# STATISTICAL ANALYSIS PLAN

Version 3.0

| | |
|---|---|
| Name of Investigational Product | TAS-303 |
| Title of Study | A phase IIa study of TAS-303 in female patients with stress urinary incontinence |
| Protocol number | 10060050 |
| Date of preparation or revision | September 4, 2018 |
| Author | [redacted] |

History of preparation and revision

| Date | Author | Reason for preparation or revision |
|---|---|---|
| March 20, 2018 | [redacted] | Preparation of the first version |
| May 29, 2018 | [redacted] | Prior to the key code breaking, details of analyses were changed according to results of blinded review. |
| September 4, 2018 | [redacted] | Contents were added because it was an additional analysis. |

Date and Signature

| | | |
|---|---|---|
| [redacted] Person in charge of statistical analysis | MM DD, YYYY | |
| [redacted] Responsible person for statistical analysis | MM DD, YYYY | |
| Taiho Pharmaceutical Co., Ltd. Statistical Team Leader | MM DD, YYYY | |

Changes in Statistical Analysis Plan

Version 1.0 (March 20, 2018) to Version 2.0 (May 29, 2018)

| Item | Before change | After change | Reason for change |
|---|---|---|---|
| 5 | Blinded Review | Blinded Review | Japanese notation was changed. |
| 6.2.1<br>Table 6-2 | – | Age Category 2<br>< 60 years, ≥ 60 years | Category was added according to results of blinded review. |
| 7.3<br>8.1.1<br>8.1.7<br>8.1.8 | Baseline values and allocation adjustment factors (status of urinary incontinence, mean incontinence episode frequency (IEF) per 24 hours [≥ 2 per day, < 2 per day]) were used as covariates. | Allocation adjustment factors (status of urinary incontinence, mean IEF per 24 hours [≥ 2 per day, < 2 per day]), and age (< 60 years and ≥ 60 years) were used as covariates. | Covariates were changed according to results of blinded review. |
| 7.3 | Intergroup comparison will be conducted using the placebo group as the control by the van Elteren test[5)] using allocation adjustment factors as strata. | Intergroup comparison will be conducted using the placebo group as the control by the van Elteren test[5)] using allocation adjustment factors and age (< 60 years and ≥ 60 years) as strata. If any stratum includes a group with 0 patient, the test will be conducted by excluding age from the strata. The van Elteren test will be performed in two patterns, i.e., a pattern in which allocation adjustment factors at the time of enrollment are used as strata (analysis in the FAS only) and a pattern in which allocation adjustment factors recorded in CRF are used as strata. | Stratification factors were added according to results of blinded review. Moreover, details of stratification factors were added. |
| 8.1.1 | The van Elteren test using allocation adjustment factors as strata will be conducted. | The van Elteren test using allocation adjustment factors and age (< 60 years and ≥ 60 years) as strata will be conducted. If any stratum includes a group with 0 patient, the test will be conducted by excluding age from the strata. The van Elteren test will be performed in two patterns, i.e., a pattern in which allocation adjustment factors at the time of enrollment are used as strata (analysis in the FAS only) and a pattern in which allocation adjustment factors recorded in CRF are used as strata. | Stratification factors were added according to results of blinded review. Moreover, details of stratification factors were added. |

| Item | Before change | After change | Reason for change |
|---|---|---|---|
| 8.1.7 | The van Elteren test using allocation adjustment factors as strata will be conducted. The van Elteren test will be performed in two patterns, i.e., a pattern in which allocation adjustment factors at the time of enrollment are used as strata and a pattern in which allocation adjustment factors recorded in CRF are used as strata. | The van Elteren test using allocation adjustment factors and age (< 60 years and ≥ 60 years) as strata will be conducted. If any stratum includes a group with 0 patient, the test will be conducted by excluding age from the strata. The van Elteren test will be performed in two patterns, i.e., a pattern in which allocation adjustment factors at the time of enrollment are used as strata (analysis in the FAS only) and a pattern in which allocation adjustment factors recorded in CRF are used as strata. | Stratification factors were added according to results of blinded review. Moreover, details of stratification factors were added. |
| 8.1.8 | - | For percent changes (Weeks 4 and 8 of treatment) and changes (Weeks 4 and 8 of treatment) from baseline to each time point of evaluation, student's t-test, Wilcoxon's rank sum test, and van Elteren test using allocation adjustment factors and age (< 60 years and ≥ 60 years) as strata will be conducted using the placebo group as the control. If any stratum includes a group with 0 patient, the test will be conducted by excluding age from the strata. The van Elteren test will be performed in two patterns, i.e., a pattern in which allocation adjustment factors at the time of enrollment are used as strata (analysis in the FAS only) and a pattern in which allocation adjustment factors recorded in CRF are used as strata. | Stratification factors were added according to results of blinded review.<br>Since tests for the mean frequency of UUI episode per 24 hours were not mentioned, they were added. |

Version 2.0 (May 29, 2018) to Version 3.0 (September 4, 2018)

| Item | Before change | After change | Reason for change |
|---|---|---|---|
| 8.1.1<br>8.1.2<br>8.1.5<br>8.1.7<br>8.1.8 | Fisher's exact test will be conducted. | Fisher's exact test with placebo as control will be conducted. | Since Fisher's exact test uses the placebo as the control, the statement was added. |
| 8.1.5 | - | The proportion of improvement (Very much better and Much better) at the time of enrollment and Weeks 4 and 8 of treatment and 95% CI will be calculated, and Fisher's exact test with placebo as the control will be conducted. | The statement was added because it was an additional analysis. |
| 8.1.10<br>8.1.11<br>8.1.12<br>8.1.13<br>8.1.14<br>8.1.15<br>8.1.16<br>8.1.17 | - | (Addition of items) | Analytical items were added because it was an additional analysis. |
| 8.3.1 | - | The number of patients and the incidence will be displayed by severity (incidence ≥ 2%) for each item of adverse event. ADRs will be analyzed in the same way. | The statement was added because it was missing. |
| 8.3.3 | - | The incidence of abnormal variations in 12-lead ECG will be calculated. | The statement was added because it was missing. |
| 8.3.3 | - | For amounts of changes in measured values of the QTcF interval and heart rate from values immediately before start of treatment to each evaluation time point, student's t-test will be conducted using the placebo group as the control. | The statement was added because it was an additional analysis. |
| 8.3.4 | - | For amounts of changes in measured values of blood pressure and pulse rate from values immediately before start of treatment to each evaluation time point, student's t-test will be conducted using the placebo group as the control. | The statement was added because it was an additional analysis. |

# TABLE OF CONTENTS

# LIST OF ABBREVIATIONS AND DEFINITION OF TERMS

Abbreviations and definitions used in the Statistical Analysis Plan are as shown below.

| Abbreviation (Term) | Description (Definitions) |
|---|---|
| BMI | Body Mass Index |
| FAS | Full Analysis Set |
| ICH-E3 Guideline | "Structure and Content of Clinical Study Reports"[1)] |
| ICIQ-SF | International Consultation on Incontinence Questionnaire-Short Form |
| I-QOL | Incontinence-Quality of Life |
| LOCF | Last Observation Carried Forward |
| MedDRA | Medical Dictionary for Regulatory Activities |
| MMRM | Mixed-effect Models for Repeated Measures |
| MUI | Mixed Urinary Incontinence |
| PGI-I | Patient Global Impression-Improvement |
| POP | Pelvic Organ Prolapse |
| PPS | Per Protocol Set |
| free $Q_{ave}$ | free $Q_{ave}$ |
| free $Q_{max}$ | free Qmax |
| QOL | Quality of Life |
| QTcF | QT corrected for heart rate by Fridericia's formula |
| SUI | Stress Urinary Incontinence |
| UUI | Urgency Urinary Incontinence |

# 1. SCOPE AND PURPOSE

This plan specifies details of the statistical analysis plan for “a phase IIa study of TAS-303 in female patients with stress urinary incontinence” (Protocol No. 10060050).

# 2. STUDY DESIGN AND TARGET SAMPLE SIZE

## 2.1. Study Design

This clinical trial is a multicenter, randomized, double-blind, placebo-controlled parallel-group study with central enrollment for examination of efficacy and safety of TAS-303 in female patients with SUI. This study consists of a single-blind observation period, a double-blind treatment period, and a follow-up observation period.

### 2.1.1. Primary Endpoint

Percent change in the mean incontinence episode frequency per 24 hours (Week 8 of treatment)

## 2.2. Target Sample Size

### 2.2.1. Target Sample Size

250 to be enrolled in the treatment period

### 2.2.2. Rationale for Sample Size Determination

The number of patients was determined, assuming that efficacy comparable to that of duloxetine can be expected based on results of nonclinical studies. Based on the mean rate of change in the number of urinary incontinence of -62.28% in the duloxetine group and -41.86% in the placebo group from data of a clinical trial of duloxetine, Yentreve clinical study results, Trial ID #6231, the difference between the TAS-303 6-mg or 3 mg group and the placebo group in the rate of change in the number of urinary incontinence, which is the primary endpoint of the present study, was assumed to be -20%. From the standard deviation of 33.73% in the duloxetine group and 32.08% in the placebo group in the above study data, the standard deviation common in two groups was assumed to be 35%. Using the two-sided significance level of the entire test of 5% and the two-sided significance level of each test of 2.5% by Bonferroni’s method, when the statistical power is specified as 90% and the rate of withdrawal in the treatment period as about 5%, the number of patients necessary for confirmation of a significant difference between the TAS-303 group and the placebo group is approximately 250 (83 per group).

# 3. DATA AND ANALYSIS POPULATIONS USED FOR STATISTICAL ANALYSIS

## 3.1. Materials and Data Other Than Case Report Form

Materials and data necessary for statistical analysis other than the case report form are listed in Table 3-1.

**Table 3-1 Materials or Data Necessary for Statistical Analysis Other than Case Report Form**

| Item | Materials or data | Remarks |
|---|---|---|
| Attribution of patients | Flag for attribution of each patient to each analysis population | To identify attribution of each patient to each analysis population. |
| | List of patients excluded from analysis populations with reasons | To list patients excluded from analysis populations |
| Information on randomization | Key code, allocation adjustment factor at the time of allocation | To identify the treatment group and for analyses with the allocation adjustment factor |
| Information on deviation | Details of deviation, significant deviation flag | To list patients with deviations |
| Coding to MedDRA terms | Correspondence table between AEs by investigators and AEs in MedDRA terms (English to Japanese) | To tabulate the incidence of AEs using standardized terms. |
| | Correspondence table with MedDRA terms for medical history and complications | To list medical history and complications in MedDRA terms |
| Institutional reference value | Institutional reference values for laboratory test values | To evaluate normal/abnormal laboratory values |
| Drug name | Correspondence table between concomitant medications by investigator and WHO Drug Dictionary | To tabulate grouped concomitant medications and specify the order |

## 3.2. Analysis Populations

The name of analysis populations and their definitions are shown in Table 3-2. The handling of patients with problems in terms of their attribution apart from the criteria shown below will be determined through discussion by the sponsor, the medical experts, and other personnel. The primary efficacy analysis population is PPS while the primary safety analysis population is all treated population.

**Table 3-2 Definition of Analysis Populations**

| Analysis Populations | Definition |
|---|---|
| All screened population | All patients who gave informed consent |
| All observation period enrolled population | All patients enrolled in the observation period of this study |
| All treatment period enrolled population | All patients enrolled in the treatment period of this study |
| All treated population | Patients in the all observation period enrolled population who have received at least 1 dose of the study drug. |
| Full Analysis Set(FAS) | Patients in the all treatment period enrolled population who have received at least 1 dose of the study drug for the treatment period and have at least 1 available efficacy endpoint before start of the treatment period and during the treatment period. |
| Per Protocol Set(PPS) | Patients in FAS who have not met any of the following criteria:<br>(1) Patients that have violated any inclusion criteria<br>(2) Patients that have met any exclusion criteria (any criteria that affects the efficacy evaluation)<br>(3) Patients whose compliance rate is < 80% during the treatment period<br>(4) Patients whose rate of completion of Bladder diary (status of urinary incontinence) is < 70% at Week 8 of treatment<br>(5) Patients that have received any prohibited concomitant medications or prohibited concomitant therapies that may affect the efficacy evaluation |

# 4. STATISTICAL ANALYSIS POLICY

The purpose of the primary analysis is to evaluate efficacy of TAS-303. To compare the difference between TAS-303 6-mg group and placebo group and between the TAS-303 3 mg group and the placebo group as measured by the percent change in IEF at Week 8 of treatment.

# 5. TIMING OF STATISTICAL ANALYSIS

Statistical analysis is performed after completion of all patients' study. Before unblinding, preliminary examination (blinded review) of measurement of efficacy and safety, distribution of endpoints, relationship among items, and other aspects will be performed to make decisions regarding handling of patient data, statistical analysis methods, and other points.

# 6. COMMON ITEMS BETWEEN ENDPOINTS AND STATISTICAL ANALYSES

## 6.1. Endpoints

The definition of efficacy and safety endpoints are as shown below.

### 6.1.1. Efficacy Endpoints

#### 6.1.1.1. Evaluation of Symptom Score

Endpoints used to evaluate symptom scores are ICIQ-SF and PGI-I.

##### 6.1.1.1.1. ICIQ-SF

Subjective symptoms (frequency of urine leakage, volume of urine leakage, impact on living, and reason for urine leakage) will be evaluated using ICIQ-SF. The total score of ICIQ-SF is defined as the sum of scores for Q1, Q2, and Q3.

##### 6.1.1.1.2. PGI-I

Patients' impression of improvement of urinary incontinence will be evaluated using PGI-I. Improvement of PGI-I is defined as selection of "Very much better," "Much better," or "A little better."

#### 6.1.1.2. Evaluation of QOL Score

QOL associated with urinary incontinence will be evaluated using I-QOL.

The total score of I-QOL is defined as the sum of scores for all questions. In evaluation of I-QOL scores for individual areas, (1) the score for "avoidance and limiting behavior" is defined as the sum of scores for Questions 1, 2, 3, 4, 10, 11, 13, and 20, (2) the score for "psychosocial impacts" is defined as the sum of scores for Questions 5, 6, 7, 9, 15, 16, 17, 21, and 22, and (3) the score for "social embarrassment" is defined as the sum of scores for Questions 8, 12, 14, 18, and 19.

#### 6.1.1.3. Bladder Diary

Time of voiding, presence or absence of urgency, time of urinary incontinence, reason for leakage, and status of use of pads will be evaluated using Bladder diary.

#### 6.1.1.4. One-Hour Pad Test

Amount of urinary incontinence will be evaluated by a 1-hour pad test.

### 6.1.2. Safety Endpoints

#### 6.1.2.1. Adverse Events and Adverse Drug Reactions

An AE is any untoward medical condition that have occurred in a patient while participating in this clinical study. Among adverse events, those for which a causal relationship with study drugs is rated as being "reasonably possible" will be considered to be ADRs.

### 6.1.3. Treatment Compliance Endpoints

#### 6.1.3.1. Compliance Rate

The definition of compliance rate during observation and treatment periods is shown below.

- Observation/Treatment periods discontinued patient

  If there is no dose Modification/Interruption/Discontinuation at all, the compliance rate is 100%, and if "Yes" is selected for dose Modification/Interruption/Discontinuation, the following formula shall apply.

[1] Observation period

Observation period compliance rate (%) = 100*[Total Treatment Days (Dose Modification/ Interruption/ Discontinuation were not included)] / [The larger of 21 and (the last dose date in Observation Period - the first dose date in Observation Period + 1)]

[2] Treatment period

Treatment period compliance rate (%) = 100*[Total Treatment Days (Dose Modification/ Interruption/ Discontinuation were not included)] / [The larger of 56 and (the last dose date in Treatment Period - the first dose date in Treatment Period + 1)]

## 6.2. Common Items in Statistical Analysis

### 6.2.1. Patient Background

The patient background information investigated at the time of enrollment is listed in Table 6-1. Depending on data of each item, appropriate summaries will be provided by calculation of summary statistics or frequency tabulation. Among patient background information shown in Table 6-1, details of the case where background factor of continuous values are categorized are shown in Table 6-2.

**Table 6-1 Background Factors**

| Item | Factor |
|---|---|
| Patient Background | Gender, age, height, body weight, presence or absence of medical history, presence or absence of complication, race, ethnic group |
| Disease Characteristics | BMI, duration of SUI (years), type of urinary incontinence, menopause, delivery, number of deliveries (natural childbirth/caesarean section), menstrual cycle, use incontinence pads, have repair surgery for POP, average number of urinary incontinence episodes per 24 hours, 1-hour pad test |

**Table 6-2 Detailed Category**

| Category | Description |
|---|---|
| Age Category 1 | < 65 years, ≥ 65 years |
| Age Category 2 | < 60 years, ≥ 60 years |
| Race | Asian/Oriental, Other |
| Ethnic group | Japanese, Other |
| Status of urinary incontinence | Stress Urinary Incontinence, Mixed Urinary Incontinence |

### 6.2.2. Statistical Methods

#### 6.2.2.1. Summary Statistics of Continuous Volume

When summary statistics of continuous volume is provided, it should also show the sample size, minimum, maximum, first quartile, median, third quartile, mean, and standard deviation.

#### 6.2.2.2. Estimation Method

(1) Confidence interval of percentage

The confidence interval of the percentage will be calculated by an accurate method based on F-distribution[2)].

The following formula will be used to estimate the upper limit of confidence interval $P_U$ and lower limit $P_L$ of confidence interval. $P$ indicates the percentage, $X$ indicates the number of patients with events, $N$ indicates the analysis set, and $\alpha$ indicates the two-sided significance level.

$$P = X / N$$

$$P_U = \frac{\nu_1 F_{\nu_1,\nu_2}(\alpha/2)}{\nu_2 + \nu_1 F_{\nu_1,\nu_2}(\alpha/2)}, \quad \begin{matrix}\nu_1 = 2(X+1) \\ \nu_2 = 2(N-X)\end{matrix}$$

$$P_L = \frac{\nu_2}{\nu_2 + \nu_1 F_{\nu_1,\nu_2}(\alpha/2)}, \quad \begin{matrix}\nu_1 = 2(N-X+1) \\ \nu_2 = 2X\end{matrix}$$

### 6.2.3. Allocation Adjustment Factors Used for Analysis

In analyses that use allocation adjustment factors, allocation adjustment factors recorded in CRF will be used for analyses in all treated population/PPS. Analyses in FAS will be performed in two patterns, i.e., a pattern in which allocation adjustment factors recorded in CRF are used and a pattern in which allocation adjustment factors entered in IWRS are used.

### 6.2.4. Statistical Analysis Software

In this study, all statistical processing will be performed using SAS Version 9.4 and SAS/STAT 14.2.

# 7. ANALYSES OF PRIMARY ENDPOINT

The below shows the summary of items to be analyzed for the primary endpoint.

## 7.1. Primary Analysis for Primary Objective

In PPS, the difference between the TAS-303 6 mg group and the placebo group and the difference between the TAS-303 3 mg group and the placebo group will be compared for percent change from the baseline to Week 8 in the average number of urinary incontinence episodes per 24 hours using the student's t-test with the level of significance set at two-sided 5% for the entire test. The significance level for each comparison will be adjusted by Hochberg's method[3)], considering multiplicity

## 7.2. Sensitivity Analysis of Primary Analysis

The primary analysis will be performed in FAS. Any missing values will be imputed using the last observation carried forward (LOCF) method. Analyses will be performed in the same fashion if no imputation takes place.

## 7.3. Secondary Analyses of Primary Endpoint

Following analyses will be performed in FAS and PPS:

- Summary statistics will be calculated for each treatment group. Missing values will not be imputed.
- For percent change in the mean IEF per 24 hours from baseline at each week, mixed-effect models for repeated measures (MMRM) will be performed using treatment groups, evaluation time points (Week 4 and Week 8 of treatment), and interactions between the treatment group and the evaluation time point as fixed effects, and allocation factors (status of urinary incontinence [SUI or MUI], mean IEF per 24 hours [≥ 2 per day or < 2 per day]) and age (≥ 60 years or < 60 years) as covariates. Thus, the least mean square (LS Mean) in each treatment group and the difference in the LS mean between each dose group and the placebo group and the two-sided 95% CI will be estimated at Week 8 of treatment. Unstructured covariance structure will be assumed, and toeplitz type will be used in case of no convergence. If convergence is still not achieved, the First-order autoregressive → Compound symmetry → Variance component will be used in this order. The Kenward-Roger method (2009)[4)] will be used to calculate degrees of freedom.
- Intergroup comparison will be conducted using the placebo group as the control by the van Elteren test[5)] using Wilcoxon rank sum test and allocation adjustment factors and age (< 60 years and ≥ 60 years) as strata. If any stratum includes a group with 0 patient, the test will be conducted by excluding age from the strata. The van Elteren

test will be performed in two patterns, i.e., a pattern in which allocation adjustment factors at the time of enrollment are used as strata (analysis in the FAS only) and a pattern in which allocation adjustment factors recorded in CRF are used as strata. Missing data will be imputed using LOCF.

- The time profile of percent change from the baseline will be graphically presented by treatment group.

## 7.4. Analyses Based on Allocation Factors

Following analyses will be performed in PPS:

- Primary endpoints will be analyzed by status of urinary incontinence (SUI, MUI).
- Primary endpoints will be analyzed by mean IEF per 24 hours (≥ 2 per day, < 2 per day).

# 8. ANALYSES OF SECONDARY ENDPOINT

The below shows the summary of analyses for secondary endpoints. A confidence coefficient of 95% will be used to present the confidence interval.

## 8.1. Efficacy Analysis

Following analyses will be performed by treatment group in FAS and PPS:

### 8.1.1. Mean IEF per 24 Hours

- Summary statistics will be calculated for measured values at each evaluation time point (at the time of enrollment, Weeks 4 and 8 of treatment) and percent changes (Week 4 of treatment) as well as changes (Weeks 4 and 8 of treatment) from baseline. Summary statistics of the mean frequency of SUI episode per week will also be calculated.
- For percent changes (Weeks 4 and 8 of treatment) and changes (Weeks 4 and 8 of treatment) from baseline to each time point of evaluation, student's t-test, Wilcoxon's rank sum test, and van Elteren test using allocation adjustment factors and age (< 60 years and ≥ 60 years) as strata will be conducted using the placebo group as the control. If any stratum includes a group with 0 patient, the test will be conducted by excluding age from the strata. The van Elteren test will be performed in two patterns, i.e., a pattern in which allocation adjustment factors at the time of enrollment are used as strata (analysis in the FAS only) and a pattern in which allocation adjustment factors recorded in CRF are used as strata.
- For percent change and changes from baseline at each week, MMRM will be performed using treatment groups, evaluation time points (Weeks 4 and 8 of treatment), and interactions between the treatment group and the evaluation time point as fixed effects, and allocation factors (status of urinary incontinence [SUI or MUI], mean IEF per 24 hours [≥ 2 per day or < 2 per day]) and age (≥ 60 years or < 60 years) as covariates. Thus, the least mean square (LS Mean) in each treatment

group and the difference in the LS Mean between each dose group and the placebo group and the two-sided 95% CI will be estimated at overall evaluation time points. Unstructured covariance structure will be assumed, and toeplitz type will be used in case of no convergence. If convergence is still not achieved, the First-order autoregressive → Compound symmetry → Variance component will be used in this order. The Kenward-Roger method (2009)[4)] will be used to calculate degrees of freedom.

- The percentage of patients with at least 50% reduction from baseline at each time point (Weeks 4 and 8 of treatment) and its 95% confidence interval (CI) will be calculated, and Fisher's exact test with placebo as the control will be conducted.

### 8.1.2. Urinary Incontinence Amount in 1-Hour Pad Test

- Summary statistics will be calculated for measured values at each evaluation time point (at the time of enrollment in the treatment period and Week 8 of treatment) and percent changes (Week 8 of treatment) as well as changes (Week 8 of treatment) from baseline.
- For percent change from baseline and change from baseline to each evaluation time point (Week 8 in the treatment period), student's t-test will be conducted using the placebo group as the control.
- The percentage of patients with an observed value of ≤ 2.0 g at Week 8 of treatment, and its 95% CI will be calculated.
- The percentage of patients with at least 50% reduction from baseline at Week 8 of treatment and its 95% CI will be calculated, and Fisher's exact test with placebo as the control will be conducted.

### 8.1.3. Mean Number of Pads Used per 24 Hours (Only for Patients Who Had Ever Used Incontinence Pads)

- Summary statistics for the observed value (Weeks 4 and 8 of treatment), percent change (Weeks 4 and 8 of treatment) from baseline, and change (Weeks 4 and 8 of treatment) from baseline at each time point will be calculated.
- For percent change from baseline and change from baseline to each evaluation time point (Weeks 4 and 8 of treatment), student's t-test will be conducted using the placebo group as the control.

### 8.1.4. ICIQ-SF score

- Summary statistics for the total score (at the time of enrollment in the observation period, at the time of enrollment in the treatment period, and Weeks 4 and 8 of treatment) and change (Weeks 4 and 8 of treatment) from baseline at each time point will be calculated.
- Summary statistics for the subscale score (at the time of enrollment in the observation period, at the time of enrollment in the treatment period, and Weeks 4 and 8 of

treatment) and change (Weeks 4 and 8 of treatment) from baseline at each time point will be calculated.

- For the change in the total score from baseline to each evaluation time point (Weeks 4 and 8 of treatment), student's t-test will be conducted using the placebo group as the control.
- For the change in the subscale score from baseline to each evaluation time point (Weeks 4 and 8 of treatment), student's t-test will be conducted using the placebo group as the control.

### 8.1.5. Patient Global Impression of Improvement (PGI-I)

- The percentage of patients by score on the 7-point scale assessment at the time of enrollment in the treatment period, and Weeks 4 and 8 of treatment will be calculated.
- The percentage of patients with an improvement (patients selecting "Very much better," "Much better," or "A little better" or patients selecting "Very much better" or "Much better") at the time of enrollment in the treatment period and Weeks 4 and 8 of treatment and its 95% CI will be calculated, and Fisher's exact test with placebo as the control will be conducted.
- The proportion of improvement ("Very much better" and "Much better") at the time of enrollment in the treatment period and Weeks 4 and 8 of treatment and 95% CI will be calculated, and Fisher's exact test with placebo as the control will be conducted.

### 8.1.6. I-QOL Score

- Summary statistics for the total score (at the time of enrollment in the treatment period, and Weeks 4 and 8 of treatment) and change (Weeks 4 and 8 of treatment) from baseline at each time point will be calculated.
- Summary statistics for the domain score (at the time of enrollment in the treatment period and Weeks 4 and 8 of treatment) and change (Weeks 4 and 8 of treatment) from baseline at each time point will be calculated.
- For the change in the total score from baseline to each evaluation time point (Weeks 4 and 8 of treatment), student's t-test will be conducted using the placebo group as the control.
- For the change in the domain score from baseline to each evaluation time point (Weeks 4 and 8 of treatment), student's t-test will be conducted using the placebo group as the control.

### 8.1.7. Mean Frequency of SUI Episode per 24 Hours

Of urinary incontinence episodes documented in the Bladder diary, those excluding episodes because of "Leaks irrepressibly" will be used for following analyses.

- Summary statistics for the observed value (Weeks 4 and 8 of treatment), percent change (Weeks 4 and 8 of treatment) from baseline, and change (Weeks 4 and 8 of treatment) from baseline at each time point will be calculated.
- For percent changes (Weeks 4 and 8 of treatment) and changes (Weeks 4 and 8 of treatment) from baseline to each time point of evaluation, student's t-test, Wilcoxon's rank sum test, and van Elteren test using allocation adjustment factors and age (< 60 years and ≥ 60 years) as strata will be conducted using the placebo group as the control. If any stratum includes a group with 0 patient, the test will be conducted by excluding age from the strata. The van Elteren test will be performed in two patterns, i.e., a pattern in which allocation adjustment factors at the time of enrollment are used as strata (analysis in the FAS only) and a pattern in which allocation adjustment factors recorded in CRF are used as strata.
- For percent change and changes from baseline at each week, MMRM is performed using treatment groups, evaluation time points (Weeks 4 and 8 of treatment), and interactions between the treatment group and the evaluation time point as fixed effects, and allocation factors (status of urinary incontinence, mean IEF per 24 hours [≥ 2 per day or < 2 per day]) and age (≥ 60 years or < 60 years) as covariates. Thus, the least mean square (LS Mean) in each treatment group and the difference in LS Mean between each dose group and the placebo group and the two-sided 95% CI are estimated at overall evaluation time points. Unstructured covariance structure will be assumed, and the toeplitz type will be used in case of no convergence. If convergence is still not achieved, the First-order autoregressive → Compound symmetry → Variance component will be used in this order. The Kenward-Roger method (2009)[4)] will be used to calculate degrees of freedom.
- The percentage of patients with at least 50% reduction from baseline at each time point (Weeks 4 and 8 of treatment) and its 95% confidence interval (CI) will be calculated, and Fisher's exact test with placebo as the control will be conducted.

### 8.1.8. Mean Frequency of UUI Episode per 24 Hours

Of urinary incontinence episodes documented in the Bladder diary, those excluding episodes because of "Leaks when you do physical movements such as coughing, laughing, sneezing, standing, or doing sports" will be used for following analyses.

- Summary statistics for the observed value (Weeks 4 and 8 of treatment), percent change (Weeks 4 and 8 of treatment) from baseline, and change (Weeks 4 and 8 of treatment) from baseline at each time point will be calculated.
- For percent changes (Weeks 4 and 8 of treatment) and changes (Weeks 4 and 8 of treatment) from baseline to each time point of evaluation, student's t-test, Wilcoxon's rank sum test, and van Elteren test using allocation adjustment factors and age (< 60 years and ≥ 60 years) as strata will be conducted using the placebo group as the control. If any stratum includes a group with 0 patient, the test will be conducted by excluding age from the strata. The van Elteren test will be performed in two patterns, i.e., a pattern in which allocation adjustment factors at the time of enrollment are used

as strata (analysis in the FAS only) and a pattern in which allocation adjustment factors recorded in CRF are used as strata.

- For percent change and changes from baseline at each week, MMRM is performed using treatment groups, evaluation time points (Weeks 4 and 8 of treatment), and interactions between the treatment group and the evaluation time point as fixed effects, and allocation factors (status of urinary incontinence, mean IEF per 24 hours [≥ 2 per day or < 2 per day]) and age (≥ 60 years or < 60 years) as covariates. Thus, the least mean square (LS Mean) in each treatment group and the difference in LS Mean between each dose group and the placebo group and the two-sided 95% CI are estimated at overall evaluation time points. Unstructured covariance structure will be assumed, and toeplitz type will be used in case of no convergence. If convergence is still not achieved, the First-order autoregressive → Compound symmetry → Variance component will be used in this order. The Kenward-Roger method (2009)[4)] will be used to calculate degrees of freedom.
- The percentage of patients with at least 50% reduction from baseline at each time point (Weeks 4 and 8 of treatment) and its 95% confidence interval (CI) will be calculated, and Fisher's exact test with placebo as the control will be conducted.

### 8.1.9. Analyses Based on Allocation Factors

Following analyses will be performed in PPS:

- Secondary endpoints will be analyzed by status of urinary incontinence (SUI, MUI) (8.1.1 to 8.1.8).
- Secondary endpoints will be analyzed (8.1.1 to 8.1.8) by mean IEF per 24 hours (≥ 2 per day, < 2 per day).

### 8.1.10. Mean Incontinence Episode Frequency per 24 Hours: Subgroup Analysis

Following analyses will be performed in PPS for each subgroup specified in Table 8.

- Summary statistics for the observed value (Weeks 4 and 8 of treatment), percent change (Weeks 4 and 8 of treatment) from baseline, and change (Weeks 4 and 8 of treatment) from baseline at each time point will be calculated. Summary statistics of the mean frequency of SUI episode per week will also be calculated.
- For percent changes (Weeks 4 and 8 of treatment) and changes (Weeks 4 and 8 of treatment) from baseline to each time point of evaluation, student's t-test, Wilcoxon's rank sum test, and van Elteren test using allocation adjustment factors and age (< 60 years and ≥ 60 years) as strata will be conducted using the placebo group as the control. If any stratum includes a group with 0 patients, the test will be conducted only for the age stratum identified as a prognostic factor in the blinded review. If there is any group with 0 patient, the test will not be conducted.

**Table 8 Subgroup**

| Item | Details of subgroup |
|---|---|
| By timing of enrollment in the treatment period: Quarters and pollen season | "January to March," "April to June," "July to September," "October to December"<br>"January to February (pollen season)" |
| By timing of enrollment in the treatment period: Number of patients enrolled | Months prior to the month in which a half of the patients in the whole PPS are enrolled,<br>Month in which a half of the patients in the whole PPS are enrolled and thereafter |
| Mean IEF per 24 hours at baseline (≥ 2 per day or < 2 per day) × status of urinary incontinence | "< 2 per day and SUI," "< 2 per day and MUI," "≥ 2 per day and SUI," "≥ 2 per day and MUI |
| By measured values in the 1 hour pad test at the time of enrollment in the treatment period | ≤ 5 g, > 5 g<br>≤ 10 g, > 10 g<br>≤ 50 g, > 50 g<br>≤ 100 g, > 100 g<br>> 5 g, ≤ 10 g, > 10 g, ≤ 50 g, > 50 g, ≤ 100 g |
| By ICIQ-F score at the time of enrollment in the treatment period | < 6, ≥ 6<br>< 13, ≥ 13<br>< 19, ≥ 19<br>≥ 6 and < 13, ≥ 13 and < 19, ≥ 19 and ≤ 21 |
| By I-QOL score at the time of enrollment in the treatment period | < 65, ≥ 65<br>< 85, ≥ 85 |
| By baseline mean IEF per 24 hours at the time of enrollment in treatment | < 1.5 times/day, ≥ 1.5 times/day<br>< 2.5 times/day, ≥ 2.5 times/day<br>< 3 times/day, ≥ 3 times/day<br>< 3.5 times/day, ≥ 3.5 times/day<br>< 4 times/day, ≥ 4 times/day |

### 8.1.11. Mean Incontinence Episode Frequency per 24 Hours: Integrated TAS-303 Group

TAS-303 3 mg group and 6 mg group will be integrated, and following analyses will be performed in FAS and PPS: Similar analyses will be performed by mean IEF per 24 hours (< 2 per day, ≥ 2 per day) and status of urinary incontinence (SUI, MUI).

- Summary statistics for the observed value (Weeks 4 and 8 of treatment), percent change (Weeks 4 and 8 of treatment) from baseline, and change (Weeks 4 and 8 of

treatment) from baseline at each time point will be calculated. Summary statistics of the mean frequency of SUI episode per week will also be calculated.

- For percent changes (Weeks 4 and 8 of treatment) and changes (Weeks 4 and 8 of treatment) from baseline to each time point of evaluation, student's t-test, Wilcoxon's rank sum test, and van Elteren test using allocation adjustment factors and age (< 60 years and ≥ 60 years) as strata will be conducted using the placebo group as the control. If any stratum includes a group with 0 patients, the test will be conducted only for the age stratum identified as a prognostic factor in the blinded review. If there is any group with 0 patient, the test will not be conducted.

### 8.1.12. Mean Frequency of SUI Episode per 24 Hours 2

In PPS, data on urinary incontinence only for the reason categorized as "Leaks when you do physical movements such as coughing, laughing, sneezing, standing, or doing sports" recorded in the Bladder diary will be used for following analyses. This will also be performed for the integrated TAS-303 3 mg group and 6 mg group. Similar analyses will be performed by mean IEF per 24 hours (< 2 per day, ≥ 2 per day) and status of urinary incontinence (SUI, MUI).

- Summary statistics for the observed value (Weeks 4 and 8 of treatment), percent change (Weeks 4 and 8 of treatment) from baseline, and change (Weeks 4 and 8 of treatment) from baseline at each time point will be calculated.
- For percent changes (Weeks 4 and 8 of treatment) and changes (Weeks 4 and 8 of treatment) from baseline to each time point of evaluation, student's t-test, Wilcoxon's rank sum test, and van Elteren test using allocation adjustment factors and age (< 60 years and ≥ 60 years) as strata will be conducted using the placebo group as the control. If any stratum includes a group with 0 patients, the test will be conducted only for the age stratum identified as a prognostic factor in the blinded review. If there is any group with 0 patient, the test will not be conducted.

### 8.1.13. Mean Frequency of SUI Episode per 24 Hours 2: Subgroup Analysis

In PPS, data on urinary incontinence only for the reason categorized as "Leaks when you do physical movements such as coughing, laughing, sneezing, standing, or doing sports" recorded in the Bladder diary will be used for following analyses by subgroup specified in Table 8. Similar analyses will be performed by status of urinary incontinence (SUI, MUI).

- Summary statistics for the observed value (Weeks 4 and 8 of treatment), percent change (Weeks 4 and 8 of treatment) from baseline, and change (Weeks 4 and 8 of treatment) from baseline at each time point will be calculated.
- For percent changes (Weeks 4 and 8 of treatment) and changes (Weeks 4 and 8 of treatment) from baseline to each time point of evaluation, student's t-test, Wilcoxon's rank sum test, and van Elteren test using allocation adjustment factors and age (< 60 years and ≥ 60 years) as strata will be conducted using the placebo group as the control. If any stratum includes a group with 0 patients, the test will be conducted only for the age stratum identified as a prognostic factor in the blinded review. If there is any group with 0 patient, the test will not be conducted.

### 8.1.14. Mean Frequency of UUI Episode per 24 Hours 2

In PPS, data on urinary incontinence only for the reason categorized as "Leaks irrepressibly" recorded in the Bladder diary will be used for following analyses. This will also be performed for the integrated TAS-303 3 mg group and 6 mg group. Similar analyses will be performed by mean IEF per 24 hours (< 2 per day, ≥ 2 per day) and status of urinary incontinence (SUI, MUI).

- Summary statistics for the observed value (Weeks 4 and 8 of treatment), percent change (Weeks 4 and 8 of treatment) from baseline, and change (Weeks 4 and 8 of treatment) from baseline at each time point will be calculated.
- For percent changes (Weeks 4 and 8 of treatment) and changes (Weeks 4 and 8 of treatment) from baseline to each time point of evaluation, student's t-test, Wilcoxon's rank sum test, and van Elteren test using allocation adjustment factors and age (< 60 years and ≥ 60 years) as strata will be conducted using the placebo group as the control. If any stratum includes a group with 0 patients, the test will be conducted only for the age stratum identified as a prognostic factor in the blinded review. If there is any group with 0 patient, the test will not be conducted.

### 8.1.15. Mean Frequency of UUI Episode per 24 Hours 2: Subgroup Analysis

In PPS, data on urinary incontinence only for the reason categorized as "Leaks irrepressibly" recorded in the Bladder diary will be used for following analyses by subgroup specified in Table 8. Similar analyses will be performed by status of urinary incontinence (SUI, MUI).

- Summary statistics for the observed value (Weeks 4 and 8 of treatment), percent change (Weeks 4 and 8 of treatment) from baseline, and change (Weeks 4 and 8 of treatment) from baseline at each time point will be calculated.
- For percent changes (Weeks 4 and 8 of treatment) and changes (Weeks 4 and 8 of treatment) from baseline to each time point of evaluation, student's t-test, Wilcoxon's rank sum test, and van Elteren test using allocation adjustment factors and age (< 60 years and ≥ 60 years) as strata will be conducted using the placebo group as the control. If any stratum includes a group with 0 patients, the test will be conducted only for the age stratum identified as a prognostic factor in the blinded review. If there is any group with 0 patient, the test will not be conducted.

### 8.1.16. Mean Frequency of Episodes of Micturition Urgency per 24 Hours

In PPS, data on "urgency" recorded in the Bladder diary will be used for following analyses. Similar analyses will be performed by mean IEF per 24 hours (< 2 per day, ≥ 2 per day) and status of urinary incontinence (SUI, MUI).

- Summary statistics for the observed value (Weeks 4 and 8 of treatment), percent change (Weeks 4 and 8 of treatment) from baseline, and change (Weeks 4 and 8 of treatment) from baseline at each time point will be calculated.
- For percent changes (Weeks 4 and 8 of treatment) and changes (Weeks 4 and 8 of treatment) from baseline to each time point of evaluation, student's t-test, Wilcoxon's rank sum test, and van Elteren test using allocation adjustment factors and age (< 60 years and ≥ 60 years) as strata will be conducted using the placebo group as the

control. If any stratum includes a group with 0 patients, the test will be conducted only for the age stratum identified as a prognostic factor in the blinded review. If there is any group with 0 patient, the test will not be conducted.

### 8.1.17. Mean Leakage Amount per 24 Hours

Leakage amounts will be scored as Small = 2, Moderate = 4, and Large = 6, and summary statistics for the observed value (Weeks 4 and 8 of treatment) and change (Weeks 4 and 8 of treatment) from baseline at each time point will be calculated. For change from baseline to each evaluation time point (Weeks 4 and 8 of treatment), student's t-test will be conducted using the placebo group as the control. Similar analyses will be performed by mean IEF per 24 hours (< 2 per day, ≥ 2 per day) and status of urinary incontinence (SUI, MUI).

## 8.2. Analysis of Administration Status

In all treated population, following analyses will be performed by observation/treatment period and treatment group.

### 8.2.1. Administration Status

Summary statistics will be calculated for the status of administration (total dose and total duration of administration).

### 8.2.2. Status of Administration Completion

- Summary statistics of compliance with treatment with study drugs will be calculated for each patient.
- The presence or absence of study discontinuation (including timing) and its reason will be tabulated.

## 8.3. Safety Analysis

Following analyses will be performed in the all-treated population at each treatment group. However, analyses for the treatment period will be performed in patients in the all-treated population who have received treatment during the treatment period.

### 8.3.1. Adverse Events and Adverse Drug Reactions

Following analyses will be performed for each of the observation and treatment periods:

- The number of patients and the incidence will be calculated. ADRs will be analyzed in the same way.
- The number of patients and the incidence will be displayed by severity for each item of adverse event. ADRs will be analyzed in the same way.
- The number of patients and the incidence will be displayed by severity (incidence ≥ 2%) for each item of adverse event. ADRs will be analyzed in the same way.
- The incidence of AEs in the treatment period will be displayed by time of onset (Weeks 4 and 8 and follow-up). ADRs will be analyzed in the same way.

- All AEs that have occurred from the start of study treatment to the end date of the follow-up will be listed by AE variable of each patient, together with the AE term, severity, date of onset, action, outcome, causal relationship with the study drug, and comments.
- The incidence of serious AEs will be calculated. ADRs will be analyzed in the same way.

### 8.3.2. Laboratory Test Values

- Changes in laboratory parameters (quantitative test) from the baseline will be graphically presented.
- Summary statistics for laboratory test parameters (quantitative test) at each evaluation time point will be calculated.
- Frequency of each laboratory parameter (urine qualitative test) at each evaluation time point will be tabulated.

### 8.3.3. 12-Lead ECG

- Summary statistics for the observed value of the QTcF interval and heart rate at each evaluation time point will be calculated.
- Summary statistics for change from baseline in the observed value of the QTcF interval and heart rate at each evaluation time point will be calculated.
- For amounts of changes in measured values of the QTcF interval and heart rate from values immediately before the start of treatment to each evaluation time point, student's t-test will be conducted using the placebo group as the control.
- The observed value of the QTcF interval at each evaluation time point was divided into categories (≤ 450 msec, > 450 msec and ≤ 480 msec, > 480 msec, and ≤ 500 msec, > 500 msec), and the frequency of each category will be tabulated.
- Change from baseline in the observed value of the QTcF interval and heart rate at each evaluation time point was divided into categories (QTcF interval, ≤ 30 msec, > 30 msec, and ≤ 60 msec, > 60 msec: heart rate <5 bpm, ≤ 5 bpm), and the frequency of each category will be tabulated.
- The incidence of abnormal variations in 12-lead ECG will be calculated.
- A list of patients with abnormal 12-lead ECG changes will be prepared.

### 8.3.4. Blood Pressure, Pulse Rate, and Body Temperature

- Summary statistics for blood pressure, pulse rate, and body temperature at each evaluation time point will be calculated.
- Summary statistics for change from baseline in the observed value of blood pressure and pulse rate at each evaluation time point will be calculated.

- For amounts of changes in measured values of blood pressure and pulse rate from values immediately before the start of treatment to each evaluation time point, student's t-test will be conducted using the placebo group as the control.
- Change from baseline in the observed value of blood pressure and pulse rate at each evaluation time point will be divided into categories (blood pressure, < 10 mmHg, ≥ 10 mmHg: pulse rate, < 5 bpm, ≥ 5 bpm), and the frequency of each category will be tabulated.

### 8.3.5. Uroflowmetry

- Summary statistics for uroflowmetry parameters (free $Q_{max}$, free $Q_{ave}$, voided volume, and voiding time) at the time of enrollment and Week 8 of treatment will be calculated.
- Summary statistics for change from baseline in the observed value of uroflowmetry parameters at Week 8 of treatment will be calculated.
- For amounts of changes in measured values of uroflowmetry parameters at Week 8 from values immediately before the start of treatment to each evaluation time point, student's t-test will be conducted using the placebo group as the control.

### 8.3.6. Residual Urine Volume

- Summary statistics at the time of enrollment in the observation period, at the time of enrollment in the treatment period, and Week 8 of treatment will be calculated.
- Summary statistics for change from baseline in the observed value at Week 8 of treatment will be calculated.
- For amounts of changes in measured values at Week 8 from values immediately before the start of treatment to each evaluation time point, student's t-test will be conducted using the placebo group as the control.

### 8.3.7. Mean Frequency of Micturition per 24 Hours

- Summary statistics at the time of enrollment and Weeks 4 and 8 of treatment will be calculated.
- Summary statistics for change from baseline in the observed value at Weeks 4 and 8 of treatment will be calculated.
- For amounts of changes in measured values at Weeks 4 and 8 from values immediately before the start of treatment to each evaluation time point, student's t-test will be conducted using the placebo group as the control.
- Above analyses will be performed by status of urinary incontinence (SUI, MUI).

# 9. OTHER ANALYSES

## 9.1. Analysis Populations

The proportion of each analysis population and reasons for exclusion will be tabulated by treatment group. Patients excluded from analysis populations will be listed along with reasons.

## 9.2. Deviation

The presence or absence of deviations and deviation items will be tabulated by treatment group. A detailed list of patients with deviations will be presented.

# 10. REVIEW AND REVISION OF STATISTICAL ANALYSIS PLAN

The statistical analysis plan will be finalized before key code breaking.

If the statistical analysis plan is revised, the change log of Statistical Analysis Plan/Table and Figure Plan should be prepared to record the changed content and the reason. Analyses newly planned after key code breaking should be distinguished from those planned beforehand and recorded in the change log.

# 11. STATISTICAL ISSUES

Explanations on ICH-E3 Guidelines[1)] and other statistical issues are summarized below.

## 11.1. Adjustments for Covariates

Adjustments for covariates will be performed using MMRM. Details of covariates will be specified separately.

## 11.2. Handling of Dropouts or Missing Data

Handling of missing values in efficacy endpoints will be specified for each analysis.

Missing values in safety endpoints will not be imputed. The proportion, including that of patients with adverse events, will be calculated using the analysis population as the denominator for the observation period and using the number of patients who have taken study drugs for the treatment period at least once among those in the analysis population as the denominator for the treatment period.

Analyses will include all measured data, except for abnormal laboratory test values with obvious reasons, such as the effect of hemolysis during blood collection. If abnormal values are excluded from the analysis, they will be specified and reasons for exclusion will be stated.

In the case of any post hoc exploratory analysis, the procedure for handling of missing data will be determined appropriately for each item to be analyzed according to the method of analysis and characteristics of items to be evaluated.

## 11.3. Interim Analyses and Data Monitoring

Interim analyses will not be conducted in this study.

## 11.4. Multicenter Studies

This study is a multicenter study.

## 11.5. Multiple Comparisons/Multiplicity

In the analysis of the primary objective, the significance level for each comparison will be adjusted by the Hochberg's method, considering multiplicity.

## 11.6. Use of an "Efficacy Subset" of Patients

Although not planned in this study, additional analyses may be performed as necessary.

## 11.7. Active-Control Studies Intended to Show Equivalence

Not applicable for this study

## 11.8. Examination of Subgroups

In this study, subgroup analyses by allocation adjustment factors are planned for some analyses. Additional analyses will be performed as necessary.

# 12. PREPARATION OF TABLE AND FIGURE PLAN

The statistical analysis plan only specifies major analysis items. All analysis items to be performed will be presented separately in the Table and Figure Plan.

# 13. REFERENCES

(1) "Structure and Content of. Clinical Study Reports" (PAB/ELD Notification No. 335 dated May 1, 1996)

(2) Akira Sakuma, Medical Statistics Q&A, Tokyo, Kanehara & Co., Ltd.; 1987: p.75

(3) Hochberg Y. A sharper Bonferroni procedure for multiple tests of significance. Biometrika.1988;75:800-802.

(4) Kenward, M. G., and Roger, J. H. (2009). "An Improved Approximation to the Precision of Fixed Effects from Restricted Maximum Likelihood." Computational Statistics and Data Analysis 53:2583–2595.

(5) van Elteren, P. H. (1960). "On the combination of independent two-sample tests of Wilcoxon," Bulletin of the International Statistical Institute, 37, 351-361.