NGM Biopharmaceuticals, Inc.

15-0106

A Phase 2, Randomized, Double Blind, Placebo Controlled, Parallel Group, Multiple Center Study to Evaluate the Safety, Tolerability, and Efficacy of NGM282 Administered for 12 Weeks in Patients with Primary Sclerosing Cholangitis (PSC)

26 May 2017

Statistical Analysis Plan

Version 1.0

Prepared by:

[redacted]

SIGNATURE PAGE

This document has been prepared and reviewed by:

[redacted]

Name

Project Statistician

[redacted]

Signature

Date

This document has been reviewed and approved by:

[redacted]

Sr. Director, Clinical Development and
Medical Affairs
NGM Biopharmaceuticals, Inc.

[redacted]

TABLE OF CONTENTS

List of Abbreviations

AE	adverse event
ALP	alkaline phosphatase
ALT	alanine aminotransferase
ANCOVA	analysis of covariance
AST	aspartate aminotransferase
ATC	Anatomical Therapeutic Chemical
AUC	area under the concentration–time curve
$C_{2h\ post\text{-}dose}$	concentration at 2 hours postdose
C4	7-alpha-hydroxy-4-cholesten-3-one
CI	confidence interval
C_{max}	maximum drug concentration
C_{trough}	trough concentration
CTCAE	Common Terminology Criteria for Adverse Events
CV	Coefficient of Variation
ECG	electrocardiogram
eCRF	electronic Case Report Form
EOS	end of study
EOT	end of treatment
FAS	Full Analysis Set
FGF19	fibroblast growth factor 19
HBsAg	hepatitis B surface antigen
HCV	hepatitis C virus
HDL	high-density lipoprotein
hs-CRP	high-sensitivity C-reactive protein
IA	Interim Analysis
IBD	inflammatory bowel disease
IgG4	immunoglobulin G 4
IRB	Institutional Review Board
ISR	injection-site reaction
[redacted]	
LDL	low-density lipoprotein
LISSA	local injection-site symptom assessment
LS	least square
LLT	Lowest Level Term (MedDRA)
MedDRA	Medical Dictionary for Regulatory Activities
MMRM	mixed-effect model repeated measures
MRCP	magnetic resonance cholangiopancreatography

NGM	NGM Biopharmaceuticals, Inc.
[redacted]	
NOAEL	no-observed-adverse-effect level
NRS	Numeric Rating Scale
PBC	primary biliary cirrhosis
PD	pharmacodynamics; pharmacodynamic
PI	Principal Investigator
PK	pharmacokinetics; pharmacokinetic
PP	per protocol
PSC	primary sclerosing cholangitis
PT	preferred term
SAE	serious adverse event
SD	standard deviation
SE	standard error
SOC	system organ class
SUSAR	suspected, unexpected serious adverse reaction
T2D	type 2 diabetes
TE	transient elastography
TEAE	treatment-emergent adverse event
TSH	thyroid-stimulating hormone
UA	Urinalysis
UC	ulcerative colitis
UDCA	ursodeoxycholic acid
ULN	upper limit of normal
WHO-DD	World Health Organization Drug Dictionary

1. Introduction

Primary Sclerosing Cholangitis (PSC), a chronic cholestatic liver disease, is characterized by inflammation and onion skin-layers of connective tissue around the biliary system, which finally leads to liver cirrhosis. Ongoing destruction to intra and/or extrahepatic bile ducts often leads to cholestasis, advanced fibrosis, liver cirrhosis, and eventually liver failure. The etiology to PSC is still not clear. It may be caused by multiple factors, such as autoimmunity, portal bacteremia, absorption of toxins, ischemic injury, viral infections, and toxic bile acids in the context of a genetic predisposition. PSC can be diagnosed via a complication, such as cholangitis or hepatic occurs.

There is not any established or approved pharmacologic therapy to PSC until now. The primary treatment to this disease is endoscopic/radiologic dilatation or stenting of bile-duct strictures to slow the progression of disease. Ursodeoxycholic acid (UDCA) is considered the only pharmacologic treatment option with significant biochemical response after multiple prospective clinical studies. However, patients with high PSC and high serum alkaline phosphatase (ALP) usually fail to respond to UDCA therapy. Although orthotopic liver transplantation is the only clinically proven treatment, PSC recurs in up to one-third of deceased donor liver transplantation and up to two-thirds of living-related donor liver transplantation patients.

FGF19 is an endocrine hormone secreted in the ileum and is considered to play an important role in the regulation of bile acid (BA), glucose, and lipid synthesis, as well as energy homeostasis in adult mammals. It is has been proven that FGF19 mRNA expression and gall bladder levels are reduced in PSC patients in recent clinical data.

[redacted] With previous clinical studies, treatment with NGM282 has been proven to have favorable safety and tolerability in normal volunteers and PBC patients. By pharmacologic FGF19 signaling to reduce the synthesis of toxic bile acids, NGM282 presents a great potential treatment to PSC.
This statistical analysis plan is to describe the details of the interim statistical analysis parameters and methods for NGM's protocol 15-0106, Amendment 2, issued in 27May2016.

Objectives

1.1. Primary Objective

- Evaluate the treatment effect of NGM282 as measured by the mean change in ALP from Baseline to Week 12 in patients with PSC.

1.2. Secondary Objectives

Secondary objectives are to:

- Assess the safety and tolerability of NGM282 in patients with PSC with 12 weeks of treatment
- Evaluate the percentage change from Baseline at Week 12 in ALP
- Evaluate the absolute and percentage changes from Baseline at Week 12 in the following:

- ALT, AST, bilirubin (total, direct), and GGT
- Total cholesterol, HDL-cholesterol, LDL-cholesterol, and triglycerides
- Serum bile acids
- Bile-mediated absorption as measured by fat-soluble vitamins

- Compare the incidence and severity of IBD-associated intestinal symptoms
- Compare the incidence and severity of acute cholangitis during the study period.
- Evaluate the exposure of 1 mg and 3 mg of NGM282 in patients with PSC
- Compare the dose-related changes in safety, tolerability, and pharmacodynamic (PD) parameters

[redacted]

2. Investigational Plan

2.1. Overall Study Design and Plan

This is a multiple center, randomized, double-blind, placebo-controlled, parallel-group study of NGM282 when administered for 12 weeks as a daily subcutaneous (SC) injection in subjects with PSC. About 60 from the 90 screened subjects will be randomized to one of the three treatment arms (NGM282 1 mg, NGM282 3 mg, or Placebo) in a ratio of 1:1:1. Subjects will be assessed according to the schedule of procedures and assessments (Appendix 13.1).

2.2. Study Endpoints

The primary efficacy endpoint is defined as the mean change in ALP from Baseline at Week 12.

The secondary efficacy and PD endpoints are the following:

- Percent change from Baseline at Week 12 in ALP
- Changes and percent changes from Baseline at Week 12 in the following
  - ALT, AST, bilirubin (total, direct), and GGT
  - Total cholesterol, HDL-cholesterol, LDL-cholesterol, and triglycerides
  - Bile-mediated absorption as measured by fat-soluble vitamins
- Severity of IBD-associated intestinal symptoms

[redacted]

2.3. Treatments

Subjects will be instructed to self-administer/dose at approximately the same time each morning in the whole treatment period. Blinded study drug (NGM282 or placebo) will be administered as a subcutaneous (SC) injection in the abdomen. Subjects will be trained on self-administering a

SC injection. Self-administration will occur in clinic under the observation of clinic staff on weeks 1, 2, 4, 8, and 12.

3. General Statistical Considerations

Continuous data will be summarized using descriptive statistics (i.e. n, mean, standard deviation, median, minimum, and maximum). Categorical data will be summarized using the subject count and percentage in each category. For the summary statistics of all numerical variables, unless otherwise specified, minimum and maximum will be displayed to the same level of precision as the reported data. Mean and median will be displayed to one level of precision greater than the data collected. Standard deviation/standard error will be displayed to two levels of precision greater than the data collected. P-values will be rounded to three decimal places. If a p-value is less than 0.001 it will be reported as "<0.001." If a p-value is greater than 0.999 it will be reported as ">0.999."

Data will be displayed in all listings sorted by treatment group. Subjects will be identified in the listings by the subject identification number concatenated with the investigator number.

When count data are presented, the percentage will be suppressed when the count is zero in order to draw attention to the non-zero counts. The denominator for all percentages will be the number of subjects in that treatment within the analysis set of interest, unless otherwise specified.

Unless otherwise specified, Baseline will be defined as the last non-missing evaluation prior to or on the date of study drug SC injection (Day 1).

The study day will be calculated as follows:

If the assessment date occurs on or after the date of NGM282 SC injection (Day 1):

Study day = assessment date – date of NGM282 SC injection (Day 1) + 1.

If the date of interest occurs before the date of NGM282 SC injection (Day 1):

Study day = assessment date – date of NGM282 SC injection (Day 1)

There is no study day 0.

For all efficacy analyses, analysis visit windows will be used for assigning assessments to an analysis visit, as follows:

Scheduled Visit	Target Day	Visit Window (Days)
Screening	-42 to -1	Not applicable
Day 1	1	Not applicable
Week 1	7	4 to 10
Week 2	14	11 to 17
Week 4	28	25 to 31
Week 8	56	53 to 59
Week 12	84	81 to 87

Only those assessments which are assigned to an analysis visit will be included in the summary tables and figures which are presented by visit. For summaries of an event at or before a given time point, such as ALP at Week 12, all assessments (including unscheduled visits) on or before the upper limit of the corresponding visit window will be considered. If multiple assessments are assigned to an analysis visit, the nearest assessment to the target day will be defined as the primary analysis record for the summary analysis. All assessments will be presented in the data listings.

For tabulations of data on the Safety set, subjects who were randomized but did not receive treatment will not be included; however data for all subjects in the Safety set will be included in the data listings. In the listings, subjects who received treatment will be grouped by the actual treatment. Percentages for Safety set summaries will be based on the total number of subjects who receive treatment.

[redacted]

All analyses will be conducted using SAS Version 9.2 or higher.

3.1. Sample Size

A sample size of 20 subjects randomized per treatment group was chosen in light of logistical needs and so as to accumulate sufficient safety data on NGM282. The associated power to detect a treatment difference in the primary efficacy analysis, for either NGM282 comparison versus placebo, is at least 90%, based on the following assumptions:

A. ALP Population Mean Change from Baseline:

	NGM282 (either 1 mg or 3 mg)	Placebo
Subjects taking UDCA at Baseline	–75	0
Subjects not taking UDCA at Baseline	–475	0

B. Equal numbers of completing subjects for each of the four cells above
C. Population (not sample) standard deviation (SD) between subjects within each of the four cells above: 200
D. Analysis of variance with treatment, UDCA (yes/no), treatment with UDCA (all as classification variables), and Baseline value (as a continuous covariate) in the model
E. 20% dropout rate, from randomization to Week 12, yielding 16 completing subjects per treatment group
F. Dunnett's 1-step adjustment for multiplicity, requiring a type 1 error rate of 0.027 (2-sided) for each comparison versus placebo

Assumptions underlying the means and SDs are based on studies evaluating UDCA as a first-line therapy and in studies where secondary or combination therapies with UDCA have been utilized. The power in the primary efficacy analysis may be slightly higher, since it incorporates Dunnett's step-down adjustment and utilizes repeated measures to reduce the power losses due to dropouts. On the other hand, power may decrease slightly, to the degree that unequal numbers of subjects are randomized taking and not taking UDCA. In addition, the probability that the observed least square (LS) mean treatment difference exceeds 30% of the placebo LS mean change has not been examined.

To randomize 60 subjects, approximately 90 will be screened.

If, in contrast, the population SD is 300, the power is roughly 81% for each NGM282 comparison versus placebo.

3.2. Randomization, Stratification, and Blinding

At Day 1, all eligible subjects will be randomized [redacted] to one of the three treatment arms (NGM282 1 mg, NGM282 3 mg, or placebo) in a 1:1:1 ratio. The randomization will be stratified by ursodeoxycholic acid (UDCA) use (yes/no), to ensure an even distribution across the three treatment groups. The determination of the status will be based on the medical history and the concomitant medications history at randomization.

The unmasking of a subject's treatment will be available to the Principal Investigator (PI) in the event of a medical emergency or an AE that necessitated identification of the study drug for the welfare of that subject. Except for the medical emergency, the PI and clinic staff will remain blinded during the conduct of the study and until all discrepancies in the clinical database are resolved and the database is locked. The date and time when the PI removed the study blind for an individual subject will be documented [redacted] and an automated notification will be sent to the Sponsor. However, all attempts should be made to contact the Sponsor prior to unblinding. The contracted CRO's pharmacovigilance team may also be required to break the blind in the case of a medical emergency.

3.3. Analysis Sets

3.3.1. Randomized Set

All randomized subjects will be included into the Randomized Set. All subjects in the Randomized Set will be grouped according to the randomized treatment even if this differs from the actual treatment received.

3.3.2. Safety Set

All randomized subjects who receive at least one dose (full or partial) of study drug and have at least one post-dose safety evaluation will be included in the Safety Set. In analyses using this set, subjects will be grouped according to actual treatment received if this differs from the randomized treatment.

3.3.3. 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 parameter value will be included in the Full Analysis Set (FAS). This will be the set for the primary analyses of efficacy and PD endpoints. In FAS analyses, subjects will be grouped according to randomized treatment if this differs from actual treatment received.

4. Subject Disposition

4.1. Disposition

Subject disposition will be summarized for the randomized Set. A summary of the analysis sets includes the number and percentage of subjects for the following categories: subjects randomized, subjects who completed treatment, subjects who completed the study, subjects who discontinued

from the study, subjects in Safety Set, subjects in the FAS Set, subjects in the PP Set, and subjects in the pharmacokinetic Set. All percentages will be based on the number of subjects randomized.

Primary reason for treatment discontinuation and primary reason for study discontinuation will be summarized in the disposition table. Reasons for study discontinuation include the following: adverse event, death, lost to follow-up, trial screen failure, withdrawal of consent by subject, subject refused to return for follow-up visit, other.

A listing to subject disposition data will be presented for each subject, including enrollment date, treatment start date, treatment end date, treatment discontinuation reason and study discontinuation reason.

4.2. Protocol Deviations

The protocol deviations/violations will be identified and assessed by clinical research physician or designee following institution standard operational procedures. All protocol deviations will be summarized using the FAS. Protocol deviations will be presented in a listing.

5. Demographics and Baseline Characteristics

5.1. Demographics

A summary of demographic and baseline information will be presented for the Randomized Set, FAS and Safety Set. The demographic characteristics consist of age (years), sex, race, and ethnicity. The Baseline characteristics consist of Baseline height (cm), Baseline weight (kg), and Baseline body mass index (BMI) (kg/m^2). Body mass index is calculated as (body weight in kilograms) / (height in meters)^2.

A subject's age in years is calculated using the informed consent date and date of birth. Age (years), Baseline height (cm), Baseline weight (kg), and Baseline BMI (kg/m^2) will be summarized using descriptive statistics. The number and percentage of subjects by sex (Male, Female), race (White, Black or African American, Asian, American Indian or Alaska Native, Native Hawaiian or Other Pacific Islander, and Other), and ethnicity (Hispanic or Latino, Not Hispanic or Latino) will also be reported. Percentages will be based on the total number of subjects in the Randomized Set, FAS or Safety Set, as applicable.

Subject demographic and Baseline characteristics will be presented in a listing.

5.2. Baseline Disease Characteristics

A summary of Baseline disease information will be presented for Randomized, FAS and Safety Set. The Baseline disease characteristics consist of UDCA use and duration of PSC (years).

Duration of PSC in years is calculated as (the informed consent date - date of diagnosis of PSC + 1)/365.25. It will be analyzed using descriptive statistics. If the date of diagnosis of PSC is completely or partially missing, the following imputation rules should be utilized in the analysis:

- If year, month, and day are missing then use the earlier of the patient's first visit date or the consent date.

- If either month or month and day are missing, use January 1.
- If only day is missing, impute the first day of the month.

UDCA use (yes/no) will be summarized by number and percentage of subjects. Percentage will be based on the total number of subjects in the Randomized, FAS or Safety Set, as applicable.

Subject Baseline disease characteristics will be listed.

5.3. Medical History

Medical history will be coded using the Medical Dictionary for Regulatory Activities (MedDRA, version 18.1 or higher). The number and percentage of subjects with any medical history will be summarized overall and for each system organ class and preferred term. Percentages will be calculated based on number of subjects in the Randomized, FAS or Safety Set, as applicable.

Subject medical history data including specific details will be presented in a listing.

5.4. Inclusion and Exclusion Criteria

Prior to randomization, the investigator will assess if the subject fulfills all of the inclusion and exclusion criteria outlined in the protocol (sections 5.1 and 5.2). The specific inclusion criterion not met or exclusion criterion which was met will be recorded in the eCRF. This information will be presented in a listing.

6. Treatments and Medications

6.1. Prior and Concomitant Medications

All medications used within 4 weeks prior to the date of screening and during the study period, as well as the reason for use, will be collected on the source documents and the CRF. Non-study medications will be classified as prior and concomitant, and coded using the current version of the World Health Organization Drug Dictionary Enhanced (WHO-DD).

Prior medications are defined as those taken at least once during the 28 days before Screening. Concomitant medications are defined as those taken at least once from 28 days prior to Screening through the EOS visit. Medications stopped in the 28 days prior to the Screening time point will not be considered concomitant.

The total number of concomitant medications and the number and percentages of subjects with at least one concomitant medication will be summarized by treatment group. The number and percentages of all concomitant medications will be summarized by treatment group, Anatomical Therapeutic Chemical (ATC) class and preferred term. Subjects who take the same medication (in terms of PT) more than once will be counted only once for that medication. All summaries will be performed using the Safety Set.

Prior medications and concomitant medications will be listed.

6.2. Study Treatments

NGM282 is provided as a sterile solution for injection in a single-use, pre-filled syringe for SC administration at doses of 1 mg and 3 mg, and volume-matched placebo. Subjects will be instructed to self-administer/dose at approximately the same time each morning over the 12-weeks treatment period. A daily study-drug administration diary is required for each subject.

Data on study-drug administration will be presented in a listing.

6.2.1. Extent of Exposure

Duration of exposure is defined as the total number of days a subject is exposed to any study drug and will be presented as the total number of days from the first dose date (Day 1) to the last dose date (date of last dose minus the date of first dose + 1) as recorded on the End of Study/Early Withdrawal page on the CRF. If the last dose date on the End of Study/Early Withdrawal page is missing, or if a subject is lost to follow-up, but the drug accountability log confirms that the subject has taken study drug, the visit date following the last completed drug accountability log will be used.

The duration of exposure to study drug by treatment will be summarized for all subjects in the Safety Set and will be presented in a table by summary statistics. The duration of exposure will then be classified into one of the following categories: <= 7 days, 8 to 14 days, 14 to 28 days, 28 to 56 days, 56 to 84 days, and >= 84 days and will be presented as the number and percentage of subjects in each duration category. Percentages will be computed from the number of subjects in the Safety Set.

A summary of each subject's exposure will be presented in a listing.

6.2.2. Treatment Compliance and Modifications

Subjects are expected to take all doses of study drug as instructed. At each visit, the number of syringes taken will be based on the drug accountability form and calculated by subtracting the number of syringes returned on the visit from the number of syringes dispensed in the previous visit.

The overall study drug compliance (%) will be calculated by dividing the total number of syringes taken across all visits by the total number of syringes intended across all visits and then multiplying by 100.

Compliance (%) = [(total no. of syringes dispensed – total no. of syringes returned) / (The actual duration of treatment (Days) *No. of syringes prescribed per day)] * 100

The actual duration of treatment (Days) will be calculated as
The actual duration of treatment (Days) = the last treatment date – the first treatment date

The overall study drug compliance will be summarized by treatment group.

A subject is considered compliant if overall study drug compliance is greater than or equal to 80% and less than or equal to 120%. A categorical summary of whether subjects were compliant (yes/no) will be presented.

Summary statistics on percentage of treatment compliance as well as the number and percentage of subjects in each compliance category (<80%, 80-120%, and >120% compliant) will be presented overall. Percentages will be calculated out of the number of subjects in the Full analysis Set.

7. Efficacy Analysis

All efficacy endpoints will be analyzed using FAS.

7.1. Primary Efficacy Endpoint

The primary efficacy endpoint is the mean change in ALP from Baseline at Week 12.

7.1.1. Primary Analysis

The primary analysis will be performed on the FAS to evaluate the treatment effect of NGM282. The mean change in ALP from Baseline at Week 12 will be compared between each of the two treatment groups (NGM282 1 mg or 3 mg, as applicable) and the placebo group.

Mixed-effect model repeated measures (MMRM) analysis of covariance (ANCOVA) of mean change in ALP from Baseline at Week 12 will be used to compare the difference between treatment groups. Mathematically stated:

H_0: Mean change in ALP from Baseline at Week 12 $_{\text{NGM282 1 mg}}$ = Mean change in ALP from Baseline at Week 12 $_{\text{placebo}}$,

and

Mean change in ALP from Baseline at Week 12 $_{\text{NGM282 3 mg}}$ = Mean change in ALP from Baseline at Week 12 $_{\text{placebo}}$

H_1: Mean change in ALP from Baseline at Week 12 $_{\text{NGM282 1 mg}} \neq$ Mean change in ALP from Baseline at Week 12 $_{\text{placebo}}$,

or

Mean change in ALP from Baseline at Week 12 $_{\text{NGM282 3 mg}} \neq$ Mean in ALP change from Baseline at Week 12 $_{\text{placebo}}$

Mean change in ALP from Baseline at Week 12 will be evaluated using the MMRM analysis adjusting for treatment group, visit, treatment group by visit interaction, UDCA use, treatment group by UDCA use interaction as classification variables and the Baseline value as covariate. For each treatment and treatment comparison versus placebo, the least squares mean, associate standard error (SE), 95% confidence interval and corresponding p-value will be present. The overall type I error rate for the two primary comparisons versus placebo will be controlled using the step-down Dunnett multiple-testing procedure [Dmitrienko and D'Agostino (2013)].

With the same model, linear combinations $\mu_1 - \mu_3$ and $(\mu_1+\mu_3)/2 - \mu_p$ will be estimated, where μ_1, μ_3, and μ_p are the population LS means for the primary efficacy endpoint of the 1 mg NGM282, 3 mg NGM282, and placebo groups, respectively.

The relevance of the explanatory variables to the mean change in ALP from Baseline at Week 12 will be checked. All relevant explanatory variables will be included in the primary analysis model with FAS.

7.2. Secondary Efficacy Endpoint

The following endpoints will be analyzed for the FAS.

The endpoints among the following that are based on assessments scheduled for more than 1 post-treatment time point, will be analyzed using the FAS. They will be evaluated using MMRM analyses similar to that of the primary endpoint analysis. In contrast, the endpoints among the following based on assessments scheduled at Week 12 only will be evaluated using an ANCOVA model including treatment group, UDCA use, and treatment by UDCA use interaction as classification variables, and the Baseline value as the continuous covariate.

- Percent change from Baseline in ALP at Week 12
- Changes and percent changes from Baseline at Week 12 in the following:
    - ALT, AST, bilirubin (total, direct), and GGT
    - Total cholesterol, HDL cholesterol, LDL cholesterol, and triglycerides
    - Bile-mediated absorption as measured by fat-soluble vitamins

For each NGM282 treatment group, the LS mean rate (slope, per week) of change in ALP during weeks 1-4 will be compared to that during weeks 5-12. Changes in the slopes will be estimated with a MMRM from primary analysis. The same analysis will be done for percent change in ALP.

The relevance of the explanatory variables to each secondary efficacy endpoint will be checked. Similar to the model in the primary efficacy analysis, all relevant explanatory variables will be included into the analysis model with FAS.

The Mayo Partial IBD Score is a non-invasive 9-point partial Mayo score used as an outcome measure for clinical trials assessing therapy for ulcerative colitis. The Mayo Partial IBD Score will be assessed at all study visits. It is administered pre-dose at Day 1 and Weeks 1, 2, 4, 8, 12 and 16 (End of study). Severity of IBD-associate symptoms will be summarized with frequency and percentage.

8. Safety and Tolerability Analysis

Safety and tolerability will be assessed by clinical review and summaries of the following parameters:

1. AEs
2. Clinical laboratory test results
3. Physical examination results
4. LISSAs
5. Electrocardiogram results

6. Pregnancy Test
7. Vital signs

All safety and tolerability analyses will be conducted using the Safety Set. Percentage for the safety and tolerability analyses based on the number of subjects with a given event or attribute will be based on the number of subjects who were treated.

8.1. Adverse Events

All AE summaries will be restricted to treatment-emergent AEs (TEAEs) only. A TEAE is defined as an AE that meets any of the following conditions:

- begins on or after the first dose of study drug and before the stop of study drug +30 days;
- is completely missing an onset date and end date;
- is completely missing an onset date and the end date is on or after the first dose of study drug.

For the purpose of inclusion in TEAE tables, incomplete AE onset and end dates will be imputed as follows:

All adverse events will be classified by System Organ Class (SOC). If the SOC is 'Gastrointestinal disorders', events will be further classified using the Lowest Level Term (LLT). For other SOC's, events will be further classified using the Preferred Term (PT). Adverse event coding is done according to the Medical Dictionary for Regulatory Activities (MedDRA, Version 18.1 or higher).

An overview summary of the number and percentage of subjects with any TEAE, serious TEAE, study drug-related TEAE, study drug-related serious TEAE, TEAE leading to treatment discontinuation will be provided by treatment group and total NGM282 group. TEAE leading to study termination and AE leading to death will be presented in a listing.

8.1.1. Incidence of Adverse Events

Summaries of the total number of TEAEs and the number of subjects with at least one TEAE will be provided by treatment group and total group. The number of subjects and the number of events will also be presented by SOC and PT/LLT. At each level of subject summarization, a subject is counted once if the subject reported one or more events. Percentage will be calculated out of the number of subjects in the Safety Set.

A summary of TEAEs will also be presented in descending order from the SOC with the highest total incidence (that is, summed across all treatment groups) to the SOC with the lowest total incidence. If the total incidence for any two or more SOCs is equal, the SOCs will be presented in alphabetical order. Within each SOC, the PTs/LLTs will be presented in alphabetical order.

All AEs will be presented in a listing.

8.1.2. Relationship of Adverse Events to Study Drug

A summary of TEAEs by relationship to study drug will be presented in a table. The investigator will provide an assessment of the relationship of the event to the study drug. The possible relationships are "Not Related", "Possible", "Probable", and "Definite". In the TEAE relationship table, if a subject reports multiple occurrences of the same TEAE, only the most closely related occurrence will be presented. Treatment-emergent AEs that are missing a relationship will be presented in the summary table as "Definite" but will be presented in the data listing with a missing relationship. Percentages will be calculated based on the number of subjects in the Safety Set.

The TEAE data will be categorized and presented by SOC, PT/LLT, and relationship in a manner similar to that described in Section 8.1.1.

8.1.3. Severity of Adverse Event

A summary of TEAEs by severity will be presented in a table. The severity that will be presented represents the most extreme severity captured on the CRF page. The possible severities are "Mild", "Moderate", "and "Severe". In the TEAE severity table, if a subject reported multiple occurrences of the same TEAE, only the most severe will be presented. Treatment-emegent AEs that are missing severity will be presented in the table as "Severe" but will be present in the data listing with a missing severity. Percentages will be calculated out of the number of subjects in the Safety Set.

The TEAE data will be categorized and presented by SOC, PT/LLT, and severity in a manner similar to that described in Section 8.1.1.

AE's severity will be rated using CTCAE 4.03. Each CTCAE 4.03 term is a Medical Dictionary for Regulatory Activities (MedDRA) Lowest Level Term (LLT). The CTCAE displays Grades

1-5 with unique clinical descriptions of severity of each AE. TEAE will be summarized by maximum CTCAE grade. Percentages will be calculated out of the number of subjects in the Safety Set.

The TEAE data will be categorized and presented by SOC, PT/LLT, and CTCAE grade in a manner similar to that described in Section 8.1.1.

8.1.4. Serious Adverse Events

The seriousness of an AE should be assessed by the Investigator independently from the severity of the AE. A serious AE (SAE) is defined as any untoward medical occurrence that at any dose results in death, is life-threatening, is a congenital anomaly/birth defect, requires in-patient hospitalization or prolongation, or results in significant disability/incapacity.

Important medical events that may not result in death, be life-threatening, or require hospitalization may be considered SAEs when, based upon medical judgment, they may jeopardize the subject and may require medical or surgical intervention to prevent one of the outcomes listed above.

Serious treatment-emergent adverse events (SAEs) will be presented in a table. Treatment-emergent SAEs by relationship to study drug will be presented in a table. A treatment-related treatment-emergent SAE is a treatment-emergent SAE with any relation to study drug other than "Not Related". Treatment-emergent SAEs that are missing a relationship will be presented in the table as "Related" but will be presented in the data listing with a missing relationship. At each level of subject summarization, a subject is counted once if the subject reported one or more events. Percentages will be calculated out of the number of subjects in the Safety Set.

The treatment-emergent SAE data will be categorized and presented by SOC and PT/LLT in a manner similar to that described in Section 8.1.1 and by severity as described in Section 8.1.3.

8.1.5. Adverse Events Leading to Treatment Discontinuation

A summary of TEAEs with a study drug action taken of "Drug Withdrawal" will be presented in a table. At each level of subject summarization, a subject is counted once if the subject reported one or more events. Percentages will be calculated out of the number of subjects in the Safety Set within the subgroup category.

The summary of TEAEs with a study drug action taken of "Drug Withdrawal" will also be presented in descending order of frequency from the SOC with the highest total incidence (that is, summed across all treatment groups) to the SOC with the lowest total incidence. If the total incidence for any two or more SOCs is equal, the SOCs will be presented in alphabetical order. Within each SOC, the PTs/LLTs will be presented in alphabetical order.

8.1.6. Adverse Events Leading to Study Discontinuation

All subjects who have an AE with the answer to "Caused Study Discontinuation" is "Yes" will be presented in a listing.

8.1.7. Death

All subjects who have an AE with an outcome of "Death Related to Adverse Event" will be presented in a listing.

8.2. Clinical Laboratory Evaluations

Observed values and absolute change from Baseline for routine chemistry, hematology, and urinalysis values will be summarized by treatment group and overall at each per-protocol time point for subjects in the Safety Set.

All relevant clinical laboratory tests in chemistry and hematology will be classified as Low, Normal, and High, or Normal/Abnormal according to the normal ranges. This categorical data will be summarized in shift tables comparing the extreme results at each scheduled post-baseline visit with those at the Baseline visit. Extreme post-baseline results will also be summarized. When there are multiple values within a visit for a particular laboratory variable, the worst value will be taken (worst being the smallest value for criteria below a certain threshold or the largest value for criteria above a certain threshold). If a subject has a value below the threshold and above the threshold, the value furthest from the threshold will be chosen.

In data listings, laboratory values will be compared to normal ranges; out-of-range and clinically significant laboratory values will be identified.

8.3. Vital Sign Measurements

Observed values and changes from Baseline for systolic blood pressure (mmHg), diastolic blood pressure (mmHg), temperature (°C), pulse (bpm), and respiratory rate (breaths/minute), will be summarized by treatment group and overall at each per-protocol time point for subjects in the Safety Set.

All vital sign measurement data by subject will be presented in a listing.

8.4. Physical Examination

Physical examination will be performed at Screening, Day 1, and Weeks 2, 4, 8, 12 and 16. All findings during the physical examination for all subjects will be presented in a listing.

8.5. Electrocardiogram

12-lead ECGs will be performed at Screening, Day 1, and weeks 12 and 16. ECG results are interpreted as Normal, Abnormal not clinically significant, or Abnormal clinically significant. This categorical data will be summarized in shift tables comparing the interpretations at each scheduled post-baseline visit with those at the Baseline visit.

Observed results at each visit will be presented in a listing.

8.6. Pregnancy Test

Urine and serum pregnancy test data will be listed for each female subject.

8.7. Local Injection-Site Symptom Assessments

The injection-site evaluation will be made using LISSA at Day 1, Weeks 1, 2, 4, 8, 12 and 16 (End of study). LISSA scores will be documented on the subject's eCRF.

A summary to the evaluated ISR will be presented with number and percentage of subjects at each time point by treatment group using Safety Set.

[redacted]

10. Data Monitoring Committee

A Data Monitoring Committee (DMC) will be established in order to protect subject welfare, preserve study integrity, and provide recommendations as needed regarding study conduct. The DMC will be comprised of two external liver disease clinical experts as well as a statistician and a medical team member from the sponsor. The DMC will meet at predefined time points and on an as-needed basis based on enrollment, treatment milestones, and safety. Formal minutes and recommendations will provided by the DMC to the sponsor regarding additional data requests and the continuation of the conduct of the study as outlined in the current protocol. The DMC will operate under the guidance of an agreed-upon charter.

The analyses and outputs provided to the DMC are a subset of the IA and Full Study SAP.

11. Changes in the Planned Analysis

The descriptive summarization and MMRM ANCOVA model analyses for Change from Baseline in (5-D Pruritus Scale and Pruritus and Fatigue) were removed from the current SAP. At this time only the listings to display the data are planned.

The analyses for acute cholangitis were removed from the current SAP. At this time only the listings to display the data are planned.

The C4 and Serum bile acids, Fecal microbiome composition and PK data will not be transferred to [redacted] the Contract Research Organization performing the analyses of this SAP. Therefore, the analyses of these data were removed from this SAP.

12. References

Chuck Kincaid (2005). Guidelines for Selecting the Covariance Structure in Mixed Model Analysis, Proceeding of the 30th SAS Users Group International Conference.

Dmitrienko A, D'Agostino R, Sr. (2013). Traditional multiplicity adjustment methods in clinical trials. Stat Med;32:5172–218.

Hollander, Myles, Douglas A. Wolfe. (1973),Nonparametric Statistical Methods, New York: John Wiley & Sons, 75-82.

Liu, H. (2007) Cochran-Armitage Trend Test Using SAS, PharmaSUG 2007, SP05.

13. Appendices

13.1. Schedule of Study Procedures

The visits should occur as close to the intended dates as possible. However, there is an acceptable ± 3-day window for individual scheduled visits. Subjects attending any visits out of windows from Day 1 to Week 12/Early Withdrawal (EOT) Visit should be brought back into compliance with the overall study-visit schedule as soon as possible thereafter. Subjects will then return to the clinic at Week 16 (or 4 weeks after last dose) for an EOS follow-up visit.

Study Procedure	Days –42 to –1 (Screening) [a]	Day 1	Wk 1	Wk 2	Wk 4	Wk 8	Wk 12 (EOT)/ Early Withdrawal	Wk 16 (EOS)/ Follow-up
Informed consent	X							
Demographics	X							
Medical history	X							
Inclusion/exclusion criteria	X	X						
Height	X							
Body weight	X	X			X	X	X	X
Physical exam	X	X		X	X	X	X	X
12-lead electrocardiogram	X	X					X	X
Vital signs	X	X	X	X	X	X	X	X
Colonoscopy [b]	X							
MRCP [c]	X							
Prior and concomitant medications	X	X	X	X	X	X	X	X
Randomization		X						
Study drug self-administration training		X	X	X	X			
Dispense study drug		X	X	X	X	X		
Study-drug compliance			X	X	X	X	X	
Adverse event evaluations		X	X	X	X	X	X	X
LISSA evaluations [d]		X	X	X	X	X	X	X
PK blood samples [e]		X	X	X	X	X	X	X
NRS Itch/Fatigue [f]			X		X		X	
5-D Pruritus		X	X	X	X	X	X	X
Mayo partial IBD Score[g]	X	X	X	X	X	X	X	X
Chemistry	X	X	X	X	X	X	X	X
Hematology	X	X	X	X	X	X	X	X
Urinalysis	X	X			X		X	X
Lipid panel		X					X	
TSH	X							
CA19-9[h]	X							
p-ANCA	X							
ELF panel		X					X	

Hepatitis and HIV screen	X							
Study Procedure	Days –42 to –1 (Screening) [a]	Day 1	Wk 1	Wk 2	Wk 4	Wk 8	Wk 12 (EOT)/ Early Withdrawal	Wk 16 (EOS)/ Follow-up
Urine drug screen	X							
Pregnancy test [i]	X	X					X	X
C4 and serum bile acids		X					X	
Vitamin D		X					X	X
INR	X	X					X	X
Stool sample for microbiome, fecal fat, and calprotectin		X					X	
Anti-drug antibodies		X	X	X	X	X	X	X
Neutralizing antibodies		X	X	X	X	X	X	X
Exploratory biomarkers		X	X	X	X	X	X	X

C4 = 7-alpha-hydroxy-4-cholesten-3-one; CA19-9 = carbohydrate antigen 19-9; ELF = enhanced liver fibrosis; EOS = End of Study; EOT = End of Treatment; ERCP = endoscopic retrograde cholangiopancreatography; HIV = human immunodeficiency virus; hs-CRP = high-sensitivity C-reactive protein; IBD = inflammatory bowel disease; INR = International Normalized Ratio; LISSA = local injection-site symptom assessment; MRCP = magnetic resonance cholangiopancreatography; NRS = Numeric Rating Scale; p-ANCA = perinuclear anti-neutrophil cytoplasmic antibodies; PK = pharmacokinetic; TSH = thyroid-stimulating hormone; Wk = week.

[a] There must be a minimum of 14 days between Screening and Day 1 visits for adequate separation of the repeated Chemistry (liver function tests) and INR assessments.
[b] Colonoscopy will be performed only in subjects who do not have a colonoscopy available within 12 months of Screening.
[c] MRCP will be performed in all subjects at Screening.
[d] LISSA evaluations will be performed pre-dose for all on-treatment visits.
[e] PK blood samples will be collected before subjects dose themselves in the clinic (pre-dose). At Day 1 and Week 12, an additional PK blood sample will be collected 2 hours post-dose.
[f] NRS Itch/Fatigue will be completed as a daily diary during Study Weeks 1, 4, and 12. Diary will be dispensed at the study visit prior to the collection week.
[g] Mayo Partial IBD Score will be collected in all subjects, independent of whether they have IBD or not.
[h] A second CA19-9 sample may be collected during the Screening period a minimum of 4 weeks from the original Screening sample in order to meet Inclusion Criteria 3g.
[i] A serum pregnancy test will be performed on all female subjects at Screening, Week 12, and Week 16. A urine pregnancy test will be performed on all female subjects at Day 1 (pre-dose).