This is the translated version of the Statistical Analysis Plan written in Japanese.

Statistical Analysis Plan

Study Title:	A Phase III Confirmatory Study of KLH-2109 in Uterine Fibroid Patients with Menorrhagia and Pain
Protocol Number:	KLH2302
Version:	2.0
Date of Preparation/Revision:	June 26, 2024
Sponsor:	Kissei Pharmaceutical Co., Ltd.
NCT Number	NCT05445167

Table of Contents

1. Purpose

The purpose of this statistical analysis plan is to detail the statistical analyses that were described in the study protocol, version 2.0.

2. Analysis Set

Analysis sets are defined below. When the full analysis set (FAS) and the per protocol set (PPS) are used as the analysis set, the analysis will be performed based on the assigned treatment. When the pharmacokinetic analysis set (PKS) and safety set (SS) are used as the analysis set, the analysis will be performed based on the actual treatment received.

1) FAS

Population of subjects who are randomized and received at least one dose of the investigational product for the treatment period

2) PPS

Population excluded the following subjects from FAS:

- Subjects who do not meet the inclusion criteria
- Subjects who meet the exclusion criteria that are related to efficacy evaluation
- Subjects who withdraw from the study during the treatment period
- Subjects who have a compliance rate of <75% for the investigational product in the treatment period
- Subjects who have an input rate of <75% for PBAC (bleeding or not) or NRS scores into the symptom diary during the treatment period
- Subjects who receive prohibited concomitant medications or prohibited concomitant therapy before the end of the treatment period

3) PKS

Population of subjects who receive at least one dose of KLH-2109 and have plasma concentration data at least one time point

4) SS

Population of subjects who receive at least one dose of the investigational product for the treatment period

3. Analysis Groups

The analysis sets and subjects for the analysis groups are shown in the following table.

Analysis sets and subjects

Analysis Groups	Analysis sets	Subjects
KLH-2109 200 mg group	FAS, PPS	Subjects who are assigned to KLH-2109 200 mg
	PKS, SS	Subjects who receive KLH-2109 200 mg
Placebo group	FAS, PPS	Subjects who are assigned to placebo
	SS	Subjects who receive placebo

4. General principles for statistical analysis

- The two-sided significance level of 5% will be used for the statistical test. However, the two-sided significance level of 15% will be used for the analysis of the imbalance between groups in subject demographics.

- Summary statistics will be presented as the number of subjects, mean, standard deviation, minimum, median, maximum, and quartiles.
- Unless otherwise specified, tabulation will be performed by group and by time point.
- Confidence intervals (CIs) for proportion will be presented as exact two-sided 95% CIs (Clopper–Pearson) by group and CIs for between-group differences will be presented as exact two-sided 95% CIs (Chan and Zhang).
- Analysis sets for the endpoints and analysis parameters are specified in the table below.

Analysis sets for endpoints and analysis parameters

Endpoints	Analysis parameters	Analysis sets
Demographic and other baseline characteristics	Subject demographics	FAS, PPS, SS
	Others	SS
Treatment compliance	All parameters	FAS, SS
Efficacy	Primary endpoints	FAS (primary analysis set), PPS
	Others	FAS
Safety	All parameters	SS
Pharmacokinetics	All parameters	PKS
Pharmacodynamic effects	All parameters	FAS

5. Data Handling

5.1 Presentation of digits for calculated values

- P value and statistics

  The value will be rounded down to the third decimal place. However, P values less than 0.001 will be presented as “P<0.001” or “<0.001.”
- Mean, standard deviation (SD), and CI

  The value will be rounded off to one digit after the last place of the significant digit of the data.
- Minimum, median, maximum, and quartile

  The value will be presented by the significant digits of the data.
- Proportion and CI

  The value will be rounded off to the first decimal place.

5.2 Concomitant medications

Drug names will be presented based on WHO-Drug Global (September 2022).

5.3 Complications, adverse events, and adverse drug reactions

Primary SOCs and PTs will be presented based on MedDRA, version 25.1.

5.4 Drug concentrations

Measured values below the limit of quantification will be imputed with 0. In the calculation of geometric means and geometric CVs, if a measured value before log transformation is 0, it will be handled as missing data.

5.5 Clinical laboratory tests and pharmacodynamic tests

Measured values equal to or below the limit of quantification will be imputed with the lower limit of

quantification. Measured values equal to or above the limit of quantification will be imputed with the upper limit of quantification.

5.6 Disease Duration

If the year of the previous diagnosis of uterine fibroid is missing, the disease duration will be considered missing. If only the month of the previous diagnosis of uterine fibroid is missing, it will be imputed with January for calculation.

5.7 PBAC score

If "no bleeding" is entered in the PBAC score (bleeding or not) in the symptom diary, the PBAC score will be handled as 0.

5.8 UFS-QOL score

Any missing answers to questions will be handled based on the UFS-QoL Scoring Manual.

5.9 Myoma volume and uterine volume

If measured values are available for D1 and D2 with the value for D3 missing, the value for D3 will be imputed with the measured value for D2 for calculation.

5.10 Tabulation by time point (except for the symptom diary)

- Data will be summarized by time point based on the allowable windows shown in Table 5.10-1. Data outside of the allowable windows will be excluded from summarization by time point.
- If multiple values are available within an allowable window, the value for the scheduled visit will be employed.
- Observations and tests performed at unscheduled visits will be excluded from tabulation.
- For the treatment period, any observations and tests performed later than 7 days after the last dose will be excluded from tabulation.

Table 5.10-1 Allowable windows for assessment points

Assessment point	Scheduled day [a)]	Allowable windows (days)
Start of the screening period	—	−118 to −26
Start of the run-in period	—	−42 to −21
Start of the treatment period [b)] (baseline)	1	The start date of menstruation at the start of the treatment period to 1
Week 4	29	25 to 33 (scheduled day ±4)
Week 8	57	50 to 64 (scheduled day ±7)
Week 12	85	78 to 92 (scheduled day ±7)
Week 4 of the follow-up period	28 [c)]	21 to 35 (scheduled day ±7)

a) The first day of administration of the investigational product for the treatment period is Day 1, and the day before the first day of administration of the investigational product for the treatment period is Day −1.
b) The allowable window for 3 h after administration is ±30 min.
c) For the follow-up period, the day following Week 12 (or discontinuation of the treatment period) visit is Day 1 of the follow-up period.

5.11 Tabulation at the end of the treatment period

Among the data obtained from the start of administration of the investigational product for the treatment period to Week 12 (or the discontinuation of the treatment period), the data for observations and tests performed at the last scheduled visit within 7 days after the end of administration of the investigational

product are defined as the data at the end of the treatment period.

5.12 Tabulation at the end of the follow-up period

The data for observations and tests conducted at the last scheduled visit during the period from the start to Week 4 of the follow-up period are defined as the data at the end of the follow-up period.

6. Statistical/analytical issues

6.1 Adjustments for covariates

No adjustments for covariates will be performed.

6.2 Handling of dropouts or missing data

All missing data for analysis except for those presented in "5 Data Handling" will be handled as missing data, and no statistical imputation will be performed.

6.3 Interim analyses and data monitoring

No interim analyses and data monitoring will be performed.

6.4 Multicenter study

The analysis will not be performed by site.

6.5 Multiple comparison/multiplicity

Although this study has two primary endpoints, the primary objective of the study will only be achieved if both of the two primary endpoints are confirmed; hence, there will be no multiplicity. No multiplicity will be considered for tests for other items.

6.6 Use of two different analysis sets for efficacy evaluation

The FAS will be used for the primary analysis set for efficacy evaluation. An analysis of the primary endpoint in the PPS will be performed as a supplemental analysis to evaluate the robustness of the results.

7. Definition of derived data

Items will be calculated using the following formulas:

- Disease duration (unit: years; significant digit: 0.1) = ((year of written informed consent × 12 + month of written informed consent) − (year of previous diagnosis of uterine fibroid × 12 + month of previous diagnosis of uterine fibroid))/12
- Treatment duration with the investigational product (unit: days; significant digit: 1) = date of end of administration − date of start of administration + 1
- Treatment compliance of the investigational product (unit: %; significant digit: 0.1) = number of tablets of the investigational product administered/specified number of tablets administered (Treatment duration with the investigational product × 2) × 100
- Total PBAC score (unit: points; significant digit: 1)
  - Assessment periods are defined as baseline (period of 1 menstrual cycle immediately before the start of the treatment period), from Week 2 to Week 6 after administration of the investigational product

(Day 15 to Day 42), from Week 6 to Week 12 after administration of the investigational product (Day 43 to Day 84), and 6 weeks prior to the end of the investigational product administration (42 days immediately prior to the date of the end of administration).

  - The total PBAC score will be calculated in each assessment period.

- Time to a total PBAC score of <10 (unit: days; significant digit: 1) = event onset date − date of start of administration + 1
  - For the period from the date of start of administration through date of end of administration, the total PBAC score will be calculated every 42 days (e.g., Day 1 to Day 42, Day 2 to Day 43, and Day 3 to Day 44). If a total PBAC score of <10 and it has been continually <10 in all later 42-day periods through date of end of administration, this will be considered an event onset (achievement of a total PBAC score of <10). The event onset date is defined as the first date during the first 42 days in which the event onset criterion is met.
  - If subjects discontinued within 42 days after the start of administration or confirmed no event onset, the last date confirmed bleeding based on a PBAC score until the date of end of administration will be considered the censoring date. If no data on PBAC score have been obtained after the start of administration of the investigational product, Day 1 of treatment will be considered the censoring date.
  - Time to amenorrhoea (a total PBAC score of 0) will also be calculated.
- Maximum NRS scores (unit: none; significant digit: 1) and average NRS scores (unit: none; significant digit: 0.1)
  - Assessment periods are defined as baseline (period of 1 menstrual cycle immediately before the start of the treatment period), from Day 1 to Day 28, Day 29 to Day 56, Day 57 to Day 84, and 28 days prior to the end of the investigational product administration.
  - The maximum and average NRS scores will be calculated in each assessment period.
- Proportion of days without symptoms (NRS score for pain symptoms of 0) (unit: %; significant digit: 0.1) = ((Number of days with an NRS score of 0 during the assessment period)/(Number of data on NRS score obtained during assessment period)) × 100
  - The same assessment periods as those for the maximum and average NRS scores will be used.
- Proportion of days with analgesic use (unit: %; significant digit: 0.1) = ((Number of days with analgesic use during assessment period)/(Number of data on analgesic use obtained during assessment period)) × 100
  - The same assessment periods as those for the maximum and average NRS scores will be used.
- UFS-QOL score (unit: none; significant digit: 0.1)
  - Symptom severity, HRQL total and HRQL subscale scores will be calculated according to the UFS-QoL Scoring Manual.
- Myoma volume or uterine volume (unit: cm^3; significant digit: 0.1) = D1 × D2 × D3 × π/6
- QTcF interval (unit: msec; significant digit: 1) = QT interval/RR interval$^{0.33}$
- Duration from last administration of the investigational product to menstrual recovery (unit: days;

significant digit: 1) = date of menstrual recovery − date of last administration of the investigational product

- Time to onset of adverse event (unit: days; significant digit: 1) = date of onset of adverse event − date of start of administration + 1

8. Disposition of subjects

Analyses will be performed on subjects randomly assigned.

- The numbers and proportions of subjects who were included or excluded in the FAS, PPS, SS, and PKS will be presented, and Fisher's exact test will be used to examine between-group differences.
- The numbers and proportions of subjects for presence or absence of discontinuation and completion or not completion of the follow-up period will be presented. For presence or absence of discontinuation, Fisher's exact test will be used to examine between-group differences. Reasons for discontinuation and not completion of the follow-up period will be categorized as follows, and the numbers and proportions of subjects in each category will be presented.
  - Adverse events: Discontinuation according to the subject discontinuation criterion 1) “adverse events.”
  - Lack of efficacy: Discontinuation according to the subject discontinuation criterion 2) “lack of efficacy.”
  - Withdrawal by subject: Discontinuation according to the subject discontinuation criterion 3) “subject's voluntary withdrawal from the study.”
  - Protocol deviation: Discontinuation according to the subject discontinuation criterion 4) “significant deviation from the protocol during the study period.”
  - Pregnancy: Discontinuation according to the subject discontinuation criterion 5) “pregnancy.”
  - Reduction of HGB Concentration: Discontinuation according to the subject discontinuation criterion 9) “blood hemoglobin level decreased to <8 g/dL with or without the use of iron preparation.”
  - Others: Discontinuation due to reasons other than the above categories.
- The numbers and proportions of subjects with important deviations (Inclusion or exclusion criteria/Discontinuation criteria/Dosage and administration/Prohibited concomitant therapy/Serious noncompliance) will be presented.

9. Demographic and other baseline characteristics

9.1 Subject demographics

- Analysis parameters are shown below. If data are available for multiple time points, baseline data will be used for assessment.
  - Nominal scale variables
    - Type of uterine fibroids: subserosal fibroid, intramural fibroid, submucosal fibroid, cervical fibroid
    - Birth experience: No/Yes
    - Smoking history: The subject has never smoked/The subject is a current smoker/The subject is an ex-smoker

- Any medications for uterine fibroids: No/Yes
- Any medications for uterine fibroids: GnRH agonist, GnRH antagonist, herbal medicine, and other medicines for uterine fibroids
- Any surgeries for uterine fibroids: No/Yes

– Ordinal scale variables
- Age: <40 years, ≥40 years
- Body weight: <60 kg, ≥60 kg
- BMI: <25.0 kg/m^2, ≥25.0 kg/m^2
- Myoma volume: ≤170 cm^3, >170 cm^3
- Uterine volume: ≤170 cm^3, >170 cm^3
- Total PBAC score: <200 points, ≥200 points
- Maximum NRS score: ≥4 to <7, ≥7
- UFS-QOL score (symptom severity): ≤25, >25
- UFS-QOL score (HRQL total): ≤75, >75
- Blood hemoglobin: <10 g/dL, ≥10 to <12 g/dL, ≥12 g/dL

– Continuous variables: Age, height, body weight, BMI, disease duration, myoma volume, uterine volume, total PBAC score, maximum NRS score, UFS-QOL score (symptom severity, HRQL total), and blood hemoglobin

- The numbers and proportions of subjects for the nominal and ordinal scale variables, as well as summary statistics for continuous variables, will be presented in Overall and by group.
- Between-group imbalance will be examined using two-sample Wilcoxon test for total PBAC score (continuous variable) and two-sample *t*-tests for maximum NRS score (continuous variable).

9.2 Complications

The number and proportion of subjects will be presented in total and by primary SOC and PT.

9.3 Concomitant medications

9.3.1 Oral iron preparation

For oral iron preparations used from the start of administration of the investigational product for the treatment period to the end of the follow-up period, the number and proportion of subjects will be presented by standardized medication name.

9.3.2 Analgesic drugs

For analgesic drugs used to relieve severe pain associated with uterine fibroids from the start of administration of the investigational product for the treatment period to the end of the follow-up period, the number and proportion of subjects will be presented by standardized medication name.

9.3.3 Drugs used for complications

For drugs used to treat complications from the start of administration of the investigational product for the treatment period to the end of the follow-up period, the number and proportion of subjects will be presented

by standardized medication name.

10. Treatment Compliance

10.1 Treatment compliance

- Summary statistics will be presented.
- Treatment compliance will be classified into <75%, ≥75% to <90%, and ≥90%, and the number, and proportion of subjects in each category will be presented.

10.2 Treatment Duration

- Summary statistics will be presented.
- Treatment duration will be classified into ≥1 day to ≤14 days, ≥15 days to ≤42 days, ≥43 days to ≤84 days, ≥85 days, and the number and proportion of subjects in each category will be presented.

11. Efficacy

11.1 Primary endpoint

The primary endpoints regarding menorrhagia and pain symptoms are shown below. The primary objective of the study will only be achieved if both primary endpoints are confirmed.

- Proportion of subjects with a total PBAC score of <10 from Week 6 to Week 12 after administration of the investigational product
- Proportion of subjects with the maximum NRS score for pain symptoms of ≤1 for 28 days prior to the end of the investigational product administration

11.1.1 Analysis method for the primary endpoint

- The primary endpoint regarding menorrhagia
  - Calculation will be performed using the following formula.
    Calculation formula: Proportion of subjects with a total PBAC score of <10 from Week 6 to Week 12 after administration of the investigational product = ((Number of subjects with a total PBAC score of <10 from Week 6 to Week 12 after administration of the investigational product)/(Number of subjects with an available total PBAC score from Week 6 to Week 12 after administration of the investigational product)) × 100
  - A subject with a total PBAC score of <10 from Week 6 to Week 12 after administration of the investigational product will be considered a subject with improvement. The number and proportion of subjects with improvement and its two-sided 95% CI will be presented by group, and the proportions of subjects with improvement will be graphically displayed. The point estimate of the difference between the KLH-2109 200 mg group and the placebo group (KLH-2109 200 mg group − placebo group) with its two-sided 95% CI will be presented. Fisher's exact test will be used to confirm the superiority of the KLH-2109 200 mg group to the placebo group.
- The primary endpoint regarding pain symptoms
  - Calculation will be performed using the following formula.

Calculation formula: Proportion of subjects with the maximum NRS score for pain symptoms of ≤1 for 28 days prior to the end of the investigational product administration = ((Number of subjects with a maximum NRS score of ≤1 for 28 days prior to the end of the investigational product administration)/(Number of subjects with an available maximum NRS score for 28 days prior to the end of the investigational product administration)) × 100

– A subject with the maximum NRS score for pain symptoms of ≤1 for 28 days prior to the end of the investigational product administration will be considered a subject with improvement. The number and proportion of subjects with improvement and its two-sided 95% CI will be presented by group, and the proportions of subjects with improvement will be graphically displayed. The point estimate of the difference between the KLH-2109 200 mg group and the placebo group (KLH-2109 200 mg group − placebo group) with its two-sided 95% CI will be presented. Fisher's exact test will be used to confirm the superiority of the KLH-2109 200 mg group to the placebo group.

11.2 Secondary endpoints

11.2.1 PBAC score

- Proportion of subjects with a total PBAC score of <10
  - The same formula as for the primary endpoint regarding menorrhagia will be used for calculation.
  - A subject with a total PBAC score of <10 will be considered a subject with improvement, and the number and proportion of subjects with improvement and its two-sided 95% CI will be presented. The point estimate of the difference between the KLH-2109 200 mg group and the placebo group with its two-sided 95% CI will be presented.
  - The proportions of subjects with improvement will be graphically displayed.
  - Assessment time points: Week 2 to Week 6, and 6 weeks prior to the end of the investigational product administration.
- Time to a total PBAC score of <10
  - The Kaplan–Meier method will be used to present the estimated cumulative incidences of events at Weeks 2, 6, and 12 and their 95% CIs (Greenwood's formula).
  - The cumulative incidences of events with overlapped group will be graphically displayed.
  - Estimated percentiles (25th, 50th, and 75th) for time to event onset and its 95% CI (Brookmeyer and Crowley) will be presented.
  - The log-rank test will be used to compare the KLH-2109 200 mg group and the placebo group.
- Proportion of subjects with amenorrhoea (a total PBAC score of 0)
  - The same analyses as for the proportion of subjects with a total PBAC score of <10 will be performed.
  - Assessment time points: Week 2 to Week 6, Week 6 to Week 12, and 6 weeks prior to the end of the investigational product administration.
- Time to amenorrhoea
  - The same analyses as for time to a total PBAC score of <10 will be performed.

11.2.2 NRS score

- Proportion of subjects with a maximum NRS score of ≤1
  - The same formula as for the primary endpoint regarding pain symptoms will be used for calculation.
  - A subject with a maximum NRS score of ≤1 will be considered a subject with improvement, and the number and proportion of subjects with improvement and its two-sided 95% CI will be presented. The point estimate of the difference between the KLH-2109 200 mg group and the placebo group with its two-sided 95% CI will be presented.
  - The proportions of subjects with improvement will be graphically displayed.
  - Assessment time points: Day 1 to Day 28, Day 29 to Day 56, and Day 57 to Day 84
- Proportion of subjects with a maximum NRS score of 0
  - The same analysis as for the proportion of subjects with a maximum NRS score of ≤1 will be performed.
  - Assessment time points: Day 1 to Day 28, Day 29 to Day 56, Day 57 to Day 84, and 28 days prior to the end of the investigational product administration.
- Maximum NRS score
  - Summary statistics will be presented.
  - Means and SDs will be graphically displayed.
  - Assessment time points: Baseline, Day 1 to Day 28, Day 29 to Day 56, Day 57 to Day 84, and 28 days prior to the end of the investigational product administration.
- Change from baseline in the maximum NRS score
  - Summary statistics will be presented. The point estimate of the difference between the KLH-2109 200 mg group and the placebo group with its two-sided 95% CI will be presented.
  - Means and SDs will be graphically displayed.
  - Assessment time points: Day 1 to Day 28, Day 29 to Day 56, Day 57 to Day 84, and 28 days prior to the end of the investigational product administration.
- Average NRS score
  - The same analyses as for the maximum NRS score will be performed.
- Change from baseline in the average NRS score
  - The same analyses as for the change from baseline in the maximum NRS score will be performed.
- Proportion of days without symptoms (days with an NRS score for pain symptoms of 0)
  - The same analyses as for the maximum NRS score will be performed.
- Change from baseline in the proportion of days without symptoms
  - The same analyses as for the change from baseline in the maximum NRS score will be performed.

11.2.3 Use of analgesics

- Proportion of days with analgesic use
  - The same analyses as for the maximum NRS score will be performed.
- Change from baseline in the proportion of days with analgesic use
  - The same analyses as for the change from baseline in the maximum NRS score will be performed.

11.2.4 Blood hemoglobin, blood hematocrit, serum iron, and serum ferritin

- Blood hemoglobin, blood hematocrit, serum iron, and serum ferritin
  - Summary statistics will be presented.
  - Means and SDs will be graphically displayed.
  - Assessment time points (table): The start of the screening period, the start of the run-in period, baseline, Weeks 4, 8, and 12, the end of the treatment period, Week 4 of the follow-up period, and the end of the follow-up period
  - Assessment time points (figure): The start of the screening period, the start of the run-in period, baseline, Weeks 4, 8, and 12, and Week 4 of the follow-up period
- Change from baseline in blood hemoglobin, blood hematocrit, serum iron, and serum ferritin
  - Summary statistics will be presented. For blood hemoglobin, the point estimate of the difference between the KLH-2109 200 mg group and the placebo group with its two-sided 95% CI will be presented.
  - Means and SDs will be graphically displayed.
  - Assessment time points (table): Weeks 4, 8, and 12, the end of the treatment period, Week 4 of the follow-up period, and the end of the follow-up period
  - Assessment time points (figure): Weeks 4, 8, and 12, and Week 4 of the follow-up period
- Frequency tabulation of blood hemoglobin
  - Blood hemoglobin will be classified into <12 and ≥ 12 (unit: g/dL) at each assessment time point, and the number and proportion of subjects and its two-sided 95% CI will be presented.
  - Assessment time points: The start of the screening period, the start of the run-in period, baseline, Weeks 4, 8, and 12, the end of the treatment period, Week 4 of the follow-up period, and the end of the follow-up period
- Frequency tabulation of changes from baseline in blood hemoglobin
  - Changes from baseline in blood hemoglobin will be classified into <1 and ≥ 1 (unit: g/dL) at each assessment time point, and the same analyses as for the frequency tabulation of blood hemoglobin will be performed.
  - Assessment time points: Weeks 4, 8, and 12, the end of the treatment period, Week 4 of the follow-up period, and the end of the follow-up period

11.2.5 Myoma volume and uterine volume

- Myoma volume and uterine volume
  - Summary statistics will be presented.
  - Assessment time points: The start of the screening period, Baseline, Weeks 4, 8, and 12, and the end of the treatment period
- Percent changes from baseline in myoma volume and uterine volume
  - Summary statistics will be presented. The point estimate of the difference between the KLH-2109 200 mg group and the placebo group with its two-sided 95% CI will be presented.
  - Means and SDs will be graphically displayed.

– Assessment time points: Weeks 4, 8, and 12, and the end of the treatment period

11.2.6 UFS-QOL score

Analyses will be performed for symptom severity, HRQL total, and HRQL subscales (concern, activities, energy/mood, control, self-conscious, and sexual function).

- UFS-QOL score
  - Summary statistics will be presented.
  - Assessment time points: Baseline, Week 12, and the end of the treatment period
- Change from baseline in UFS-QOL score
  - Summary statistics will be presented. The point estimate of the difference between the KLH-2109 200 mg group and the placebo group with its two-sided 95% CI will be presented.
  - Assessment time points: Week 12 and the end of the treatment period

11.2.7 PGIC

- General symptoms
  - The numbers and proportions of subjects will be presented. The two-sample Wilcoxon test will be used to compare the KLH-2109 200 mg group and the placebo group.
  - The proportions of subjects will be graphically displayed.
  - Assessment time points: Week 12 and the end of the treatment period

11.3 Subgroup analyses for efficacy

11.3.1 Subgroup analyses for the primary endpoint regarding menorrhagia

The number and proportion of subjects with improvement and its two-sided 95% CI will be provided for each subgroup of age (<40 years, ≥40 years), BMI (<25.0 kg/m^2, ≥25.0 kg/m^2), total PBAC score at baseline (<200 points, ≥200 points), birth experience (No/Yes), uterine volume (≤170 cm^3, >170 cm^3), and type of uterine fibroids (subserosal fibroid, intramural fibroid, submucosal fibroid, and cervical fibroid).

11.3.2 Subgroup analyses for the primary endpoint regarding pain symptoms

The number and proportion of subjects with improvement and its two-sided 95% CI will be provided for each subgroup of age (<40 years, ≥40 years), BMI (<25.0 kg/m^2, ≥25.0 kg/m^2), maximum NRS score at baseline (≥4 to <7, ≥7), birth experience (No/Yes), uterine volume (≤170 cm^3, >170 cm^3), and type of uterine fibroids (subserosal fibroid, intramural fibroid, submucosal fibroid, and cervical fibroid).

11.3.3 Subgroup analyses for blood hemoglobin

The same analyses as in "11.2.4 Blood hemoglobin, blood hematocrit, serum iron, and serum ferritin" will be performed in each subgroup of use of an oral iron preparation (No/Yes), blood hemoglobin level at baseline (<12 g/dL, ≥12 g/dL), and use of an oral iron preparation in subjects with a blood hemoglobin level at baseline of <12 g/dL (No/Yes).

12. Safety

12.1 Adverse events and adverse drug reactions

Events occurring from the start of administration of the investigational product for the treatment period to the end of the follow-up period will be tabulated.

- Incidences of adverse events and adverse drug reactions
  - The number of events, the number of subjects with events, the incidence of events, and its two-sided 95% CIs will be presented.
  - Fisher's exact test will be used to compare the KLH-2109 200 mg group and the placebo group, and between-group differences in incidence and its two-sided 95% CIs will be presented.
- Incidences of adverse events and adverse drug reactions (serious events, events leading to treatment discontinuation, and events leading to treatment interruption)
  - The number of events, the number of subjects with events, and incidences of events will be presented for all events, events leading to death, serious events excluding death, events leading to treatment discontinuation, and events leading to treatment interruption.
- Occurrence of adverse events and adverse drug reactions
  - The numbers of subjects with events and the incidences of events will be presented in total and by primary SOC and PT for all events, events leading to death, serious events excluding death, events leading to treatment discontinuation, and events leading to treatment interruption.
- Occurrence of adverse events and adverse drug reactions (by severity)
  - The numbers of subjects with events and the incidence of events in total and by primary SOC and PT will be presented by highest severity.
- Incidences of adverse events and adverse drug reactions (by time of onset)
  - Time to the onset of an adverse event will be categorized into ≥1 day to ≤28 days, ≥29 days to ≤56 days, ≥57 days to ≤84 days, and ≥85 days, and the numbers of subjects with events and the incidences of events will be presented in all events, events that leading to death, serious events excluding death, events leading to treatment discontinuation, events leading to treatment interruption and by primary SOC and PT for all events. The overall number of subjects with events and the overall incidence of events, irrespective of the time of onset, will also be presented. The denominator for calculation of the incidence will include time category including the end of the follow-up period.

12.2 Clinical laboratory tests

Analyses will be performed for the quantitative variables of red blood cell count, white blood cell count, platelet count, APTT, INR, PT, Na, K, Ca, P, creatinine, total bilirubin, total protein, albumin, AST, ALT, γ-GTP, ALP, creatine kinase, uric acid, urea nitrogen, LDH, HDL-cholesterol, LDL-cholesterol, total cholesterol, triglyceride, urine specific gravity, and urine pH, and the qualitative variables of urine protein, urine sugar, urine urobilinogen, and urine ketone bodies.

- Clinical laboratory tests (quantitative variables)
  - Summary statistics will be presented.

  - Assessment time points: The start of the screening period; the start of the run-in period; baseline; Weeks 4, 8, and 12; the end of the treatment period; Week 4 of the follow-up period; and the end of the follow-up period
- Changes from baseline in clinical laboratory tests (quantitative variables)
  - Summary statistics will be presented.
  - Assessment time points: Weeks 4, 8, and 12; the end of the treatment period; Week 4 of the follow-up period; and the end of the follow-up period
- Clinical laboratory tests (qualitative variables)
  - The numbers and proportions of subjects will be presented for measured values.
  - Assessment time points: The start of the screening period; the start of the run-in period; baseline; Weeks 4, 8, and 12; the end of the treatment period; Week 4 of the follow-up period; and the end of the follow-up period
- Shift tables of clinical laboratory tests
  - The numbers and proportions of subjects will be presented for measured values before and after administration (lower, normal, and higher quantitative values and normal and higher qualitative values) for items with a reference value.
  - Assessment time points: Baseline and the end of the treatment period
- Scatter plots of clinical laboratory tests (quantitative variables)
  - Scatter plots of measured values before and after the administration of the investigational product will be presented.
  - Assessment time points: Baseline and the end of the treatment period

12.3 Vital signs and body weight

Analyses will be performed for vital signs (systolic blood pressure, diastolic blood pressure, pulse rate, and body temperature) and body weight.

- Vital signs and body weight
  - Summary statistics will be presented.
  - Assessment time points: The start of the screening period; the start of the run-in period; baseline; Weeks 4, 8, and 12; the end of the treatment period; Week 4 of the follow-up period; and the end of the follow-up period
- Changes from baseline in vital signs and body weight
  - Summary statistics will be presented.
  - Assessment time points: Weeks 4, 8, and 12; the end of the treatment period; Week 4 of the follow-up period; and the end of the follow-up period
- Scatter plots of vital signs and body weight
  - Scatter plots of measured values before and after the administration of the investigational product will be presented.
  - Assessment time points: Baseline and the end of the treatment period

12.4 12-lead ECG

Analyses will be performed for ECG parameters (RR interval, PR interval, QRS interval, QT interval, QTcF interval, and heart rate).

- ECG parameters
  - Summary statistics will be presented.
  - Assessment time points: The start of the screening period; the start of the run-in period, baseline, the start of the treatment period (3 h after administration); Weeks 4, 8, and 12; the end of the treatment period; Week 4 of the follow-up period; and the end of the follow-up period
- Changes from baseline in ECG parameters
  - Summary statistics will be presented.
  - Assessment time points: The start of the treatment period (3 h after administration); Weeks 4, 8, and 12; the end of the treatment period; Week 4 of the follow-up period; and the end of the follow-up period
- Frequency tabulation of QT interval and QTcF interval
  - For QT interval and QTcF interval, baseline and maximum values after administration of the investigational product will be classified into ≤450, >450 to ≤480, >480 to ≤500, and >500 (unit, msec) for each subject. Cross-tabulation of baseline and the maximum values after administration of the investigational product will be performed to present the numbers and proportions of subjects.
- Frequency tabulation of maximum prolonged time of the QT interval and QTcF interval
  - For QT interval and QTcF interval, the maximum change from baseline (maximum prolonged time) will be determined for each subject, and the numbers and proportions of subjects will be presented for each evaluation category of ≤30, >30 to ≤60, and >60 (unit, msec).
- Scatter plots of QT interval and QTcF interval and plasma drug concentrations
  - Scatter plots of changes from baseline in QT interval and QTcF interval and plasma drug concentrations (KLH-2109 and KP017) will be presented.
  - Assessment time points: The start of the treatment period (3 h after administration) and Weeks 4, 8, and 12

12.5 Duration from the last administration of the investigational product to menstrual recovery

Summary statistics will be presented.

12.6 Subgroup analyses for safety

No analyses will be performed.

12.7 Listings for safety

Listings of deaths, serious adverse events excluding deaths, adverse events leading to treatment discontinuation, adverse events leading to treatment interruption, and abnormal laboratory values by subject will be prepared. A listing, however, will not be prepared if no above-mentioned events have been reported.

13. Pharmacokinetics

Plasma drug concentrations (plasma KLH-2109 [unchanged form] and KP017 [major metabolite] concentrations) will be analyzed.

- Plasma drug concentrations
  - Summary statistics, geometric means, and geometric CVs will be presented.
  - Geometric means and individual values will be graphically displayed.
  - Assessment time points: The start of the treatment period (3 h after administration); Weeks 4, 8, and 12
- Time elapsed after administration
  - Summary statistics will be presented.
  - Assessment time points: The start of the treatment period (3 h after administration); Weeks 4, 8, and 12
- Scatter plots of plasma drug concentrations and times elapsed after administration
  - Scatter plots will be presented using plasma drug concentration on the vertical axis and time elapsed after the most recent administration (hours) on the horizontal axis.
  - Assessment time points: All time points at which plasma drug concentrations are measured

14. Pharmacodynamic effects

Analyses will be performed for pharmacodynamic tests of E2, LH, FSH, and progesterone.

- Pharmacodynamic tests
  - Summary statistics will be presented.
  - Means and SDs will be graphically displayed.
  - Assessment time points (table): Baseline; Weeks 4, 8, and 12; the end of the treatment period; Week 4 of the follow-up period; and the end of the follow-up period
  - Assessment time points (figure): Baseline; Weeks 4, 8, and 12; and Week 4 of the follow-up period
- Changes from baseline in pharmacodynamic tests
  - Summary statistics will be presented.
  - Assessment time points: Weeks 4, 8, and 12; the end of the treatment period; Week 4 of the follow-up period; and the end of the follow-up period
- Time elapsed after administration
  - For KLH-2109 200 mg group, summary statistics will be presented.
  - Assessment time points: Weeks 4, 8, and 12; the end of the treatment period
- Changes over time in pharmacodynamic tests by subject
  - Measured values in individual subjects will be graphically displayed by group.
  - Assessment time points: Baseline; Weeks 4, 8, and 12; and Week 4 of the follow-up period

15. Software for Analyses

SAS System, release 9.4 or later, for Windows (SAS Institute Inc.) will be used for analyses. Other statistical software will be used as necessary.

16. Mock-ups of tables, figures and listing

Details are specified in a separate document.

17. Changes from the Clinical Study Protocol

1) Changes in the statistical analysis plan version 1.0

No changes from the protocol (version 2.0) have been made.

2) Changes in the statistical analysis plan version 2.0

Section	Details of changes	Reason
5.10 Tabulation by time point (except for the symptom diary)	Change of the allowable window for the start of the treatment period	The allowable window was revised.
7. Definition of derived data	Addition of units and significant digits for each item	To clarify units and significant digits.
	Addition of description on how to derive "Time to a total PBAC score of <10"	To clarify the handling of subjects who discontinued within 42 days after the start of administration.
8. Disposition of subjects	Addition of analysis content	To perform the analysis of important deviations.
9.1 Subject demographics	Change of tabulation category for smoking history	To perform the analysis based on the data collection category of smoking history.
	Change of tabulation category for maximum NRS score	Tabulation category was revised.
	Deletion of the description regarding the analysis content of any medications for uterine fibroids	To display the proportions of any medications for uterine fibroids as well.
	Change of the statistical test method used to examine the between-group imbalance in the total PBAC score (continuous variable)	The statistical test method was revised based on blind review.
11.2.2 NRS score	Addition of "Maximum NRS score" and "Change from baseline in the maximum NRS score"	To perform a more detailed evaluation of the maximum NRS score.
11.2.5 Myoma volume and uterine volume	Addition of assessment time point for "Myoma volume and uterine volume"	To perform the analysis based on the time points of data collection.
11.3.2 Subgroup analyses for the primary endpoint regarding pain symptoms	Change of subgroup of maximum NRS score	To correct for the change of tabulation category for the maximum NRS score in “9.1 Subject demographics”
12.1 Adverse events and adverse drug reactions	Change of analysis content of "Occurrence of adverse events and adverse drug reactions (by severity)"	The analysis content was revised.
12.4 12-lead ECG	Addition of analysis content of "Frequency tabulation of QT interval and QTcF interval" and "Frequency tabulation of maximum prolonged time of the QT interval and QTcF interval"	To display the proportions of each evaluation category.
14. Pharmacodynamic effects	Specify the group to be analyzed for "Time elapsed after administration"	To evaluate in the KLH-2109 200 mg group only.
	Change of assessment time points for "Time elapsed after administration"	The assessment time points were revised.

18. Revision history

Version	Date of preparation/revision	Prepared by	Details
1.0	September 28, 2022	[redacted]	Preparation of the first version
2.0	June 26, 2024	[redacted]	Change of analysis plan (refer to 17.2))

19. Appendices

- UFS-QoL Scoring Manual