STATISTICAL ANALYSIS PLAN

A Phase 2b, Randomized, Double-blind, Placebo-controlled, Multi-center Study to Evaluate the Efficacy, Safety, and Tolerability of Three Doses of Aldafermin Administered for 24 Weeks for the Treatment of Histologically Confirmed Nonalcoholic Steatohepatitis (NASH)

NCT03912532

Inverstigational Product:	Aldafermin (NGM282)
Protocol Number:	18-0108 (ALPINE 2/3)
Development Phase:	2b
Sponsor:	NGM Biopharmaceuticals, Inc. 333 Oyster Point Boulevard South San Francisco, CA 94080
Original SAP (Version 1.0)	19 November 2020
Amendment 1 (Version 2.0)	10 May 2021

SIGNATURE PAGE

STUDY TITLE: A Phase 2b, Randomized, Double-blind, Placebo-controlled, Multi-center Study to Evaluate the Efficacy, Safety, and Tolerability of Three Doses of Aldafermin Administered for 24 Weeks for the Treatment of Histologically Confirmed Nonalcoholic Steatohepatitis (NASH)

Prepared by **Date**

[redacted]
Sr. Director, Biometrics
NGM Biopharmaceuticals, Inc.

10 May 2021

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

Signature **Date**

[redacted]

10 MAY 2021

Sr. Director, Biometrics
NGM Biopharmaceuticals, Inc.

[redacted]

10 MAY 2021

Sr. Director, Clinical Development
NGM Biopharmaceuticals, Inc.

GLOSSARY OF ABBREVIATIONS

Abbreviation	Definition
ADA	Antidrug antibody
AE	Adverse event
ALP	Alkaline phosphatase
ALT	Alanine aminotransferase
ANCOVA	Analysis of covariance
ANOVA	Analysis of variance
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
FAS	Full analysis set
GGT	Gamma-glutamyl transferase
HDL-C	High density lipoprotein cholesterol
HOMA IR	Homeostasis model assessment–estimated insulin resistance
[redacted]	[redacted]
ITT	Intent to treat
IxRS	Interactive Web/Voice Response System
LDL-C	Low density lipoprotein cholesterol
LFC	Liver fat content
LISSA	Local injection-site symptom assessment
MAR	Missing at random
MCP-Mod	Multiple comparison procedure – modelling
MedDRA	Medical Dictionary for Regulatory Activities

Abbreviation	Definition
MI	Multiple imputation
MRI PDFF	magnetic resonance imaging proton density fat fraction
NAB	Neutralizing antibody
NAFLD	Nonalcoholic fatty liver disease
NAS	NAFLD activity score
NASH	Nonalcoholic steatohepatitis
NCI	National Cancer Institute
NIH	National Institutes of Health
PK	Pharmacokinetic
PP	Per protocol
Pro-C3	N-terminal Type III collagen
PT	Preferred term
[redacted]	[redacted]
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

1. Introduction

This Statistical Analysis Plan (SAP) provides the details of the statistical analyses to be performed for NGM Biopharmaceuticals clinical study 18-0108 (ALPINE 2/3) entitled "A Phase 2b, Randomized, Double-blind, Placebo-controlled, Multi-center Study to Evaluate the Efficacy, Safety, and Tolerability of Three Doses of Aldafermin Administered for 24 Weeks for the Treatment of Histologically Confirmed Nonalcoholic Steatohepatitis (NASH)".

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

1.1 Primary Study Objectives and Design

The primary objectives of this study are to evaluate the efficacy, safety and tolerability of Aldafermin administered for 24 weeks of treatment of histologically confirmed NASH compared to placebo.

This is a multiple center evaluation of Aldafermin administered for 24 weeks as a daily subcutaneous (SC) injection in a randomized, double-blind, placebo-controlled study in subjects with histologically confirmed NASH. Approximately 152 subjects, of which no more than 15% (approximately 22 subjects) will be self-identified Asian subjects, will be randomized at approximately 40 sites in the United States (US).

Subjects to be studied will have histologically confirmed NASH as defined by the National Institutes of Health (NIH) NASH Clinical Research Network (CRN) and determined by a qualified central pathologist's reading. Subjects must have a nonalcoholic fatty liver disease (NAFLD) activity score (NAS) of at least 4, with a minimum of 1 point in each of the three components (steatosis, inflammation and ballooning), along with the presence of either Stage 2 or 3 fibrosis. Subjects with no (Stage 0) or minimal (Stage 1) fibrosis and cirrhosis (Stage 4) will be excluded from this study. Subjects will undergo magnetic resonance imaging (MRI) during the Screening period, which must demonstrate ≥8% total liver fat content (LFC) by MRI proton density fat fraction (MRI PDFF).

On Day 1, eligible subjects will be randomized into one of the four treatment groups (Aldafermin 0.3 mg, Aldafermin 1 mg, Aldafermin 3 mg, or placebo) in a 1:1:1:1 ratio (n=38 subjects in each treatment group). The randomization will be stratified by baseline stage of fibrosis [redacted]

The first dose of Aldafermin (Day 1) and doses at Weeks 2, 4, 8, 12, 18, and 24 study visits will be self administered in the clinic, with all other daily doses through Week 24 self administered daily at a similar time at home. Subjects will return to the clinic on Weeks 2, 4, 8, 12, and 18 for on treatment assessments and to receive Aldafermin study drug kits.

Low density lipoprotein cholesterol (LDL-C) will be evaluated at Weeks 2, 4, 8, 12 and 18 in all subjects for possible increases in lipid levels associated with Aldafermin administration. In statin naïve subjects, rosuvastatin will be started in subjects meeting specific LDL-C level criteria at Week 2. In statin experienced subjects, rosuvastatin (10mg QD in non-Asians and 5mg in Asians) will be started at Week 2. Rosuvastatin and matched placebo are over encapsulated to maintain the blind at the sites and with the sponsor study teams. Initiation and ongoing dose adjustments of rosuvastatin/matched placebo will be managed by an unblinded third-party medical monitor and through the [redacted]. The sponsor and study sites will be blinded to LDL-C values and the specific rosuvastatin dosing decisions as outlined in Section 9.4.14 of the protocol. Subjects who have not achieved an adequate response or cannot tolerate rosuvastatin as defined in Section 9.4.14 of the protocol will be considered for the addition of ezetimibe as second-line lipid management therapy. Ezetimibe 10 mg tablets can be administered with or without food at the same time as rosuvastatin and Aldafermin or as monotherapy. Subjects whose LDL-C values are not adequately managed by their maximum tolerated rosuvastatin dose plus ezetimibe will be discontinued from study and an early withdrawal study visit will be completed. The follow-up/end-of-study (EOS) visit should be scheduled for 6 weeks after the early withdrawal study visit.

Rosuvastatin or matching placebo will be dispensed at each study visit through Week 18 (statin-naïve) or Week 24 (statin experienced). If the subject is taking over-encapsulated rosuvastatin or matched placebo, evaluate tolerability and continue therapy to Week 24 if statin naïve or Week 30 if statin experienced.

All subjects will return to the clinic at Week 25 for post-treatment LDL-C assessment. Subjects who have an LDL-C value of >100 mg/dL and > 15 mg/dL above their Day 1 value at Week 25 will be re-assessed at Week 30. Subjects who have an LDL-C value of > 100 mg/dL and >15 mg/dL above their Day 1 value at Week 30 will have an additional 4-week follow-up visit to confirm their LDL-C value is ≤ 100 mg/dL or ≤ 15 mg/dL above their Day 1 value.

The study visit at Week 24 will be the Aldafermin end-of-treatment (EOT) clinic visit and subjects will return to the clinic at Weeks 25 and 30 (or at 1 and 6 weeks after last dose of Aldafermin) for a post-treatment response and EOS follow-up visits. Subjects are required to undergo MRI PDFF on Screening Day -28 and at Weeks 12, 24/Early Withdrawal (EOT), and 30 (EOS) visits.

All subjects will have blood samples drawn for [redacted]

1.2 Secondary and Other Objectives

The secondary objectives of this study are to evaluate the effect of Aldafermin on the pharmacokinetics, biomarkers of target engagement, fibrogenesis, and imaging.

[redacted]

1.3 Sample Size

The sample size for this study is determined based on power simulations for the dose-response testing for the primary efficacy endpoint – the liver fibrosis response (Section 5.1.1). The simulations were originally performed using a generalized MCP-Mod (multiple comparison procedure - modelling) approach (Pinheiro 2014) and those results are presented in the study protocol. The simulation results (Table 1) in this SAP were obtained using a different MCP-Mod approach (Klingenberg 2009) as this approach will be used for both dose-response testing and estimation for the primary efficacy endpoint.

The assumptions about the liver fibrosis response rate for the 4 treatment groups are the same as those in the protocol. Specifically, assuming that the liver fibrosis response rate is 15% for placebo and 50% for Aldafermin 3mg, respectively, it is demonstrated that a total sample size of 152 subjects (38 subjects per treatment group) will provide at least 85% power to detect a positive dose-response trend (i.e., an upward slope) at the 5% significance level (two-sided) under 6 different dose-response scenarios as shown in Table 1 (the last 6 rows). To account for potential premature discontinuation of the study subjects, a 15% drop-out rate at Week 24 is also assumed in all these simulations (i.e., only 32 subjects per group were used in these simulations).

In addition, simulations were also performed under the assumption of no dose effect (i.e., response rate=15% for all 4 treatment groups) to assess the type I error rate control of the MCP-Mod approach (first row in Table 1).

Table 1. Power Simulations for Dose-Response Testing

Liver Fibrosis Response Rate				
Placebo	**Aldafermin 0.3 mg**	**Aldafermin 1 mg**	**Aldafermin 3 mg**	**Power**
15%	15%	15%	15%	5%
15%	16%	19%	50%	92%
15%	18%	30%	50%	90%
15%	20%	42%	50%	93%
15%	32%	39%	50%	85%
15%	36%	45%	50%	87%
15%	43%	48%	50%	90%

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.
- **Safety Analysis Set**: All subjects who receive at least one dose (full or partial) of study drug will be included in the safety analysis set. All safety endpoints will be summarized using the safety analysis set and will be based on the actual treatment received if it differs from the randomized treatment.
- **Full Analysis Set**: All randomized subjects who receive at least one dose (full or partial) of study drug and have at least one valid, non-missing post-dose efficacy/PD value will be included in the full analysis set (FAS). The FAS will be used for sensitivity analyses to support the ITT analyses and will be based on the randomized treatment if it differs from the actual treatment received.
- **Per Protocol Analysis Set**: The Per Protocol (PP) analysis set is a subset of the FAS and will include subjects who have at least one valid, non-missing baseline and post-dose liver biopsy results and do not have protocol deviations or inter-current events that impact the liver biopsy assessments. The PP analysis set will be used for sensitivity analyses to support the ITT analyses, and this analysis set will be determined by the study team prior to the database lock or unblinding.

- [redacted]

2.2 Data Conventions

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

- All subject-level data will be provided in data listings, unless stated otherwise.
- 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 follow-up – baseline. Percent (or relative) change from baseline (percent/relative change) is calculated as 100 × (change from baseline/baseline).
- 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 and percentage distribution will be presented for categorical endpoints. Kaplan-Meier estimates will be presented for time-to-event endpoints if applicable.
- Continuous endpoints will be analyzed using an analysis of variance (ANOVA) or analysis of covariance (ANCOVA) model. Ordinal categorical endpoints will be analyzed using a Wilcoxon rank-sum test. Nominal categorical endpoints will be analyzed using a Pearson's Chi-square test or Cochran-Mantel-Haenszel (CMH) test (A Fisher's exact test will be used if ≥ 25% of the cells have expected counts < 5). Time-to-event endpoints will be analyzed using a (stratified) log-rank test and Cox's proportional hazards model.
- Type III sums of squares will be used for all ANOVA and ANCOVA models.

- The level of statistical significance used for all statistical tests will be 0.05, unless stated otherwise. Statistical hypotheses for the primary efficacy analyses and the associated sensitivity analyses will be tested for statistical significance with two-sided alternatives, and one-sided alternatives will be used for all secondary and other efficacy analyses. All confidence intervals (CIs) will be two-sided, unless stated otherwise.
- Data from unscheduled visits will be presented in data listings, but will not be used for statistical analysis/summaries.
- Missing data for selected histological endpoints (e.g., primary and 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

No analysis visit windows will be defined for this study, i.e., the nominal visits /time-points will be used for all statistical summaries/analyses.

3. Disposition and Exit Status

3.1 Disposition and Exit Status

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

3.2 Protocol Deviations

Subject data will be examined prior to database lock or unblinding for evidence of protocol deviations in order to identify subjects with protocol violations that could exclude such subjects from the PP analysis set. Classification of possible protocol deviations will be reviewed and approved by the study team based on predefined criteria. Protocol deviations will be provided in data listings.

4. Demographics and Other Baseline Characteristics

4.1 Demographics

Demographics will be summarized with descriptive statistics by treatment group and overall for each analysis set. These variables include age (years), sex (Male or Female), race (Asian, Black, 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 Screening.

4.2 Disease Characteristics

Baseline disease characteristics will be summarized with descriptive statistics by treatment group and overall for each analysis set. These variables include NAS score, fibrosis stage (2 or 3), type 2 diabetes status (Yes/No), LFC (%) by MRI PDFF, ALT, AST, cholesterol, HDL-C, LDL-C, triglycerides, Pro-C3, ELF, and C4.

4.3 Prior and Concomitant Medications

Medications taken at least once within 42 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 (WHODrug Global, B3, March 2018). 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 ITT 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 ITT analysis set. Surgical procedures recorded for the corresponding events and concurrent procedures will be provided in the data listings.

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. Summaries for continuous endpoints will be provided for the actual value, change and percent change from baseline (if applicable), respectively.

5.1 Primary Efficacy Analyses

5.1.1 Primary Efficacy Endpoint

The primary efficacy endpoint for this study is a binary response derived from the liver biopsy assessments of the NAFLD activity score (NAS) and fibrosis score defined by NIH NASH CRN (Appendix A). Liver biopsies are collected at Screening (prior to the study randomization) and Week 24, and will be assessed by a qualified central study pathologist.

The primary efficacy endpoint is the liver fibrosis response at Week 24 (hereafter referred to as “fibrosis response”), defined as a subject achieving an improvement in liver fibrosis greater than or equal to one stage (NASH CRN fibrosis score) with no worsening of steatohepatitis after 24 weeks of treatment. In this definition, fibrosis stages 1a, 1b and 1c are all considered stage 1, and no worsening of steatohepatitis is defined as no increase in NAS for ballooning, inflammation, or steatosis.

5.1.2 Primary Estimand

[redacted]

[redacted]

[redacted]

5.1.3 Statistical Hypotheses

The statistical hypotheses (null and alternative) to be tested with respect to the fibrosis response are:

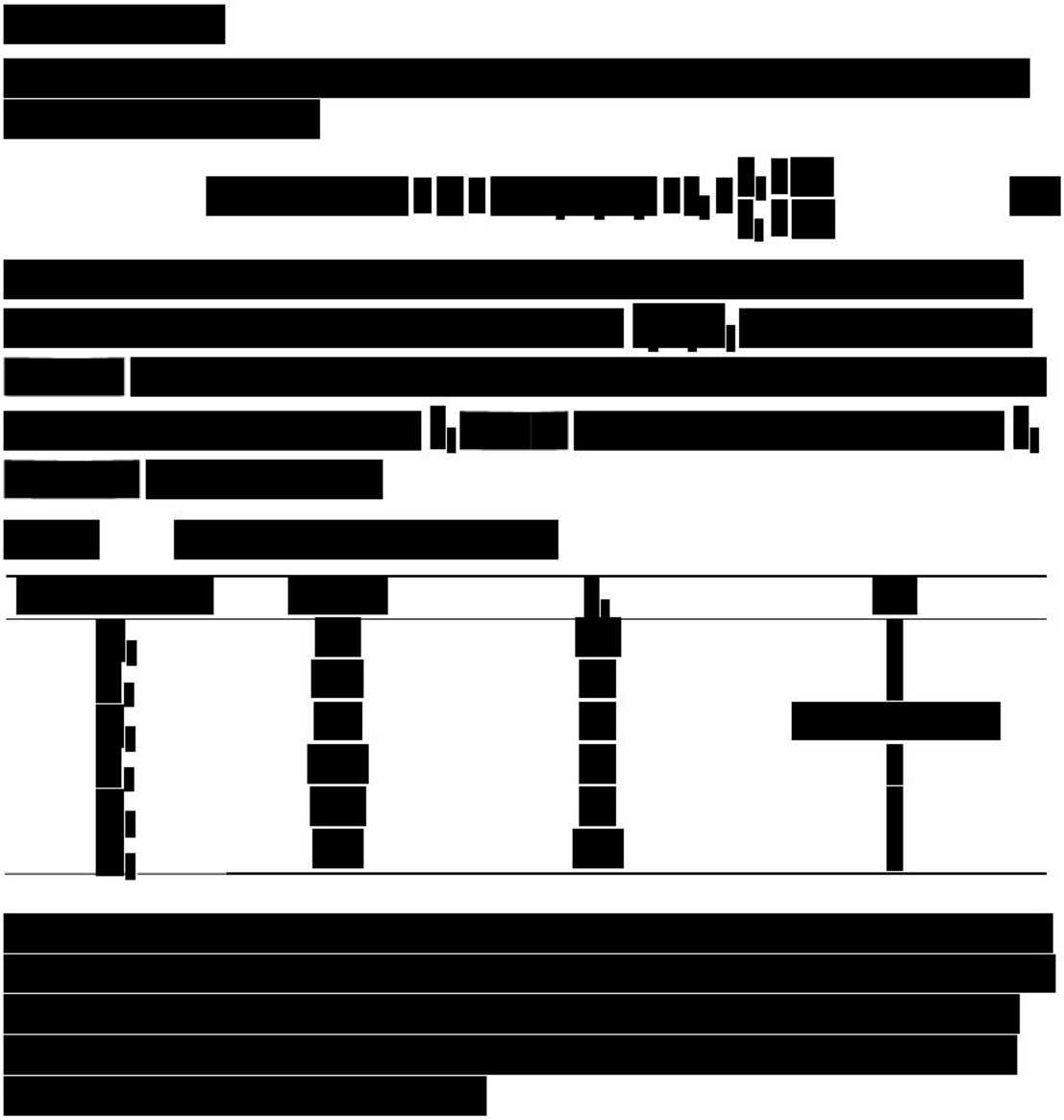
- **H0 (Null Hypothesis)**: there is no dose-response relationship between Aldafermin dose and the fibrosis response.
- **H1 (Alternative Hypothesis)**: there is a dose-response relationship between Aldafermin dose and the fibrosis response.

These hypotheses will be tested using the MCP-Mod (multiple comparison procedure - modelling) approach proposed by Klingenberg (2009). The details of this approach are provided in Section 5.1.3.

5.1.4 The MCP-Mod Approach

The MCP-Mod approach is a two-stage procedure in which a set of plausible dose-response models are selected at the design stage. These candidate models will then be used for both dose-response testing (MCP step) and estimation (Mod step) at the analysis stage. The dose-response testing will be performed using an appropriate multiple comparison method to account for the multiplicity issue associated with the multiple candidate models.

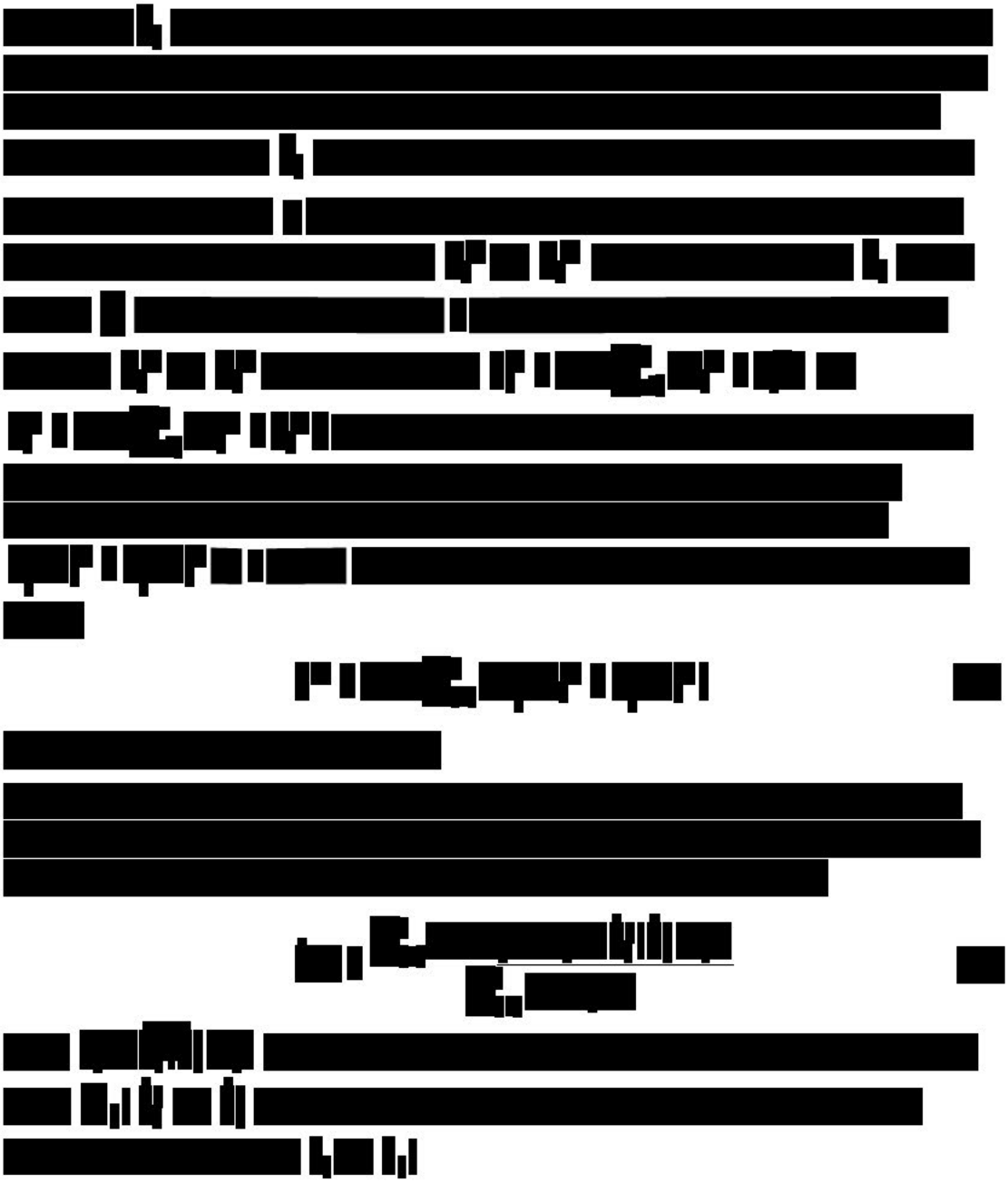
This section provides the statistical details of the MCP-Mod approach to be used in the data analysis for this study. The SAS implementation of this approach is provided in Appendix B.



[redacted]

[redacted]

[redacted]



[redacted]

5.1.5 Data Analyses

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

Imputation

Missing fibrosis responses are imputed 25 times to generate 25 complete data sets.
The imputed values will be obtained using a logistic model with fixed effects for treatment group and baseline fibrosis stage. At the end of this step, all the 25 complete data sets contain the same non-missing values while the missing values are replaced by imputed values.
The following is an example of SAS programming code to be implemented.

The random seeds to be used for this and all subsequent analyses are provided in Appendix C.

Analysis

Each of the 25 competed data sets will be analyzed using the MCP-Mod approach described in Section 5.1.3. The SAS programming code to be implemented for this approach is provided in Appendix B. Results to be obtained from each data set include the multiplicity adjusted p-value for each candidate model (macro MCP), the point estimate of the response rate $\hat{F}(x)$ (E4) for each dose level and the standard error for $\hat{F}(x)$ (macro Mod). The standard error will be estimated by the sample standard deviation obtained from the bootstrap distribution of the point estimate (E4).

Pooling

The results from the 25 completed data sets obtained in the analysis step are combined for the inference. The pooled estimate of the multiplicity adjusted p-value for testing H0 vs H1 will be the median of the individual multiplicity adjusted p-values from the 25 completed data sets. The pooled estimate of the response rate along with its standard error and 95% CI for each dose level will be obtained using the method developed by Rubin (1987). The following is an example of SAS programming code to be implemented.

The null hypothesis H0 (Section 5.1.2) will be rejected and a statistically significant trend (i.e., an upward slope) of the fibrosis response across Aldafermin doses will be established if

- the pooled estimate of the multiplicity adjusted p-value obtained from the MI inference is ≤ 0.05 and;
- the pooled estimate of the response rate for Aldafermin 3 mg is higher than that for placebo.

The ITT analysis set will be used for these analyses.

5.1.6 Sensitivity Analyses

[redacted]

5.2 Secondary Efficacy Analyses

Secondary efficacy analyses will include treatment group comparisons of the fibrosis response (the primary efficacy endpoint) between each Aldafermin dose and placebo as well as the analyses of the following secondary efficacy endpoints:

- Subjects with resolution of NASH defined as an NAS score of 0 or 1 for inflammation and 0 for ballooning, with no worsening (compared to baseline) of fibrosis as determined by the NASH CRN criteria at Week 24
- Subjects with combined response of fibrosis improvement and NASH resolution, defined as subjects with an improvement in liver fibrosis by ≥ 1 stage with no worsening of steatohepatitis (no increase in NAS for ballooning, inflammation, or steatosis) and resolution of NASH as determined by the NASH CRN criteria at Week 24
- Subjects with improved, no change, and worsening of the following histological assessments at 24 weeks compared to baseline
  - Fibrosis stage
  - NAS (total score and individual components)
- Subjects with the following changes in LFC by MRI-PDFF at Week 24:
  - Normalization of LFC (defined as LFC < 5%)
  - ≥ 5% decrease from baseline in LFC
  - ≥ 30% relative decrease from baseline in LFC
- Change and percent change from baseline to Week 24 in the following:
  - LFC by MRI-PDFF
  - ALT, AST, total bilirubin, ALP, and gamma glutamyl transpeptidase (GGT)
  - Total cholesterol, high density lipoprotein cholesterol (HDL-C), LDL-C, and triglycerides
  - Lipoprotein particles [redacted]
  - Homeostasis model assessment–estimated insulin resistance (HOMA IR)
  - N-terminal Type III collagen (Pro-C3)
  - Total ELF Score and individual components (hyaluronic acid, PIIINP, TIMP-1)
  - 7 alpha hydroxy 4 cholesten 3 one (C4) and serum bile acids (total and individual bile acids)
  - [redacted]
- Subjects with normalization (i.e., values within the normal range) of ALT and AST at Week 24

For binary efficacy endpoints (i.e., responders), treatment group comparisons between each Aldafermin dose and placebo will be performed using the Cochran-Mantel-Haenszel (CMH) test stratified by baseline fibrosis stage. The treatment group difference of the response rates along with the corresponding 90% CI will be provided. The 90% CI will be constructed using the normal approximation method.

For continuous efficacy endpoints, change and percent change from baseline will be analyzed using an analysis of covariance (ANCOVA) model with effects for treatment, baseline fibrosis stage and baseline outcome value as a covariate. Within the framework of the ANCOVA model, a point estimate of the treatment group difference between each Aldafermin dose and placebo will be provide along with the corresponding one-sided p-value and 90% CI.

All secondary efficacy analyses will be performed using the ITT analysis set. FAS and PP analysis sets will be used as sensitivity analyses. The missing data methods and sensitivity analyses described for the primary efficacy endpoint in Section 5.1 (for ITT and FAS analysis sets) will be used for the following histological endpoints:

- Fibrosis response
- NASH resolution
- Combined response of fibrosis improvement and NASH resolution

All statistical tests for the secondary efficacy analyses will be performed using a one-side test at the significance level of 5%, and no multiplicity adjustment will be made among these secondary efficacy analyses.

5.3 Other Efficacy Analyses

The analyses of the secondary efficacy endpoints at Week 24 in Section 5.2 will also be performed for all other applicable visits.

[redacted]

5.4 Subgroup Analyses for Efficacy Endpoints

The primary and secondary efficacy endpoints will be summarized with descriptive statistics by baseline fibrosis stage (F2 or F3) and by baseline statin use (experienced or naive).
No inferential statistics will be provided for the subgroups.

6. Safety Analyses

6.1 Study Treatment Exposure and Compliance

6.1.1 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: <2 weeks, 2 to <4 weeks, 4 to <8 weeks, 8 to <12 weeks, 12 to <16 weeks, 16 to <20 weeks, and 20 to <24 weeks, etc.

6.1.2 Study Treatment Compliance

Study treatment compliance is defined as the number of injections actually administered (i.e., number prescribed – number returned) by a patient during the 24-week period divided by the number of injections prescribed multiplied by 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%.

6.2 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-related TEAEs, any Rosuvastatin-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
- TEAEs by primary SOC, PT and relationship to Rosuvastatin
- TEAEs leading to study treatment withdrawal by primary SOC and PT
- Aldafermin-related TEAEs leading to study treatment withdrawal by primary SOC and PT
- Rosuvastatin-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-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 separate listings.

6.3 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.

6.4 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.

6.5 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. Frequency and percentage distribution of the overall ECG interpretation results (Abnormal, Indeterminate, Normal, Not evaluable, Unknown per CRF) will be presented by treatment group for each visit.

6.6 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.

[redacted]

6.8 Pregnancy Test

Urine and serum pregnancy test for female subjects will be presented in data listings.

7. Pharmacokinetic, Biomarker, and Immunogenicity Data

[redacted]

8. Health Outcomes Data Analyses

Not applicable

9. Interim Analyses and Data Monitoring

No interim analyses are planned for this study.

10. Deviations from the Protocol

[redacted]

[redacted]

11. Amendments

[redacted]

12. Literature References

Klingenberg B. Proof of concept and dose estimation with binary responses under model uncertainty. *Statistics in Medicine* 2009; 28: 274-292.

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

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.

13. Apendices

13.1 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).

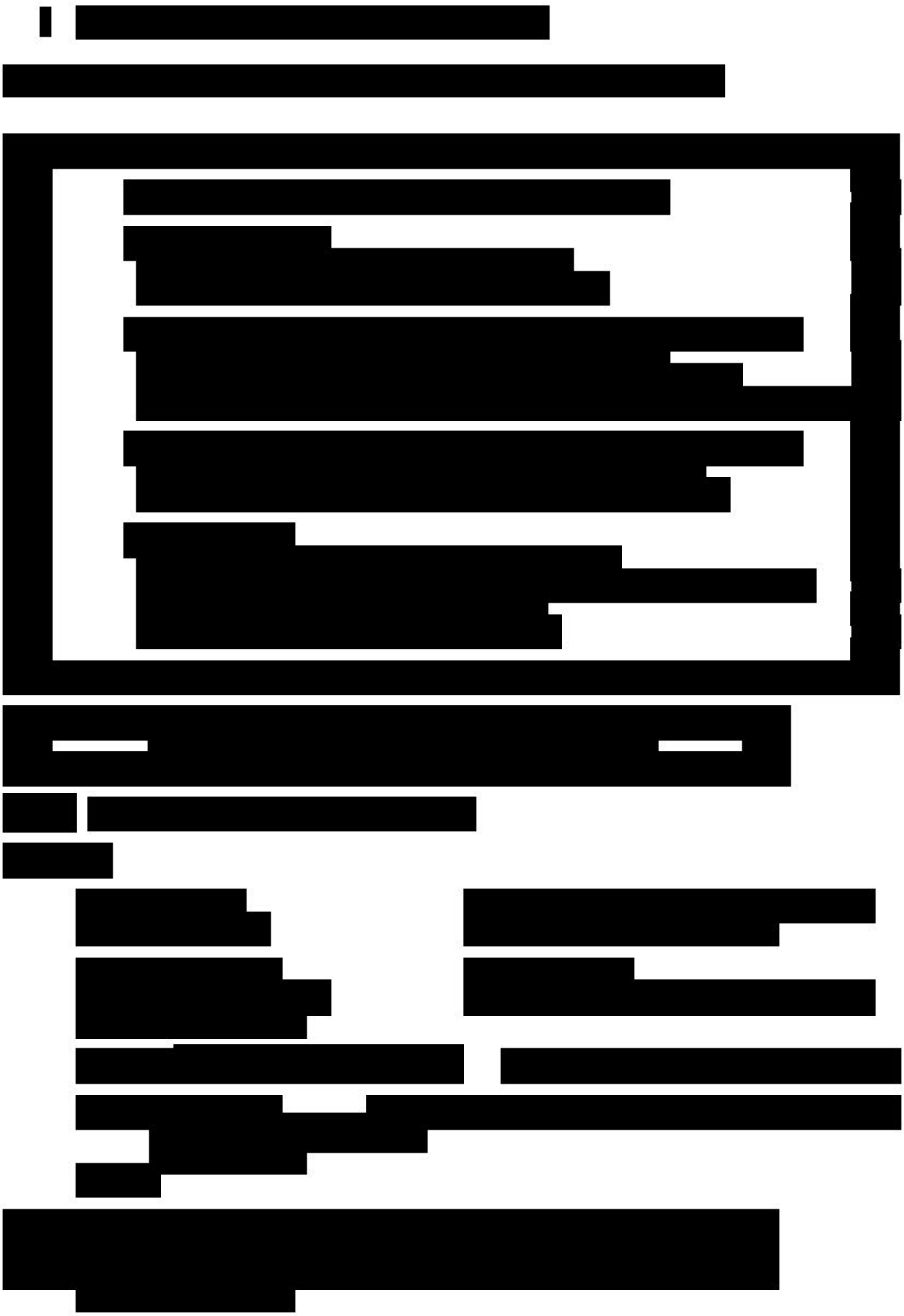
Fibrosis Stage	Score
0	No Fibrosis
1a	Zone 3, Mild
1b	Zone 3, Moderate
1c	Periportal Only
2	Zone 3 and Periportal
3	Bridging
4	Cirrhosis

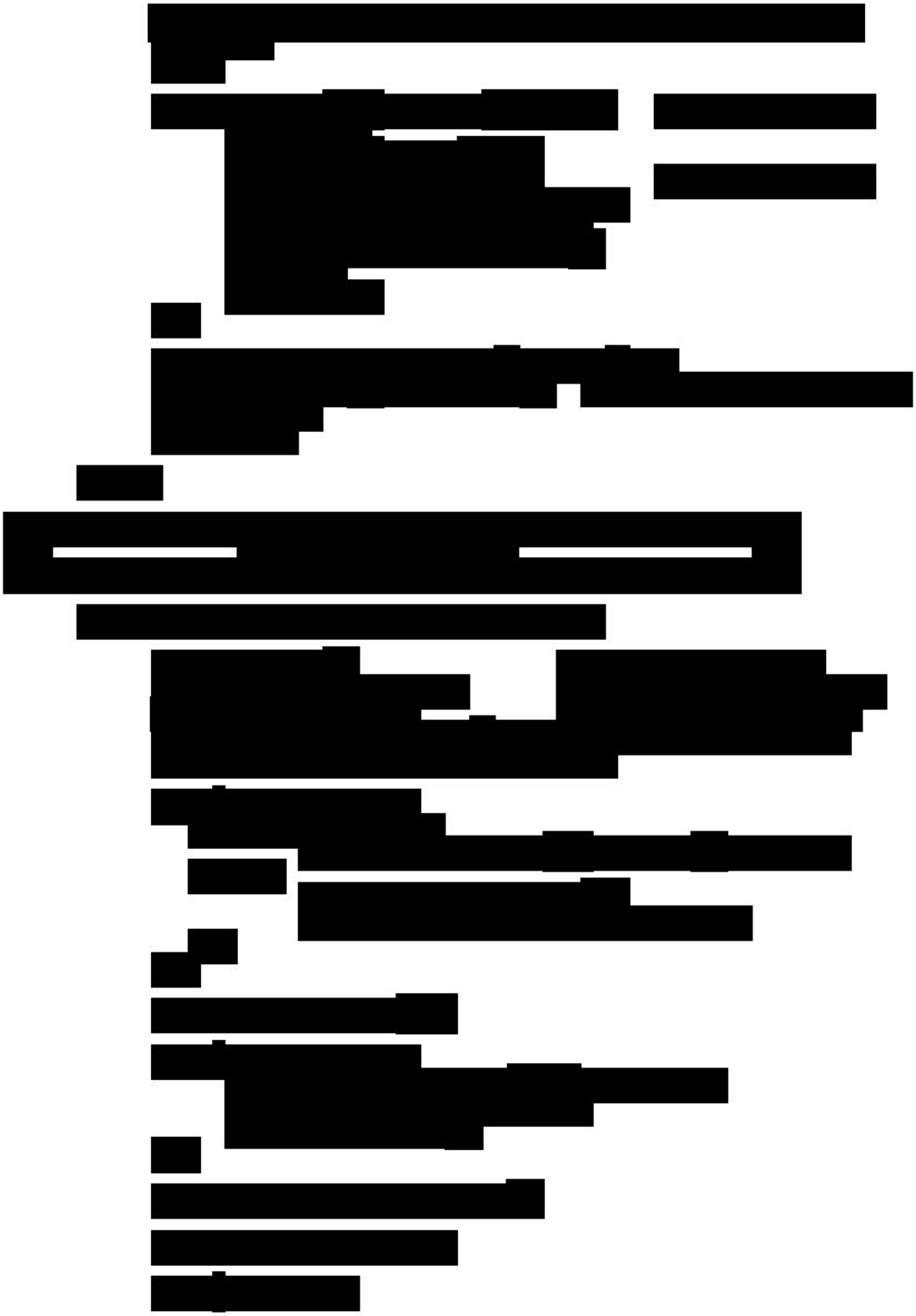
13.2 Appendix B: SAS Programming Code for MCP-Mod Analysis

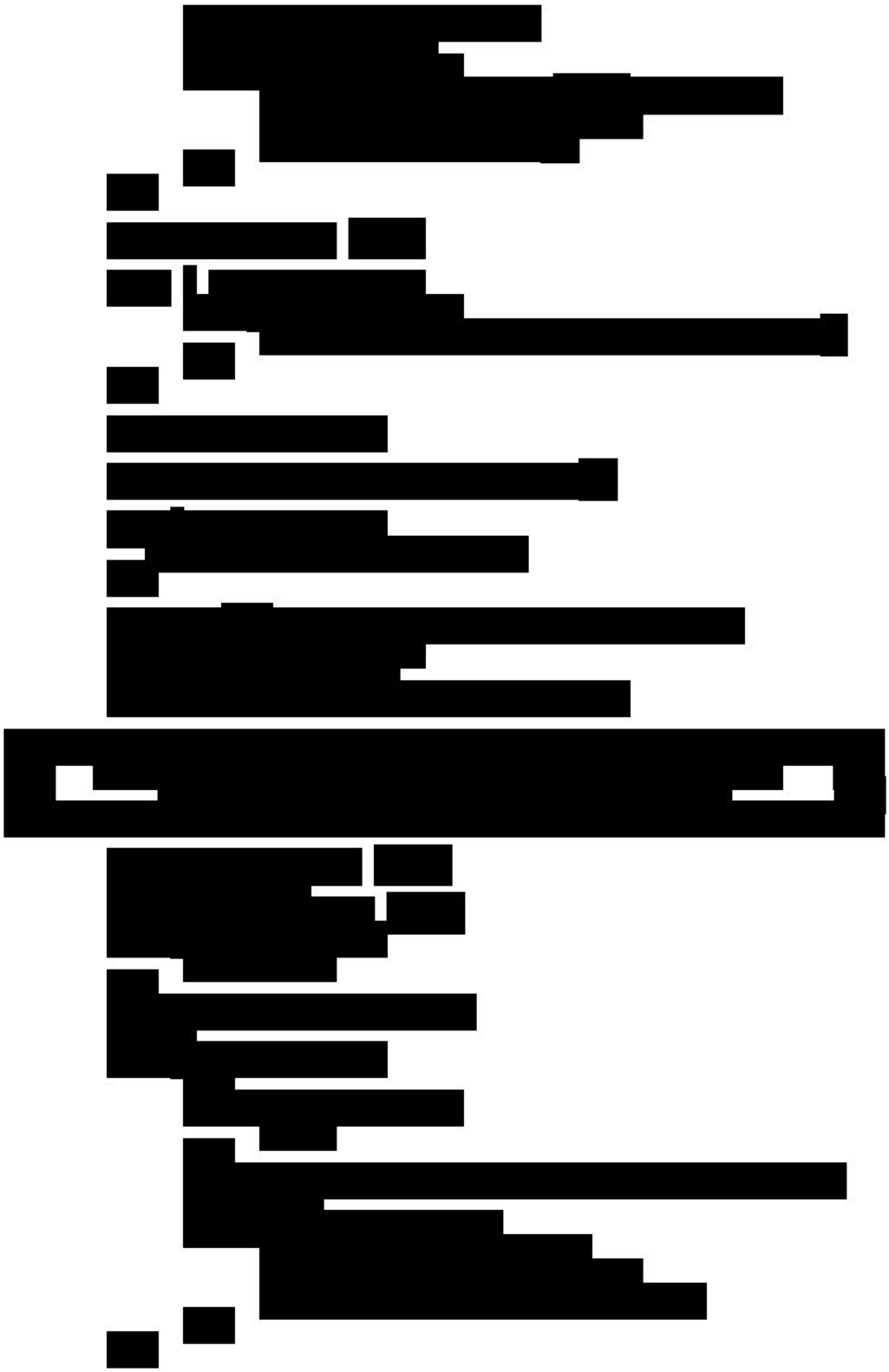
[redacted]

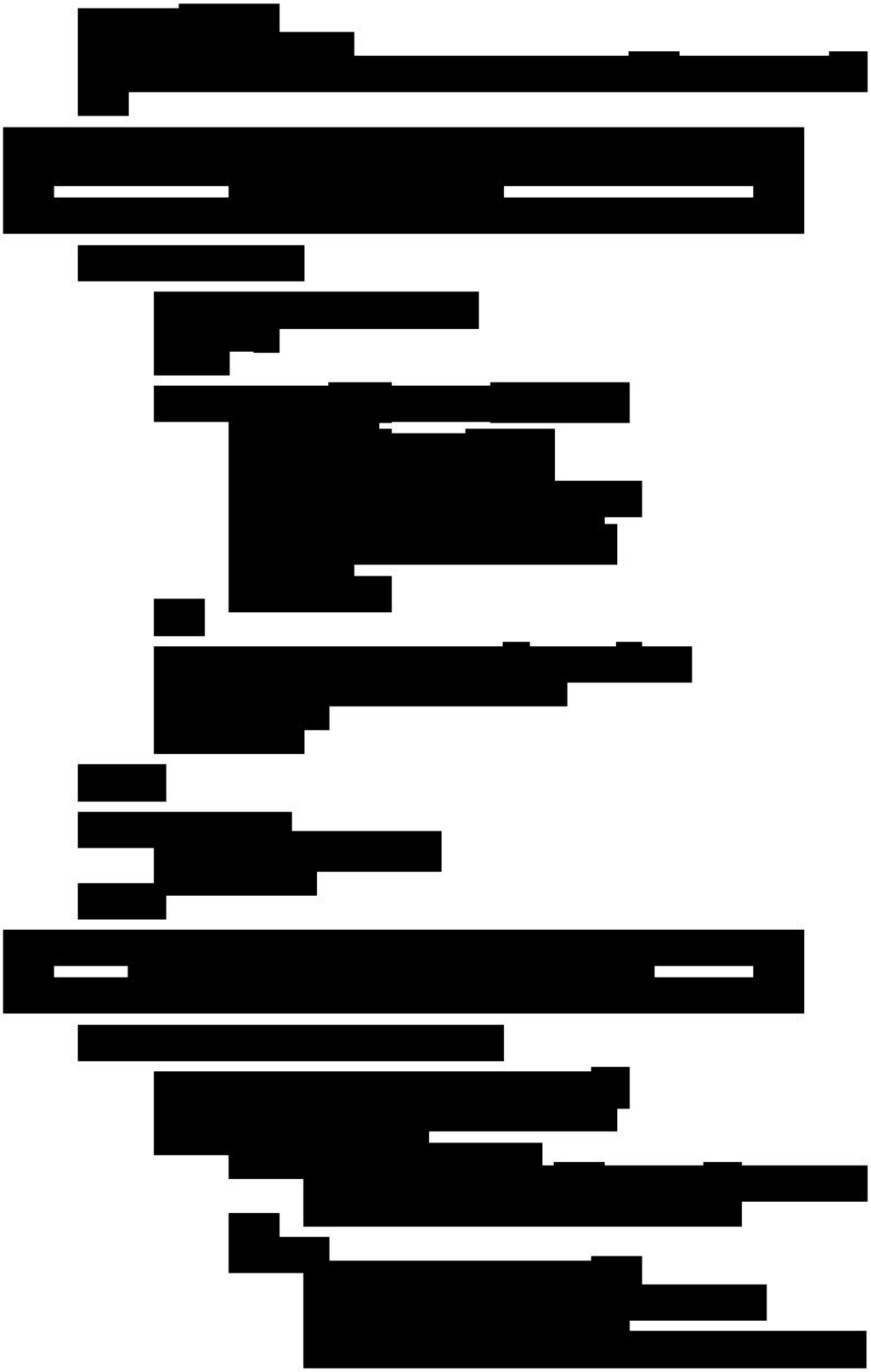
Obs	COL1	COL2
1	0.0	0
2	0.0	0
3	0.3	1

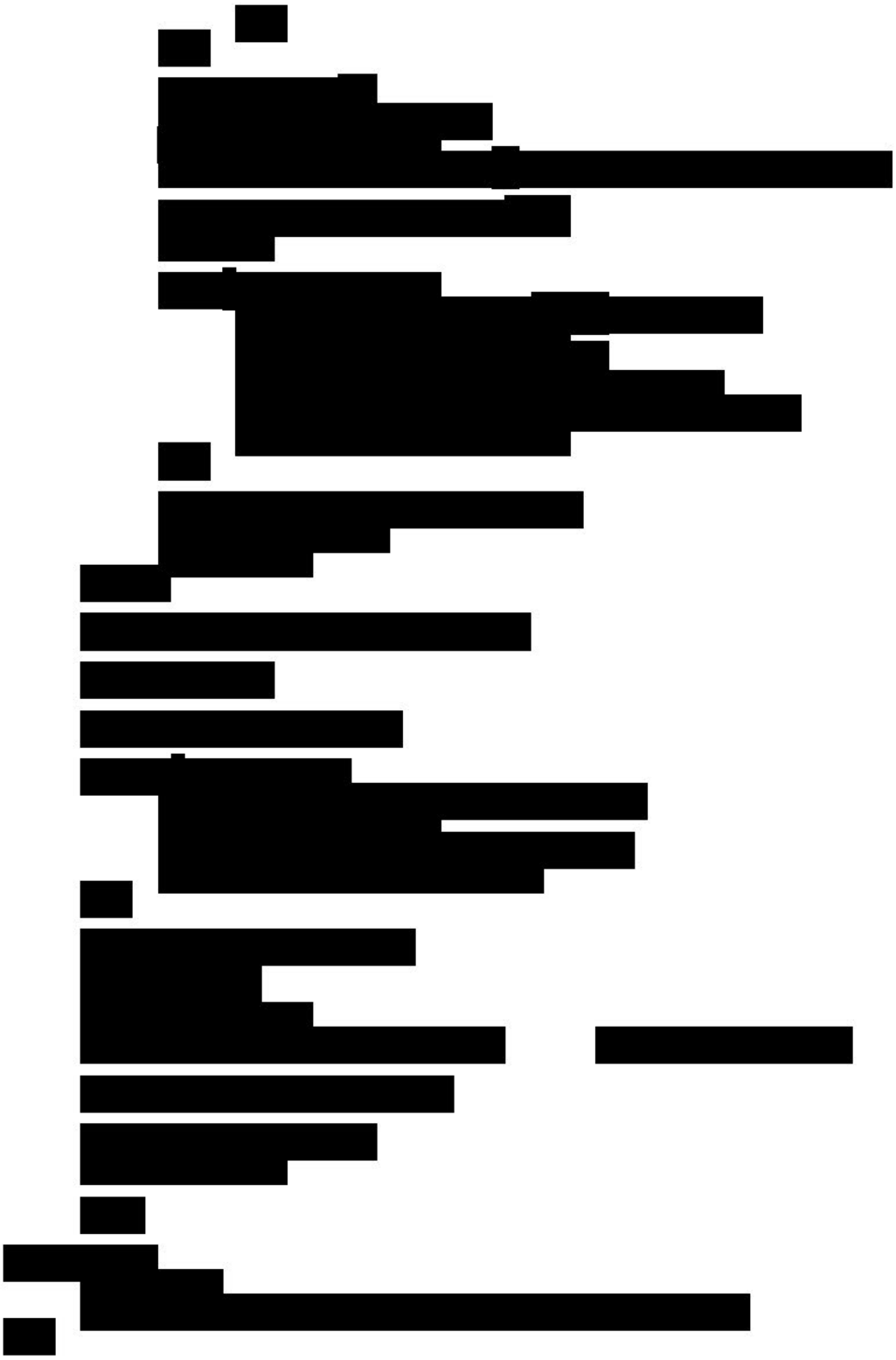
[redacted]











[redacted]

13.3 Appendix C: Simulation Parameters

[redacted]

[redacted]