Statistical Analysis Plan

Title of Study:	A phase III clinical study of MR13A9 in previously treated hemodialysis patients with pruritus
Protocol No.:	MR13A9-5
Version:	3.0
Date of Creation/Amendment:	October 24, 2022
Sponsor:	Kissei Pharmaceutical Co., Ltd.

Table of Contents

1. Study Objectives

The purpose of this statistical analysis plan is to provide details of the analyses described in "14. Statistical Analysis" of the clinical study protocol (Version, 1.0).

2. Analysis Sets

Analysis sets are defined as follows: When the full analysis set (FAS) and the per protocol set (PPS) are used as analysis sets, analyses will be performed based on the treatment to which each subject has been assigned. When the safety set (SS) is used as an analysis set, analyses will be performed based on the treatment provided to each subject actually.

1) FAS
Population of subjects given at least one dose of the study drug, meeting the inclusion criterion related to NRS score, and having a baseline mean NRS score.

2) PPS (only in the double-blind period)
Population of subjects excluding the following subjects from the FAS:

- Subjects not meeting the inclusion criteria
- Subjects meeting any of the exclusion criteria related to efficacy evaluation
- Subjects withdrawing from the study before Week 4 in the double-blind period
- Subjects given < 9 or > 15 doses of the study drug before Week 4 in the double-blind period
- Subjects changing the dosage and administration of a conditionally permitted concomitant medication or newly using a conditionally permitted concomitant medication before Week 4 in the double-blind period
- Subjects using a prohibited concomitant drug or therapy before Week 4 in the double-blind period
- Subjects changing the frequency of dialysis per week or the hemodialysis method before Week 4 in the double-blind period

3) SS
Population of subjects given at least one dose of the study drug

Unless otherwise specified, the analysis sets for endpoints and analysis items in the double-blind period and entire study period are as specified in Table 2-1 and Table 2-2.

Table 2-1 Analysis sets for endpoints and analysis items in the double-blind period

Endpoints	Analysis items	Analysis sets
Demographic and other baseline characteristics	Subject demographics	FAS, PPS, SS
	Other	SS
Treatment compliance	All items	FAS, SS
Efficacy	Primary variable (primary analysis)	FAS (primary analysis set), PPS
	Other	FAS
Safety	All items	SS

Table 2-2 Analysis sets for endpoints and analysis items in the entire study period

Endpoints	Analysis items	Analysis sets
Demographic and other baseline characteristics	Subject demographics	FAS, SS
	Other	SS
Treatment compliance	All items	FAS, SS
Efficacy	All items	FAS
Safety	All items	SS

3. Analysis Groups

The analysis sets and subjects included in the analysis sets for the analysis groups in the double-blind period are shown in Table 3-1. The analysis sets and subjects included in the analysis sets for the analysis groups in the entire study period are shown in Table 3-2.

Table 3-1 Analysis sets and subjects included in the analysis sets for the analysis groups in the double-blind period

Analysis groups	Analysis sets	Subjects included in the analysis set
MR13A9 0.5 µg/kg group	FAS, PPS	Subjects randomized to MR13A9 0.5 µg/kg
	SS	Subjects given MR13A9 0.5 µg/kg
Placebo group	FAS, PPS	Subjects randomized to placebo
	SS	Subjects given placebo

Table 3-2 Analysis sets and subjects included in the analysis sets for the analysis groups in the entire study period

Analysis groups	Analysis sets	Subjects included in the analysis set
MR-MR group	FAS	Subjects randomized to MR13A9 0.5 µg/kg in the double-blind period and entering into the extension period
	SS	Subjects given MR13A9 0.5 µg/kg in the double-blind period and entering into the extension period
P-MR group	FAS	Subjects randomized to placebo in the double-blind period and entering into the extension period
	SS	Subjects given placebo in the double-blind period and entering into the extension period

4. General Principles of Statistical Analysis

- All statistical tests will be performed using a two-sided significance level of 5%; however, a two-sided significance level of 15% will be used for the analysis of between-group demographic imbalance.
- Summary statistics to be presented include the number of subjects, mean, standard deviation, minimum, median, maximum, and quartile.
- For the double-blind period, data from the start of the screening period to the end of the double-blind period will be analyzed. Over the entire study period, data from the start of the screening period to the end of the follow-up period will be analyzed.
- Unless otherwise specified, data will be summarized by group and time point for each of the double-blind and entire study periods.

5. Handling of Data

5.1 Number of Digits to be Reported for Calculated Values

1) P values and statistics
Values will be rounded down to the third decimal place. However, P values less than 0.001 will be reported as "P < 0.001" or "< 0.001."

2) Mean, standard deviation, standard error, adjusted mean, estimate of difference, and confidence interval
Values will be rounded off to one digit after the last place of the significant digits of the data.

3) Minimum, median, maximum, and quartiles
Values will be presented by the significant digits of the data.

4) Proportion, estimate and confidence interval of incidence
Values will be rounded off to the first decimal place.

5.2 Reporting of Prior and Concomitant Medications

Drug names will be reported according to the Iyakuhinmei Data File. Administration routes will be reported by code names (ROUTE) according to the CDISC SDTM Terminology.

5.3 Reporting of Complications, Adverse Events, and Adverse Drug Reactions

Primary SOCs and PTs will be reported according to the MedDRA.

5.4 Handling of Laboratory Test Values

Values below the limit of quantification will be handled as the value of the quantification limit.

5.5 Handling of Symptom Diary Data (NRS Score and Shiratori's Severity Criteria)

If the number of days of observation during the evaluation period is less than 4 days as of the time of calculation of the mean NRS score and the mean itch score based on the Shiratori's severity criteria at each time point, the score at that time point will be handled as missing.

5.6 Handling of Skindex-16 and 5-D Itch Scale Score

If the responses to the questions (components) used for calculation of the overall score or the subscores (symptoms score, emotions score, and functioning score) of the Skindex-16 or the total 5-D itch scale score are missing, the score will be handled as missing.

5.7 Time Windows for Assessment Time Points

Data for observations and tests will be handled according to the following criteria.

5.7.1 Handling of data in tabulation at each time point (excluding symptom diary

data)

- Data will be summarized by time point according to the time windows shown in Table 5.7.1-1. Data obtained outside the time windows will be excluded from tabulation at each time point.
- If multiple data are available within the time window, the data obtained at scheduled visit will be employed.
- Data for observations and tests performed at unscheduled visits or 10 days after the day of the final dose will be excluded from the tabulation.
- For dependency assessment, all data collected will be employed regardless of the time window.

Table 5.7.1-1 Time windows for assessment time points

Assessment time point	Scheduled date[a)]	Time window
At Week −2	−14	Scheduled date ± 3 days
At baseline	1	Day −3 to Day 1
At Weeks 1 to 12	7 × x + 1[b)]	Scheduled date ± 3 days
At Weeks 18 to 58	7 × x + 1[b)]	Scheduled date ± 7 days

a) The start of study treatment in the double-blind period is defined as Day 1 and the day before the start of study treatment as Day −1.
b) x: Number of weeks

5.7.2 Handling of data in tabulation at the final assessment point

- In the double-blind period, efficacy data (Skindex-16, 5-D itch scale, and PGIC) obtained at Weeks 1 to 4 (or at the time of withdrawal during the double-blind period before Week 4) and safety data obtained at Weeks 1 to 6 (or at the time of withdrawal during the double-blind period) will be included in the tabulation. Data for observations and tests performed at the latest time point within 10 days after the day of the final dose during the double-blind period will be used as the data at the end of the double-blind period. Data for observations and tests performed at unscheduled visits will be excluded from the tabulation.
- For the entire study period, data from Weeks 1 to 58 (or at the time of withdrawal during the extension period) will be included in the tabulation. Data for observations and tests performed at the latest time point within 10 days after the day of the final dose during the extension period will be used as the data at the end of the extension period. Data for observations and tests performed at unscheduled visits will be excluded from the tabulation.

6. Statistical/Analytical Issues

6.1 Adjustments for Covariates

The change from baseline in the mean NRS score at Week 4 of the double-blind period, the primary variable, will be analyzed using the baseline mean NRS score and the presence of prior treatment with nalfurafine hydrochloride, a dynamic allocation factor, as a covariate. For details, refer to "11.1.1 Primary variable." If an imbalance is observed in the subject demographics of the FAS, the primary analysis set for efficacy, a supplementary analysis will be performed with

additional items showing the imbalance as the covariates in the primary analysis.

For the changes from baseline in the itch score based on the Shiratori's severity criteria, changes from baseline in Skindex-16, and changes from baseline in 5-D Itch Scale, an analysis will be performed using the baseline as a covariate for each item. Details are shown in "11.2.1.1 Change from baseline in the mean itch score," "11.2.2.1 Changes from baseline in Skindex-16," and "11.2.3.1 Changes from baseline in 5-D itch scale," respectively.

6.2 Handling of Dropouts or Missing Data

When there is a lack in the data used for analysis, all items, except for those listed in "5 Handling of Data" and below, will be handled as missing values, and no imputation by statistical methods will be performed.

- Sensitivity analyses of the primary variable will use multiple imputation. Details are shown in "11.1.1.2 Sensitivity analysis."

6.3 Interim Analyses and Data Monitoring

6.3.1 Interim analysis

When all data obtained in the double-blind period are locked, an interim analysis of the data in the double-blind period will be performed. The interim analysis will be performed on all analysis items except for those listed in "12.6 Listing of Safety Variables."

6.3.2 Interim tabulation

An interim tabulation will be performed at the time of application for approval of MR13A9, in order to attach the study results obtained up to that time. When data obtained up to Week 34 of the extension period, excluding dependency assessment data, are locked, a first interim tabulation of these data will be performed. The first interim tabulation will be performed on all analysis items except for dependency assessment data. When data obtained up to Week 34 of the extension period, including dependency assessment data, are locked, a second interim tabulation of these data will be performed. The second interim tabulation will be performed on all analysis items. Each interim tabulation will include data obtained up to Week 34.
For adverse events and adverse drug reactions, events occurring before the Week 34 visit will be included in the tabulation. For subjects who discontinued before Week 34 and entered the follow-up period, events occurring during the follow-up period will also be included in the tabulation. Information on the events collected by the time of interim tabulation will be used for tabulation.
For concomitant medications, those used prior to the Week 34 visit will be included in the tabulation.

6.4 Multicenter Study

Since the number of subjects per site is not large, analysis by site will not be performed.

6.5 Multiple Comparison and Multiplicity

6.5.1 Multiplicity of multiple analyses and multiple assessment time points

The interim analysis of the double-blind period is defined as the primary analysis, and Week 4 of the double-blind period is defined as the primary assessment time point. Since evaluations at other time points will be handled as references, multiplicity of multiple analyses and multiple assessment time points will not be adjusted.

6.5.2 Multiplicity among multiple items

Since the change from baseline in the mean NRS score at Week 4 of the double-blind period is handled as the primary variable, and other items are handled as references, multiplicity among multiple items will not be adjusted.

6.5.3 Multiplicity of multiple analysis sets

For the double-blind period, the FAS will be used as the primary analysis set for efficacy, and the PPS will be used for the supplemental analysis to examine the robustness of the results, so multiplicity among multiple analysis sets will not be adjusted. It is not applicable for safety because only the SS is used as the analysis set.
For the entire study period, it is not applicable for efficacy because only the FAS is used as the analysis set. It is not applicable for safety because only the SS is used as the analysis set.

6.6 Use of "Two Different Analysis Sets for Efficacy Evaluation" of Subjects

For the double-blind period, the primary efficacy analysis set will be the FAS. On the other hand, the analysis will also be performed in the PPS as a supplemental analysis to examine the robustness of the results. The definitions of FAS and PPS are detailed in "2 Analysis Sets."

6.7 Active-controlled Studies Intended to Show Non-inferiority

Not applicable.

6.8 Examination of Subgroups

As the subgroup analyses of efficacy endpoints, the matters shown in "11.3 Efficacy Subgroups" will be performed. As the subgroup analyses of safety endpoints, the matters shown in "12.5 Safety Subgroups" will be performed.

7. Definition of Derived Data

Each variable will be calculated using the following formula:

- Disease duration of itch (years) = Date of written informed consent (year) – Time of onset of itching (year)
- Hemodialysis history (years) = Date of written informed consent (year) – Date of initiation of hemodialysis (year)
- Urea reduction ratio at baseline (%) = ([BUN before dialysis − BUN after dialysis] / BUN before dialysis) × 100

- Treatment compliance (%) = Number of doses of the study drug / Number of dialysis sessions performed during the double-blind and extension periods × 100
- Treatment duration (days) = Day of the final dose – Start of the study treatment + 1
- Average dose per dry weight (μg/kg) = Total dose of the study drug per dry weight / Number of doses
- NRS score and itch score based on Shiratori's severity criteria
  - Mean score at each time point = Sum of scores observed during the assessment time window / Number of days when the score was observed during the assessment time window
    *The details of the assessment time window are shown in Table 11.1-1 and Table 11.1-2.
  - Change from baseline in the mean score at each time point = Mean score at each time point − Mean score at baseline
  - Change from baseline in daily score = Daily score − Mean score at baseline
- Skindex-16[1)]
  - Score for each question = Response for each question / 6 × 100
  - Total score = Sum of scores for all questions / 16
  - Subscores
    - Symptoms score = (Sum of scores for Questions 1 to 4) / 4
    - Emotions score = (Sum of scores for Questions 5 to 11) / 7
    - Functioning score = (Sum of scores for Questions 12 to 16) / 5
- 5-D itch scale
  The total score will be calculated as the sum of scores for the following questions (components)[2)]:

  (1) Duration: 1 to 5 points

  (2) Degree: 1 to 5 points

  (3) Direction: 1 to 5 points

  (4) Disability: 1 to 5 points
  The highest score among sleep, leisure/social, housework/errands, and work/school will be used.

  (5) Distribution: 1 to 5 points
  The score is determined according to the total number of body parts with itching as follows:

  - 0 to 2: 1 point
  - 3 to 5: 2 points
  - 6 to 10: 3 points
  - 11 to 13: 4 points

  - 14 to 16: 5 points
- Time to onset of an adverse event (days) = Date of onset of the adverse event – Start of study treatment + 1

8. Disposition of Subjects

Analysis for the double-blind period will include subjects who entered the double-blind period. Analysis for the entire study period will include subjects who entered the extension period.

8.1 Analysis set

The number and percentage of subjects who were included or excluded in the FAS, PPS (only for the double-blind period), and SS will be presented. For the double-blind period, the between-group imbalance will be examined by Fisher's exact test.

8.2 Presence or Absence and Reasons of Discontinuation

The number and percentage of subjects for presence or absence of discontinuation will be presented. For the double-blind period, the between-group imbalance will be examined by Fisher's exact test. Reasons for discontinuation will be classified into the following categories, and the number and percentage of subjects will be presented for each category.

- Adverse event: When the discontinuation is due to discontinuation criteria (discontinuation of individual subjects) 1) "adverse events"
- Lack of efficacy: When the discontinuation is due to discontinuation criteria (discontinuation of individual subjects) 2) "lack of efficacy"
- Withdrawal by subject: When the discontinuation is due to discontinuation criteria (discontinuation of individual subjects) 3) " voluntary request for study discontinuation by the subject"
- Significant deviation: When the discontinuation is due to discontinuation criteria (discontinuation of individual subjects) 4) "significant deviation from the protocol during the study"
- Other: When the reason for discontinuation does not fall under the categories of "adverse events," "lack of efficacy," "voluntary request for study discontinuation by the subject," or "significant deviation."

9. Demographic and Other Baseline Characteristics

9.1 Subject Demographics

Analysis items are shown below. If there are data at more than one time point, the data at baseline will be used for evaluation.

- Nominal scale variables
  - Sex: Male, female
  - Primary disease of dialysis: Diabetic nephropathy, glomerulonephritis chronic, nephrosclerosis, polycystic kidney, pyelonephritis chronic, other, unspecified

- Type of dialysis: HD, off-line HDF, on-line HDF, I-HDF
- Remaining renal function: No, yes
- Prior treatment with nalfurafine hydrochloride: No, yes
- Specific signs or symptoms occurring during the screening period*: No, yes

  *Specific signs or symptoms: Dizziness, syncope, palpitations, tachycardia, fall, seizure, gait disturbance, mental status changes, somnolence, mood changes

- Ordinal scale variables
  - Age: < 65 years, ≥ 65 years
  - Dry weight at the start of the screening period: < 45 kg, ≥ 45 and < 65 kg, ≥ 65 and < 85 kg, ≥ 85 kg
  - Mean NRS score: < 6, ≥ 6
- Continuous variables: Age, dry weight at the start of the screening period, hemodialysis history, single-pool Kt/V, urea reduction ratio, disease duration of itch, mean NRS score, mean itch score based on the Shiratori's severity criteria (higher of daytime or night-time), Skindex-16 (overall score), 5-D Itch Scale (total score)

The number and percentage of subjects for nominal and ordinal scale variables, and summary statistics for continuous variables will be presented for the overall population and by group. However, only the number of subjects is shown for the primary disease of dialysis.

For the double-blind period, the between-group imbalance will be examined by tests for the prior treatment with nalfurafine hydrochloride, dry weight at the start of the screening period (ordinal scale variable), and mean NRS score (continuous variable). Depending on the characteristics of the data, Fisher's exact test will be used for nominal scale variables, two-sample Wilcoxon test for ordinal scale variables, and two-sample t-test for continuous variables.

9.2 Complications

The numbers and percentages of subjects are presented for all complications and by primary SOC and PT.

9.3 Concomitant Medications for Target Disease (Itching)

The numbers and percentages of subjects will be presented by generic name and route of administration.
The medications used during the following periods will be tabulated:

- Double-blind period: From the start of study treatment to the end of the double-blind period
- Entire study period: From the start of study treatment to the end of the extension period

9.4 Medications Used to Treat Complications

The numbers and percentages of subjects will be presented by generic name and route of administration.

The medications used during the following periods will be tabulated:

- Double-blind period: From the start of study treatment to the end of the double-blind period
- Entire study period: From the start of study treatment to the end of the extension period

10. Treatment Compliance

10.1 Number of Doses

The number of doses of the study drug will be classified into the following categories (unit: doses), and the number and percentage of subjects will be presented for each category.

- Double-blind period: ≤3, 4≤ ≤6, 7≤ ≤9, 10≤ ≤12, 13≤ ≤15, 16≤ ≤18, 19≤
- Entire study period: ≤36, 37≤ ≤72, 73≤ ≤108, 109≤ ≤144, 145≤ ≤180, 181≤

10.2 Treatment Compliance

Summary statistics will be presented.

10.3 Treatment Duration

Summary statistics will be presented.

10.4 Average Dose per Dry Weight

Summary statistics will be presented.
The assessment periods are as follows:

- Double-blind period: Overall
- Entire study period
  - MR-MR group: Weeks 1 to 6, Weeks 7 to 58, overall
  - P-MR group: Weeks 1 to 6, Weeks 7 to 58

10.5 Change in Dry Weight Category

The status of change in the dry weight category (< 45 kg, ≥ 45 and < 65 kg, ≥ 65 and < 85 kg, ≥ 85 kg) at the time of determination of the study drug dose will be classified into the following categories, and the number and percentage of subjects will be presented for each category. Subjects falling under the dose increase and decrease category will only be presented for overall of the entire study period.

- No change: Subjects with no change in category
- Dose increase: Subjects whose dry weight increased, resulting in category change (excluding those in the dose increase and decrease category)
- Dose decrease: Subjects whose dry weight decreased, resulting in category change (excluding those in the dose increase and decrease category)
- Dose increase and decrease: Subjects whose dry weight increased and decreased, resulting

in category change

The assessment periods are as follows:

- Entire study period
  - Weeks 7 to 34: Status of change from the time of determination of study drug dose for Weeks 1 to 6 (start of the screening period)
  - Weeks 35 to 58: Status of change from the time of determination of study drug dose for Weeks 7 to 34 (Week 6)
  - Overall: Status of change from the time of determination of study drug dose for Weeks 1 to 6 (start of the screening period)

11. Efficacy

11.1 Primary Endpoint

The primary endpoint is the NRS score. Assessment time points and their time windows in the screening period, the double-blind period, and the extension period are shown in Table 11.1-1. The assessment time point and its time window in the follow-up period are shown in Table 11.1-2.

Table 11.1-1 Assessment time points and their time windows in the screening period, the double-blind period, and the extension period

Period	Assessment time point	Time window[a)]
Screening period	At Week −1	From Day −13 to Day −7
Double-blind period	At baseline	From Day −6 to Day 1
	At Weeks 1 to 6	From Day 7 × x − 5 to Day 7 × x + 1[b)]
Extension period	At Weeks 7 to 58	From Day 7 × x − 5 to Day 7 × x + 1[b)]

a) The start date of the double-blind period will be regarded as Day 1 and the previous day of the start date of the double-blind period will be regarded as Day −1.
b) x: Number of weeks

Table 11.1-2 Assessment time point and its time window in the follow-up period

Period	Assessment time point	Time window[a)]
Follow-up period	At Week 1	From Day 1 to Day 7

a) The start date of the follow-up period (the completion date of the extension period or the next day of the date of discontinuation) will be regarded as Day 1.

11.1.1 Primary variable

Change from baseline in the mean NRS score at Week 4 of the double-blind period

11.1.1.1 Primary analysis

Using the data obtained up to Week 4 of the double-blind period, the superiority of the MR13A9 0.5 μg/kg group to the placebo group will be confirmed by performing the analysis using a mixed-effects model for repeated measures (MMRM) with change from baseline in the mean NRS score at each time point as an objective variable; treatment group, time point, and treatment group-by-time point interaction as fixed effects; baseline mean NRS score and the

presence or absence of prior treatment with nalfurafine hydrochloride as a dynamic allocation factor as a covariate; and subject as a random effect.
The analysis will include all available data obtained at each time point from Week 1 to Week 4 of the double-blind period as objective variables. Estimation will be performed using a restricted maximum likelihood method. An unstructured covariance structure will be used to estimate error variance. If the unstructured covariance structure fails to provide convergence, the structure that minimizes Akaike Information Criterion (AIC), among the Toeplitz, first-order autoregressive, and compound symmetry structures, will be used. The degree of freedom will be adjusted with the Kenward-Roger method.
The number of subjects, adjusted mean of the change, standard error, and two-sided 95% confidence interval in each group at each time point will be presented, and the adjusted mean and standard error will be presented in graphs. The adjusted mean between-group difference in the change at each time point for the placebo group and the MR13A9 0.5 μg/kg groups and its two-sided 95% confidence interval as well as the P value will be presented.
The assessment time points are Weeks 1 to 4.

11.1.1.2 Sensitivity analysis

Sensitivity analyses will be performed to determine the robustness of the analysis results. In the individual sensitivity analyses, treatment effects will be estimated as with primary analysis. However, this is not presented in graphs.
Multiple imputation (MI) will be performed 100 times.

1) MI - MMRM

   Multiple imputed data will be generated in the procedures shown below. Each imputation data will be analyzed using the same MMRM model as that for the primary analysis, and the results of the analyses will be combined. Imputation will be performed for all available data obtained at each time point between the baseline and Week 4 of the double-blind period.

   (1) Non-monotone missing data will be imputed with a Markov Chain Monte Carlo (MCMC) method.

   (2) Monotone missing data will be imputed with an imputation formula generated using a regression model developed based on the data from all the groups.

2) Placebo MI - MMRM

   Multiple imputed data will be generated in the procedures shown below. Each imputation data will be analyzed using the same MMRM model as that for the primary analysis, and the results of the analyses will be combined. Imputation will be performed for all available data obtained at each time point between the baseline and Week 4 of the double-blind period.

   (1) Non-monotone missing data will be imputed with an MCMC method.

   (2) Monotone missing data will be imputed with an imputation formula generated using a regression model developed based on the data from the placebo group.

11.1.2 Other assessment variables

11.1.2.1 Change from baseline in the mean NRS score

Summary statistics will be presented. A one-sample t-test will be used for comparisons within each group. For the double-blind period, a two-sample t-test will be used for comparisons between the placebo group and the MR13A9 0.5 μg/kg group. Means and standard deviations will be presented graphically.
The assessment time points are as follows:

- Double-blind period: Weeks 1 to 6
- Entire study period: Weeks 1 to 58, Week 1 of the follow-up period

11.1.2.2 Mean NRS score

Summary statistics will be presented. Means and standard deviations will be presented graphically.
The assessment time points are as follows:

- Double-blind period: Week −1, baseline, Weeks 1 to 6
- Entire study period: Week −1, baseline, Weeks 1 to 58, Week 1 of the follow-up period

11.1.2.3 Percentage of subjects with improvement in the mean NRS score

The number and percentage of subjects with 3-point improvement (change from baseline in the mean NRS score is ≤ −3) and subjects with 4-point improvement (change from baseline in the mean NRS score is ≤ −4) will be presented. For the double-blind period, a Fisher's exact test will be used for comparisons between the placebo group and the MR13A9 0.5 μg/kg group. The percentages of responders will be presented graphically.
The assessment time points are as follows:

- Double-blind period: Weeks 1 to 6
- Entire study period:
  - Table: Weeks 1 to 58, Week 1 of the follow-up period
  - Figure: Weeks 4, 6, 10, 18, 26, 34, 46, and 58

11.1.2.4 Change from baseline in the daily NRS score

The same analysis as described in "11.1.1.1 Primary analysis" will be performed using the data obtained from Days 2 to 15 of the double-blind period. The analysis will include all available data obtained at each time point from Days 2 to 15 of the double-blind period as objective variables.
The assessment time points are as follows:

- Double-blind period: Days 2 to 15

11.1.2.5 Daily NRS score

Summary statistics will be presented. Means and standard deviations will be presented graphically.

The assessment time points are as follows:

- Double-blind period: Days −6 to 15

11.2 Secondary Endpoints

11.2.1 Itch score based on the Shiratori's severity criteria

Higher of daytime or night-time score, daytime score, and night-time score will be analyzed.
The mean itch score and the change from baseline at each time point and the change from baseline in daily itch score will be calculated using the same calculation formulas as those for the mean NRS score, change from baseline in NRS score, and change from baseline in daily NRS score as described in "11.1 Primary Endpoint."

11.2.1.1 Change from baseline in the mean itch score

The same analysis as "11.1.1.1 Primary analysis" will be performed. However, the covariate of the model will be the baseline of the data to be analyzed.
The same analysis as "11.1.2.1 Change from baseline in the mean NRS score" will be performed.
The assessment time points are as follows:

- Analysis using a model: Weeks 1 to 4
- Other
  - Double-blind period: Weeks 1 to 6
  - Entire study period: Weeks 1 to 58, Week 1 of the follow-up period

11.2.1.2 Mean itch score

The same analysis as "11.1.2.2 Mean NRS score" will be performed.
The assessment time points are as follows:

- Double-blind period: Week −1, baseline, Weeks 1 to 6
- Entire study period: Week −1, baseline, Weeks 1 to 58, Week 1 of the follow-up period

11.2.1.3 Change from baseline in daily itch score

The same analysis as "11.1.2.4 Change from baseline in the daily NRS score" will be performed. However, the covariate of the model will be the baseline of the data to be analyzed.
The assessment time points are as follows:

- Double-blind period: Days 2 to 15

11.2.1.4 Daily itch score

The same analysis as "11.1.2.5 Daily NRS score" will be performed.
The assessment time points are as follows:

- Double-blind period: Days −6 to 15

11.2.1.5 Percentage of subjects with improvement in night-time mean itch score

Night-time mean itch score will be analyzed.

For the subjects with a baseline mean itch score of 3 or more, the number and percentage of subjects with improvement (the mean itch score is ≤ 2) will be presented. For the double-blind period, a Fisher's exact test will be used for comparisons between the placebo group and the MR13A9 0.5 µg/kg group. The percentages of responders will be presented graphically.
The assessment time points are as follows:

- Double-blind period: Weeks 1 to 6
- Entire study period:
  - Table: Weeks 1 to 58, Week 1 of the follow-up period
  - Figure: Weeks 4, 6, 10, 18, 26, 34, 46, and 58

11.2.2 Skindex-16

The overall score and subscores (symptoms score, emotions score, and functioning score) will be analyzed.

11.2.2.1 Changes from baseline in Skindex-16

Analysis of covariance with treatment group as a fixed effect and the baseline score as a covariate will be performed on the data obtained at the end of the double-blind period. The number of subjects and adjusted mean change for each group, with standard error and two-sided 95% confidence interval, will be presented. The adjusted mean between-group difference in the change for the placebo group and the MR13A9 0.5 µg/kg groups, with standard error and two-sided 95% confidence interval, as well as the P value will be presented.
The same analysis as "11.1.2.1 Change from baseline in the mean NRS score" will be performed.
The assessment time points are as follows:

- Analysis of covariance: At the end of the double-blind period
- Other
  - Double-blind period: Week 4, at the end the double-blind period
  - Entire study period: Weeks 4, 10, 18, 26, 34, 46, and 58

11.2.2.2 Skindex-16

The same analysis as "11.1.2.2 Mean NRS score" will be performed.
The assessment time points are as follows:

- Double-blind period: Baseline, Week 4, at the end the double-blind period
- Entire study period: Baseline, Weeks 4, 10, 18, 26, 34, 46, and 58

11.2.3 5-D Itch Scale

The overall score and component scores (duration, degree, direction, disability, and distribution scores) will be analyzed.

11.2.3.1 Changes from baseline in 5-D itch scale

The same analysis as "11.2.2.1 Changes from baseline in Skindex-16" will be performed.
The assessment time points are as follows:

- Analysis of covariance: At the end of the double-blind period
- Other
  - Double-blind period: Week 4, at the end the double-blind period
  - Entire study period: Weeks 4, 10, 18, 26, 34, 46, and 58

11.2.3.2 5-D Itch Scale

The same analysis as "11.1.2.2 Mean NRS score" will be performed.
The assessment time points are as follows:

- Double-blind period: Baseline, Week 4, at the end the double-blind period
- Entire study period: Baseline, Weeks 4, 10, 18, 26, 34, 46, and 58

11.2.4 PGIC

The number and percentage of subjects with global symptoms will be presented. For the double-blind period, a two-sample Wilcoxon test will be used for comparisons between the placebo group and the MR13A9 0.5 µg/kg group. The percentage of global symptoms will be presented graphically.
The assessment time points are as follows:

- Double-blind period: Week 4, at the end the double-blind period
- Entire study period: Weeks 4, 10, 18, 26, 34, 46, and 58

11.3 Efficacy Subgroups

The same analyses as described in "11.1.1.1 Primary analysis" and "11.1.2 Other assessment variables" will be performed on data from the double-blind period in each of the subgroups defined by prior treatment with nalfurafine hydrochloride (no, yes) and mean NRS score at baseline ($< 6, \geq 6$). For the subgroup defined by prior treatment with nalfurafine hydrochloride, only the mean NRS score at baseline will be used as a covariate in the model.

12. Safety

12.1 Adverse Events and Adverse Drug Reactions

Unless otherwise specified, the following events will be tabulated.

- Double-blind period
  - Events occurring between the start of study treatment and the end of the double-blind period
  - Adverse events of special interest occurring between the start of study treatment and the end of the double-blind period
- Entire study period
  - Events occurring between the start of study treatment and the end of the follow-up period

- Events occurring between the start of study treatment and the end of the extension period
- Adverse events of special interest occurring between the start of study treatment and the end of the follow-up period
- Adverse events of special interest occurring between the start of study treatment and the end of the extension period

12.1.1 Incidence of adverse events and adverse drug reactions

The number of events, the number of subjects experiencing events, the incidence of events and its two-sided 95% confidence interval will be presented. For the double-blind period, a Fisher's exact test will be used for comparisons between the placebo group and the MR13A9 0.5 µg/kg group. The difference in the incidence between the placebo group and the MR13A9 0.5 µg/kg groups and its two-sided 95% confidence interval will be presented.

12.1.2 Incidences of adverse events and adverse drug reactions (serious adverse events, adverse events leading to interruption/discontinuation of treatment, and adverse events of special interest)

The events to be tabulated are as follows.

- Double-blind period: Events occurring between the start of study treatment and the end of the double-blind period
- Entire study period
  - Events occurring between the start of study treatment and the end of the follow-up period
  - Events occurring between the start of study treatment and the end of the extension period

The number of events, the number of subjects experiencing events, and the incidence of events will be presented for all events, events leading to death, serious events other than death, events leading to interruption, and events leading to discontinuation.

12.1.3 Occurrence of adverse events and adverse drug reactions

The number of subjects experiencing events and the incidence of events will be presented for all events and by primary SOC and PT.

12.1.4 Occurrence of adverse events and adverse drug reactions (by severity)

The numbers of events by severity will be presented for all events and by primary SOC and PT.

12.1.5 Occurrence of adverse events and adverse drug reactions (by time of onset)

The events to be tabulated are shown in Table 12.1.5-1.
The time to onset of adverse events will be classified into the time categories as shown in Table 12.1.5-2, and the number of subjects experiencing events and the incidences of events by primary SOC and PT will be presented for each category for all events, events leading to death, serious events other than death, events leading to interruption, and events leading to discontinuation. The number of subjects experiencing events and the incidence of events

throughout the entire period, regardless of the time of onset, will also be presented. Discontinued subjects will be included in the denominator when calculating incidence for the time categories included in the tabulation periods shown in Table 12.1.5-1, but will be excluded from the denominator for the later time categories.

Table 12.1.5-1 Events to be tabulated

Period	Tabulation time periods	Events to be tabulated
Entire study period	From the start of study treatment to the end of the follow-up period	• Events occurring between the start of study treatment and the end of the follow-up period • Adverse events of special interest occurring between the start of study treatment and the end of the follow-up period
	From the start of study treatment to the end of the extension period	• Events occurring between the start of study treatment and the end of the extension period • Adverse events of special interest occurring between the start of study treatment and the end of the extension period

Table 12.1.5-2 Time category

Time category (unit: weeks)	Time to onset of adverse event (unit: days)
< 6	< 43
< 12	< 85
12 ≤ < 24	85 ≤ < 169
24 ≤ < 36	169 ≤ < 253
36 ≤ < 48	253 ≤ < 337
48 ≤	337 ≤

12.1.6 Details of adverse events of special interest

The events to be tabulated are as follows.

- Double-blind period: Adverse events of special interest occurring between the start of study treatment and the end of the double-blind period
- Entire study period
  - Adverse events of special interest occurring between the start of study treatment and the end of the follow-up period
  - Adverse events of special interest occurring between the start of study treatment and the end of the extension period

The number of events, the number of subjects experiencing events, and incidence of adverse events and adverse drug reactions will be presented. The number of events will be presented by severity (mild, moderate, or severe), action taken with study treatment (drug withdrawn, drug interrupted, dose not changed, unknown, or not applicable), outcome (recovered/resolved, recovering/resolving, not recovered/not resolved, recovered/resolved with sequelae, results in death, or unknown), and time of resolution (during the double-blind period, during the extension period, or after entering the follow-up period).

12.2 Laboratory Tests

Red blood cell count, hemoglobin, hematocrit, white blood cell count, differential white blood cell count (neutrophil, eosinophil, basophil, monocyte, lymphocyte), platelet count, AST, ALT,

γ-GTP, CRP, LDH, ALP, total protein, albumin, glycoalbumin, total cholesterol, total bilirubin, direct bilirubin, creatinine, BUN (before dialysis), Na, K, Cl, Ca, P, serum iron, UIBC, TIBC, TSAT, ferritin, testosterone, free testosterone, prolactin, TSH, FT3, FT4, intact-PTH, and antidiuretic hormone will be analyzed.

12.2.1 Laboratory tests

Summary statistics will be presented.
The assessment time points are as follows:

- Double-blind period: Week −2, baseline, Weeks 1, 2, 4, and 6, and at the end of the double-blind period
- Entire study period: Week −2, baseline, Weeks 1, 2, 4, 6, 7, 8, 10, 12, 18, 26, 34, 46, and 58, and at the end of the extension period

12.2.2 Shift table of laboratory data

The numbers and percentages of subjects with values (low, normal, or high) before and after treatment will be presented.
The assessment time points are as follows:

- Double-blind period: Baseline and at the end the double-blind period
- Entire study period: Baseline and at the end of the extension period

12.2.3 Scatterplots of laboratory data

Scatterplots of measurements before and after treatment will be presented.
The assessment time points are as follows:

- Double-blind period: Baseline and at the end the double-blind period
- Entire study period: Baseline and at the end of the extension period

12.3 Vital Signs and Body Weight

Analysis items are shown below.

- Vital signs: Systolic blood pressure, diastolic blood pressure, pulse rate, and body temperature
- Body weight (before dialysis)

12.3.1 Vital signs and body weight

Summary statistics will be presented.
The assessment time points are as follows:

- Vital signs
  - Double-blind period: Week −2, baseline, Weeks 1, 2, 4, and 6, and at the end of the double-blind period
  - Entire study period: Week −2, baseline, Weeks 1, 2, 4, 6, 7, 8, 10, 12, 18, 26, 34, 46, and 58, and at the end of the extension period

- Body weight
  - Double-blind period: Week −2, baseline, Week 6, and at the end of the double-blind period
  - Entire study period: Week −2, baseline, Weeks 6, 12, 34, and 58, and at the end of the extension period

12.3.2 Scatter plots of vital signs and body weight

Scatterplots of measurements before and after treatment will be presented.
The assessment time points are as follows:

- Double-blind period: Baseline and at the end the double-blind period
- Entire study period: Baseline and at the end of the extension period

12.4 Dependency Assessment

12.4.1 Evaluation of dependency based on judgment by dependency assessment members

The number and percentage of subjects will be presented for dependency (negative, positive, not evaluable) and severity (mild, moderate, severe) at the interim and final assessments. However, only subjects with dependency will be included in the tabulation of severity.

12.5 Safety Subgroups

The same analyses as described in "12.1.1 Incidence of adverse events and adverse drug reactions" will be performed for each of the subgroups defined by sex (male, female), age (< 65 years, ≥ 65 years), prior treatment with nalfurafine hydrochloride (no, yes), and specific signs or symptoms occurring during the screening period (no, yes).

12.6 Tabulated Lists of Safety Variables

For subjects who entered the double-blind period, lists of death, serious adverse events other than death, adverse events leading to interruption, adverse events leading to discontinuation, adverse events of special interest, and abnormal laboratory values by subject will be prepared. However, the lists will not be prepared if there is no applicable event.

13. Software Used for Analysis

SAS System Release 9.4 for Windows (SAS Institute Inc.) will be used to conduct the analysis. Other statistical analysis software will also be used as needed.

14. Tables, Figures and Listing Shells

The TFL shells will be provided separately.

15. Changes from the Clinical Study Protocol

1) Changes in the statistical analysis plan version 1.0

Not applicable.

2) Changes in the statistical analysis plan version 2.0

(1) It has been decided to conduct 2 interim tabulations.
Reason: To clarify the contents and timing of analysis.

(2) In "9.1 Subject Demographics," the mean NRS score (ordinal scale variable) has been added to the analysis items.
Reason: To perform a more detailed evaluation of the distribution of the mean NRS score at baseline.

(3) In "10.5 Change in Dry Weight Category," the status of change from the last time of determination of study drug dose has been added, and the text has been modified to clarify the content of the analysis.
Reason: To perform a more detailed evaluation of the status of change in dry weight category.

(4) In "11.1.2.1 Change from baseline in the mean NRS score," "11.2.2.1 Changes from baseline in Skindex-16," and "11.2.3.1 Changes from baseline in 5-D itch scale," the text has been revised to add graphical presentations for the double-blind period.
Reason: To perform a more detailed visual assessment of the efficacy endpoints in the double-blind period.

(5) In "11.1.2.3 Percentage of subjects with improvement in the mean NRS score," the assessment time points in the figure for the entire study period have been limited to some of the time points.
Reason: To perform visual assessment at the limited time points.

(6) In "11.1.2.4 Change from baseline in the daily NRS score," "11.1.2.5 Daily NRS score," "11.2.1.3 Change from baseline in daily itch score," and "11.2.1.4 Daily itch score," the assessment time periods have been extended up to Day 15.
Reason: To perform a more detailed evaluation of the efficacy in the early stage of treatment.

(7) In "11.1.2.4 Change from baseline in the daily NRS score" and "11.2.1.3 Change from baseline in daily itch score," the text has been updated to include MMRM analysis as performed in the primary analysis.
Reason: To perform an analysis that takes into account the influence of covariates and correlations between time points.

(8) In "11.2.3 5-D Itch Scale," component scores (duration, degree, direction, disability, and distribution scores) have been added to the analysis items.
Reason: To perform a more detailed evaluation of the 5-D Itch Scale data.

(9) In "11.2.2.1 Changes from baseline in Skindex-16" and "11.2.3.1 Changes from baseline in 5-D itch scale," the variables to be displayed for analysis of covariance have been described.
Reason: To clarify the variables to be displayed.

(10) In "11.3 Efficacy Subgroups," the same analysis as described in "11.1.2 Other assessment variables" has been added.
Reason: To perform a more detailed evaluation of the NRS score in subgroups.

(11) In "11.3 Efficacy Subgroups," a subgroup defined by the mean NRS score at baseline

(< 6, ≥ 6) has been added.
Reason: To evaluate whether efficacy varies depending on the mean baseline NRS score.

(12) In "12.1.6 Details of adverse events of special interest," "adverse events of special interest occurring between the start of study treatment and the end of the extension period" has been added to the events to be tabulated in the entire study period.
Reason: To perform a more detailed evaluation of adverse events of special interest.

(13) In "12.4.1 Evaluation of dependency based on judgment by dependency assessment members," the description of the target assessments (interim and final) has been added.
Reason: To clarify the target assessments.

3) Changes in the statistical analysis plan version 3.0

(1) "11.2.1.5 Percentage of subjects with improvement in night-time mean itch score" has been added.
Reason: To evaluate in more detail of the improvement in night-time itching.

16. History of Revision of Statistical Analysis Plan

Version	Date of creation/revision	Author	Description
1.0	January 7, 2021	Shota Okamura	First version prepared
2.0	November 16, 2021	Shota Okamura	Revised version prepared
3.0	October 24, 2022	Shota Okamura	Revised version prepared

17. Literature References

1) Higaki Y. Skindex-16 (Japanese version): A skin disease-specific quality of life scale. Tokyo. Medical Professional Relations Inc. 2002.

2) Ebata T, Ishiuji Y, Saeki H. et al. Development of the Japanese Version of the 5-D itch scale, Japanese Journal of Dermatology 2015; 125, 1035-40.

CERTIFICATE OF TRANSLATION

This document is the translated version of the source document below written in Japanese.
It is hereby certified that this document is a true and accurate translation of the source document from Japanese into English.

Source document:	MR13A9-5 Statistical Analysis Plan (Ver 3.0) File name: mr13a9-5-s-jp-p-1.pdf
Name of representative:	[redacted]
Signature and Date of signature:	Refer to Signature Page

SIGNATURE PAGE

Document ID: 090186a1801a35ba

User	Date(UTC)	Reason
[redacted]	2023/06/22 03:21:50	治験総括責任者として承認する。