STATISTICAL ANALYSIS PLAN

ALK8700-A302 / NCT03093324

Study Title:	A Phase 3 Study in Subjects with Relapsing Remitting Multiple Sclerosis to Evaluate the Tolerability of ALKS 8700 and Dimethyl Fumarate
Document Status:	Final Version 1.0
Document Date:	February 20, 2019
Based on:	Alkermes ALK8700-A302 study protocol Germany-specific Amendment 2.2 (incorporated Administrative Change Memo 2): 14 Dec 2018 Amendment 2.0 (incorporated Administrative Change Memo 1): 20 Sep 2018 v 2.1 (incorporates Poland-specific Amendment 1.1): 23 Aug 2017 Amendment 1.0: 15 Dec 2016
Sponsor:	Alkermes, Inc. 852 Winter Street Waltham, MA 02451 USA

TABLE OF CONTENTS

ABBREVIATIONS

The following abbreviations are used in the statistical analysis plan.

Abbreviation or Term	Explanation or Definition
AE	Adverse Event
AESI	Adverse Events of Special Interest
ALK-P	Alkaline Phosphatase
ALT	Alanine Aminotransferase
ANCOVA	Analysis Of Covariance
AST	Aspartate Aminotransferase
ATC	Anatomical Therapeutic Chemical [Classification System]
BID	Bis In Die [Twice A Day]
BMI	Body Mass Index
BUN	Blood Urea Nitrogen
CI	Confidence Interval
C-SSRS	Columbia Suicide Severity Rating Scale
DMC	Data Monitoring Committee
DMF	Dimethyl fumarate (Tecfidera®)
DMT	Disease Modifying Treatment
ECG	Electrocardiogram
eCRF	Electronic Case Report Form
[redacted]	[redacted]
[redacted]	[redacted]
[redacted]	[redacted]
ET	Early Termination
FAS	Full Analysis Set
FS	Functional System
GdE	Gadolinium-Enhancing
GGISIS	Global Gastrointestinal Symptom And Impact Scale
GI	Gastrointestinal
HDL	High Density Lipoprotein
IGISIS	Individual Gastrointestinal Symptom And Impact Scale
LDH	Lactic Dehydrogenase

Abbreviation or Term	Explanation or Definition
LDL	Low Density Lipoprotein
LLN	Lower Limit Of Normal Range
MedDRA	Medical Dictionary For Regulatory Activities
MMF	Monomethyl Fumarate
MRI	Magnetic Resonance Imaging
MS	Multiple Sclerosis
NCI-CTCAE	National Cancer Institute Common Terminology Criteria For Adverse Events
PCS	Potentially Clinically Significant
PDEAE	Post-discontinuation Emergent Adverse Event
[redacted]	[redacted]
PT	Preferred Term
QOL	Quality Of Life
QTcB	QT Interval Corrected Using Bazett's Formula
QTcF	QT Interval Corrected Using Fridericia's Formula
RRMS	Relapsing-Remitting Multiple Sclerosis
SAE	Serious Adverse Event
SD	Standard Deviation
[redacted]	[redacted]
[redacted]	[redacted]
[redacted]	[redacted]
SOC	System Organ Class
[redacted]	[redacted]
TEAE	Treatment-Emergent Adverse Event
ULN	Upper Limit Of Normal Range
WBC	White Blood Cells
WHO	World Health Organization

1. INTRODUCTION

This statistical analysis plan (SAP) describes the statistical methods and data presentations to be used for analyzing and reporting efficacy and safety data for study ALK8700-A302. This study evaluating the GI tolerability of ALKS 8700 and DMF was composed of two parts, Parts A and B. Both Parts A and B were identical in study design and included a 5-week, double-blind treatment period with 2 blinded treatment arms (ALKS 8700 and DMF). Part A was exploratory. Part A evaluated the utility of 2 gastrointestinal (GI) symptom scales (Individual GI Symptom and Impact Scale [IGISIS] and Global GI Symptom and Impact Scale [GGISIS]) and endpoints derived from the scales in assessing GI tolerability in adult subjects with relapsing-remitting multiple sclerosis (RRMS) after administration of ALKS 8700 or dimethyl fumarate (DMF). Part A also evaluated the safety and tolerability of ALKS 8700 in adult subjects with RRMS.

A total of 120 subjects were randomized to treatment in Part A of the study. A separate SAP and Clinical Study Report (CSR) for Part A study were finalized, while Part B of the study was ongoing as planned.

The primary, secondary and exploratory endpoints for the study were revised based on exploratory analyses of the results from Part A. Major changes to the protocol included:

- Change in primary, secondary, and exploratory endpoints
- Change in the analysis of primary, secondary, and exploratory endpoints
- Change in the analysis of the safety data
- Increase in number of subjects in Part B and for the study

This document has been prepared based on Protocol ALK8700-A302 v2.0 (Amendment 1.0) (dated 15 Dec 2016) [1], v 2.1 (incorporates Polish-specific Amendment 1.1) (dated 23 Aug 2017) [2], Amendment 2.0 (incorporated Administrative Change Memo 1) (dated 20 Sep 2018) [3] and Germany-specific Amendment 2.2 (incorporated Administrative Change Memo 2) (dated 14 Dec 2018) [4].

1.1. Study Objectives

The objectives of this study are to:

- Evaluate the utility of two gastrointestinal (GI) symptom scales (Individual GI Symptom and Intensity Scale [IGISIS] and Global GI Symptom and Intensity Scale [GGISIS]) and endpoints derived from the scales in assessing GI tolerability in adult subjects with relapsing remitting multiple sclerosis (RRMS) after administration of ALKS 8700 or dimethyl fumarate (DMF) in Part A
- Compare the GI tolerability of ALKS 8700 and DMF in adult subjects with RRMS using two GI symptom scales (IGISIS and GGISIS) with endpoints informed from Part A
- Evaluate the safety and tolerability of ALKS 8700 in adult subjects with RRMS in Parts A and B

Part A CSR presented the results of evaluation of utility of two GI Symptoms scales, IGIGIS and GGISIS and endpoints derived from these scales in assessing GI tolerability in adult subjects with relapsing remitting multiple sclerosis (RRMS) after administration of ALKS 8700 or dimethyl fumarate (DMF) in Part A.

In this study, GI tolerability of ALKS 8700 and DMF in adult subjects with RRMS will be compared with revised endpoints and safety and tolerability of ALKS 8700 will be evaluated based on pooled data from Part A and B.

1.2. Summary of the Study Design

This is a randomized, double-blind, multicenter study to evaluate the tolerability of ALKS 8700 and DMF in adult subjects with RRMS. The study was conducted in two parts (Parts A and B). The study design is an adaptive approach where data from Part A was used to modify the GI tolerability endpoint(s) in the study.

Both Parts A and B are identical in study design and include a 5-week, double-blind treatment period with two blinded treatment arms (ALKS 8700 and DMF). Part A is exploratory. The first 120 eligible subjects will be randomized to one of the two treatment groups in Part A (n=60 per group). Once the randomization for Part A is complete, the next 380 eligible subjects will be randomized into one of the two treatment groups in Part B (n=190 per group). Accordingly, subjects 1-120 will be randomized into Part A and subjects 121-500 will be randomized into Part B. Subjects randomized in Part A will not be eligible to participate in Part B.

Following completion of Part A, the sponsor conducted a planned, unblinded analysis of the Part A GI tolerability and safety data. Part B will be ongoing during the analysis of the Part A data. Data from Part A can be used to modify the Part B GI tolerability endpoint(s). Pooled GI tolerability data from Part A and Part B will be analyzed for the primary and secondary endpoints.

All study visits in Parts A and B will be completed on an outpatient basis. The screening period is up to 4 weeks (28 days) including a 1-week lead-in period (prior to randomization) during which subjects will complete two self-administered GI symptom scales daily using e-diaries.

In Part A and Part B, subjects meeting the eligibility criteria will be randomized in a 1:1 ratio to one of two blinded treatment groups as follows:

- Group 1: ALKS 8700 462 mg BID (with 1-week titration)
- Group 2: DMF 240 mg BID (with 1-week titration)

In each part, both ALKS 8700 and DMF will be administered orally BID as blinded study drug; either ALKS 8700 462 mg BID or DMF 240 mg BID with an initial 1-week dose titration period. Subjects will be instructed to take study drug with or without food. However, subjects will be instructed to avoid taking study drug with a high-fat, high-calorie meal. Dosing is described in detail in Protocol Section 9.1.

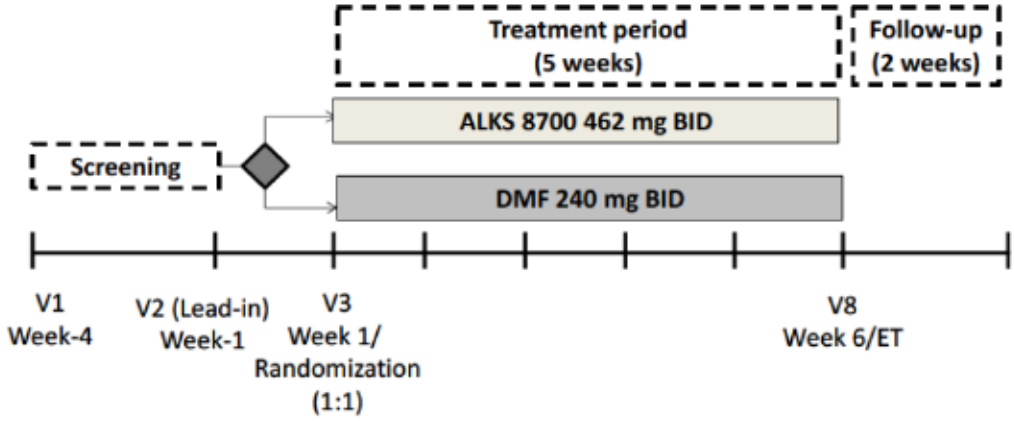
The study duration for each part is approximately 11 weeks, which includes up to 4 weeks for screening (including a 1-week lead-in period), a 5-week double-blind treatment period (with 1-week titration), and a 2-week follow-up period (for subjects not continuing into the ALK8700-A301 long-term safety study). During the 5-week treatment period (Visits 3-8),

subjects will return to the clinic for weekly scheduled visits for assessments of safety, tolerability, pharmacokinetics (in a subset of subjects in Part A at selected sites), and clinical status (see Figure 1). Any subject who prematurely discontinues from the study will return to the clinic for an ET visit (Visit 8) and a follow-up visit (Visit 9).

Subjects completing the entire 5 weeks of double-blind treatment in either Part A or Part B of the study will be eligible to enroll in a long-term, open-label safety study (Study ALK8700-A301) and receive ALKS 8700 for up to 96 weeks. Subjects not enrolling in the long-term safety study will return to the clinic 2 weeks following Visit 8 for a follow-up visit (Visit 9).

A schematic of the study design is provided in Figure 1 below.

Figure 1: Study Design Schematic (Part A and Part B)



*Subjects completing the 5-week treatment period will either continue into the ALK8700-A301 long-term safety study or enter the safety follow-up period and return to Visit 9

[redacted]

Table 1: Schedule of Visits and Assessments

	Screening		5-Week Randomized Treatment[1]						Follow-up
Visit	**1**[2]	**2**[3]	**3**	**4**	**5**	**6**	**7**	**8 (ET)**	**9**
Week	**-4**	**-1 (±2 days)**	**1 (±3 days)**	**2 (±3 days)**	**3 (±3 days)**	**4 (±3 days)**	**5 (±3 days)**	**6 (±3 days)**	**8 (±4 days)**
Informed Consent	X								
Eligibility Criteria Review	X		X[4]						
Demographics	X								
Medical History	X								
Height	X								
Weight	X		X[4]					X	
Physical Exam[5]	X		X[4]					X	X
Concomitant Medication Review	X	X	X	X	X	X	X	X	X
Adverse Event Monitoring	X	X	X	X	X	X	X	X	X
Serology Testing[6]	X								
Pregnancy Test[7]	X		X[4]						
Vital Signs[8]	X		X[4]	X	X	X	X	X	X
Standard 12-Lead ECG[9]	X		X[4]		X			X	X
[redacted]									
Biochemistry, Urinalysis, & Hematology Samples[11]	X		X[4]		X			X	X

Table 1: Schedule of Visits and Assessments (Continued)

	Screening		5-Week Randomized Treatment[1]						Follow-up
Visit	**1**[2]	**2**[3]	**3**	**4**	**5**	**6**	**7**	**8 (ET)**	**9**
Week	**-4**	**-1 (±2 days)**	**1 (±3 days)**	**2 (±3 days)**	**3 (±3 days)**	**4 (±3 days)**	**5 (±3 days)**	**6 (±3 days)**	**8 (±4 days)**
Brain MRI	X[12]								
IGISIS		X[13]	X[14]	X[14]	X[14]	X[14]	X[14]	X[14]	
GGISIS		X[13]	X[14]	X[14]	X[14]	X[14]	X[14]	X[14]	
C-SSRS[15]	X		X[4]		X			X	X
[redacted]									
Randomization			X						
Study Drug Dispensation/Accountability			X	X	X	X	X	X	

Abbreviations: C-SSRS=Columbia Suicide Severity Rating Scale; ECG=electrocardiogram; [redacted]; ET=early termination; [redacted]; GI=gastrointestinal; GGISIS=Global GI Symptom and Impact Scale; IGISS=Individual GI Symptom and Impact Scale; MRI=magnetic resonance imaging; [redacted]

[2] To be conducted within 4 weeks (-28 days) of randomization (Visit 3).

[3] To be conducted 1 week before Visit 3 (Week 1/Randomization). The screening period (Visit 1 to Visit 3) may last up to 28 days.

[redacted]

A complete physical examination at screening (Visit 1); a brief physical examination at all other scheduled visits.

[6] Serology testing includes hepatitis B surface antigen, hepatitis C antibody, and human immunodeficiency virus antibody.

[7] Serum pregnancy at screening (Visit 1); urine pregnancy at randomization (Visit 3). [redacted]

[8] Vital sign measurements include temperature, respiratory rate, blood pressure, and heart rate. Vital signs to be collected predose at randomization (Visit 3) and at any time at all other scheduled visits. Blood pressure, respiratory rate, and heart rate will be measured after the subject has been in a seated or supine position for at least 5 minutes.

[9] ECGs to be collected predose at randomization (Visit 3) and at any time at all other scheduled visits. ECGs will be collected after the subject has been in a supine position for at least 5 minutes.

[redacted]

[11] Urinalysis includes urine dipstick, urine microscopy (as applicable), urine beta-2-microglobulin, urine albumin, and urine creatinine.

[12] MRI to be performed at any time between screening (Visit 1) and randomization (Visit 3) provided all other screening criteria have been checked for subject eligibility. MRI must not be performed within 30 days of receiving a course of steroids.

[13] The IGISIS will be completed twice daily (with at least 6 hours in between each completion) and the GGISIS will be completed once daily using e-diaries during the 1-week lead-in period prior to randomization (Visit 3).

[14] The IGISIS will be completed twice daily (should be completed within 9 hours after taking each dose of study drug) and the GGISIS will be completed once daily (each morning prior to taking the morning dose of study drug) using e-diaries during the randomized treatment period (Visit 3 to Visit 8).

[15] Use "Screening" version at screening (Visit 1) and "Since-Last-Visit" version for all subsequent time points.

2. SAMPLE SIZE AND STATISTICAL POWER CONSIDERATION

Sample Size Considerations (Part A and Part B): The sample size calculation for Part B is based on the secondary endpoint of the mean number of days with any IGISIS individual symptom intensity score ≥2.

Assuming the mean number of days with any IGISIS individual symptom intensity score ≥2 is 3.5 and 2.0 for the DMF and ALKS 8700 treatment groups, respectively, a sample size of 500, including 60 randomized subjects per group (120 subjects in total) in Part A and 190 randomized subjects per group (380 subjects in total) in Part B, will provide at least 80% power to detect about a 42% reduction in the number of days with any IGISIS individual symptom intensity score ≥2 for the ALKS 8700 treatment group relative to the DMF treatment group, at an alpha level of 0.05 (two-sided) using a negative binomial regression approach. The sample size calculation assumes a standard deviation of 5.9 in the number of days with any IGISIS individual symptom intensity score ≥2 relative to exposure days for the DMF treatment group in Part A and Part B.

Assuming the mean number of days with any IGISIS individual symptom intensity score ≥2 is 3.5 and 2.0 for the DMF and ALKS 8700 treatment groups, respectively, a sample size of 380 randomized subjects (190 subjects per treatment group) in Part B will provide at least 80% power to detect about a 42% reduction in the number of days with any IGISIS individual symptom intensity score ≥2 for the ALKS 8700 treatment group relative to the DMF treatment group, at an alpha level of 0.05 (two-sided) using a negative binomial regression approach. The sample size calculation assumes a standard deviation of 5.9 in the number of days with any IGISIS individual symptom intensity score ≥2 relative to exposure days for the DMF treatment group in Part B.

3. DATA ANALYSIS

3.1. General Statistical Methodology

The safety and efficacy endpoints will be summarized by treatment group and overall as described below. In general, summary statistics (n, mean, standard deviation [SD], median, minimum, and maximum for continuous variables, and number and percentage of subjects in each category for categorical variables) will be provided by treatment group and overall. All summary tables will be based on observed data, and missing values will not be imputed unless otherwise indicated. Measurements collected from unscheduled visits or repeated assessments will not be included in the by-visit summary tables or figures, but will be included in the analyses for the potentially clinically significant (PCS) post-baseline values and subject listings. Source data for the summary tables and statistical analyses will be presented as by-subject data listings.

Baseline value is defined as the last non-missing assessment prior to the first dose of study drug.

All statistical tests and confidence intervals (CIs), unless stated otherwise, will be 2-sided and will be set at an alpha level of 0.05.

All the efficacy analyses will be summarized by the planned treatment assignment, and all the safety analyses will be summarized by actual treatment.

The data will be summarized for the following two study periods as appropriate.

- Treatment Period, which is defined as the period between the first dose date after randomization and the last dose date plus 1 day, inclusive.
- Follow-up Period, which is defined as the period between the date of the last dose date plus 2 day and the last study date, inclusive.

3.2. Study Populations

3.2.1. Definitions of Analysis Populations (Analysis Sets)

3.2.1.1. Safety Population

The safety population will include all randomized subjects who receive at least one dose of study drug (ALKS 8700 or DMF).

3.2.1.2. Full Analysis Set (FAS)

The full analysis set (FAS) population will include all subjects who receive at least one dose of study drug and who complete at least one post baseline GI tolerability assessment (GGISIS or IGISIS).

[redacted]

3.2.2. Disposition

Subject disposition will be summarized by treatment group in terms of the following:

- Subjects enrolled in the study
- Subjects in the Safety Population
- Subjects in the FAS population
- [redacted]
- Subjects who completed the treatment period
- Subjects who completed the study
- Subjects who rolled over to A301 study
- Subjects who discontinued from the study during the treatment period
- Subjects by the reason for early discontinuation

Percentages of disposition will be based on the subjects in the Safety Population. A listing of disposition will be provided for all subjects.

3.2.3. Protocol Deviations

Subjects with major protocol deviations in the following categories will be summarized by treatment group and overall, along with supportive listings for each category as listed below:

- Did not meet the inclusion/exclusion criteria
- Received prohibited medication (as determined by the sponsor medical monitor)
- Lack of adherence with study medication, as defined by subjects taking less than 70% of the protocol-specified amount of study medication
- Randomization or dosing error
- Other major protocol deviation.

3.3. Demographics and Baseline Characteristics

Summary statistics will be provided for all demographic variables and baseline characteristics by treatment group and overall. Demographic data including age, age category (<40, ≥40 years), gender, race, weight, body mass index, ethnicity and region (US and non-US) will be summarized by treatment group and overall.

Baseline characteristics will be summarized by treatment group, using descriptive statistics. These include years since disease (multiple sclerosis [MS]) onset, years since MS diagnosis, number of MS relapses in the past 12 months, prior MS treatment, [redacted].

Baseline magnetic resonance imaging (MRI) parameters (number of gadolinium-enhancing [GdE] lesions, GdE lesion category [0, 1-4, 5-8, and >8], T2 lesion volume and Normalized Brain Volume [NBV]) will also be summarized by treatment group and overall.

Medical history and number of prior disease modifying treatments (DMTs) will be summarized for the safety population using the number of observations and percentage of subjects reporting each category by treatment group and overall.

Demographic, baseline characteristics, medical history and prior MS treatment history listings will be provided for all subjects.

3.4. Prior and Concomitant Medications

Prior medications will be defined as medications taken prior to the first dose of the study medication. Concomitant medications will be defined as medications taken during the period between the first dose date and the last dose date of study drug, inclusive. All medications as documented by the investigator will be coded using the World Health Organization Anatomical Therapeutic Chemical (WHO-ATC) drug dictionary herbal enhanced version (2016, 1st quarter).

Prior and concomitant medications will be summarized by drug class, preferred term, treatment group and overall. Prior DMTs will also be summarized by treatment groups and overall. For the summary tables, if a subject has taken a prior or concomitant medication more than once, the subject will be counted only once for the medication. In addition, a separate listing for concomitant medications used to treat GI related Adverse Events will also be provided.

All reported medications will be presented in a subject data listing.

3.5. Treatment Adherence Rate and Extent of Exposure

The treatment adherence rate will be derived directly using the following formula:

(number of capsules taken) / (number of capsules expected to be taken) * 100.

When a subject discontinues treatment early, the number of capsules that should have been taken is based upon the duration the subject was on treatment.

Study drug adherence will be summarized for each treatment group and overall.

The total number of days the subject is on study drug (time on treatment) will be summarized as a continuous variable and categorized into 1-week intervals. Days on study drug will be calculated as the number of days from date of first dose to date of last dose inclusive.

3.6. GI Tolerability and Efficacy Analyses

3.6.1. General Considerations

GI tolerability assessments will be performed using the FAS Population unless otherwise specified.

All statistical tests will be 2-sided with a type I error rate of 5%, unless otherwise specified. All confidence intervals will be 2-sided 95% confidence intervals. No corrections for multiple comparisons will be applied.

3.6.2. GI Tolerability Assessment

The primary, secondary, and exploratory endpoints are as follows:

Primary Endpoint

- The number of days with any IGISIS individual symptom intensity score ≥2 relative to exposure days in Part A and Part B.

Secondary Endpoints

- Number of days with any IGISIS individual symptom intensity score ≥2 relative to exposure days in Part B
- Number of days with any IGISIS individual symptom intensity score ≥1 relative to exposure days in Part A and Part B
- Number of days with any IGISIS individual symptom intensity score ≥3 relative to exposure days in Part A and Part B
- Number of days with a GGISIS symptom intensity score ≥1 relative to exposure days in Part A and Part B
- Number of days with a GGISIS symptom intensity score ≥2 relative to exposure days in Part A and Part B
- Number of days with a GGISIS symptom intensity score ≥3 relative to exposure days in Part A and Part B
- Worst IGISIS individual symptom intensity score by week during the 5 week treatment period in Part A and Part B

Exploratory Endpoints

[redacted]

[redacted]

3.6.2.1. Individual GI Symptom and Impact Scale

The IGISIS is designed to assess the incidence, intensity, onset, duration, and functional impact of 5 individual GI symptoms: nausea, vomiting, upper abdominal pain, lower abdominal pain, and diarrhea. This scale will be subject-completed using e-diaries. These 5 symptoms were included in the IGISIS as they were the most commonly reported GI symptoms in the Tecfidera® Phase 3 placebo-controlled trials. Subjects will be asked to rate the severity of each individual symptom via an 11-point numeric rating scale ranging from 0 (did not have) to 10 (extreme). Subjects will also be asked to indicate the start and stop times of each GI symptom. Additionally, subjects will be asked to indicate how much each symptom has interfered with their ability to accomplish their regular daily activities using a 5-point Likert scale (not at all, slightly, moderately, quite a bit, extremely).

GI tolerability assessment utilizing the IGISIS will focus on intensity scores (11 point numeric severity rating scale) of each individual symptom. If an IGISIS symptom is recorded as ongoing, only the initial intensity score will be collected (i.e., logically skipped item) and the initial intensity score collected at the symptom start will be carried over until symptom end.

<u>Number of days with any IGISIS individual symptom intensity score ≥2 relative to exposure days</u>

Number of days with any IGISIS individual symptom intensity score ≥2 (event days) is counted among observed diaries. Exposure days are the number of days between a subject's first dose date and last dose date during treatment period, inclusive.

Let $\theta^{(1)}$ and $\theta^{(2)}$ denote the mean number of days with any IGISIS individual symptom intensity scores ≥ 2, for ALKS 8700 and DMF, respectively. Comparison of GI tolerability will be tested at 5% level of significance in a two-sided test, through the following hypothesis of no treatment effect:

$$H_0: \theta^{(1)} = \theta^{(2)} \quad \text{vs. } H_a: \theta^{(1)} \neq \theta^{(2)}$$

ALKS 8700 will be *claimed superior* to DMF if *the estimated* $\theta^{(1)}$ *is less than the estimated* $\theta^{(2)}$ *and p-value is less than 0.05 (in a 2-sided test)*. DMF will be *claimed superior* to ALKS 8700 if *the estimated* $\theta^{(2)}$ *is less than the estimated* $\theta^{(1)}$ *and p-value is less than 0.05 (in a 2-sided test)*.

Event days will be analyzed by the negative binomial regression model, including treatment as factor and adjusting for study parts (not included in part B only analyses), region (US and non-US), age and Body Mass Index (BMI). Additional covariates may also be included in the model. The logarithmic transformation of exposure days will be included in the model as the "offset"

parameter. If the data are underdispersed, or if the negative binomial regression model does not converge, a Poisson regression model with the same covariates will be used instead of the negative binomial regression model. Dispersion will be evaluated from the Pearson Chi-square statistic. If the ratio of the Pearson Chi-square statistic to the degrees of freedom is ≤ 1 which indicates no overdispersion, then a Poisson regression model with adjustment for underdispersion will be used. Count distributions in which the number of intervals with zero events is higher than predicted by a Negative binomial or Poisson model may be modeled using a Zero-inflated model.

This method will also be used for the secondary endpoints by setting the GI symptom intensity score cut off as 1 and 3. Empirical cumulative distribution function plot of number of event days by treatment groups will be presented.

Worst IGISIS symptom intensity score by week during the 5 week treatment period

Worst individual IGISIS symptom intensity score is defined as a subject's worst (i.e., highest) symptom intensity score for any symptom out of all completed IGISIS diaries during a given period. Worst IGISIS symptom intensity scores by week during the 5 week treatment period for each subject will be summarized by treatment groups and will be analyzed using an Analysis of covariance (ANCOVA) model. The model will include treatment as factors, and adjusting for study parts, region (US and non-US), age and BMI. Additional covariates may also be included.

Worst IGISIS individual symptom interference level

Worst IGISIS individual symptom interference level is defined as a subject's worst symptom interference level (No symptom < Not at all < Slightly < Moderately < Quite a bit < Extremely) for a particular symptom out of all completed IGISIS diaries during the treatment period. All five GI symptoms (nausea, vomiting, upper abdominal pain, lower abdominal pain and diarrhea) will be reported separately and summarized by treatment groups.

Sensitivity Analysis of IGISIS Endpoints

An evaluable diary is defined as an instance when a subject took study drug prior to completing the diary and completed the diary within the recommended reporting window (2 to 9 hours post-dose). Summary statistics will be presented for IGISIS primary and secondary endpoint by evaluable and non-evaluable dairies. Event days will be analyzed controlling for evaluable and non-evaluable dairies.

3.6.2.2. Global GI Symptom and Impact Scale

The GGISIS is a global scale to assess the overall intensity, bothersomeness, and functional impact of GI symptoms experienced during the previous 24 hours. This scale will be subject-completed using e-diaries. Subjects will be asked to rate the intensity and bothersomeness of GI symptoms experienced using an 11-point numeric rating scale. Additionally, subjects will be asked to indicate:

- How much their GI symptoms have interfered with their ability to accomplish their regular daily activities using a 5-point Likert scale (not at all, slightly, moderately, quite a bit, extremely),
- Number of hours of work missed due to GI symptoms,
- Impact of GI symptoms on work productivity using a 5-point Likert scale (not at all, slightly, moderately, quite a bit, extremely).

Number of days with a GGISIS symptom intensity score ≥2 relative to exposure days

Number of days with a GGISIS symptom intensity score ≥2 (event days) is counted among available diaries. Exposure days are the number of days between a subject's first dose date plus one and last dose date during treatment period, inclusive.

Event days will be analyzed by the negative binomial regression model, using treatment as factors and adjusting for study parts (not included in part B only analyses), region (US and non-US), age and BMI. Additional covariates may also be included in the model. The logarithmic transformation of exposure days will be included in the model as the "offset" parameter. If the data are underdispersed, or if the negative binomial regression model does not converge, a Poisson regression model with the same covariates will be used instead of the negative binomial regression model. Dispersion will be evaluated from the Pearson Chi-square statistic. If the ratio of the Pearson Chi-square statistic to the degrees of freedom is ≤ 1 which indicates no overdispersion, then a Poisson regression model with adjustment for underdispersion will be used. Count distributions in which the number of intervals with zero events is higher than predicted by a Negative binomial or Poisson model may be modeled using a Zero-inflated model.

This method will also be used for the secondary endpoints by setting the GI symptom intensity score cut off as 1 and 3. Empirical cumulative distribution function plot of number of event days by treatment groups will be presented.

Worst GGISIS bothersome score and interference level

Subject's worst GGISIS bothersome score and worst interference level (No Symptom < Not at all < Slightly < Moderately < Quite a bit < Extremely) out of all completed IGISIS diaries during the treatment period will be summarized by treatment groups.

3.6.3. Other Exploratory Endpoints

[redacted]

[redacted]

[redacted]

3.7. Safety Analyses (Part A and Part B)

The safety analysis will be carried out using the Safety Population that includes all randomized subjects who receive at least one dose of study drug (ALKS 8700 or DMF).. Safety assessments

will be summarized using descriptive statistics along with supportive listings based on observed values. Listings will be provided for all safety parameters.

3.7.1. Adverse Events

All adverse events (AEs) will be coded by System Organ Class (SOC) and Preferred Term (PT) using the Medical Dictionary for Regulatory Activities MedDRA version 20.1. The verbatim term will be included in the AE listings.

AEs will be identified as emerging in the treatment period and the follow up period. During the treatment period, an AE will be considered a treatment-emergent AE (TEAE) if the event starts or worsens on or after the date of first dose of study drug in the current trial. The Treatment Period is defined as the period from the first dose date to last dose date plus 1 day, inclusive.

During the follow up period, post-discontinuation emergent AEs (PDEAEs), defined as AEs that started or worsened after the last dose date plus 1 day, will be summarized. For PDEAEs, the greatest severity on or before the last dose date plus 1 day will be used as the benchmark for the comparison of the AEs occurring during the follow up period. The Follow-up Period is defined as the period between the date of the last dose date plus 2 day and the last study date, inclusive.

An overview table, including number of subjects with TEAEs, serious AEs (SAEs), AEs leading to study discontinuation, and study drug-related TEAEs will be provided. Adverse events leading to study discontinuation and SAEs leading to death will be summarized in the period when the discontinuation or death occurred.

The following summary tables will be provided for the treatment period by treatment group and overall:

- TEAEs by SOC and PT
- TEAEs by SOC, PT, and severity
- TEAEs by SOC, PT, and relationship
- Study drug related TEAEs by SOC and PT
- TEAEs experienced by ≥5% of subjects by SOC and PT
- TEAEs experienced by ≥2% of subjects by SOC and PT
- SAEs by SOC and PT
- AEs leading to study discontinuation by SOC and PT

All AE tables will be sorted by SOC and then PT in decreasing frequency based on all subjects in the Safety Population.

A subject having the same AE (as determined by the coded MedDRA PT) more than once will be counted only once in the number and percentage of subjects' calculation for that AE. Similarly, if a subject has more than one AE in an SOC, the subject will be counted only once in the total number of subjects with an AE for that SOC. If a subject has the same AE on multiple occasions, the highest severity (severe > moderate > mild) recorded for the event will be presented in the AE by severity summary. Similarly, if a subject has the same AE on multiple occasions, the closest relationship to study drug (related > not related, where related includes

definitely related, probably related, and possibly related; and not related includes probably not related and definitely not related) recorded for the event will be presented in the AE by relationship summary.

The following summary tables will be provided for flushing and flushing-related TEAEs and GI SOC only by treatment group and overall:

- TEAEs by PT and outcome
- Duration of TEAEs among those resolved by PT
- Time to onset of TEAEs by PT

PTs for flushing and related AEs include: flushing; hot flush; erythema, generalized erythema, burning sensation, skin burning sensation, feeling hot, pruritus; pruritus generalized; rash, rash pruritic, rash maculo-papular; rash macular; rash papular.

If a subject has the same AE on multiple occasions, the worst outcome (not resolved > resolved, where resolved includes recovered/resolved and recovered/resolved with sequelae; and not resolved includes not recovered/not resolved and fatal) recorded for the event will be presented in the AE by outcome summary. For duration and time to onset, if a subject has the same AE on multiple occasions, the first presentation will be used in summarizing time to onset and the mean duration of the multiple AE events will be used in summarizing duration.

All AEs will be included in the listings. Supporting listings of SAEs, AEs leading to study discontinuation, and AEs leading to death will be provided. Subjects with PDEAEs in the Follow-up Period will also be listed.

3.7.2. Deaths, Serious and Other Significant Adverse Events

The number and percentage of subjects with at least one SAE (regardless treatment emergent or not) will be summarized by SOC and PT for the Safety Population. Deaths and AEs leading to discontinuation (regardless treatment emergent or not) will be summarized similarly.

Supporting listings of serious AEs and AEs leading to study discontinuations will be provided. Subjects who died during the study will also be listed.

3.7.3. Adverse Events of Special Interest

Adverse events in the following categories of interest will be presented by following categories and PTs:

- Anaphylaxis and Angioedema (serious adverse events)
- Lymphopenia
- Liver Injury
- Malignancies and Pre-malignant Conditions
- Opportunistic Infections (AEs and SAEs) and All Serious Infections (Including Opportunistic Infections SAEs)
- Pancreatitis

- Renal Injury
- Cardiac Disorders
- Gastrointestinal Tolerability AEs (serious adverse events and adverse events leading to discontinuation)
- Abuse potential

Supportive AE listings including serious and non-serious AEs will be provided for each special interest group. Additionally, listing for "non-specific" AEs in the abuse potential category will be provided. Narratives will be provided for "euphoria related" and "abuse behavior" AEs in the category of abuse potential.

3.7.4. Clinical Laboratory Parameters

Laboratory parameters will be presented in conventional units. Only scheduled laboratory parameters will be included in the summaries of laboratory data. All laboratory data, including those collected at unscheduled visits, will be included in the listings. Laboratory results (baseline and change from baseline) for chemistry, hematology and urinalysis parameters for each visit during the entire study will be summarized by treatment group and overall.

For urinalysis, the number of abnormalities at any post-baseline visit will be summarized. In addition, the number and percentage of subjects with values considered potentially clinically significant (PCS) occurring at any post-baseline visit for selected parameters will be summarized by treatment group and overall.

Clinical laboratory test values, scheduled or unscheduled during the treatment period, will be considered PCS if they meet the PCS criteria listed in Table 2. The percentages will be calculated relative to the number of subjects with available non-PCS baseline values with respect to the specific criterion and at least 1 post-baseline assessment. The denominator is the total number of subjects with a non-PCS baseline value with respect to the specific criterion and at least 1 postbaseline PCS value. A supportive listing will present all values for a parameter for subjects with at least one PCS value for that parameter.

Shift tables for shift from within normal reference limits / high to low and shift from within normal reference limits / low to high during the treatment period will be provided for chemistry and hematology parameters with conventional reference limits (limits provided by the performing laboratories) by treatment group and overall for the Safety Population. For parameters which have gender and / or age specific limits, the categories will be based on the specific limits. High is defined as values greater than the upper limit of the reference range (ULRR). Low is defined as values less than the lower limit of the reference range (LLRR). Normal reference limits is defined as values within normal range.

The percentages will be calculated for the shift summary as follows:

- For the shifts from within normal reference limits / high to low: the numerator is defined as the number of subjects with a post-baseline result below the LLRR, and the denominator is defined as the number of subjects with a baseline value within the limits or above the ULRR.

- For the shifts from within normal reference limits / low to high: the numerator is defined as the number of subjects with a post-baseline result above the ULRR, and the denominator is defined as the number of subjects with a baseline value within the limits or below the LLRR.

The following figures will be presented by treatment group and overall:

- Mean along with the corresponding SE for change from baseline values for selected chemistry and hematology laboratory parameters.
- Boxplots will be used to display summary statistics for the maximum and minimum change from baseline values for selected chemistry and hematology laboratory parameters.

Pregnancy test data will be listed.

Table 2: Potentially Clinically Significant (PCS) Criteria for Laboratory Parameters

Category	Parameter	Criteria
Hematology	Hematocrit	≤32% and 3 point decrease from baseline (Female) ≤37% and 3 point decrease from baseline (Male)
	Hemoglobin	≤9.5 g/dL (Female) ≤11.5 g/dL (Male)
	Neutrophils	<1.5 × 10^3/uL
	Platelets	<75.1 × 10^3/uL ≥700 × 10^3/uL
	WBCs -total, differential (absolute)	≤2.8 × 10^3/uL ≥16 × 10^3/uL
	Eosinophils	>1.0 × 10^3 cells/uL
	Lymphocytes	< 0.5 x 10^9/L
Chemistry	ALT	≥3 × ULN
	Albumin	<2.5 g/dL
	ALK-P	≥3 × ULN
	AST	≥3 × ULN
	Bicarbonate	<15 mmol/L >31 mmol/L
	BUN	> 30 mg/dl
	Calcium	<8.2 mg/dL >12 mg/dL
	Chloride	≤ 90 mmol/L ≥ 118 mmol/L
	Creatine Kinase	>3 × ULN

Category	Parameter	Criteria
	Creatinine	≥2 mg/dL
	Glucose	< 50 mg/dL > 200 mg/dL
	HDL Cholesterol	≤ 30 mg/dL
	Lactate Dehydrogenase	>3 × ULN
	LDL Cholesterol	≥160 mg/dL
	Phosphorus	<2 mg/dL >5 mg/dL
	Potassium	<3 mmol/L >5.5 mmol/L
	Sodium	<130 mmol/L >150 mmol/L
	Total Bilirubin	≥2 mg/dL
	Total Cholesterol	>300 mg/dL
	Triglycerides	≥120 mg/dL (Female) ≥160 mg/dL (Male)
	Uric Acid	> 9 mg/dL > 8 mg/dL (Female) > 10 mg/dL (Male)
Urinalysis	Glucose	at least 2+
	Protein	at least 2+
	Albumin/Creatinine	>200 g/kg
	Beta-2 Microglobulin	>0.300 mg/L

For abnormal liver function tests, the number and percentage of subjects will be summarized by upper limit of normal (ULN) category as follows:

- ALT (alanine aminotransferase), AST (aspartate aminotransferase) and ALK-P (alkaline phosphatase) will be summarized by normal, >3xULN, >5xULN, >10xULN, total bilirubin by >1.5xULN and >2xULN.

Number of subjects who meet Hy's Law criteria (total Bilirubin >=2xULN and ALT or AST >=3xULN) will be summarized and corresponding listing will be presented.

Number of Subjects with Lymphopenia by Maximum NCI-CTCAE Toxicity Grade 1-4 will be summarized. Subjects with Lymphopenia Toxicity Grade 2 or higher will also be listed.

3.7.5. Vital Signs and ECG

3.7.5.1. Vital Signs

Vital signs for each visit during the entire study will be summarized by treatment group and overall. All vital sign data will be presented in the subject data listings.

Number and percentage of subjects with vital sign values considered PCS occurring at any post-baseline visit will be summarized by treatment group and overall. Criteria for PCS are presented in Table 3 and will be presented for each parameter. The percentages will be calculated relative to the number of subjects with non-PCS baseline values with respect to the specific criterion and at least 1 post-baseline assessment. A supportive listing will present all values for a parameter for subjects with at least one PCS value for that parameter.

In addition, the figure of mean change from baseline along with the corresponding SE will be summarized by visit and treatment group and overall for systolic blood pressure, diastolic blood pressure, and heart rate.

Table 3: Potentially Clinically Significant (PCS) Criteria for Vital Signs

Parameter	Criteria
Systolic Blood Pressure	Low: ≤90 mm Hg and decrease ≥20 mm Hg High: ≥180 mm Hg and increase ≥20 mm Hg
Diastolic Blood Pressure	Low: ≤50 mm Hg and decrease ≥15 mm Hg High: ≥105 mm Hg and increase ≥15 mm Hg
Pulse Rate	Low: ≤50 mm bpm and decrease ≥15 bpm High: ≥120 mm bpm and increase ≥15 bpm

3.7.5.2. Electrocardiograms

ECG parameters (heart rate, RR interval, PR interval, QRS interval, QT interval, QTcB [QT interval corrected using Bazett's method], QTcF [QT interval corrected using Fridericia's method]) will be summarized for each visit during the study by treatment group and overall.

Number and percentage of subjects with QTcF parameter values considered PCS occurring at any post-baseline visit will be summarized by treatment group and overall. Criteria for PCS are presented in Table 4 and will be presented for each parameter. A subject will be counted only once in the highest category for a given parameter based on the largest post-baseline value. The percentages will be calculated relative to the number of subjects with baseline value ≤ 450 msec and at least 1 post-baseline assessment in the safety analysis set.

A supportive listing will present all values for a parameter for subjects with at least one PCS value for that parameter.

Table 4: Potentially Clinically Significant (PCS) Criteria for QTcF

Parameter	Criteria
QTcF	>450 to ≤480 msec

Parameter	Criteria
	>480 to ≤500 msec
	>500 msec
	Change from baseline >30 to ≤60 msec
	Change from baseline >60 msec

3.7.6. Columbia Suicide Severity Rating Scale (C-SSRS)

The C-SSRS is a questionnaire used to measure the presence and intensity of suicidal ideation and behavior. Suicidal behavior and suicidal ideation will be summarized by treatment group and overall. The proportion of subjects who meet the criterion for each of these categories will be summarized as described in Table 5 .

Table 5: C-SSRS Categories for Analysis

Category	C-SSRS Item response is "YES"
Suicidal behavior	Preparatory acts or behavior Aborted attempt Interrupted attempt Actual attempt Complete suicide
Suicidal ideation	Wish to be dead Non-specific active suicidal thoughts Active suicidal ideation with any methods (not plan) without intent to act Active suicidal ideation with some intent to act, without specific plan Active suicidal ideation with specific plan and intent
Non-Suicidal Self-Injurious Behavior	Non-Suicidal Self-Injurious Behavior

4. INTERIM ANALYSES AND DATA MONITORING COMMITTEE (DMC)

No interim analysis was planned for this study.

5. CHANGES IN CONDUCT OR PLANNED ANALYSES FROM THE PROTOCOL

None.

6. DEFINITIONS AND CONVENTIONS FOR HANDLING OF THE DATA

Dataset specifications will be provided in a separate document.

6.1. Analysis Visit Windows

Scheduled analysis visits are visits upon scheduled time points as specified in Table 1 Schedule of Visits and Assessments.

Scheduled analysis visits during the study period will be the same as the nominal visits collected in the eCRF. There will be one valid value of the assessment kept for each scheduled analysis visit in summary/ analysis statistics.

Unscheduled visits are visits with data not collected on the scheduled time point. Unscheduled visits will not be used for by-visit summary/analysis statistics unless specified otherwise. All unscheduled visits will be included as collected in eCRF in listings.

6.2. Handling of Partial Dates of Concomitant Medication

Partial start dates of prior and concomitant medications will be assumed to be the earliest possible date consistent with the partial date. Partial stop dates of prior and concomitant medications will be assumed to be the latest possible date consistent with the partial date. In the case of completely missing stop dates, medication will be assumed to be ongoing.

6.3. Safety Data Handling

All efforts should be made to obtain the missing information from the Investigator. For C-SSRS, vital signs, laboratory testing (chemistry, hematology, urinalysis), and 12-lead ECGs, only observed data will be used for analyses, and missing data will not be imputed.

7. GENERAL STATISTICAL METHODOLOGY

In general, summary statistics (n, mean, SD, median, minimum, and maximum for continuous variables, and number and percentage of subjects in each category for categorical variables) will be provided by treatment group. All summary tables will be based on observed data, and missing values will not be imputed unless otherwise indicated. Measurements collected from unscheduled visits or repeated assessments will not be included in the by-visit summary tables or figures, but will be included in the derivation of the last post-baseline value during treatment, the analyses of PCS post-baseline values, and subject listings. Source data for the summary tables and statistical analyses will be presented as subject data listings.

7.1. Reporting Precision

Summary statistics will be presented to the degree of precision in Table 6, unless otherwise specified:

Table 6: Degree of Precision

Statistics	Degree of Precision
Mean, Geometric mean, Median, Quartiles, Confidence limit boundaries	One more than the raw data, up to 3 decimal places.
Standard deviation, Standard error	One more than the mean, up to 3 decimal places
Minimum, Maximum	The same as the raw data, up to 2 decimal places
p-value	Rounded to 3 decimal places and therefore presented as 0.xxx; p-values smaller than 0.001 as '<0.001'; p-values greater than 0.999 as '>0.999'.
Percentage	One decimal place. A percentage of 100% will be reported as 100%. Percentages of zero will be reported as 0.

Fractional numeric values will be presented with a zero to the left of the decimal point (for example, 0.12–0.30).

For weight, height, and body mass index (BMI), one decimal place will be used for summary statistics.

8. PROGRAMMING SPECIFICATIONS

Programming specifications will be provided in a separate document.

9. MOCK TABLES, LISTINGS AND GRAPHS (TLGS)

Mock-up tables and listings will be provided in a separate document.

10. REFERENCES

1. Alkermes ALK8700-A302 Clinical Study Protocol Amendment: 15DEC 2016
2. Alkermes ALK8700-A302 Clinical Study Protocol Amendment 1.1 for sites in Poland only: 23 Aug 2017
3. Alkermes ALK8700-A302 Clinical Study Protocol Amendment 2.0 (incorporated Administrative Change Memo 1): 20 Sep 2018
4. Alkermes ALK8700-A302 Clinical Study Protocol Germany-specific Amendment 2.2 (incorporated Administrative Change Memo 2): 14 Dec 2018

APPENDIX I ALK8700-A302 PART A SAP

APPENDIX II AESI SEARCH CRITERIA