistical tests for the secondary efficacy analyses will be performed using a two-sided test at the significance level of 10% with 90% CI. In addition, two-sided 95% CI will also be presented to aid the result interpretation. No multiplicity adjustment will be made between the primary and secondary efficacy analyses, or among secondary efficacy analyses to control the experiment-wise type I error rate.

5.3. Other Efficacy Analyses

The analyses of the secondary efficacy endpoints at Week 48 in Section 5.2 will also be performed for all other applicable visits.

[redacted]

[redacted]

5.4. Subgroup Analyses for Efficacy Endpoints

Subgroup analyses will be performed for the primary endpoint and Fibrosis response endpoint. The group variables include:

- ELF levels (<9.8, 9.8–11.3, >11.3)
- Clinical diagnosis of T2D (Yes/No),
- Fibrosis stage (F3, F4)

The subgroup analysis for each endpoint will be using the same method described for the respective endpoint. If the sample size of a subgroup is too small and lead to convergence issues, only summary statistics will be provided.

[redacted]

All safety analyses will be conducted using the safety analysis set. All analyses are descriptive in nature without statistical hypotheses testing. In general, the safety analysis results will be presented by individual treatment groups and an overall aldafermin group.

7.1. Study Treatment Exposure and Compliance

7.1.1. Exposure to Study Treatment

Treatment exposure duration will be calculated as the date of the last dose of study drug minus the date of first dose of study drug plus one day. Treatment exposure duration will be summarized with descriptive statistics by treatment group. In addition, the number and percentage of subjects with treatment exposure duration of the following categories will be summarized by treatment group for the following time periods:

- ≤ 2 weeks

- 2 - ≤ 4 weeks
- 4 - ≤12 weeks
- 12 - ≤24 weeks
- 24 - ≤ 48 weeks
- > 48 weeks

7.2. Adverse Events

Adverse events are collected both for the screening/baseline period pre-treatment (which are referred to as pre-treatment AEs) and for the follow-up period after treatment is initiated (which are referred to as post-treatment AEs). All adverse events will be coded using the Medical Dictionary for Regulatory Activities (MedDRA, v23.0).

Treatment emergent adverse events (TEAEs) are defined as AEs that commence on or after the date (and time if available) of first study drug administration. AEs without an onset date or time will be defined as treatment emergent, except if an incomplete date (e.g., month and year) clearly indicates that the event started prior to the start of first administration of study drug or if the AE stop date indicates that the event started or stopped prior to the start of first administration of study drug.

The severity of an AE is rated by National Cancer Institute (NCI) Common Terminology Criteria for Adverse Events (CTCAE) v5. The relationship of an AE to study drug is made using a four-category system: not related, possible, probable, or definite.

The following summary of TEAEs will be provided.

- An overall summary of any TEAEs, any TEAEs leading to study treatment withdrawal, any serious TEAEs, any Aldafermin/Placebo-related TEAEs, any Rosuvastatin/Placebo-related TEAEs, and death
- TEAEs by PT in descending order of incidence
- TEAEs by primary SOC and PT
- TEAEs by primary SOC, PT and severity (CTCAE grade)
- TEAEs by primary SOC, PT and relationship to Aldafermin/Placebo
- TEAEs by primary SOC, PT and relationship to Rosuvastatin/Placebo
- TEAEs leading to study treatment withdrawal by primary SOC and PT
- Aldafermin/Placebo-related TEAEs leading to study treatment withdrawal by primary SOC and PT
- Rosuvastatin/Placebo-related TEAEs leading to study treatment withdrawal by primary SOC and PT
- Serious TEAEs by primary SOC and PT

A subject will be counted only once within each preferred term and system organ class in the summary tables. The most severe occurrence of a repeat AE, as well as the most extreme relationship of the AE to the study treatment will be used for these analyses. In addition, for the rosuvastatin/Placebo-related summary tables, only subjects who received rosuvastatin will be included. All SAEs and AEs leading to study drug withdrawal and death will be provided in listings.

7.3. Clinical Laboratory Evaluations

Clinical laboratory evaluations include biochemistry, hematology and urinalysis. Actual values and change from baseline values of continuous laboratory tests will be summarized by treatment group for each visit. For categorical lab tests, frequency and percentage distribution will be presented by treatment group for each visit.

7.4. Vital Signs, Body Weight, BMI and Waist Circumference

Vital sign assessments include seated systolic/diastolic blood pressure, respiration rate, pulse and temperature. Actual values and change from baseline values for vital signs, body weight, BMI and waist circumference will be summarized by treatment group for each visit.

7.5. 12-Lead Electrocardiograms

The 12-lead ECG assessments include QT interval, QTcF interval, P-R interval, R-R interval and QRS duration. Actual values and change from baseline values for 12-lead ECG assessments will be summarized by treatment group for each visit. Shift table of ECG interpretation results (Abnormal, Indeterminate, Normal, Not evaluable, Unknown per CRF) will be presented by treatment group for each visit.

7.6. Injection Site Reactions

Injection site reactions will be evaluated using a local injection-site symptom assessment (LISSA). Frequency and percentage distribution of the severity of each pre-specified symptom (pain, tenderness, erythema/redness and induration/swelling) will be presented by treatment group for each visit.

7.6.1. Study Treatment Compliance

Study treatment compliance is defined:

$$\text{Compliance rate for each patient} = \frac{\text{number of injections administered}}{\text{number of injections prescribed}} * 100$$

Descriptive statistics for study treatment compliance will be presented by treatment group. In addition, the number and percentage of subjects with treatment compliance of the following categories will be summarized by treatment group: <80%, 80% to 100%, and >100%. If the last prescribed syringes were not returned, then it was assumed that injections were taken daily until the date of last dose and the remaining syringes were then returned. Otherwise, the not-returned syringes were assumed to be all taken.

8. BIOMARKER, AND IMMUNOGENICITY DATA

The analysis of biomarker (other than Pro-C3, C4 and bile acids), and anti-drug antibody (ADA)/ neutralizing antibody (NAb) testing data will be provided in separate reports.

9. HEALTH OUTCOMES DATA ANALYSES

A chronic liver disease questionnaire (CLDQ-NASH) will be administered at Screening, Weeks 24, 48/early drop out, and Week 54. There are 36 questions in the CLDQ-NASH, and each question has 2 stages. The first stage asks whether the condition has been present at all in the past 2 weeks. If the answer to this first stage is "Yes," then the second stage of the question asks the rater to choose in how much of the past 2 weeks the condition has been present. The answers to each second stage question have a range of 1 to 7, with the scores shown in the table below. For example, question 1 is "How much of the time during the past 2 weeks have you been troubled by a feeling of abdominal bloating?", and the possible answers include the following:

Score	Description
1	All of the time
2	Most of the time
3	A good bit of the time
4	Some of the time
5	A little of the time
6	Hardly any of the time
7	None of the time

CLDQ-NASH includes 6 domains:

1. AS – Abdominal symptoms; includes 3 questions (1, 5, 17).
2. FA – Fatigue; includes 6 questions (2, 4, 8, 11, 13, 35).
3. SS – Systemic symptoms; includes 6 questions (3, 6, 21, 23, 27, 36).
4. AC – Activity; includes 5 questions (7, 9, 14, 30, 31).
5. EF – Emotion function; includes 9 questions (10, 12, 15, 16, 19, 20, 24, 26, 34).
6. WO – Worry; includes 7 questions (18, 22, 25, 28, 29, 32, 33).

The domain score will be calculated as the average of the scores of the non-missing domain's items. The total score will be calculated as the average of the 6 domain scores. A summary table including descriptive statistics (n, mean, SD, median, minimum and maximum) for 6 domain scores and the total score will be presented by time and treatment group. Change from baseline (screening) to each post-baseline time point will be summarized for 6 domain scores and the total scores.

Higher scores for each item and/or total score indicate better quality of life for each activity. Therefore, a positive change from baseline in the total score indicates improvement in a patient's quality of life.

Between-group comparisons (aldafermin doses vs placebo) of the change from baseline in total score at post-baseline time points will be performed using mixed model repeated measures (MMRM) model with fixed effects for treatment, baseline T2D status (Yes/No) and the baseline outcome value as a covariate. The within-subject covariance structure is assumed to be

unstructured. The treatment group LSMEANS difference of the change from baseline along with the corresponding 95% CI will be provided from the MMRM model. The analysis will be done using ITT analysis set.

10. INTERIM ANALYSES AND DATA MONITORING

No formal interim analysis is planned for the study.

The study will be performed under review by an external independent Data Safety and Monitoring Board (DSMB) to closely monitor the safety and welfare of the study subjects. The DSMB will be comprised of experienced hepatologists and other relevant specialists. The DSMB will review study data at regular intervals and determine whether it is safe to continue the study according to the protocol. The DSMB members will be independent of the sponsor, the clinical study sites, and investigators. The event definitions information relevant to the operation of the committee will be described in a DSMB Charter.

11. DEVIATIONS FROM THE PROTOCOL

The following changes have been made from the protocol:

- The primary endpoint analysis method has been changed from MCP-MOD to MMRM – the 0.3 mg dose has been dropped from the randomization, MCP-MOD requires a minimum of 4 arms.
- The term “analysis population” has been replaced by “analysis set” - to be consistent with the language used in the ICH E9 guideline.
- mITT analysis set definition has been changed.
- Removed sensitivity analysis based on MAR and LOCF for the primary and key secondary endpoint.
- No formal interim analysis was conducted.

12. AMENDMENTS

Not applicable

13. LITERATURE REFERENCES

Bornkamp B (2006). Comparison of Model-Based and Model-Free Approaches for the Analysis of Dose-Response Studies. Master's thesis, Fachbereich Statistik, Universit¨at Dortmund.

Bornkamp B, Pinheiro J, Bretz F. DoseFinding: planning and analyzing dose finding experiments. R package, http://www.cran.rproject. org/web/packages/DoseFinding/

Bretz F, Pinheiro JC, Branson M. Combining multiple comparisons and modeling techniques in dose-response studies. *Biometrics*. 2005;61(3):738-748.

Buckland ST, Burnham KP, Augustin NH. "Model Selection an Integral Part of Inference." *Biometrics*. 1997;53, 603–618.

Committee for Medicinal Products for Human Use (CHMP). Qualification Opinion of MCP-Mod as an efficient statistical methodology for model-based design and analysis of Phase II dose finding studies under model uncertainty. Jan 2014. Available from: https://www.ema.europa.eu/en/documents/regulatory-procedural-guideline/qualification-opinion-mcp-mod-efficient-statistical-methodology-model-based-design-analysis-phase-ii_en.pdf

Golub G, Pereyra V. "Separable nonlinear least squares: The variable projection method and its applications." *Inverse Problems*. 2003;19, R1–R26.

Goria MN. ON THE FOURTH ROOT TRANSFORMATION OF CHI-SQUARE. *Australian Journal of Statistics*. 1992;34(1):55-64.

Pinheiro J, Bornkamp B, Glimm E et al. Model-based dose finding under model uncertainty using general parametric models. *Statistics in Medicine* 2014; 33: 1646-1661.

Ratitch B, Lipkovich I, & O'Kelly M. Combining analysis results from multiply imputed categorical data. PharmaSUG 2013-Paper SP03, 2013, 1-19. Available from: https://www.lexjansen.com/pharmasug/2013/SP/PharmaSUG-2013-SP03.pdf

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

Westfall, PH and Young, SS. *Resampling Based Multiple Testing* 1993. John Wiley & Sons Inc., New York.

Wilson EB, Hilferty MM. The Distribution of Chi-Square. Proc Natl Acad Sci U S A. 1931 Dec;17(12):684-8.

14. APPENDICES

APPENDIX A. NASH CLINICAL RESEARCH NETWORK NAFLD ACTIVITY SCORE AND FIBROSIS SCORE

Steatosis	S Score	Lobular Inflammation	L Score	Hepatocyte Ballooning	B Score
< 5%	0	None	0	None	0
5%–33%	1	< 2	1	Few ballooned cells	1
34%–66%	2	2–4	2	Many ballooned cells	2
> 66%	3	> 4	3	—	—

NAFLD = nonalcoholic fatty liver disease.
Note: NAFLD activity grade score = total score: S + L + B (range 0–8).

Fibrosis Stage	Score
0	No Fibrosis
1a	Zone 3, Mild
1b	Zone 3, Moderate
1c	Periportal Only
2	Zone 3 and Periportal
3	Bridging
4	Cirrhosis

APPENDIX B. SAMPLE SAS CODE

Primary analysis sample code:

```
proc mixed data=ngm282;
class trt usubjid visit;
model Chg ELF = trt visit trt*visit baseline_ELF baseline_T2D_status (Yes/No)
randomization ratio/ddfm=kr2;
repeated visit/type=un subject=usubjid;
LSMEANS trt*visit/diff cl;
run;
```

Missing data imputation for the sensitivity analysis No 1. for the primary analysis sample code:

```
PROC MI DATA= NIMPUTE=100 OUT=mi_monotone;
 BY trt baseline_T2D_status randomization ratio;
 MCMC IMPUTE=MONOTONE;
 VAR baseline_ELF ELF1 ELF2;
 RUN;

 PROC MI DATA=mi_monotone NIMPUTE=1 OUT=mi_complete;
 BY _imputation_;
 CLASS trt baselne_T2D_status randomization ratio;
 VAR trt baselne_T2D_status randomization ratio baseline_ELF ELF1
ELF2;
 MONOTONE method=reg;
 RUN;

<Imputed datasets will be analyzed using proc mixed as described in
5.1.4 and the results will be combined as below>

PROC MIANALYZE  data=;
class trt baseline_T2D_status randomization ratio;
MODELEFFECTS trt;
Stderr stderr;
RUN;
```

Missing data imputation for the sensitivity analysis No 2. for the primary endpoint sample code:

```
proc mi data = non_mono out = monotone nimpute = 100 seed = 123;
by trt baseline_T2D_status randomization ratio;
var  baseline_ELF ELF1 ELF2;
mcmc impute = monotone;
run;

proc mi data=monotone out=imputed1 seed=124 nimpute=1;
 by _imputation_;
 class trt baseline_T2D_status randomization ratio;
 var baselne_T2D_status randomization ratio baseline_ELF ELF1 ELF2;
```

```
 monotone method=reg;
 mnar model(baseline_ELF ELF1 ELF2/ modelobs= (trt='Placebo'));
run;
```

Missing data imputation for the fibrosis response analysis sample code:

```
proc mi data=data1 seed=53156635 nimpute=100 out=outmi1 noprint;
class trt bslfib rsp randomization ratio;
monotone logistic(rsp=trt bslfib randomization ratio);
var trt bslfib randomization ratio rsp;
run;
```

Cochran-Mantel-Haenszel (CMH) test stratified by baseline T2D status and randomization ratio sample code:

```
proc freq data=;
   tables baseline_T2D_status*randomization ratio*trt*Response
/relrisk cmh commonriskdiff;
run;
```

Missing data imputation for the fibrosis response sensitivity analysis sample code:

```
proc mi seed=98765 nimpute=100;
  class trt RSP baseline_T2D_status randomization ratio;
  monotone logistic(RSP=trt baseline_T2D_status randomization ratio);
  mnar model(RSP/modelobs=(trt='PLACEBO'));
  var trt baseline_T2D_status randomization ratio RSP;
run;
```

Apply Wilson-Hilferty transformation to the CMH statistic from each imputation (CMH statistics saved in dataset CMH):

```
data cmh_wh;
  set cmh(WHERE=(AltHypothesis="General Association"));
  cmh_value_wh=((VALUE/DF)**(1/3) - (1-2/(9*DF)))/SQRT(2/(9*DF));
  cmh_sterr_wh = 1.0;
run;
```

ELECTRONIC SIGNATURE PAGE

Placeholder for electronic signature page

Signature Page for VV-CLIN-001357 v1.0

Approval Task	[redacted] Clinical 07-Feb-2023 05:39:59 GMT+0000
Approval Task	[redacted] Clinical 08-Feb-2023 19:41:00 GMT+0000

Signature Page for VV-CLIN-001357 v1.0