STATISTICAL ANALYSIS PLAN

Protocol Number:	URO-901-3005
Protocol Title:	A Phase 3 Double-Blind, Randomized, Placebo-Controlled, Multi-Center Study to Evaluate the Efficacy, Safety and Tolerability of Vibegron in Men with Overactive Bladder (OAB) Symptoms on Pharmacological Therapy for Benign Prostatic Hyperplasia (BPH)
Protocol Version and Date:	Version 4.0; 25 FEB 2022
Investigational Product:	Vibegron (URO-901, RVT-901, MK-4618, KRP-114V)
Indication	OAB symptoms for men with BPH
Development Phase:	Phase 3
US IND Number:	106410
EudraCT Number:	2018-003135-30
Sponsor:	Urovant Sciences, GmbH Gotthardstrasse 3 6300 Zug Switzerland
SAP Version **Effective Date**	v2.0 07-Aug-2023
SAP Author	[redacted], MS Sr. Principal Biostatistician/Director, Urovant Sciences, Inc.

NCT Number: NCT03902080
This NCT number has been applied to the document for purposes of posting on Clinicaltrials.gov

SIGNATURE PAGE

This document has been prepared and/or reviewed by:

DocuSigned by:
[redacted]
Signer Name: [redacted]
Signing Reason: I am the author of this document
Signing Time: 07-Aug-2023 | 06:44 PDT
[redacted], MS — 07-Aug-2023 | 06:44 PDT
Sr. Principal Biostatistician/Director — Date (dd-MMM-yyyy)
Urovant Sciences, Inc

DocuSigned by:
[redacted]
Signer Name: [redacted]
Signing Reason: I approve this document
Signing Time: 07-Aug-2023 | 08:27 PDT
[redacted], MD — 07-Aug-2023 | 08:27 PDT
SVP Clinical Development — Date (dd-MMM-yyyy)
Urovant Sciences, Inc

DocuSigned by:
[redacted]
Signer Name: [redacted]
Signing Reason: I approve this document
Signing Time: 07-Aug-2023 | 07:58 PDT
[redacted], PhD, MS — 07-Aug-2023 | 07:59 PDT
Sr. Director Clinical Development — Date (dd-MMM-yyyy)
Urovant Sciences, Inc

DocuSigned by:
[redacted]
Signer Name: [redacted]
Signing Reason: I approve this document
Signing Time: 07-Aug-2023 | 11:27 PDT
[redacted], BS — 07-Aug-2023 | 11:27 PDT
Director Clinical Operation — Date (dd-MMM-yyyy)
Urovant Sciences, Inc

DocuSigned by:
[redacted]
Signer Name: [redacted]
Signing Reason: I approve this document
Signing Time: 07-Aug-2023 | 09:51 PDT
[redacted], PharmD — 07-Aug-2023 | 09:51 PDT
ssoc ate rector, Pharmacovigilance — Date (dd-MMM-yyyy)
Urovant Sciences, Inc

This document has been approved by

DocuSigned by:
[redacted]
Signer Name: [redacted]
Signing Reason: I approve this document
Signing Time: 07-Aug-2023 | 11:07 PDT
71339BB7B4E54647B77379787D22398F

07-Aug-2023 | 11:07 PDT

[redacted], MS
Sr. Director Biostatistics
Urovant Sciences, Inc

Date (dd-MMM-yyyy)

TABLE OF CONTENTS

TABLE OF CONTENTS....4
LIST OF TABLES....7
SAP VERSION HISTORY....10
1. INTRODUCTION....11
1.1. Study Objectives and Endpoints....11
1.1.1. Primary Efficacy Objectives....11
1.1.2. Secondary Efficacy Objectives....11
1.1.3. Safety Objectives....11
1.1.4. Other Efficacy Objectives....12
1.2. Study Design....14
1.2.1. Overall Study Design....14
1.2.2. Randomization and Blinding....15
1.2.3. Statistical Hypotheses....16
1.2.3.1. Statistical Hypotheses for Co-Primary Efficacy Endpoints....16
1.2.3.2. Statistical Hypotheses for Secondary Efficacy Endpoints....16
1.2.4. Sample Size Justification....16
1.2.4.1. Power in Baseline BPH Treatment Subgroups....18
1.2.4.2. Urodynamics Substudy....18
2. PLANNED ANALYSES....19
2.1. Interim Analysis....19
2.2. Final Analysis....19
3. ANALYSIS POPULATION....20
3.1. Analysis Sets....20
3.1.1. Screened Set....20
3.1.2. Placebo Run-in Set....20
3.1.3. Randomized Set....20
3.1.4. Safety Analysis Set....20
3.1.5. Urodyamics Evaluable Set....21
3.1.6. Full Analysis Set....21
3.1.7. Per-Protocol Set....21
3.2. Violations and Deviation....22

4. GENERAL STATISTICAL CONSIDERATIONS 23
4.1. General Principles for Data Analysis 23
4.1.1. Multicenter Study 23
4.1.2. Testing Strategy and Multiplicity Adjustments 23
4.1.3. Examination of Subgroups 23
4.2. General Data Handling Conventions 24
4.2.1. Study Treatment Description 24
4.2.2. Reporting Conventions 24
4.2.3. Premature Withdrawal and Missing Data 25
4.2.4. Assessment Windows 27
4.2.4.1. Study Reporting Periods 27
4.2.4.2. Analysis Visit Windows 28
4.3. Data Definitions and Derivations of Endpoints 29
4.3.1. Bladder Diary Endpoints 29
4.3.1.1. Co-primary Efficacy Endpoint 30
4.3.1.2. Secondary Efficacy Endpoint (Bladder Diary) 30
4.3.1.3. Other Exploratory Efficacy Endpoints 31
4.3.2. Patient Reporting Questionnaire Scoring 31
4.3.2.1. International Prostate Symptom Score 31
4.3.2.2. Overactive Bladder Questionnaire 32
4.3.2.3. EQ-5D-5L 33
4.3.2.4. International Index of Erectile Function 34
4.3.3. Study Day and Duration 34
4.3.4. Baseline and Change from Baseline 35
4.3.5. Pre-existing and Baseline Hypertension 35
5. STUDY SUBJECTS 36
5.1. Subject Disposition and Withdrawals 36
5.2. Demographic and Baseline Characteristics 36
5.3. Other Baseline Characteristics 37
5.4. Medical History and Concomitant Disease 37
5.5. Procedures 37
5.6. Prior and Concomitant Medication 38
5.6.1. Prior Medication 38

5.6.2. Concomitant Medication 38
5.7. ECG 39
5.8. Prostate Volume Measurements 39
5.9. Treatment Exposure 39
5.10. Treatment Compliance 39
6. EFFICACY ANALYSIS 40
6.1. Co-Primary Endpoint Analysis 40
6.1.1. Primary Analysis 40
6.1.2. Sensitivity Analysis 42
6.1.2.1. Per-Protocol Analysis 42
6.1.2.2. Analysis of Covariance with Multiple Imputed data 42
6.1.2.3. MMRM with Stratification Factors as Observed 44
6.1.3. Subgroup Analysis 44
6.2. Secondary Efficacy Analyses 44
6.3. Other/Exploratory Endpoints 45
6.3.1. Efficacy 45
6.3.2. Quality of Life 47
6.3.2.1. CFB in [redacted] Scores 47
6.3.2.2. [redacted] 47
6.3.2.3. Patient Global Impression Scores 47
6.3.2.4. [redacted] 48
7. SAFETY ANALYSIS 48
7.1. Adverse Events 48
7.2. Laboratory Evaluations 51
7.3. Vital Sign Assessments 51
7.3.1. Height, Weight, and Sitting Vital Signs 51
7.3.1.1. Vital Sign Subgroups 52
7.3.1.2. Categorical/Threshold Analysis 52
7.3.2. Orthostatic Blood Pressure and Pulse Rate (Part 1 only) 52
7.4. Physical Examination 54
7.5. Post-Void Residual (PVR) Urine Volume 54
7.6. Urodynamics Assessments (Urodynamics Sub-study Only) 54
7.7. Uroflow Measures 55

7.8. IPSS Total Score....55
8. COVID-19 CONSIDERATIONS....56
9. CHANGE FROM ANALYSIS PLANNED IN PROTOCOL....57
9.1. Exploratory endpoints....57
9.2. Responder Analysis....57
9.3. COVID-19 Considerations....57
10. REFERENCES....58
11. APPENDIX....59
11.1. Appendix 1. Multiple Imputation for the Response Efficacy Endpoints....59
11.2. Table of Contents for Data Display Specifications....61
11.2.1. Output Tables....61
11.2.2. Output Figures....74
11.2.3. Output Listings....76

LIST OF TABLES

Table 1: Data and Sources Supporting Study Sample Size Calculations....17
Table 2: Definition of Study Reporting Periods....27
Table 3: Analysis Window Slotting....28
Table 4: [redacted] Domain Questions....32
Table 5: [redacted] Domain Questions....34
Table 6: OAB Medication Selection....38
Table 7: PGI Scoring....47
Table 8: At Risk Subgroup for Vital Sign Summaries....52
Table 9: List of Output Tables....61
Table 10: List of Output Figures....74
Table 11: List of Output Listings....76

LIST OF ABBREVIATIONS

Abbreviation	Term
5-ARI	5α-reductase inhibitor
ADaM	analysis data model
AE	adverse event
AESI	adverse events of special interest
BMI	body mass index
BP	blood pressure
BPH	benign prostatic hyperplasia
CFB	change from baseline
CI	confidence interval
CMH	Cochran-Mantel-Haenszel
CRF	case report form
DBP	diastolic blood pressure
DSMB	data safety monitoring board
ECG	Electrocardiogram
eCRF	electronic case report form
FAS	full analysis set
FAS-I	full analysis set for incontinence
HRQL	health-related quality of life
[redacted]	[redacted]
IPSS	International Prostate Symptom Score
IWRS	interactive web response system
LS	least squares
MAR	missing at random
MCC	maximum cystometric capacity
MCMC	Markov chain Monte Carlo
MedDRA	Medical Dictionary for Regulatory Activities
MMRM	mixed model for repeated measure
OAB	overactive bladder
[redacted]	[redacted]
PdetQmax	detrusor pressure at maximum urinary flow
PGI	patient global impression
PPS	per-protocol set
PT	preferred term
PRO	patient reported outcome
PVR	post-void residual
Qmax	peak flow rate during voiding
REML	restricted (or residual) maximum likelihood
SAE	serious adverse event
SAF	safety set

Abbreviation	Term
SAP	statistical analysis plan
SBP	systolic blood pressure
SD	standard deviation
SDTM	study data tabulation model
SE	standard error
SOC	system organ class
TEAE	treatment emergent adverse event
TLF	table, listing, figure
UUI	urge urinary incontinence
VAS	visual analog scale
VV	volume voided
WHO	World Health Organization
β3-AR	beta-3 adrenergic receptor

SAP VERSION HISTORY

Version	Date	Description of Changes
1.0	**02FEB2023**	**Original Document**
2.0	**04AUG2023**	**Updated Per Protocol exclusionary Protocol Deviations** **Updated Urodynamics, Uroflow, Post-Void Residual, and Vital Signs analyses** **Added Tipping Point Sensitivity analysis of primary efficacy endpoints**

1. INTRODUCTION

This Statistical Analysis Plan (SAP) provides a description of the statistical methodology to be implemented for the analyses of data from Study URO-901-3005 per Protocol Version 4.0. This SAP is to be finalized, approved by the sponsor, and placed on file before the study database is locked. Changes to the final SAP will be documented in the final clinical study report.

1.1. Study Objectives and Endpoints

1.1.1. Primary Efficacy Objectives

Co-Primary Efficacy Objectives	Endpoints
To assess the efficacy of vibegron compared with placebo in men with OAB symptoms on pharmacological therapy for BPH as defined by micturition and urgency episodes	• Change from baseline (CFB) to Week 12 in the average number of micturition episodes per day • CFB to Week 12 in the average number of urgency episodes (urgency: need to urinate immediately) per day

1.1.2. Secondary Efficacy Objectives

Secondary Efficacy Objectives	Endpoints
To assess the efficacy of vibegron compared with placebo in men with OAB symptoms on pharmacological therapy for BPH as defined by other key measures	• CFB to Week 12 in the average number of nocturia episodes per night • CFB to Week 12 in the average number of urge urinary incontinence episodes per day for subjects with urinary incontinence at baseline • CFB to Week 12 in the International Prostate Symptom Score (IPSS) Storage score (1-week recall) • CFB to Week 12 in the average volume voided per micturition

1.1.3. Safety Objectives

Safety Objectives	Endpoints
To assess the safety and tolerability of vibegron compared with placebo in men with OAB symptoms on pharmacological therapy for BPH	• Adverse Event (AE) • Clinical laboratory • Vital sign assessments • Orthostatic blood pressure (Part 1 only) • Physical exam (PE)

	• Post-Void residual (PVR) urine volume • Urodynamics (Part 2 sub-study only) • Uroflow measures • Prostate volume measurements • IPSS total score

1.1.4. Other Efficacy Objectives

Other Efficacy Objectives	Endpoints
To assess the efficacy of vibegron compared with placebo in men with OAB symptoms on pharmacological therapy for BPH as defined by other measures	<u>CFB to Week 2, Week 4, Week 8, Week 16, Week 20 and Week 24 in the following 5 parameters</u> • Average number of micturition episodes per day • Average number of urgency episodes per day • Average number of nocturia episodes per night • Average number of urge urinary incontinence episodes per day for subjects with urinary incontinence at baseline • Average volume voided per micturition • CFB to Week 4, Week 8, Week 20, and Week 24 in IPSS Storage score (1-week recall) <u>CFB to Week 4, Week 8, Week 12, Week 20 and Week 24 in the following 3 parameters</u> • IPSS Quality of Life score (1-week recall) • IPSS Voiding score (1-week recall) • IPSS total score • CFB to Week 2, Week 4, Week 8, Week 12, Week 16, Week 20, and Week 24 in average number of total incontinence episodes per day for subjects with urinary incontinence at baseline <u>Responder analysis at Week 2, Week 4, Week 8, Week 12, Week 16, Week 20 and Week 24 measured by:</u> • Percent of all subjects with a 50% reduction from baseline in average urgency episodes (urgency: need to urinate immediately) per day • Percent of subjects with urinary incontinence at baseline with a 75% reduction from baseline in average urge urinary incontinence episodes per day

To assess the effects of vibegron compared with placebo on quality-of-life assessments and patient-reported outcomes (PROs) in men with OAB symptoms on pharmacological therapy for BPH	<u>CFB to Week 12 and Week 24</u>
	• Symptom Bother Score as assessed by [redacted], 1-week recall)
	• HRQL scores as assessed by [redacted].
	  • HRQL subscale Coping score
	  • HRQL subscale Concern score
	  • HRQL subscale Sleep score
	  • HRQL subscale Social Interaction score
	  • HRQL total score
	• [redacted] as assessed by measure of health status questionnaire
	• [redacted] overall satisfaction and function domains
	
	<u>CFB to Week 4, Week 12, and Week 24</u>
	• Overall bladder symptoms based on Patient Global Impression of [redacted] (PGI-[redacted])
	• Overall control over bladder symptoms based on Patient Global Impression of [redacted])
	• Overall symptom frequency based on Patient Global Impression of Symptom [redacted]
	• Overall urgency-related leakage over bladder symptoms based on Patient Global Impression of Urgency-Related [redacted] (PGI-[redacted]) in subjects with urinary incontinence at baseline
	• Overall change of bladder symptoms based on Patient Global Impression of [redacted] ([redacted]

1.2. Study Design

1.2.1. Overall Study Design

This is a Phase 3, randomized, double-blind, placebo-controlled, 2-part, parallel-group, multicenter study to evaluate the safety, tolerability, and efficacy of vibegron 75 mg in men with symptoms of OAB on stable doses of pharmacological therapy for BPH. At Baseline, subjects who meet all eligibility criteria are randomized 1:1 to receive either vibegron 75 mg or placebo in a double-blind fashion.

The study consists of two parts: Part 1 (approximately 80 subjects) and Part 2 (approximately 1008 subjects). Part 1 includes additional orthostatic blood pressure and heart rate measurements at Screening, Run-in, Baseline, Week 2, and Week 4 to assess for potential orthostatic changes (pre-dose through 6 hours post dose, except pre-dose for Screening and Run-In), but is otherwise consistent with the schedule for Part 2. Part 2 will proceed following review of 4-week safety data (including orthostatic blood pressure results) of all 80 subjects from Part 1 by an independent Data Safety Monitoring Board (DSMB). Details on tasks and responsibilities, assessments of safety parameters, including orthostatic blood pressure and criteria for defining clinically significant orthostatic events, will be provided in the DSMB Charter. Part 2 will also include a urodynamics sub-study (approximately 60 subjects). The independent DSMB will review the safety data throughout the entire study.

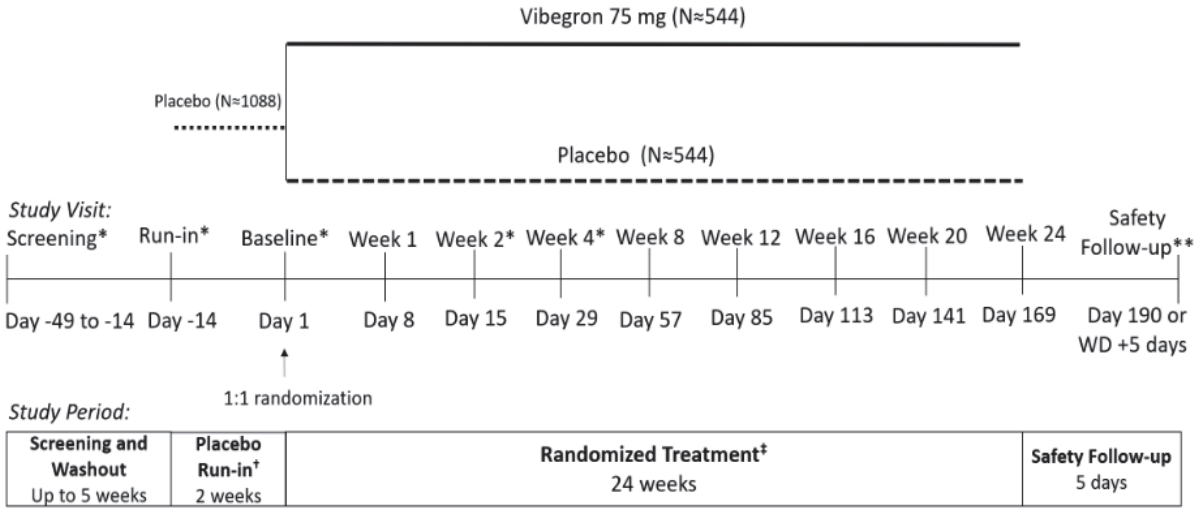
Both study parts will consist of a Screening Period (1 to 4 weeks), a single-blind therapy plus placebo Run-in Period (2 weeks), and a randomized double-blind Treatment Period (24 weeks).

Study visit schedules: Visit 1 (Screening), Visit 2 (Run-in), Visit 3 (Baseline), Visit 4 (Week 1), Visit 5 (Week 2), Visit 6 (Week 4), Visit 7 (Week 8), Visit 8 (Week 12), Visit 9 (Week 16), Visit 10 (Week 20), and Visit 11 (Week 24) (End of the Treatment).

Subjects who complete the Week 24 Visit may be offered the opportunity to enroll in a 28-week extension study URO-901-3006, which will be conducted under a separate study protocol. Subjects who do not enroll into the extension study will have a Safety Follow-up contact approximately 5 days (± up to 2 days) after the subject's last dose of study treatment (i.e., or approximately 5 days (± up to 2 days) after withdrawal for subjects who discontinue the study early). Additionally, Unscheduled Visit(s) may be arranged as needed.

Approximately 1088 subjects (544 per treatment group: 80 total in Part 1 and 1008 total in Part 2) will be randomized at up to approximately 135 sites in North America (United States and Canada) and Europe to have 924 subjects complete the study (assuming a 15% dropout rate). The study schema is shown in Figure 1.

Figure 1: Study Schema



*: Screening, Run-in, Baseline, Week 2, and Week 4 include extended blood pressure and heart rate assessments for Part 1 participants only; safety data from Part 1 will be evaluated prior to initiating Part 2.
**: The Safety Follow-up occurs 5 days after the last dose for subjects who completed the Week 24 visit but do not enroll in the optional 28-week extension study or 5 days after withdrawal (WD) for subjects who withdraw early from the study.
†Single-blind Run-in (subjects will not know they are receiving placebo)
‡Double-blind

1.2.2. Randomization and Blinding

Randomization will occur centrally using an interactive web response system (IWRS) using central, stratified block randomization. Randomization will be stratified based on the following stratification factors:

- Baseline average micturition episodes per day (≤ 12 vs > 12), and
- Alpha blocker use with or without 5-ARI (with vs without), and
- Urinary Incontinence (yes vs no)

Part 1 and Part 2 excluding the urodynamics sub-study will use the same stratified block randomization list. In order to ensure balanced treatment assignment within the urodynamics sub-study a separate randomization list with the same stratification factors and the block size will be used for subjects enrolled in the sub-study.

For this study in the double-blind treatment phase a double-blind/masking technique will be used: vibegron and its matching placebo will be packaged identically so that treatment blind/masking is maintained. The patient, the Investigator, and Sponsor personnel or delegate(s) who are involved in the treatment or clinical evaluation of the patients will not be aware of the treatment group assignments.

1.2.3. Statistical Hypotheses

1.2.3.1. Statistical Hypotheses for Co-Primary Efficacy Endpoints

The co-primary efficacy objectives of the study are to demonstrate that vibegron 75 mg is superior to placebo with respect to mean CFB in the following co-primary endpoints after 12 weeks treatment.

- **Daily Micturition**
  - **Co-Primary H_0**: In men with OAB-BPH, vibegron 75 mg will have the same mean CFB to Week 12 in the average number of daily micturition episodes as in men treated with placebo.
  - **Co-Primary H_1**: In men with OAB-BPH, vibegron 75 mg will have different CFB to Week 12 in the average number of daily micturition episodes from placebo.
- **Daily Urgency Episode**
  - **Co-Primary H_0**: In men with OAB-BPH, vibegron 75 mg will have the same mean CFB to Week 12 in the average number of daily urgency episodes as in men treated with placebo.
  - **Co-Primary H_1**: In men with OAB-BPH, vibegron 75 mg will have different mean CFB to Week 12 in the average number of daily urgency episodes from placebo.

1.2.3.2. Statistical Hypotheses for Secondary Efficacy Endpoints

The secondary efficacy objectives of the study are to demonstrate that vibegron 75 mg is superior to placebo with respect to mean CFB to Week 12 in the following 4 endpoints. The null hypothesis is that there is no difference in mean CFB to Week 12 between vibegron 75 mg and placebo. The alternative hypothesis is that vibegron 75 mg is different (i.e., two-sided hypothesis) from placebo with respect to CFB to Week 12 in the following parameters, respectively. These tests will be performed sequentially each at an alpha level of 0.05 after the co-primary endpoints are tested and both demonstrate statistical superiority of vibegron 75 mg over placebo at a significance level of 0.05, see details described in section 4.1.2.

- Average number of nocturia episodes per night
- Average number of daily urge urinary incontinence episodes for subjects with urinary incontinence at baseline
- IPSS Storage score (1-week recall)
- Average volume voided per micturition

1.2.4. Sample Size Justification

Approximately 1088 subjects will be randomized in a 1:1 ratio to receive one of the following study treatments:

- Vibegron 75 mg tablet (N = 544)
- Matching placebo tablet (N = 544)

Approximately 544 subjects each will be assigned to the vibegron and placebo treatment groups. Assuming a total of 15% of subjects will discontinue study treatment prior to Week 12 (for any reason), there will be approximately 462 evaluable subjects in the vibegron and placebo treatment groups (924 subjects total) at the end of Week 12. The study has approximately 94.3% power to detect a between-group treatment difference of 0.46 in change from baseline in the average number of daily micturition episodes at a 2-sided 0.05 level. This calculation assumes a standard deviation (SD) of 1.972 based on vibegron Study MRK-008 data and mirabegron NDA data showing that the micturition treatment effect in men taking 50 mg of mirabegron is 70% of the overall 50 mg of mirabegron population. Additional details are provided at the end of this section in Table 1.

For the second co-primary endpoint, 462 evaluable subjects per treatment group provides approximately 90% power to detect a between-group treatment difference of 0.60 in change from baseline in urgency episodes at a 2-sided 0.05 level. This calculation assumes a SD of 2.811 based on vibegron Study MRK-008 data and mirabegron NDA data. Additional details are provided at the end of this section in Table 1.

The overall study power will be greater than 84% with the co-primary endpoints based on the Bonferroni inequality (i.e., for any events A and B, $P(A \cap B) \geq P(A) + P(B) - 1$).

Derivation of Estimated Treatment Effect in Men

Based on the published results of the Phase 2 trial of mirabegron and the FDA statistical review of the OAB NDA, we estimated the following:

- The Phase 3 results will have at least 85% of the treatment difference of the Phase 2 results
- The treatment difference for men is at least 70% of the effect of the general OAB population

Table 1 summarizes this information and sources.

Table 1: Data and Sources Supporting Study Sample Size Calculations

Mirabegon Results for Mean Change in Average Number of Daily Micturitions			
	N		**PBO Adj CFB**
Study	Overall	Men	Overall
Phase 2 178-CL-046	473	133	-0.6
Phase 3 178-CL-047	425	116	-0.61
		Simple Average	-0.605

Mirabegon Results for Mean Change in Average Number of Daily Micturitions

	Phase 2 (178-CL-046) Micturition PBO Adj CFB	Phase 3 (178-CL-047) Micturition PBO Adj CFB	Phase 3 as a % of Phase 2
Comparing Phase 3 to Phase 2	-0.64	-0.605	94.5%

Vibegron Study MRK-008 Results			
Measure	**Vibegron 50 mg**	**Vibegron 100 mg**	**SD**
Micturition	-0.64	-0.91	1.972
Urgency Episodes	-0.75	-1.24	2.811

Ph3 Projections

Midpoint between 50 mg and 100 mg	**Micturition**	**Urgency Episodes**
Ph 2 Projected Effect of 75 mg	-0.78	-1.00
Ph3 Projected Effect of 75 mg	-0.66	-0.85
Ph 3 Projected Effect of 75 mg in Men*	-0.46	-0.60

*Used for URO-901-3005 Sample Size Calculations

1.2.4.1. Power in Baseline BPH Treatment Subgroups

Although this study will not formally test the treatment effect in subgroups by baseline BPH treatments, it is sufficiently sized to estimate treatment differences within these subgroups. Assuming that approximately half of the subjects are on alpha blockers alone and half on alpha blockers plus 5-ARI per stratified randomization, then approximately 544 subjects will be randomized to each subgroup with 462 evaluable (231 per arm). Using the above assumptions with a two-sided test with a type-I error of 0.10, each subgroup will have 97.1% power for micturition episodes and 94.5% power for urgency episodes. Assuming that the endpoints and subgroups are independent, then the power for each subgroup to be significant on both endpoints is 91.7%, and the power for both subgroups to show significant treatment difference on both endpoints is 84.0%.

1.2.4.2. Urodynamics Substudy

Approximately 75 subjects will be enrolled in the Urodynamics substudy, randomized 1:1 to placebo or vibegron 75mg. Assuming a 20% drop-out rate, 60 subjects will be evaluable, 30 in each treatment arm. Noninferiority margins of -3 mL/second for maximum urinary flow (Qmax) and 15 cm H_2O for detrusor pressure at maximum urinary flow (PdetQmax) from Abrams et al [1] were used to characterize power. Standard Deviation of each parameter was estimated from Nitti et al [2], which used the same noninferiority margins.

Using the noninferiority margins described above and a one-sided test of noninferiority with type-I error of 0.05:

30 evaluable subjects per treatment group provides approximately 97% power against the alternative hypothesis of equality for Qmax assuming standard deviation of 3.2 mL/second.

For PdetQmax, 30 evaluable subjects per treatment group provides approximately 79% power against the alternative hypothesis of equality assuming standard deviation of 23.5 cm H_2O. Assuming independence of these tests and requiring both to be within the non-inferiority margin, the power of the composite outcome is approximately 77%.

2. PLANNED ANALYSES

2.1. Interim Analysis

No interim analysis is planned for this study.

An external independent DSMB will be utilized in the study to ensure external objective medical and/or statistical review of safety issues in order to protect the ethical and safety interests of subjects and to protect the scientific validity of the study.
An external statistician will provide the summaries required to the DSMB. No adjustment to the type I error will be made because there is no efficacy data to be reviewed. The schedule of any planned analysis and analysis plan for DSMB review is described in the charter and separate DSMB SAP.

2.2. Final Analysis

The study consists of two parts: Part 1 (approximately 80 subjects) and Part 2 (approximately 1008 subjects). The two study parts are consistent with respect to the study schedule; the only exception is that Part 1 includes additional orthostatic blood pressure and heart rate measurements up to Week 4 to assess for potential orthostatic changes. Part 2 also contains a urodynamics sub-study (approximately 60 subjects). Thus, all data except for orthostatic blood pressure for Part 1 and urodynamic data from the subgroup of Part 2 will be pooled for analyses.

The production and quality control of all tables, figures and listings will be provided by a clinical research organization (eClinical Solutions, LLC) under the supervision of Urovant Sciences, Inc.

Statistical programming will start after data have been collected and are available in the database. Blinded dry-runs using dummy treatment code will be performed prior to database lock to ensure programming displays and algorithms are developed as planned.

The planned final analysis will be performed once the clinical database lock has taken place and treatment codes have been unblinded.

3. ANALYSIS POPULATION

3.1. Analysis Sets

All analysis sets defined below include subjects from both Part 1 and Part 2 except for the Safety Analysis Set for orthostatic blood pressure for subjects in Part 1, and the Urodynamics Evaluable Set and Per-Protocol Set for Urodynamics which include subjects who participate in the urodynamics sub-study in Part 2 only.

3.1.1. Screened Set

The Screened Set consists of all subjects who are screened for the study. This population is used primarily for subject accounting purposes and will generally not be used for summary or analysis. This set includes all subjects that signed the informed consent form and have screening data entered into the database, including screen failures (subjects who did not receive any run-in treatment), run-in failures (subjects who discontinued after receiving run-in treatment and prior to randomization), and randomized subjects.

3.1.2. Placebo Run-in Set

The Placebo Run-In Set consists of all subjects who receive placebo run-in treatment. Subjects will be considered run-in failures if they receive run-in treatment but are not randomized to receive double-blind medication. Subjects who inadvertently received vibegron during run-in are not included in the placebo run-in set.

3.1.3. Randomized Set

The Randomized Set consists of all subjects who are randomized to receive any double-blind study medication regardless of whether they took a dose or not.

3.1.4. Safety Analysis Set

The Safety Analysis Set will serve as the primary population for the analysis of safety data in this trial. Since Part 1 includes additional orthostatic blood pressure assessments while Part 2 includes an additional sub-study with urodynamics measurements, it is necessary to have separate FAS definitions for these assessments. The following SAF populations are defined in the study:

- **Safety Analysis Set (SAF):** consists of all randomized subjects who receive at least one dose of double-blind study medication. Subjects will be classified according to the treatment they actually received; subjects who receive any dose of vibegron will be classified as vibegron subjects, subjects who receive only doses of placebo will be classified as placebo subjects.
- **Safety Analysis Set for Orthostatic Blood Pressure (SAF – Orth)**: a subset of SAF to include all subjects in Part 1 who received at least one dose of double-blinded study medication and have at least one evaluable change from baseline orthostatic blood pressure measurement.

3.1.5. Urodyamics Evaluable Set

The Urodynamics Evaluable Set (UES) consists of all subjects in Part 2 who were randomized to the urodynamics sub-study, received at least one dose of double-blind study medication, had both Qmax and PdetQmax measurements at baseline and at least one Qmax or PdetQmax measurement at the post-baseline timepoint.

3.1.6. Full Analysis Set

The Full Analysis Set (FAS) population will serve as the primary population for the analysis of efficacy data in this trial. Since the endpoints related to incontinence only apply to subjects who meet the definition of incontinence at trial entry, it is necessary to have a separate FAS definition with an additional criterion to define the primary analysis population for incontinence endpoints.

The following FAS populations are defined in the study:

- **Full Analysis Set (FAS):** consists of all randomized subjects who took at least one dose of double-blind study medication and have at least one evaluable change from baseline micturition measurement. Subjects will be analyzed according to their randomized treatment, irrespective of premature discontinuation, according to the Intent-to-Treat principle.
- **Full Analysis Set for Incontinence (FAS-I):** consists of all randomized subjects with incontinence at baseline (as calculated from the baseline diary) who took at least one dose of double-blind study treatment and have at least one evaluable change from baseline urinary incontinence measurement.

Note that "subjects with urinary incontinence at baseline" are defined as meeting the following criteria based on the 3-day bladder diary completed by the subject during the Run-in period:

- An average of ≥ 8.0 micturitions per Diary Day. A "Diary Day" is defined as the time between when the subject gets up for the day each morning (i.e., the time the subject got up for the day yesterday to the time the subject got up for the day today; approximately a 24-hour period).
- An average of ≥ 1.0 urinary incontinence episodes per Diary Day

3.1.7. Per-Protocol Set

The Per-Protocol Set (PPS) and Per-Protocol Set for Incontinence (PPS-I) exclude patients from the FAS and FAS-I due to important deviations from the protocol that may substantially affect the results of the efficacy endpoints (i.e., Major PDs associated with efficacy). The Per-Protocol set for Urodynamics (PPS-UES) exclude subjects from UES due to important deviations from the protocol that may substantially affect the results of the urodynamics endpoints. A supportive analysis using the PPS and PPS-I will be performed for the co-primary and secondary efficacy endpoints, and a supportive analysis using the PPS-UES will be performed for the primary urodynamics parameters (Qmax and PdetQmax). The final determination on protocol deviations, and thereby the composition of the Per-Protocol Sets, will be made prior to the unblinding of the database and will be documented per the Protocol Deviation Plan during the blinded data review

meeting. Subjects will be included in the treatment group to which they are randomized, regardless of which treatment was actually received, for the analyses of efficacy data using the PPS and PPS-I.

3.2. Violations and Deviation

Subjects who do not meet eligibility criteria but are still randomized will be analyzed according to the analysis sets described in Section 3.1.
Major protocol deviations are considered to have major impact on subject safety, efficacy or the validity of the study data.

Patients with major efficacy protocol deviations will be excluded from the PPS and PPS-I under the assumption that the deviation may have an impact on the efficacy analysis, and from PPS-UES under the assumption that the deviation may have an impact on the urodynamics analysis. This may include, but is not limited to the following:

- Subjects who do not meet the following inclusion criteria:
  - On a stable dose of alpha blocker
  - Average frequency of ≥ 8 and ≤20 micturitions per day in the baseline diary
  - Average frequency of ≥ 3 urgency episodes per day in the baseline diary
  - Average frequency of ≥ 2 nocturia episodes per night in the baseline diary
- Subjects who meet any of the following exclusion criteria:
  - Total 24-hour urine volume output of >3000 mL based on any baseline diary day
  - History of prostate surgery
  - Subject started using diuretics within 28 days prior to screening
  - Nocturnal Polyuria (> 1/3 of total urine output per 24 hours occurring at nighttime) assessed on any day of the baseline diary
  - Has received an intravesical or intraprostatic treatment with any botulinum toxin, resiniferatoxin, or capsaicin within 6 months prior to screening
  - Use of any prohibited anticholinergics or beta-3 adrenergic agonists as detailed in protocol section 10.6 within 28 days prior to screening
- Concomitant use of prohibited medications as detailed in protocol section 10.6
- <70% Investigational Product [IP] compliance from Baseline through Week 12
- Visit delay at or prior to Week 12 causing more than 14 days without IP
- Failed to complete any diary day for an expected Week 12 diary within Study Day 68-102 [PPS and PPS-I only]

Exclusion from per-protocol analysis is noted for each protocol deviation in the Protocol Deviation specifications.

The final list of major protocol deviations will be finalized and documented before database lock except for the deviation of wrong treatment which will be confirmed upon study unblinding. All major protocol deviations related to efficacy will be finalized prior to unblinding. Only major protocol deviations will be summarized and listed in the Clinical Study Report (CSR).

4. GENERAL STATISTICAL CONSIDERATIONS

4.1. General Principles for Data Analysis

4.1.1. Multicenter Study

In this study the stratified permuted block randomization is not done within centers. The analyses will be conducted by pooling data from all study centers and will not include study center as a covariate in the statistical modeling.

4.1.2. Testing Strategy and Multiplicity Adjustments

A hierarchical gate-keeping procedure will be used to control the overall family-wise Type-I error rate at 2-sided α=0.05 level over the co-primary and key secondary hypotheses stated in Section 1.2.3.

If statistical significance for both co-primary efficacy endpoints is achieved (i.e., if two-sided p-value for each endpoint is less than 0.05), then the secondary efficacy endpoints will be tested sequentially in the pre-defined order below.

1. Average number of nocturia episodes per night
2. Average number of daily urge urinary incontinence episodes for subjects with urinary incontinence at baseline
3. IPSS Storage score (1-week recall)
4. Average volume voided per micturition

Once a secondary efficacy endpoint is found to be insignificant (i.e., two-sided p-value $\geq$ 0.05), the testing procedure will stop. For all subsequent secondary efficacy endpoints, nominal p-values will be provided but will not be considered a formal test of the hypotheses. Unless otherwise stated, all statistical tests will be conducted at the 2-sided α=0.05 level of significance.

All other additional and exploratory efficacy endpoints will be considered supportive and no multiplicity adjustments will be performed for these other efficacy endpoints. Nominal p-values will be computed for other efficacy endpoints as a measure of the strength of the treatment effect rather than formal tests of hypotheses, if applicable.

All confidence intervals will be 2-sided 95% confidence intervals, unless stated otherwise.

4.1.3. Examination of Subgroups

To determine whether the treatment effect is consistent across various subgroups, the co-primary endpoints will be summarized descriptively for each of the following subgroups:

- Alpha blocker use with or without 5-ARI (with vs. without)
- Baseline urinary incontinence (yes vs. no)
- Region (US vs. Non-US)

- Age category 1 (<55, ≥55 to <65, ≥65 to < 75, ≥75 years)
- Age category 2 (<65, ≥65 years)
- Age category 3 (<65, ≥65 to <75, ≥75 to <85, ≥85 years)
- Race (White, Black or African American, Asian and Other)
- Prior anticholinergic use (yes vs. no)
- Prior beta-3 agonist use (yes vs. no)

If any single subgroup constitutes <5% of total subjects, categories may be combined.

4.2. General Data Handling Conventions

4.2.1. Study Treatment Description

Randomized treatment groups will be displayed as shown in the following table:

Data Displays for Reporting	
Description	Order in TLF
Placebo	1
Vibegron 75mg	2
Overall (Demographic and baseline data)	3

4.2.2. Reporting Conventions

General rules

In general, all collected safety data and any derived efficacy and patient reported outcome (PRO) data will be presented in subject data listings for all enrolled subjects. Listings will be ordered by treatment group, subject number, and assessment week or event date. The treatment group presented in listings will be based on the planned assignment, unless otherwise noted.

Summary tables will be provided for all randomized subjects. Unless otherwise specified, all demographic and baseline data will be presented by treatment arm and overall. Efficacy, PRO and safety data will be presented by treatment arm. In general, continuous variables will be summarized to indicate the population sample size (N), number of subjects with available data (n), arithmetic mean, SD, median, minimum, Q1, Q3 and maximum values. Categorical variables will be summarized by the population size (N), number of subjects with available data (n), number of subjects in each category, and the percentage of subjects in each category. Unless otherwise noted, the denominator to determine the percentage of subjects in each category will be based on the number of subjects with available data (n) in the analysis set of interest. Select ordinal data may be summarized using both descriptive statistics and counts and percentages of subjects in each category, as appropriate.

For selected continuous safety endpoints, least squares means and 95% CIs will be derived from an analysis of covariance (ANCOVA) model with restricted maximum likelihood estimation. All models will include the baseline value of the parameter and any additional covariates will be noted; separate models will be fit for each timepoint. Example SAS code of a change from baseline at Week 12 is provided below:

```
proc mixed data = datain method = reml;
   class TRTA;
   model CHG = TRTA BASE / solution;
   where avisit = "Week 12";
   lsmeans TRTA / pdiff cl alpha=0.05;
run;
```

In the above code, CHG represents change from baseline, TRTA is the actual treatment received, and BASE is baseline value of the parameter.

The data analyses will be conducted using the SAS® System (SAS Institute Inc., Cary, NC) version 9.4 or above. The eClinical Solutions standard operating procedures will be followed for the creation and validation of all SAS programs and outputs.

Formats

Numeric data will be reported at the precision collected on the eCRF. The reported precision from non eCRF sources may be adjusted to a clinically interpretable number or decimal places. P-values will be reported to 4 decimal places, with p-values less than 0.0001 noted as <0.0001.

Unscheduled Visits

All visits, including unscheduled visits will be assigned to an analysis study visit using the all-inclusive windows defined in Section 4.2.4. However, data summaries will only report the visit nearest to the planned assessment time points for each parameter if multiple visits are within the analysis visit window, according to the Time and Events table. Assessments at unscheduled visits will be included for "any time On-treatment" timepoints and in data listings, as well as algorithms to determine the maximum or minimum.

4.2.3. Premature Withdrawal and Missing Data

All data collected in the study database after the subject's cessation of study treatment should be included in summary and analysis. Data of subjects who withdraw after the screening examination or are not treated will only be listed. Assessments taken after the End of Study visit date for subjects continuing on to study URO-901-3006 will not be included in analysis or listings for study URO-901-3005; it will be included in datasets to support inclusion in analysis of study URO-901-3006.

If any visits are missing diary data, no explicit missing data imputation will be performed for the primary analysis of change from baseline since a mixed model for repeated measures (MMRM) will be applied to change from baseline analysis. If time to bed or time to awaken the next morning is missing from a diary day, the average of non-missing days for that subject at that visit

will be used. If no visit average bedtime or time of awakening is calculable, no nocturia measurements will be calculated for that subject at that visit.

Missing items from the subject reported outcomes (PROs) will be handled according to the respective measure instructions as described in Section 4.3.2.

Missing return dates from the study drug log will be imputed as the day of the subsequent drug dispensation, end of treatment visit, or subject discontinuation in the study, whichever is first. Non-return of IP will be considered taken per protocol for calculation of treatment compliance but will be noted in the listing of study medication.

If the first dose of double-blind study medication is not witnessed, the start of medication will be imputed as the day after the drug was dispensed.

In general, missing safety data will not be imputed and only observed values will be analyzed.

If the relationship of an AE record ("Relationship to investigational product" on AE CRF) is missing this AE will be considered "Probably Related" to the study treatment. If the AE intensity is missing every effort should be made to acquire the information from the investigator. "Severe" will be assigned to a missing intensity for reporting purposes.
The general imputation rules of partial missing date and time for both AE and concomitant medication is detailed below:

<u>Dates missing the day of the year will adhere to the following conventions:</u>

- The missing day of onset date will be set to:
  - First day of the month that the event occurred, if the onset YYYY-MM is after the YYYY-MM of first double-blind treatment
  - The day of the first double-blind study treatment, if the onset YYYY-MM is the same as YYYY- MM of the first double-blind treatment
  - The day of the first run-in treatment, if the onset YYYY-MM is the same as YYYY-MM of the first run-in treatment, but different from YYYY-MM of the first double-blind treatment
  - The date of informed consent, if the onset YYYY-MM is before the YYYY-MM of the first run-in treatment.
- The missing day of end date will be set to:
  - The death date, if the end date YYYY-MM is the same as the YYYY-MM of the death date
  - The end of study participation, if the end date YYYY-MM is the same as the YYYY-MM of the end of study participation date
  - The last day of the month of the occurrence otherwise.

<u>Dates missing both the day and month of the year will adhere to the following conventions:</u>

- Missing onset date will be set to:

  - January 1 of the year of the onset, if the onset YYYY is after the YYYY of the first double-blind treatment
  - The date of the first run-in treatment, if the onset YYYY is the same as YYYY of the first run-in treatment, but different from YYYY of the first double-blind treatment
  - The date of the first double-blind treatment, if the onset YYYY is the same as YYYY of the first double-blind treatment
  - The date of informed consent, if the onset YYYY is before YYYY of the first run-in treatment
- The missing end date will be set to:
  - The death date, if the end date YYYY is the same as the YYYY of the death date
  - The end of study participation, if the end date YYYY is the same as the YYYY of the end of study participation date
  - December 31 of the year of occurrence otherwise.

4.2.4. Assessment Windows

4.2.4.1. Study Reporting Periods

Based on study design and variables under consideration, study time periods are defined as below.

Table 2: Definition of Study Reporting Periods

Analysis description	Analyzing Study Period	Start Date	End Date
General	Screening	Date of informed consent	Date of first run-in dose
	Run-in	Date of first run-in single-blind dose	Date of last run-in single-blind dose
	Treatment (Day 1 – Week 24)	Date of first double-blind dose	Date of last double-blind dose + 5 days
	Follow-up (up to Week 27)	>Date of last dose + 5 days	Date of EOS visit
Adverse events (based on AE start date)	Prior	Date of informed consent	Date of first run-in dose
	Run-in	Date of first single-blind run-in dose	Date of first double-blind dose
	TEAE	Date of first double-blind dose*	Date of last dose + 5 days
	Post-Treatment	Date of last dose + 5 days	N/A
Concomitant Medication/ Procedure (any overlap with period)	Prior medication/ procedure	N/A	Date of first double-blind dose

Analysis description	Analyzing Study Period	Start Date	End Date
	Concomitant medication/ procedure	Date of first double-blind dose	Date of last double-blind dose + 5 days
	Post medication/ procedure	>Date of last double-blind dose + 5 days	N/A
EOS: End of study is defined as the date when the subject has completed one of the following: completes Week 24 Visit for subjects who enroll into Study URO-901-3006 or Safety Follow-up for subjects who do not enroll into Study URO-901-3006, discontinued from the study, or is lost to follow-up. *AEs with a start date equal to the date of first double-blind dose will be considered treatment-emergent unless the relatedness to study drug is marked 'Not applicable'. Overlapping period: Medication start date<= period start date<medication end date, medication start date<period end date<=medication end date, or period start date<=medication start or end date<period end date, then the medication is considered overlapping with the period.			

4.2.4.2. Analysis Visit Windows

Analysis visit windows for both diary and non-diary data are defined as shown in Table 3. All observations will be given an analysis visit, assigned based on relative study day, as defined in section 4.3.3. If fewer than three diary days are within the analysis window, only the data within the window will be used to calculate visit-level averages. Unless otherwise noted, all data will be listed.

Table 3: Analysis Window Slotting

Analysis window label	Nominal visit	Nominal day	Non-Diary Visit Window	Diary Visit Window
Screening	Visit 1	-49 to - 21	date of informed consent to date of first run-in dose	
Run-in	Visit 2	-14	date of first run-in dose to day -1	7 days prior to date of first run-in dose
Baseline	Visit 3	1	[-1,1]	date of first run-in dose to day -1
Week 1	Visit 4	8	[2, 11]	
Week 2	Visit 5	15	[12, 22]	[7, 18]
Week 4	Visit 6	29	[23, 42]	[19, 39]
Week 8	Visit 7	57	[43, 70]	[47, 67]
Week 12	Visit 8	85	[71, 98]*	[68, 102]
Week 16	Visit 9	113	[99, 126]	[103, 123]
Week 20	Visit 10	141	[127, 154]	[131, 151]

Analysis window label	Nominal visit	Nominal day	Non-Diary Visit Window	Diary Visit Window
Week 24	Visit 11	169	[155, last dose date + 5]	[159, 179]
Follow-Up for Early Withdrawals	NA	NA	>Last dose date + 5	
Follow-Up for Completers	NA	NA	>Last dose date + 5	

*Urodynamics window for Week 12 is [57, 113]

For subjects who withdraw early, assessments taken at the end of study visit will also be summarized with scheduled data if they fall within an on-study visit window and are within 21 days after the last treatment date.

For parameters which were **not** scheduled to be collected at all visits, still use the all-inclusive visit intervals defined for all visits (i.e., the window defined above) to slot records. However, if a value is slotted to a visit unscheduled for a parameter, it will not be summarized (and not be included in the efficacy datasets) but will be included in data listings.

1. If after all records have been slotted, there are multiple valid records within a visit, use the record by the following the rules below:
   - If there are more than three bladder diary days within the window, the three days closest to the nominal visit day will be used to calculate analysis visit averages. If two days are equidistant to the nominal visit day, the earlier day will be selected first.
   - If there are multiple values for other efficacy endpoints in one window, the visit closest to the nominal day will be selected for assessing endpoints at a particular visit and for by-visit displays. If the visits are equidistant from the nominal day, then the earlier visit will be selected. All values will be stored in analysis datasets.
   - The same visit windows will be used for safety and laboratory assessments. If there are multiple safety or laboratory values in one time window on-treatment, the worst value (furthest from the center of the normal range if both high and low abnormalities exist) will be selected for the by-visit analyses.

4.3. Data Definitions and Derivations of Endpoints

4.3.1. Bladder Diary Endpoints

The bladder diary recorded what time subjects woke up for the day, what time they went to bed, and asked the subjects to record every time they had a urination event or leakage. For each event, subjects recorded the time, if they had a need to urinate immediately, if they urinated in the toilet, if they had accidental urine leakage and if they had leakage the main reason for the leakage (urge or other) for 3 consecutive days, and volume voided per micturition (over one 24-hour period) prior to each required timepoint. The following efficacy parameters will be derived based on the Diary data for baseline, Weeks 2, 4, 8, 12, 16, 20 and 24. Parameters will be derived

for the visit only if there is at least one valid diary day within the analysis window. If a subject answered 'no' to all three questions for every record on a diary day, with no volume recorded on any record, that day will be treated as missing for analysis purposes but will be included in the listings.

4.3.1.1. Co-primary Efficacy Endpoint

Although these endpoints are automatically derived in eCRF Day 3 Bladder Diary page, the final statistical analysis will use the value derived by the biometrics team based on the algorithm described below.

CFB in average number of daily micturition episodes to Week 12 (co-primary endpoint)

The number of micturition episodes will be defined as the number of times a subject has voided in the toilet as indicated on the Bladder Diary, either by marking the 'urinated in toilet' question as yes or recording a non-zero volume voided. This will be derived from individual log lines within each diary day; automatically derived daily totals on the eCRF Bladder Diary pages will not be used. Average number of micturition episodes per day at each study visit will be calculated using records within the diary analysis visit windows described in Table 3 divided by non-missing diary days (diary days with at least one void reported).

CFB in average number of daily urgency episodes to Week 12 (co-primary endpoint)

The number of urgency episodes will be defined as the number of times a patient has checked that he had the need to urinate immediately as indicated on the Bladder Diary. Average number of daily urgency episodes at each study visit will be calculated as the total number of urgency episodes within the diary analysis visit windows divided by non-missing diary days.

4.3.1.2. Secondary Efficacy Endpoint (Bladder Diary)

CFB in average number of daily nocturia episodes to Week 12 (secondary endpoint #1)

Nocturia episodes will be calculated in the same manner as average number of daily micturition episodes but only include episodes occurring during sleep (after bedtime on the current diary day and prior to wake-up time on the subsequent day). Although the endpoint is automatically derived in eCRF Day 3 Bladder Diary page (and daily totals are automatically derived on the eCRF Bladder Diary pages), the final statistical analysis will use the value derived by the biometrics team based on the algorithm described above.

CFB in average number of daily urge urinary incontinence (UUI) episodes to Week 12 for subjects with UI at baseline (secondary endpoint #2)

The number of UUI episodes will be defined as the number of times a subject has checked that they had "urge" as the main reason for the leakage. Average UUI episodes per day at each study visit will be calculated in the same manner as described above for the micturition endpoint. The UUI endpoint will be analyzed for subjects with UI at baseline.

CFB in average volume voided per micturition to Week 12 (secondary endpoint #4)

The total volume voided will be calculated using all urinary volumes collected regardless of whether patients checked "Urinated in Toilet or not". Average volume voided per micturition at each study visit will be calculated as the total volume voided over diary days within the analysis visit window divided by the total number of micturition episodes with recorded volume.

4.3.1.3. Other Exploratory Efficacy Endpoints

Average number of daily urinary incontinence (UI) episodes for subjects with UI at baseline (exploratory)

Average number of urinary incontinence episodes per day will be calculated as the total number of "Accidental Urine Leakage" regardless of the choice of "Main Reason for Leakage" over diary days within the analysis visit window divided by non-missing diary days.

Urgency episodes 50% daily responder (exploratory)

The achievement of at least a 50% reduction from baseline in average daily urgency episodes is a binary variable of 0 (endpoint is not attained) or 1 (endpoint is attained). Subjects will be considered as responders when the endpoint is attained.

Urgency urinary incontinence (UUI) episodes daily 75% responder (exploratory)
The achievement of at least a 75% reduction from baseline in average daily UUI episodes is a binary variable of 0 (endpoint is not attained) or 1 (endpoint is attained). Subjects will be considered as responders when the endpoint is attained. The endpoint will be analyzed for subjects with UI at baseline.

4.3.2. Patient Reporting Questionnaire Scoring

4.3.2.1. International Prostate Symptom Score

The International Prostate Symptom Score (IPSS) is based on the responses to 7 questions concerning urinary symptoms and 1 question concerning quality of life. Each question concerning urinary symptoms allows the subject to choose 1 out of 6 answers indicating increasing severity of the particular symptom. The responses are assigned points from 0 to 5. The total score can therefore range from 0 to 35 (asymptomatic to very symptomatic).

The questions refer to the following urinary symptoms:

Questions	Symptom	Subscale Category
1	Incomplete emptying	Voiding
2	Frequency	Storage
3	Intermittency	Voiding
4	Urgency	Storage
5	Weak Stream	Voiding
6	Straining	Voiding
7	Nocturia	Storage

Question 8 refers to the subject's perceived quality of life (QoL). The answers to the single question to assess the quality of life this question range from "delighted" to "terrible" or 0 to 6.

Raw scores for each scale will be calculated as the sum of responses to all the questions within the scale, with missing values imputed as the average value over the scale if <50% of the scale items are missing. If ≥ 50% of the items are missing, no scale score will be calculated, and the subscale score will be considered missing.

To transform the raw scores to a unified score ranging from 0 to 100, the following algorithms are used to derive the corresponding score at baseline, Week 12 and Week 24.

- For the Symptom Bother Score, the transformed score is:

$$score = \frac{(sum\ of\ actual\ score - lowest\ possible\ score)}{range} \times 100 = \frac{sum\ of\ actual\ score - 8}{40} \times 100$$

For Symptom Bother, higher scores correspond to the symptoms having a larger bother and lower scores represent a lower amount of bother due to symptoms.

- For the Coping, Concern, Sleep, Social Interaction, and Total HRQL Scores, the transformed score is:

$$score = \frac{(highest\ possible\ score - \ sum\ of\ actual\ score)}{range} \times 100$$

In this case, higher scores correspond to a higher quality of life and lower scores represent a lower quality of life.

4.3.2.3. [redacted] covers 5 dimensions of quality of life: mobility, self-care, usual activities, pain/discomfort, anxiety/depression, with respondents selecting one of 5 responses for each dimension. Each dimension has 5 levels: no problems, slight problems, moderate problems, severe problems and extreme problems. The subject is asked to indicate his/her health state by ticking the box next to the most appropriate statement in each of the five dimensions. This decision results in a 1-digit number that expresses the level selected for that dimension. The digits for the five dimensions can be combined into a 5-digit number that describes the patient's health state. The 5-digit number will be mapped to an index value using the crosswalk link function for the US value set based on the published [redacted] [4] at baseline, Week 12 and Week 24. At Week 12 and Week 24, the Pareto Classification of Health Change will also be assessed, which classifies health state relative to baseline:

- Health state is better if it is better in at least one dimension and no worse in any other dimension
- Health state is worse if it is worse in at least one dimension and no better in any other dimension
- Health state is unchanged if it all dimensions remain the same

- Health state is mixed if it is better in at least one dimension and worse in at least one dimension.

[redacted] also includes a visual analog scale (VAS) to capture a patient's overall health rating.

4.3.3. Study Day and Duration

Study day is relative to the start date of the double-blind treatment. This is used to describe the relative time of an event or assessment that happened during the study. The first day of the study is defined as the day a patient first receives either vibegron or placebo in the double-blind treatment period. This is expected to be the same day as the randomization as Day 1. There is no study Day 0 defined in the study.

- For event or assessment occurring on or after the first dose of double-blind treatment date:
  Study day = Date of event or examination – date of first double-blind treatment + 1
- For event or assessment occurring prior to the first dose of double-blind treatment date:
  Study day = Date of event or examination – date of first double-blind treatment

A duration between any two dates (such as AE duration) expressed in days will be calculated using the following convention:

- Duration = Later date – earlier date + 1

4.3.4. Baseline and Change from Baseline

In general, the last recorded value on or prior to the date of randomization will serve as the baseline measurement for efficacy endpoints while the last recorded value prior to first double-blind dose of study treatment will serve as the baseline measurement for safety endpoints. The mean of multiple values will be used as baseline for the following situations:

- If multiple measurements are scheduled on the same baseline day (i.e., blood pressure)
- If multiple measurements are collected on the same baseline day without the time or "repeat" status to differentiate the records

For efficacy endpoints derived from the Bladder Diary, the baseline will be the average value over the diary days in the baseline analysis visit window.

Change from baseline will be calculated as the post-baseline value minus the baseline value. Percentage change from baseline will be calculated as the change from baseline divided by the baseline value, multiplied by 100. If either the baseline or post-baseline value is missing, then change from baseline and percentage change from baseline will be set to missing.

4.3.5. Pre-existing and Baseline Hypertension

Pre-existing Hypertension will be assigned based on a search of a subject's medical history for the presence (Yes) of any or absence (No) of all the following coded terms:

- Accelerated hypertension
- Diastolic hypertension
- Essential hypertension
- Hypertension
- Hypertensive crisis
- Hypertensive emergency
- Hypertensive heart disease
- Malignant hypertension
- Malignant hypertensive heart disease
- Secondary hypertension
- Supine hypertension
- Systolic hypertension

Baseline Hypertension will be assigned based on a subject's baseline vital signs:

- Visit average SBP $\geq$ 140
- Visit average DBP $\geq$ 90

5. STUDY SUBJECTS

5.1. Subject Disposition and Withdrawals

Subject disposition will be summarized by treatment arm for the screened, run-in and randomized subjects. The summary table will present the frequency and percentage of subjects in each of the analysis sets and those who discontinued the study prematurely along with the primary reasons for discontinuation.

For the summary for the randomized subjects, the following additional categories will be presented as well; randomized, received treatment with study drug, did not receive treatment with study drug, completed treatment with study drug, discontinued treatment with study drug (and reason), discontinued treatment with study drug but completed study follow-up, completed study, and withdrawn from study (and reason).

The frequency and percentage of subjects with at least one major Protocol Deviation (PD), major PD by classification and reasons/category for PD will be summarized by treatment arm for the FAS. Inclusion in each of the analysis sets (SAF, SAF-Orth, SAF-UDS, UES, FAS, FAS-I, PPS, PPS-I), and any reasons for exclusion will be summarized by treatment arm for the Randomized Set.

Screen Failure, Run-in and Double-Blind Period disposition, with reasons for discontinuation of study will also be listed if applicable, including the date of discontinuation.

Eligibility criteria, screening failures (including date and primary reason for failure), and informed consent (protocol version, informed consent version date and date signed) will be listed for all patients screened.

A summary of randomized subjects by country and investigator will be provided. Randomization details will also be listed, including the date of randomization, randomization number and randomization strata.

5.2. Demographic and Baseline Characteristics

All demographic and baseline characteristic data will be summarized by treatment group using descriptive statistics for all subjects for each of the following analysis sets: SAF, SAF-Orth, UES, PPS-UES, FAS, FAS-I, PPS and PPS-I.

The summary table will include Sex, Ethnicity and Race, Region (US and non-US), Prior Anticholinergic Use (Yes/No), Prior Beta-3 agonist Use (Yes/No), Diabetes Mellitus (Yes/No), Baseline Hypertension (Yes/No), Pre-existing Hypertension (Yes/No), Age category 1 (<55, ≥55 to <65, ≥65 to <75, ≥75 years), Age category 2 (<65, ≥65 years), Age category 3 (<65, ≥65 to <75, ≥75 to <85, ≥85 years), treatment with alpha blocker, treatment with 5-ARI, baseline average micturition episodes per day (≤ 12 vs > 12), Alpha blocker use with or without 5-ARI (with vs without), Urinary Incontinence (yes or no) and IPSS total score. Stratification factors will be summarized both as randomized and corrected if any misstratification was present.

Prior Anticholinergic Use and Prior Beta-3 agonist Use will be calculated from all prior medications documented on the Prior and Concomitant Medications electronic case report form (eCRF).

Pre-existing and Baseline hypertension are defined in section 4.3.5.

Age (years), height (cm), weight (kg), BMI (kg/m^2) captured at Screening, and baseline IPSS total score will be summarized as continuous variables.

Unless otherwise stated, percentages will be calculated out of the number of patients in the given analysis set.

All demographic data will be listed.

5.3. Other Baseline Characteristics

The data from the 3-day bladder diary during run-in period (reviewed at Visit 3 Baseline visit) prior to first dose of double-blind medication will be used as baseline for each subject. This includes average number of daily micturition episodes, average number of daily urgency episodes, average number of nocturia per night, average number of daily urge urinary incontinence episodes, and average volume voided per micturition. In addition, the baseline IPSS Storage score, prostate volume, Qmax (maximal urinary flow) at baseline as determined by uroflowmetry (see section 7.7 for details), and baseline post-void residual urine will be presented. These will be summarized by treatment group and overall using descriptive statistics for continuous data for all subjects in FAS, FAS-I, PPS and PPS-I analysis sets.

5.4. Medical History and Concomitant Disease

Descriptions of medical history findings will be coded using Medical Dictionary for Regulatory Activities (MedDRA) version 22.0 or higher.

A disease or illness reported as medical history without a start date will be included in medical history without a date assigned. Medical history will be sorted by descending overall frequency, by system organ class (SOC) and preferred term (PT) in the summary table. Medical history data listings will be sorted by treatment, subject number, start date, SOC and PT. The SAF will be the analysis set for medical history data.

5.5. Procedures

Procedures will be coded using Medical Dictionary for Regulatory Activities (MedDRA) version 22.0 or higher.

Prior procedures are defined as procedures with a start date prior to the first dose of double-blind treatment. Concomitant procedures will be defined as procedures started on or after the first dose of double-blind medication but prior to the last dose of the double-blind medication + 5 days, or started prior to double-blind medication and were ongoing during the double-blind period. Post procedure is any procedure initiated after last dose of double-blind medication + 5 days (See Table 2). Partial procedure start dates will be imputed as detailed in 4.2.3.

All procedures will be listed for the SAF population, sorted by treatment, subject number, start date, SOC and PT, with indicators of prior and concomitant status.

5.6. Prior and Concomitant Medication

Prior and concomitant medications will be coded using the World Health Organization (WHO) Drug B3-Sept Format, 2018 version or higher. Except for Prior OAB medication, the number and percentage of subjects taking prior medications and concomitant medications will be summarized overall by ATC (Anatomical Therapeutical Chemical) Levels 2 and 4 for all subjects in the SAF. Prior OAB medications will be summarized by ATC Levels 2, 4 and Preferred Term in the SAF. Prior medications OAB medication and concomitant medications will be listed for all subjects in the SAF.

5.6.1. Prior Medication

Prior medications are defined as those medications taken prior to the first dose of double-blind treatment. The prior non-OAB and prior OAB medications will be summarized separately.

The following criterion will be used for selecting prior OAB medication:

Table 6: OAB Medication Selection

Class	Variable Selected	Selection
Anticholinergics	Preferred Term	darifenacin, fesoterodine, festoterodine fumarate hyoscyamine, oxybutynin, oxybutynin hydrochloride, propantheline, solifenacin, solifenacin succinate, tolterodine, tolterodine l-tartrate, trospium, trospium chloride
Beta-3 adrenergic agonists	Preferred Term	mirabegron, vibegron, solabegron

5.6.2. Concomitant Medication

Concomitant medications will be defined as medications started on or after the first dose of double-blind medication but prior to the last dose of the double-blind medication + 5 days, or started prior to double-blind medication and ongoing during the double-blind period. Post medication is the medication taken after the last dose of double-blind medication + 5 days (See Table 2). Partial medication start dates will be imputed as detailed in 4.2.3.

Non-OAB and OAB concomitant medications will be summarized separately.

5.7. ECG

12-Lead ECG data will be collected at the screening visit only. All data collected will be listed.

5.8. Prostate Volume Measurements

Prostate volume is measured at Screening or historically within 12 months of screening via ultrasound. All prostate volume measurements will be listed.

5.9. Treatment Exposure

The duration of exposure during the double-blind treatment period will be expressed as the time in days from the first dose as recorded on the Study Drug Administration CRF page through to last treatment day (inclusive) as recorded on the End of Treatment CRF page, excluding any days where it is recorded that an interruption has occurred on the AE CRF page or gaps in the study drug log. If no dose was witnessed at the baseline visit, the date of baseline study drug dispensation + 1 will be used. If no last treatment date is recorded, it will be imputed as the minimum of the date of the last visit on the study and the date when last dispensed study treatment was scheduled to be exhausted (last dispensation date + number of doses dispensed).

Gaps in the study drug log will be determined by comparing the dispensation date of a record to the previous dispensation date (where multiple bottles may have been dispensed). If one bottle was dispensed, then the gap is measured from study day of previous dispensation + 32; if two bottles were dispensed then the gap is measured from study day of previous dispensation + total number of pills dispensed.

Interruption in days = date interruption stated – date interruption stopped + 1

The duration is calculated by the following formula:

$$Duration\ (days) = date\ last\ double\ blind\ dose - date\ first\ double\ blind\ dose - interruption\ days + 1$$

Duration of exposure will be summarized by treatment group and overall for the SAF using summary statistics for continuous variables. Total drug interruption days due to AE and gaps in the study drug log will also be summarized. A listing will present the treatment start and end date together with the date of interruption, and the overall duration of exposure.

5.10. Treatment Compliance

Study treatment compliance (%) will be calculated as the actual number of doses divided by the expected number of doses, multiplied by 100 and summarized by treatment group and double-blind treatment period.

These numbers will be determined by the number of tablets dispensed and returned unused by the subject. Where no treatment bottle is returned, and thus the actual number of doses is

unknown, it will be assumed that the subject took medication as directed until their supply was exhausted. Imputed values will be clearly noted in the listing.

The expected number of doses is the same as the calculated duration of exposure detailed in section 7.1.

Overall compliance will be calculated for run-in and the double-blind treatment period, respectively.

$$Overall\ Compliance\ (\%) = \frac{Number\ of\ doses\ taken}{Expected\ number\ of\ doses} \times 100\%$$

The number of doses taken is summed from values recorded on the drug accountability CRF. If number of doses taken is missing for a log line within the drug accountability CRF, it will be imputed as the minimum of (number of doses dispensed, subsequent dispensation date – current dispensation date).

Treatment Compliance will be summarized for the SAF and FAS population. Additionally, the number and percentage of subjects within each treatment with compliance in the following categories will be provided: <80%, 80 – 120%, and >120%. All data will be listed.

6. EFFICACY ANALYSIS

Subjects were asked to fill out a Bladder Diary for 3 consecutive days prior to the Run-in visit, Baseline visit, and at Weeks 2, 4, 8, 12, 16, 20, and 24.

In general, the FAS will be used for all analyses of efficacy endpoints except for incontinence efficacy endpoints where the FAS-I will be used.

All outputs that present a CFB will also present descriptive statistics on the baseline value of the parameter.

6.1. Co-Primary Endpoint Analysis

6.1.1. Primary Analysis

The primary analysis will be based on the FAS. Descriptive summary for observed and CFB will be provided by visit and treatment for both co-primary endpoints. Refer to parameter derivations from Bladder Diary data on Section 4.3.1.

Co-primary 1: CFB to Week 12 in average number of daily micturition episodes

For the assessment of difference between the treatment groups in change from baseline to Week 12 in average number of daily micturition episodes, a mixed model for repeated measures (MMRM) will be used incorporating on-treatment values at all time points (Weeks 2, 4, 8, 12, 16, 20, and 24). The model takes account of any correlation within the same subject over time by using all available data on the subjects. With the assumption of missing at random (MAR), this approach will provide unbiased estimates of treatment effects derived from the model.

The MMRM model will include terms for treatment, visit, baseline value (average number of daily micturition episodes), alpha blocker use with or without 5-ARI (with vs. without), baseline urinary incontinence (yes vs. no), region (US vs. non-US), and interaction of visit by treatment. Stratification factors as randomized will be used in the model. The Kenward-Roger adjustment will be used with restricted (or residual) maximum likelihood (REML) to make statistical inference. An unstructured covariance matrix will be used to model the correlation among repeated measurements. If the unstructured covariance model fails to converge with the default Newton-Raphson algorithm, the Fisher scoring algorithm will be used to provide initial values of the covariance parameters. In the rare event that none of the above methods yield convergence, the following structures will be investigated: heterogeneous Toeplitz, Toeplitz, heterogeneous First-Order Autoregressive [AR (1)], heterogeneous compound symmetry (HCS), and compound symmetry (CS). The covariance structure converging to the best fit, as determined by Akaike's information criterion (AIC), will be used.

Primary inference will be drawn from the treatment difference between vibegron and placebo at Week 12. Results will be presented in terms of least square means (LSMEANS), treatment differences in LSMEANS, 95% confidence intervals for the treatment difference for all timepoints. Although nominal p-values for the treatment differences will be presented for all visits, the primary inference is at Week 12. The estimated LSMEANS (Standard Error (SE)) of CFB over time from the model will be plotted by treatment.

An example of the SAS code for the base procedure is given below:

```
proc mixed data = datain method = reml;
    class TRTP AVISITN USUBJID ALPHAB UI REGION;
    model CHG = TRTP AVISITN ALPHAB UI REGION BASE TRTP*AVISITN
                / ddfm=KR solution chisq;
    repeated AVISITN / subject=USUBJID type=UN r rcorr;
    lsmeans TRTP*AVISITN / pdiff=all cl alpha=0.05;
run;
```

Where TRTP is the planned treatment, AVISITN is the visit number, USUBJID is the unique subject identifier, ALPHAB indicates if the subject used 5-ARI in addition to alpha blockers at baseline, UI indicates baseline urinary incontinence, REGION indicates if the subject is in the United States or in the non-US countries and BASE indicates baseline value. The default Newton-Raphson algorithm is employed for type = UN. If the default algorithm fails to converge, use Fisher scoring algorithm to include option scoring = 5 in the model.

The normality assumptions for the primary analysis model will be assessed by inspection of the residuals from the model and normal probability plots. If assumptions of normality are not met, additional supportive analysis of the data will be performed in order to assess the robustness of the conclusions drawn from the primary analysis.

Co-primary 2: CFB to Week 12 average number of daily urgency episodes

The primary analysis will be identical to the analyses specified for the first co-primary described above with the addition of fitting in the categorical covariate of baseline average number of daily micturition episodes (≤ 12 vs. > 12, stratification factor as randomized).

Multiplicity of Co-Primary Endpoints

No multiplicity adjustment is required for the co-primary endpoints since both endpoints need to be statistically significant at the 0.05 level in order to claim trial success.

6.1.2. Sensitivity Analysis

6.1.2.1. Per-Protocol Analysis

To assess the robustness of the results from primary analysis, the same MMRM models will be conducted to both co-primary endpoints based on PPS.

6.1.2.2. Analysis of Covariance with Multiple Imputed data

An ANCOVA model will be used as a sensitivity analysis. For patients with missing Week 12 assessment due to any reason, the Week 12 value will be imputed using multiple imputation (see Appendix 11.1 for further details). The model will include terms for treatment, baseline value of the parameter, region (US vs. non-US), and stratification factors as randomized: alpha blocker use (with vs. without 5-ARI), baseline urinary incontinence (yes vs. no), baseline number of average daily micturition episodes (≤ 12 vs. > 12, only included for the urgency parameter). The analysis will be performed on FAS and PP analysis sets.

An example of the SAS code for the base procedure is given below:

```
proc mixed data = datain method = reml;
     class TRTP ALPHAB UI REGION;
     model CHG = TRTP AVISITN ALPHAB UI REGION BASE
                 / ddfm=KR solution chisq;
     Where AVISIT = "Week 12";
     lsmeans TRTP/ pdiff cl alpha=0.05;
run;
```

A tipping point analysis will accompany the multiply imputed ANCOVA results, and the estimated shift parameter will be presented to two decimal places.

The normality assumptions for the ANCOVA analysis model will be assessed by inspection of the residuals from the model and normal probability plots. If assumptions of normality are not met, additional supportive analysis of the data will be performed in order to assess the robustness of the conclusions drawn from the primary analysis.

This supportive analysis will use the multiple imputation (MI) data at Week 12 using a non-parametric analysis of covariance within the framework of the extended Cochran-Mantel-Haenszel test. Specifically, a Cochran-Mantel Haenszel mean score test will be used on the standardized mid-ranks of the residuals from an ordinary least squares regression with Week 12 MI response as a linear function of the baseline score (as a continuous variable). P-value for the between treatment comparison for each primary endpoint will be derived for the model.

This methodology will be carried out as follows:

1. The observed value of the primary endpoint (i.e., the last on treatment average number of daily micturition episodes) will be fitted using an ordinary least squares regression as a function of the baseline value at each visit

2. Standardized mid-ranks of the residuals from the regression will be calculated.

3. The standardized mid-rank from the last observation will be carried forward for missing values at week 12

4. Standardized mid-ranks will be recalculated

5. A CMH mean score statistic and p-value adjusting for the stratification groupings will be calculated for between treatment comparisons

An example of the SAS code for the base procedure is given below:

```
 proc glm Data=datain;
    by Avisitn;
    model aval = Base;
   output out=glmout residual=Residual;
run;

 proc rank data=glmout out=Rank1 nplus1 ties=mean;
   by Avisitn;
   var Residual;
   ranks RANKOBS;
 run;

 proc sort data=rank1;
    by subject avisitn;
 run;


data rankmi;
   set rank1;
   by subject avisitn;
   if last.subject;
run;

proc rank data=rankmi out=ADRANK nplus1 ties=mean;
   by avisitn;
   var RANKOBS;
   ranks Rankana;
run;

proc freq data = ADRANK;
    tables stratified vars*TRTP*Rankana / cmh2 scores = table;
run;
```

6.1.2.3. MMRM with Stratification Factors as Observed

If more than 7% of subjects are misstratified, the co-primary MMRM efficacy analysis will be repeated using stratification factors as calculated from the baseline diary (baseline average number of daily micturition episodes, baseline incontinence) and prior medications (alpha blocker use with or without 5-ARI).

6.1.3. Subgroup Analysis

For following selected subgroups listed in Section 4.1.3, a separate MMRM model will be fit on co-primary endpoints, respectively, based on the FAS.

- Alpha blocker use with or without 5-ARI (with vs. without, stratification factor as randomized)
- Baseline urinary incontinence (yes vs. no, stratification factor as randomized)
- Region (US vs. Non-US)
- Age category 2 (<65, ≥65 years)
- Race (White, African American, Asian and Other)

The same model terms as used for the primary analysis specified in Section 6.1.1 will be applied with additional terms for subgroup main effect, treatment by subgroup interaction, and treatment by subgroup by visit interaction. Only the estimate and p-value for the interaction term from the MMRM will be provided. With the exception of Age category 2 and Race subgroups, if less than 20% of randomized subjects are in a subgroup the corresponding subgroup analysis will not be performed. Descriptive summary statistics will be provided for all subgroups.

A forest plot will be provided to present the results from subgroup analyses. Side-by-side 95% confidence intervals will be plotted for each comparison to control within each subgroup, where the 95% confidence intervals are taken from the analysis table described above. The confidence intervals are stacked vertically on the page, with a reference line at zero for mean. All the subgroups should be plotted on a single figure (using multiple pages, if necessary).

The following subgroups will be used for descriptive statistics only: Age category 1, Age category 3, Prior anticholinergic, and Prior beta-3 agonist use.

6.2. Secondary Efficacy Analyses

Descriptive summary for observed and CFB will be provided by treatment and visit for each endpoint. Refer to Bladder Diary related parameter derivations in Section 4.3.1 and average IPSS Storage score in Section 4.3.2.1.

The following secondary efficacy endpoints will be analyzed based on the FAS:

- CFB to Week 12 in the average number of nocturia episodes per day
- CFB to Week 12 in the International Prostate Symptom Score (IPSS) Storage score (1-week recall)

- CFB to Week 12 in the average volume voided per micturition

The following secondary efficacy endpoint will be analyzed based on the FAS-I:

- CFB to Week 12 in the average number of daily urge urinary incontinence episodes for subjects with urinary incontinence at baseline

All four secondary endpoints will be analyzed and presented similarly as described in the primary analysis of co-primary efficacy endpoints in Section 6.1.1. The same MMRM model used for the CFB to Week 12 in average number of daily urgency episodes will be applied to all secondary endpoints except for CFB to Week 12 in the average number of daily urge urinary incontinence episodes. For this incontinence endpoint, the baseline urinary incontinence (yes vs. no) strata will not be fitted in the MMRM model.

As described in Section 4.1.2, a hierarchical gate-keeping procedure will be used to claim statistical significance for the secondary efficacy endpoints sequentially if statistical significance is achieved at the 0.05 level for both co-primary endpoints.

6.3. Other/Exploratory Endpoints

Refer to diary related parameter derivations in Section 4.3.1 and quality of life related parameters in Section 4.3.2. These analyses will be based on FAS, unless stated otherwise.

6.3.1. Efficacy

The CFB at Weeks 4, 8, 16 (where collected), 20, and 24 for all primary and secondary efficacy variables will be included into the MMRM models along with Week 12 data as described in Sections 6.1 and 6.2. The additional continuous exploratory endpoints will be summarized descriptively.

- CFB to Weeks 4, 8, 12, 20, and 24 in the IPSS total score and IPSS Quality of Life, Storage, and Voiding scores (FAS)

- CFB to Weeks 2, 4, 8, 12, 16, 20, and 24 in the average number of total daily incontinence episodes per day for subjects with urinary incontinence at baseline (FAS-I)

The number and percentage of responders and non-responders for the following exploratory endpoints will be tabulated by treatment and visit.

- Urgency episodes 50% responder (exploratory) at Weeks 4, 8, 12, 16, 20, and 24

- Urgency urinary incontinence (UUI) episodes 75% responder (exploratory) at Weeks 4, 8, 12, 16, 20, and 24

The treatment comparison of responder will use a Cochran-Mantel-Haenszel (CMH) common risk difference estimate stratified by appropriate strata per randomization stratification [6] with weights proposed by Greenland and Robins [7], which is calculated as follows. For the urgency episode responder analysis, the strata groups include baseline average micturition episodes per day (≤ 12 vs > 12), alpha blocker use with or without 5-ARI (with or without) and urinary

Incontinence (yes or no). For the incontinence responder analysis, urinary Incontinence (yes or no) will not be included in.

$\hat{\delta}_{MH} = \frac{\sum_{i=1}^{u} w_i \cdot \hat{\delta}_i}{\sum_{i=1}^{u} w_i}$, where

$\hat{\delta}_i = \frac{x_i}{n_i} - \frac{y_i}{m_i}$ denotes the risk difference in stratum $i, i = 1, \cdots, u$

$w_i = \frac{n_i \cdot m_i}{n_i + m_i}$ denotes the weight of stratum $i, i = 1, \cdots, u$

x_i denotes the number of subjects with event in treatment_1 in stratum i, $i = 1, \cdots, u$

y_i denotes the number of subjects with event in treatment_2 in stratum i, $i = 1, \cdots, u$

n_i denotes the number of subjects on treatment_1 in stratum i, $i = 1, \cdots, u$

m_i denotes the number of subjects on treatment_2 in stratum i, $i = 1, \cdots, u$

The estimated variance of $\hat{\delta}_{MH}$ is calculated as:

$$\widehat{var}\left(\hat{\delta}_{MH}\right) = \frac{\sum_{i=1}^{u} L_i}{\left(\sum_{i=1}^{u} w_i\right)^2}$$

where $L_i = \frac{x_i(n_i - x_i)\, m_i^3 + y_i(m_i - y_i)\, n_i^3}{n_i \cdot m_i \cdot (n_i + m_i)^2}$, $i = 1, \cdots, u$

Assuming a normal distribution of $\hat{\delta}_{MH}$, an approximate 95% CI is given as follows, where $z_{0.975}$ is the 97.5% quantile of the standard normal distribution:

$$CI = \left[\hat{\delta}_{MH} \pm z_{0.975} \cdot \sqrt{\widehat{var}\left(\hat{\delta}_{MH}\right)}\right]$$

Also, the approximate p-value can be calculated using the following:

$$\text{p} - \text{value} = 2 \cdot \Pr\left[Z > \left|\frac{\hat{\delta}_{MH}}{\sqrt{\widehat{var}\left(\hat{\delta}_{MH}\right)}}\right|\right], \text{where } Z \sim N(0,1)$$

If there is a stratum for a treatment group that has 0 subjects in it, the 0 count will be replaced by 0.5 in order to prevent dividing by 0 in the above equations, as suggested in Greenland and Robins.

The estimated common risk difference, and associated p-value and 2-sided 95% confidence interval will be tabulated.

6.3.2. Quality of Life

The FAS will be the analysis population for the PRO data unless stated otherwise.

6.3.2.1. CFB in [redacted] Scores

The questionnaire is assessed at baseline, Week 12 and Week 24. The derivation of [redacted] scores for Symptom Bother, Coping, Concern, Sleep, Social Interaction and total HRQL are detailed in Section 4.3.2.2. These derived scores and CFB will be descriptively summarized by treatment and visit.

6.3.2.2. [redacted]

[redacted] questionnaire is assessed at baseline, Week 12 and Week 24. The 5 questions related to mobility, self-care, usual activities, pain/discomfort and anxiety/depression will be summarized descriptively as categorical variables by treatment and visit. Each domain includes 5 level answers: no problems, slight problems, moderate problems, severe problems and extreme problems. At Week 12 and Week 24 the Pareto Classification of Health Change will be presented relative to baseline (better, worse, same, or mixed). The mapped Crosswalk index value and change from baseline will be summarized as a continuous variable by treatment and visit.

Observed VAS value, change from baseline and percent change from baseline will be summarized as a continuous variable by treatment and visit.

6.3.2.3. Patient Global Impression Scores

The questionnaire is assessed at baseline, Week 4, Week 12 and Week 24. The Patient Global Impression (PGI) questions are designed to assess a subject's overall impression of OAB in 5 categories: [redacted]

Each of the responses to the PGI questions will be assigned a numerical value starting from 1. A higher score represents worse outcome in each category.

Table 7: PGI Scoring

PGI Category	Score
PGI-[redacted]	1 = None 2 = Mild 3 = Moderate 4 = Severe
PGI-[redacted]	1 = Complete control 2 = A lot of control 3 = Some control 4 = Only a little control 5 = No control

PGI Category	Score
PGI-[redacted] PGI[redacted]	1 = Never 2 = Rarely 3 = Sometimes 4 = Often 5 = Very often
PGI-[redacted]	1 = Much better 2 = Moderately better 3 = A little better 4 = No change 5 = A little worse 6 = Moderately worse 7 = Much worse

The questions PGI-[redacted], PGI-[redacted] PGI-[redacted] and PGI-[redacted] ordinal assessments going from lower impact of disease with score of ranging from 1 to 4 for [redacted] and 1 to highest impact of disease with score of 5 for PGI[redacted] PGI [redacted] and PGI-[redacted] These questions will be summarized descriptively as numerical variables of score and change from baseline by treatment and visit. Additionally, these questions will be analyzed as categorical variables with count and percent for each category by treatment and visit.

The PGI[redacted] question is asked relative to baseline. The score will only be summarized as a categorical variable with count and percent for each category by treatment and visit.

The FAS will be the analysis set for all categories except for PGI[redacted] The FAS-I will be the analysis set for the PGI[redacted] analysis.

6.3.2.4. International Index of Erectile Function

The questionnaire is assessed at baseline, Week 4 and Week 24. The sum of scores and CFB for each of 4 main domains of male sexual function: erectile function, orgasmic function, sexual desire and intercourse satisfaction plus an overall satisfaction will be descriptively summarized by treatment and visit.

7. SAFETY ANALYSIS

The SAF will be used for all safety analyses unless noted otherwise. Safety will be assessed on the basis of AE reports, clinical laboratory data, physical examinations, vital signs, orthostatic blood pressure (Part 1 only), post-void residual volume, urodynamics assessments (Urodynamics substudy only), and prostate volume measurement.

7.1. Adverse Events

AEs will be coded using MedDRA version 22.0 or later.

All reported AEs (whether treatment emergent or not) will be included in by-subject AE listings. Sorting will be by country, site, subject, date of event, SOC, PT and then verbatim description.

An AE will be considered treatment emergent (TEAE) if it begins or worsens in severity after the first dose of the double-blind Study Treatment through 5 days after the last dose of Study Treatment or rollover to the extension study, whichever occurs first. Partial AE start dates will be imputed as detailed in Section 4.2.3.

Summary tables will be based on TEAEs. The incidence of TEAEs will be presented using counts and percentages of subjects with TEAEs and tabulated by SOC and PT. SOC will be sorted in descending frequency and PT within SOC will be sorted by descending frequency based on the incidence across subjects overall. If a subject has multiple occurrences (start and stop) of an event associated with a specific SOC or PT within a SOC, a subject will only be counted once in the incidence count for the SOC or PT within SOC respectively.

An overall summary table of AEs by treatment group will be presented detailing the number and percentage of subjects, and number of events for the following categories:

- At least one TEAE;
- At least one Treatment-Related TEAE;
- At least one Grade ≥ 3 TEAE (Mild = Grade 1, Moderate = Grade 2, Severe or Medically Significant = Grade 3, Life-Threatening = Grade 4, Death = Grade 5)
- At least one Grade ≥ 3 Treatment-Related TEAE;
- At least one Serious TEAE;
- At least one Serious Treatment-Related TEAE;
- At least one TEAE leading to Discontinuation from Study Medication;
- At least one TEAE of Special Interest;
- At least one Treatment-Related TEAE of Special Interest

Where not otherwise noted, all the following summaries are presented by PT in descending overall frequency:

- All TEAEs (by SOC and PT in descending overall frequency);
- All TEAEs (by PT in descending overall frequency)
- Treatment-Related TEAEs (i.e., possibly or probably related);
- All TEAEs by SOC, PT, and maximum severity (where the maximum intensity per patient will be counted at each level of summarization);
- TEAES with Grade ≥ 3;
- Treatment-Related TEAES with Grade ≥ 3;

- Serious TEAEs;
- Treatment-Related Serious TEAEs;
- Fatal TEAEs;
- TEAEs leading to Discontinuation from Study Treatment;
- TEAE of Special Interest;
- Treatment-Related TEAE of Special Interest;
- Non-fatal TEAEs;
- Hypertension TEAEs by Pre-existing Hypertension (Yes vs No) and Baseline Hypertension (Yes vs. No). Hypertension TEAEs will be selected as any TEAE with preferred term of Hypertension.
- Non-serious TEAEs occurring in 2% or more of any treatment group.

In addition, a summary of all TEAEs by PT occurring in at least 2% of subjects in the vibegron arm and greater than the placebo arm will be created and sorted by descending frequency in the vibegron arm.

Adverse events of special interest for this study include:

- Adverse events consistent with urinary retention
- Adverse events suggestive of cystitis or urinary tract infection (UTI)
- Potential major cardiac and cerebrovascular events, including death (or any event with fatal outcome), myocardial infarction, cerebrovascular accident, hospitalization for unstable angina or chest pain, hospitalization for heart failure requiring hospitalization, and coronary revascularization/angioplasty/stent
- Adverse events consistent with hypertension (see protocol section 8.4.6)
- Elevated serum aspartate aminotransferase (AST) or alanine aminotransferase (ALT) lab value requiring that study drug be temporarily withheld or permanently discontinued
- Neoplasms

Treatment listings will include the treatment arm, start and stop dates of the AE, and days on study relative to the day of first dose of study treatment. A Treatment related AE is defined as an AE for which the investigator classifies the AE as being “Probably Related” or “Possibly Related” to study treatment on Adverse Event CRF. Missing relationship and severity (intensity) will be imputed per Section 4.2.3.

The following additional listings will be provided:

- Listing of deaths
- Listing of Serious TEAEs

- Listing of treatment-emergent AESIs
- Listing of TEAEs leading to withdrawal or Interruption of study treatment

Listing of all AEs with a flag for TEAEs and onset (Prior = prior to first dose of single-blind medication, Run-in = on or after first dose of single-blind medication but prior to first dose of double-blind medication, or Treatment = on or after first dose of double-blind medication.)

A listing of all medical history and pre-double-blind treatment adverse events with coded preferred terms belonging to the standard MedDRA query of hypertension will be created.

7.2. Laboratory Evaluations

All continuous laboratory parameters will be summarized descriptively by absolute value at each visit by treatment group, together with the corresponding changes from baseline. All parameters will be summarized in SI units.

The number and percentage of subjects with laboratory measurements outside of the central laboratory normal range will also be summarized by treatment group and visit. Shift tables from baseline to maximum post-baseline value, to minimum post-baseline value, last post-baseline value, and at each post-baseline visit will be provided for hematology and chemistry parameters to display low, normal, high, and missing values by treatment group in a 3-by-3 contingency table. Denominators for percentages will be the number of subjects with non-missing data at the specific assessment and baseline.

Maximum post-baseline total bilirubin will be presented (<2 and ≥2 x ULN) and plotted against maximum post-baseline ALT (<3, ≥3 - <5, ≥5-<10, and ≥10 x ULN), expressed as multiples of ULN. This will be repeated to show maximum post-baseline total bilirubin against maximum post-baseline AST.

Data for subjects with ALT or AST ≥3 x ULN, and bilirubin ≥2 x ULN will be presented, which will include all visits for this subset of subjects. A line plot of liver biochemistry test results (including ALP, ALT, AST, total bilirubin, and GGT) over time will also be presented for this subset of patients.

A sample for the urinalysis and urine culture will be sent to the central laboratory only if the urine dipstick performed at the site tests positive for the presence of leukocytes, nitrites, or blood cells. Dipstick results and central urinalysis data will be provided only in the listing.

Any data outside the central laboratory normal reference ranges will be explicitly noted on the listings that are produced.

7.3. Vital Sign Assessments

7.3.1. Height, Weight, and Sitting Vital Signs

Vital sign data including systolic and diastolic blood pressure, pulse rate, respiration rate, body temperature and weight will be collected at all study visits except for the safety follow-up visit. Height will be measured at Screening only. Blood pressure will be measured in triplicate at each

visit. The average of triplicates will be used for summary; if fewer than three measurements were collected the average of available measurements will be used. For subjects in Study Part 1 and visits with orthostatic blood pressure assessments, blood pressure (in triplicate) and pulse rate taken pre-dose in sitting position are equivalent to pre-dose regular vital sign values and will be used.

Vital signs will be summarized using absolute value and change from baseline. A summary by-visit according to the visit window rule in Section 4.2.4.2, the maximum, and last post-baseline observation for each subject will be summarized for each of the vital signs. For change from baseline in blood pressure and pulse rate, least squares mean estimates and 95% CIs as described in section 4.2.2 will be presented from an ANCOVA model including terms for treatment, age, baseline parameter value, and pre-existing hypertension (Yes vs. No).

The line plot of mean (SE) of change from baseline in blood pressure and pulse rate over time will be prepared. A by-subject listing, sorted by subject identifier, will be presented including all vital sign results (scheduled or unscheduled).

7.3.1.1. Vital Sign Subgroups

The vital sign data for blood pressure and pulse rate will be summarized by subgroups that may be at a higher risk, in Table 8

Table 8: At Risk Subgroup for Vital Sign Summaries

Subgroup	Definition of At Risk Subgroup
Age	≥ 75 years
Pre-Existing Hypertension	Pre-Existing Hypertension will be based on medical history and/or baseline hypertension, as defined in Section 4.3.5
High BMI at baseline	BMI ≥ 35 kg/m^2
At Risk eGFR Group	eGFR ≤ 60 mL/min/SA

7.3.1.2. Categorical/Threshold Analysis

Categorical analyses (e.g. percentage of patients with an absolute change from baseline of ≥ 5, 10, or 15 mmHg for blood pressure and ≥ 5, 10, or 15 bpm for heart rate) of any post-baseline change and at 2 consecutive post-baseline visits timepoints.

The categorical analysis will be produced for the safety population and for subgroups that might be at a higher risk (e.g., pre-existing hypertension, older patients [age ≥ 75 years], renal impairment [≤ 60 mL/min/SA], and high BMI ([≥ 35 kg/m^2]), as well as their complements.

7.3.2. Orthostatic Blood Pressure and Pulse Rate (Part 1 only)

Blood pressure and pulse rate measurements are collected at Screening, Run-in, Baseline, Week 2 and 4 to assess for potential orthostatic vital sign changes. Triplicate measurements are

obtained for both sitting and standing blood pressure assessments at Hour 0 (pre-dose) and Hours 1, 2, 4, and 6 post-dose. A singular pulse rate is also obtained in sitting and standing positions at the same hourly timepoint. The average blood pressure of triplicates will be used for summary. Within each timepoint, orthostatic blood pressure and pulse rate will be derived by subtracting time-matched standing average values from sitting average values. Then time-matched change from Day 1 will be derived as follows:

- Step 1: calculate change from pre-dose at each visit

  CFPre at baseline visit: (Baseline 1 hr – Baseline pre); (Baseline 2hr – Baseline pre) ….

  (Baseline 6 hr – Baseline pre)

  CFPre at Week 2 visit: (Week 2 1 hr – Week 2 pre); (Week 2 2hr – Week 2 pre) ….

  (Week 2 6 hr – Week 2 pre)

  CFPre at Week 4 visit: (Week 4 1 hr – Week 4 pre); (Week 4 2hr – Week 4 pre) ….

  (Week4 6 hr – Week4 pre)

- Step 2: calculate time-matched from Day 1 by subtracting time-matched CFPr of Baseline visit from CFPr of Week 2 and Week 4

  [(Week 2 1 hr – Week 2 pre) - (Baseline 1 hr – Baseline pre)]

  [(Week 2 2hr – Week 2 pre) - (Baseline 2 hr – Baseline pre)]

  …

  [(Week2 6hr – Week2 pre) - (Baseline 6hr – Baseline pre)]

Observed SBP, DBP, and pulse rate will be summarized by treatment, visit and assessment time including corresponding change from baseline value. Baseline is the average of pre-dose measurement at Day 1. The following summary tables will be prepared for derived orthostatic SBP, DBP and pulse rate:

- Time-matched orthostatic parameters by treatment, visit and assessment time
- Time-matched change from Day 1 in orthostatic parameters by treatment, visit and assessment time
- Maximum CFB over 6 hours in in orthostatic parameters by treatment and visit
- Average CFB over 6 hours in orthostatic parameters by treatment and visit

In addition, the following plots will be prepared for orthostatic vital sign parameters

- Line plot over time of time-matched change from Day 1 of orthostatic parameters
- Empirical cumulative distribution function of maximum CFB over 6 Hours

For SBP and DBP, the number and percentage of subjects with a decrease of ≥20 mmHg for SBP and ≥10 mmHg for DBP from sitting to standing position will also be presented by treatment, visit and assessment time. All orthostatic blood pressure and pulse rate measurements will be listed.

7.4. Physical Examination

Brief physical examination data will be collected at Screening, Baseline, Week 12 and 24. A listing of all physical exam results for subjects with at least one abnormal result will be produced.

7.5. Post-Void Residual (PVR) Urine Volume

PVR urine volume data will be summarized at Baseline and every visit, including change from baseline for post-baseline visits. Least squares mean estimates and 95% CIs as described in section 4.2.2 will be presented from an ANCOVA model including terms for treatment and baseline parameter value.

Categorical summaries of PVR at the following categories: <100 mL, ≥ 100 and < 200 mL, ≥ 200 and < 300 mL, ≥ 300 mL will be presented. Line plots of least squares means change from baseline and 95% CIs by visit and treatment will be presented. All PVR data will be listed.

7.6. Urodynamics Assessments (Urodynamics Sub-study Only)

Urodynamics assessments are collected at or within 6 months prior to Baseline and at Week 12 for Part 2 subjects who consented to participate in the Urodynamics sub-study. Urodynamics parameter estimates will be assessed by a central reviewer. The following parameters will be measured:

- Peak flow rate during voiding (Qmax) (mL/s)
- Detrusor pressure at peak flow rate (Corrected PdetQmax) (cmH_2O)
- Instilled volume at first involuntary detrusor contraction (IDC) (Vol@1st IDC) (mL), if an IDC occurs
- Maximum detrusor pressure during the first involuntary detrusor contraction (Corrected Pdet@1st IDC) (cmH_2O), if an IDC occurs
- Maximum cystometric capacity (Corrected MCC) (mL)
- Maximum detrusor pressure during the storage phase (Corrected PdetQmax Storage) (cmH_2O)
- Voided volume (VV) (mL)

Additional parameters will be derived from the measured values:

- Bladder Outlet Obstruction Index (BOOI)=PdetQmax-2·Qmax
- Bladder Contractility Index (BCI)=PdetQmax + 5·Qmax
- Bladder Voiding Efficiency (BVE)=(VV·100)/MCC

Descriptive statistics of observed values at Baseline, Week 12, any time post-baseline,and change from baseline to Week 12 and to any time post-baseline will be presented for each of the parameters for the UES set. Least squares mean estimates and 95% CIs as described in section 4.2.2 will be presented from an ANCOVA model including terms for treatment and baseline parameter value. One-sided 95% CIs will be provided for Qmax and PdetQmax.

Categorical summaries of BOOI [Unobstructed (<20), Equivocal (20-40), and Obstructed (>40)] and BCI [Strong (>150), Normal (100-150), Weak (<100)] will also be presented.

If PP-UES differs from UES by more than 5 subjects, Qmax and PdetQmax will be summarized as above for the PP-UES population.

All urodynamics assessments will be listed, including investigator derived parameters. If a urodynamics measurement was taken after the end of study date in URO-901-3005 (implying that the subject initiated URO-901-3006 before the evaluation), that measurement will not be included in any tabular summaries but will be listed.

7.7. Uroflow Measures

Uroflowmetry is conducted at Screening and Week 24 (or early withdrawal). Measured parameters include maximum flow rate (Qmax) (mL/s), average flow rate (Qave) (mL/s), and volume voided (VV) (mL).

Uroflowmetry measures require a minimum void of 125 mL to be interpretable. Summary statistics will only be provided for uroflow tests where VV≥125 mL. The summary will comprise a continuous summary at baseline and post-baseline, including change from baseline. Least squares mean estimates and 95% CIs as described in section 4.2.2 will be presented from an ANCOVA model including terms for treatment and baseline parameter value.

If a uroflow measurement was taken after the end of study date in URO-901-3005 (implying that the subject initiated URO-901-3006 before the evaluation), that measurement will not be included in any tabular summaries but will be listed. All Uroflow measurements will be listed.

7.8. IPSS Total Score

IPSS total score at Screening, Day 1, Weeks 4, 8, 12, and 24 (or early withdrawal) will be summarized by treatment and visit, including change from baseline.

8. COVID-19 CONSIDERATIONS

This study was conducted during the COVID-19 global pandemic (starting around March 2020). During the time period of the pandemic, it is anticipated that changes in study visit schedules, missed visits, or patient discontinuations may lead to missing information (e.g., for protocol-specified procedures). In accordance with the FDA Guidance titled Statistical Considerations for Clinical Trials During the COVID-19 Public Health Emergency (first issued in June 2020), it is important to capture specific information in the case report form that explains the basis of the missing data, including the relationship to COVID-19 for missing protocol-specified information (e.g., from missed study visits or study discontinuations due to COVID-19).

The proportion of subjects with COVID-19 impact on visits will be summarized by treatment group overall and by visit. The reasons (e.g. subject acquired COVID-19, subject unable to travel due to COVID-19, investigative site closure due to COVID-19, etc) will also be summarized. A listing of subjects who had COVID-19 information collected will be generated which will document the date of contact, visit, visit impact, reason for visit impact, and if any doses were missed due to IP availability.

9. CHANGE FROM ANALYSIS PLANNED IN PROTOCOL

9.1. Exploratory endpoints

In protocol Section 3, the changes from baseline at Week 12 for the following endpoints are specified as other/exploratory endpoints. In addition, the CFB at Week 24 will be included.

- Symptom Bother Score as assessed by [redacted]
- HRQL total score as assessed by [redacted]

IPSS total score is included as an exploratory efficacy endpoint as well as a safety endpoint.

9.2. Responder Analysis

Protocol Section 9.4.4 mentions using CMH risk difference estimate for response efficacy endpoints. Missing Week 12 data will be analyzed using multiple imputation (MI). Considering the responder endpoints are exploratory in the study, the proposed MI analysis will not be carried out.

9.3. COVID-19 Considerations

The COVID-19 global pandemic developed during the conduct of the study. The recommended impact analysis was therefore added to the SAP, as the operational change of increased visit windows had been initially communicated via study memo with no protocol amendment. COVID-19 expanded visit windows are spelled out in Protocol Amendment 2; 6-week visit extensions due to COVID-19 were available starting in March 2020.

10.REFERENCES

[1] Abrams P, Kaplan S, De Koning Gans HJ et al: Safety and tolerability of tolterodine for the treatment of overactive bladder in men with bladder outlet obstruction. J Urol 2006; 175: 999.

[2] Nitti V, Mitcheson HD, Rosenberg S, et al. Urodynamics and safety of the B3-adrenoceptor agonist, Mirabegron, in males with lower urinary tract symptoms and bladder outlet obstruction. J Urol 2013; 190: 1320–7.

[3] Jiang YH, Lin VC, Liao CH, Kuo HC. International Prostatic Symptom Score-voiding/storage subscore ratio in association with total prostatic volume and maximum flow rate is diagnostic of bladder outlet-related lower urinary tract dysfunction in men with lower urinary tract symptoms. PLoS One. 2013;8(3):e59176.

[4] EQ-5D-5L crosswalk value set https://euroqol.org/eq-5d-instruments/eq-5d-5l-about/valuation-standard-value-sets/crosswalk-index-value-calculator/

[5] "Guideline on adjustment for baseline covariates in clinical trials." EMA, 26-Feb-2015

[6] Chunlei Ke, Jianming Wang, Charlie Zhang, Qi Jiang & Steven Snapinn (2017): On Errors in Stratified Randomization, Statistics in Biopharmaceutical Research, DOI:10.1080/19466315.2016.1270229

[7] S. Greenland and J. M. Robins, "Estimation of a Common Effect Parameter from Sparse Follow-Up Data," *Biometrics*, vol. 41, no. 1, p. 55, Mar. 1985.

[8] Example 80.13 Sensitivity Analysis with the Tipping-Point Approach https://documentation.sas.com/doc/en/pgmsascdc/9.4_3.4/statug/statug_mianalyze_examples13.htm

11.APPENDIX

11.1. Appendix 1. Multiple Imputation for the Response Efficacy Endpoints

This section summarizes the multiple imputation approach for the sensitivity analysis of primary efficacy endpoints to assess the robustness of the treatment effect under different underlying assumptions to account for missing data. The goal is to impute missing values as follows:

- Average number of daily micturition episodes
- Average number of daily urgency episodes

First the average daily numbers will be derived for all post-baseline visits as described in Section 4.3.1 respectively, then the following two-stage imputation will be applied. The following steps will be followed to perform this analysis.

Step 1: Create a horizontal dataset for the efficacy data where each patient has 1 row per parameter and a column for each visit from baseline through week 12, inclusive.

Step 2: Impute missing average daily numbers at Weeks 2, 4, and 8 based on MCMC to obtain a monotone missing data pattern with 30 replicates. Example code is:

```
proc mi data = datain1 out = dataout1 seed = 123 nimpute=30 noprint;
    var baseline WEEK_02 WEEK_04 WEEK_08;
    mcmc chain = multiple impute=monotone;
    em maxiter = 1000;
run;
```

Step 3: For each of the 30 replicates from Stage 1, impute the missing data following a monotone missing data pattern so that there are 30 replicates of the complete dataset. Example code is:

```
proc mi data = dataout1 out = dataout2 seed = 321 nimpute=1 noprint;
   class TRTP OABTYPE REGION ;
   monotone regression (WEEK_12 = baseline WEEK_02 WEEK_04 WEEK_08);
   var WEEK_12 baseline WEEK_02 WEEK_04 WEEK_08;
run;
```

Step 4: Transpose the dataset out2 to the usual ADaM-like dataset with 1 row per patient, per parameter, per analysis visit.

Step 5: Fit the ANCOVA model from Section 6.1.1 for each of the replicate complete datasets.

Step 6: Combine the results from Step 4 using PROC MIANALYZE.

```
proc mianalyze data = MIdatain;
    modeleffects ESTIMATE;
```

```
    stderr STDERR;
    ods output ParameterEstimates = MIout;
run;
```

The tipping point analysis will use the same approach and follow Example 80.13 from SAS/STAT User's Guide [8].

11.2. Table of Contents for Data Display Specifications

11.2.1. Output Tables

Table 9: List of Output Tables

Title		Population	Programming notes
Study Population			
14.1.1.1	Subject Disposition	Screened	Include three sections: screened, run-in, randomized
14.1.1.2	Subject Randomization by Country and Investigator	Randomized	
14.1.2.1	Major Protocol Deviations	FAS	
14.1.2.2	Reasons for Exclusion from Analysis Sets	Randomized	
14.1.3.1	Subject Demographic and Baseline Characteristics	SAF	
14.1.3.2	Subject Demographic and Baseline Characteristics	SAF-Orth	
14.1.3.3	Subject Demographic and Baseline Characteristics	FAS	
14.1.3.4	Subject Demographic and Baseline Characteristics	FAS-I	
14.1.3.5	Subject Demographic and Baseline Characteristics	PPS	

Title		Population	Programming notes
14.1.3.6	Subject Demographic and Baseline Characteristics	PPS-I	
14.1.3.7	Subject Demographic and Baseline Characteristics	SAF-UDS	
14.1.3.8	Subject Demographic and Baseline Characteristics	UES	
14.1.3.9	Other Baseline Characteristics	FAS	
14.1.3.10	Other Baseline Characteristics	FAS-I	
14.1.3.11	Other Baseline Characteristics	PPS	
14.1.3.12	Other Baseline Characteristics	PPS-I	
14.1.4	Medical History	SAF	
14.1.5.1	Prior Non-Overactive-Bladder (OAB) Medication	SAF	
14.1.5.2	Prior Overactive Bladder (OAB) Medication	SAF	
14.1.6.1	Treatment Exposure	SAF	Double-blind period
14.1.6.2	Treatment Exposure	FAS	Double-blind period
14.1.6.3	Treatment Compliance Double-Blind Period	SAF	
14.1.6.4	Treatment Compliance Double-Blind Period	FAS	
14.1.7.1	COVID-19 Impact	SAF	

Title		Population	Programming notes
Efficacy Endpoints			
Co-Primary Endpoints			
14.2.1.1	Average Number of Daily Micturition Episodes	FAS	Descriptive statistics, include CFB
14.2.1.2	Change from Baseline in Average Number of Daily Micturition Episodes Primary Efficacy Analysis (MMRM)	FAS	Include all intervals but Week 12 in the front
14.2.1.3	Change from Baseline at Week 12 in Average Number of Daily Micturition Episodes Primary Efficacy Sensitivity Analysis (ANCOVA with MI)	FAS	
14.2.1.4	Change from Baseline in Average Number of Daily Micturition Episodes Primary Efficacy Analysis	PPS	Descriptive statistics, include CFB
14.2.1.5	Change from Baseline at Week 12 in Average Number of Daily Micturition Episodes Primary Efficacy Sensitivity Analysis (ANCOVA with MI)	PPS	Include all intervals but Week 12 in the front
14.2.1.6	Average Number of Daily Micturition Episodes by Subgroup	FAS	Descriptive statistics, include CFB
14.2.1.7	Change from Baseline in Average Number of Daily Micturition Episodes Primary Efficacy Analysis (MMRM) by Subgroup	FAS	Include all intervals but Week 12 in the front
14.2.2.1	Average Number of Daily Urgency Episodes	FAS	Descriptive statistics, include CFB

Title		Population	Programming notes
14.2.2.2	Change from Baseline in Average Number of Daily Urgency Episodes Primary Efficacy Analysis (MMRM)	FAS	Include all intervals but Week 12 in the front
14.2.2.3	Change from Baseline at Week 12 in Average Number of Daily Urgency Episodes Primary Efficacy Sensitivity Analysis (ANCOVA with MI)	FAS	
14.2.2.4	Change from Baseline in Average Number of Daily Urgency Episodes Primary Efficacy Analysis (MMRM)	PPS	Include all intervals but Week 12 in the front
14.2.2.5	Change from Baseline in Average Number of Daily Urgency Episodes Primary Efficacy Sensitivity Analysis (ANCOVA with MI)	PPS	Include all intervals but Week 12 in the front
14.2.2.6	Average Number of Daily Urgency Episodes by Subgroup	FAS	Descriptive statistics, include CFB
14.2.2.7	Change from Baseline in Average Number of Daily Urgency Episodes Primary Efficacy Analysis (MMRM) by Subgroup	FAS	Include all intervals but Week 12 in the front
Secondary Efficacy Endpoints			
14.2.3.1	Average Number of Nocturia Episodes per Day	FAS	Descriptive statistics, include CFB
14.2.3.2	Change from Baseline in Average Number of Nocturia Episodes per Day (MMRM)	FAS	Include all intervals but Week 12 in the front

Title		Population	Programming notes
14.2.3.3	Change from Baseline in Average Number of Nocturia Episodes per Day (MMRM)	PPS	Include all intervals but Week 12 in the front
14.2.4.1	Average Number of Daily Urge Urinary incontinence Episodes	FAS-I	Descriptive statistics, include CFB
14.2.4.2	Change from Baseline in Average Number of Daily Urge Urinary incontinence Episodes (MMRM)	FAS-I	Include all intervals but Week 12 in the front
14.2.4.3	Change from Baseline in Average Number of Daily Urge Urinary incontinence Episodes (MMRM)	PPS-I	Include all intervals but Week 12 in the front
14.2.5.1	International Prostate Symptom Score (IPSS) Storage Score	FAS	Descriptive statistics, include CFB
14.2.5.2	Change from Baseline in International Prostate Symptom Score (IPSS) Storage Score (MMRM)	FAS	Include all intervals but Week 12 in the front
14.2.5.3	Change from Baseline in International Prostate Symptom Score (IPSS) Storage Score (MMRM)	PPS	Include all intervals but Week 12 in the front
14.2.6.1	Average Volume Voided per Micturition	FAS	Descriptive statistics, include CFB
14.2.6.2	Change from Baseline in Average Volume Voided per Micturition (MMRM)	FAS	Include all intervals but Week 12 in the front
14.2.6.3	Change from Baseline in Average Volume Voided per Micturition (MMRM)	PPS	Include all intervals but Week 12 in the front

Title		Population	Programming notes
Other Efficacy and PROs Endpoints			
14.2.7.1	International Prostate Symptom Score (IPSS) Quality of Life Score	FAS	Descriptive statistics, include CFB
14.2.7.2	International Prostate Symptom Score (IPSS) Voiding Score	FAS	Descriptive statistics, include CFB
14.2.8.1	Urgency Episodes per Day 50% Responder Analysis (CMH)	FAS	Descriptive statistics
14.2.9.1	Change from Baseline in Average Number of Total Incontinence Episodes per Day	FAS-I	Descriptive statistics, include CFB
14.2.9.2	Urge Urinary Incontinence Episodes per Day 75% Responder Analysis (CMH)	FAS-I	Descriptive statistics
14.2.10.1	Symptom Bother Score as Assessed by [redacted] (1-week recall)	FAS	Descriptive statistics, include CFB
14.2.10.2	Health-Related Quality of Life (HRQL) Subscale Coping Score as Assessed by [redacted] (1-week recall)	FAS	Descriptive statistics, include CFB
14.2.10.3	Health-Related Quality of Life (HRQL) Subscale Concern Score as Assessed by [redacted] (1-week recall)	FAS	Descriptive statistics, include CFB

Title		Population	Programming notes
14.2.10.4	Health-Related Quality of Life (HRQL) Subscale Sleep Score as Assessed by [redacted] (1-week recall)	FAS	Descriptive statistics, include CFB
14.2.10.5	Health-Related Quality of Life (HRQL) Subscale Social Interaction Score as Assessed by [redacted] (1-week recall)	FAS	Descriptive statistics, include CFB
14.2.10.6	Health-Related Quality of Life (HRQL) Total Score as Assessed by [redacted] (1-week recall)	FAS	Descriptive statistics, include CFB
14.2.11.1	[redacted] Domains	FAS	Include 5 domain categorical summaries: mobility, self-care, usual activities, pain/discomfort and anxiety/depression
14.2.11.2	[redacted] Crosswalk Index Value	FAS	Descriptive statistics, include CFB
14.2.11.3	[redacted] VAS Value	FAS	Descriptive statistics, include CFB and percent CFB
14.2.12.1	Overall Symptom Severity Bladder Symptoms Based on Patient Global Impression of [redacted] (PGI[redacted]	FAS	Let's try to include both continuous variables (score and change from baseline) and categorical variable on the same table which will facilitate review

Title		Population	Programming notes
14.2.12.2	Overall Symptom Control over Bladder Symptoms Based on Patient Global Impression of [redacted] (PGI[redacted]	FAS	Repeat 14.2.12.1
14.2.12.3	Overall Symptom Frequency Based on Patient Global Impression of [redacted] (PGI-[redacted]	FAS	Repeat 14.2.12.1
14.2.12.4	Overall Urgency-Related Leakage over Bladder Symptoms Based on Patient Global Impression of [redacted] (PGI-[redacted]	FAS-I	Repeat 14.2.12.1
14.2.12.5	Overall Change of Bladder Symptoms Based on Patient Global Impression of [redacted] (PGI[redacted]	FAS	Categorical summary
14.2.12.1	[redacted] Domain Scores and Overall Satisfaction Score	FAS	Include CFB
Safety Endpoints			
Adverse Events			
14.3.1.1	Overall Summary of Treatment Emergent Adverse Events	SAF	
14.3.1.2	Treatment Emergent Adverse Events by System Organ Class and Preferred Term	SAF	

Title		Population	Programming notes
14.3.1.3	Treatment-Related Treatment Emergent Adverse Events by Preferred Term	SAF	
14.3.1.4	Treatment-Emergent Adverse Events by Preferred Term and Maximum Intensity	SAF	
14.3.1.5	Treatment Emergent Adverse Events with Grade ≥ 3 by Preferred Term	SAF	
14.3.1.6	Treatment-Related Emergent Adverse Events with Grade ≥ 3 by Preferred Term	SAF	
14.3.1.7	Treatment-Emergent Adverse Events occurring in ≥ 2% subjects of Vibegron Arm and Greater Than the Placebo by Preferred Term	SAF	
14.3.1.8	Serious Treatment Emergent Adverse Events by Preferred Term	SAF	
14.3.1.9	Treatment-Related Serious Treatment Emergent Adverse Events by Preferred Term	SAF	
14.3.1.10	Treatment Emergent Adverse Events Leading to Discontinuation from Study Treatment by Preferred Term	SAF	
14.3.1.11	Fatal TEAEs by Preferred Term	SAF	

Title		Population	Programming notes
14.3.1.12	Treatment Emergent Adverse Events of Special Interest by Preferred Term	SAF	
14.3.1.13	Treatment-Related Treatment Emergent Adverse Events of Special Interest by Preferred Term	SAF	
14.3.1.14	Non-Fatal Treatment-Emergent Adverse Events by Preferred Term	SAF	
14.3.1.15	Hypertension Treatment Emergent Adverse Events by Preferred Term, and Pre-Existing Hypertension	SAF	
14.3.1.16	Hypertension Treatment Emergent Adverse Events by Preferred Term, and Baseline Hypertension	SAF	
14.3.1.17	Listing of Deaths	SAF	
14.3.1.18	Listing of Treatment Emergent Serious Adverse Events	SAF	
14.3.1.19	Listing of Treatment Emergent Adverse Events Leading to Withdrawal or Interruption of Study Treatment	SAF	
14.3.1.20	Listing of TEAEs of Special Interest	SAF	
Safety Labs			
14.3.2.1	Hematology Laboratory Parameters	SAF	Include observed and CFB

Title		Population	Programming notes
14.3.2.2	Clinical Chemistry Laboratory Parameters	SAF	Include observed and CFB
14.3.2.3	Urinalysis Laboratory Parameters	SAF	Include observed and CFB
14.3.2.4	Summary of Other Laboratory Parameters	SAF	
14.3.2.5	Abnormal Classification of Hematology Laboratory Parameters	SAF	
14.3.2.6	Abnormal Classification of Clinical Chemistry Laboratory Parameters	SAF	
14.3.2.7	Abnormal Classification of Urinalysis Laboratory Parameters	SAF	
14.3.2.8	Abnormal Classification Summary of Other Laboratory Parameters	SAF	
14.3.2.9	Shift Table of L/N/H Classification for Hematology Laboratory Parameters	SAF	
14.3.2.10	Shift Table of L/N/H Classification for Chemistry Laboratory Parameters	SAF	
14.3.2.11	Maximum Post-baseline ALT and AST vs. Maximum Post-baseline Bilirubin	SAF	
14.3.2.12	Listing of Subjects who Potentially Met Hy's Law	SAF	

Title		Population	Programming notes
Concomitant Medication			
14.3.3.1	Non-OAB Concomitant Medication	SAF	
14.3.3.2	OAB Concomitant Medication	SAF	
Regular and Orthostatic Vital Signs			
14.3.4.1	Vital Sign Parameters	SAF	Include all observed and CFB
14.3.4.2	Vital Sign Parameter Change from Baseline Shifts at 3 Consecutive Visits	SAF	SBP, DBP and PR
14.3.4.3	Vital Sign Parameter Change from Baseline Shifts at Week 24	SAF	SBP, DBP and PR
14.3.4.4	Vital Sign Parameter Maximum Post-Baseline Change from Baseline	SAF	SBP, DBP and PR
14.3.4.5	Orthostatic Blood Pressure and Pulse Rate	SAF-Orth	
14.3.5.1	Time-Matched Change from Sitting to Standing Position in Orthostatic Blood Pressure and Pulse Rate	SAF-Orth	
14.3.5.2	Time-Matched Change from Day 1 in Orthostatic Blood Pressure and Pulse Rate	SAF-Orth	

Title		Population	Programming notes
14.3.5.3	Maximum CFB over 6 Hours in in Orthostatic Blood Pressure and Pulse Rate	SAF-Orth	
14.3.5.4	Average CFB over 6 Hours in Orthostatic Blood Pressure and Pulse Rate	SAF-Orth	
14.3.5.5	Categorical Shift from Sitting to Standing in Blood Pressure	SAF-Orth	
Urodynamics			
14.3.6.1	Urodynamics Assessments	SAF-UDS	
Other			
14.3.7.1	Prostate Volume Measurements	SAF	
14.3.8.1	Uroflow Measurements	SAF	
14.3.9.1	Total IPSS score	SAF	
14.3.10.1	Physical Examination Shift from Baseline	SAF	

11.2.2. Output Figures

Table 10: List of Output Figures

Title		Population	Programming notes
Efficacy Endpoints			
14.2.1.1	Plot of LS Means (SE) of Change from Baseline in Average Number of Daily Micturition Episodes Primary Efficacy Analysis (MMRM)	FAS	Including all visits
14.2.1.2	Forest Plot of Vibegron Treatment Effect vs. Placebo on Average Number of Daily Micturition Episodes at Week 12 by Subgroup Primary Efficacy Analysis (MMRM)	FAS	
14.2.2.1	Plot of LS Means (SE) of Change from Baseline in Average Number of Daily Urgency Episodes Primary Efficacy Analysis (MMRM)	FAS	Including all visits
14.2.2.2	Forest Plot of Vibegron Treatment Effect vs. Placebo on Average Number of Daily Urgency Episodes at Week 12 by Subgroup Primary Efficacy Analysis (MMRM)	FAS	
14.2.3.1	Plot of LS Means (SE) of Change from Baseline in Average Number of Nocturia Episodes per Day (MMRM)	FAS	Including all visits
14.2.4.1	Plot of LS Means (SE) of Change from Baseline in Average Number of Daily Urge Urinary incontinence Episodes (MMRM)	FAS-I	Including all visits

Title		Population	Programming notes
14.2.5.1	Plot of LS Means (SE) of Change from Baseline in International Prostate Symptom Score (IPSS) Storage Score (MMRM)	FAS	Including all visits
14.2.6.1	Plot of LS Means (SE) of Change from Baseline in Volume Voided per Micturition (MMRM)	FAS	Including all visits
Safety Endpoints			
14.3.3.1	Scatter Plot of Maximum ALT Post-Baseline versus Maximum Total Bilirubin Expressed as Multiples of ULN	SAF	
14.3.3.2	Scatter Plot of Maximum AST Post-Baseline versus Maximum Total Bilirubin Expressed as Multiples of ULN	SAF	
14.3.3.3	Line Plot of Liver Chemistry Test Results over Time for Subjects with Elevated ALT or AST, and Elevated Total Bilirubin at Any Time	SAF	Only produce for subjects who potentially met Hy's law
14.3.4.1	Line Plot of Mean (SE) in CFB of Systolic Blood Pressure	SAF	X-axis starts from baseline visit
14.3.4.2	Line Plot of Mean (SE) in CFB of Diastolic Blood Pressure	SAF	X-axis starts from baseline visit
14.3.4.3	Line Plot of Mean (SE) in CFB of Pulse Rate	SAF	X-axis starts from baseline visit
14.3.5.1	Line Plot of Mean (SE) in Time-Matched Change from Day 1 of Orthostatic Systolic Blood Pressure	SAF-Orth	

Title		Population	Programming notes
14.3.5.2	Line Plot of Mean (SE) in Time-Matched Change from Day 1 of Orthostatic Diastolic Blood Pressure	SAF-Orth	
14.3.5.3	Line Plot of Mean (SE) in Time-Matched Change from Day 1 of Orthostatic Pulse Rate	SAF-Orth	
14.3.5.4	Empirical Cumulative Distribution Function of Maximum Change from Baseline over 6 Hours to Week 2 and Week 4 Orthostatic Systolic Blood Pressure	SAF-Orth	Week 2 and Week 4 CDF plots side-by-side on the same page
14.3.5.5	Empirical Cumulative Distribution Function of Maximum Change from Baseline over 6 Hours to Week 2 and Week 4 Orthostatic Diastolic Blood Pressure	SAF-Orth	Week 2 and Week 4 CDF plots side-by-side on the same page
14.3.5.6	Empirical Cumulative Distribution Function of Maximum Change from Baseline over 6 Hours to Week 2 and Week 4 Orthostatic Pulse Rate	SAF-Orth	Week 2 and Week 4 CDF plots side-by-side on the same page

11.2.3. Output Listings

Table 11: List of Output Listings

Title		Population	Programming notes
Disposition and Demographics			
16.2.1.1	Subject Disposition – Screen Failures	Screened	

Title		Population	Programming notes
16.2.1.2	Eligibility Criteria-Inclusion/Exclusion Criteria – Screen Failures	Screened	
16.2.1.3	Run-in Treatment Administration	Run-in	
16.2.1.4	Subject Disposition	Randomized	
16.2.1.5	Subject Randomization Details	Randomized	
16.2.1.6	Protocol Deviation	FAS	
16.2.1.7	Exclusion from Analysis Sets	Randomized	
16.2.1.8	Demographic and Baseline Characteristics	Randomized	
16.2.1.9	Medical History	SAF	
16.2.1.10	Procedures	SAF	
16.2.1.11	Prior Non-OAB Medications	SAF	Separate from prior OAB medication listing.
16.2.1.12	Concomitant Medications	SAF	
16.2.1.13	Prior OAB Medication	SAF	
16.2.1.14	Randomization and Treatment Administration		
16.2.1.15	Study Drug Accountability	SAF	

Title		Population	Programming notes
16.2.1.16	Study Drug Compliance	SAF	
Efficacy			
16.2.2.1	Derived Average Bladder Diary Parameters	FAS	
16.2.2.2	Derived Responder Parameters from Bladder Dairy	FAS	
16.2.2.3	Derived Parameters from International Prostate Symptom Score (IPSS) Data	FAS	
16.2.2.4	[redacted] (1-week recall) Data	FAS	
16.2.2.5	[redacted] Data	FAS	
16.2.2.6	Global Impression Data	FAS	
16.2.2.7	[redacted]	FAS	
Safety			
16.2.3.1	All Adverse Events	SAF	
16.2.3.2	All Adverse Events During Screening Period	Screened	
16.2.3.3	All Adverse Events During Run-in Period	Run-in	
16.2.3.4	All Adverse Events During Double-Blind Period	SAF	
16.2.3.5	Hematology Laboratory Parameters	SAF	
16.2.3.6	Clinical Chemistry Laboratory Parameters	SAF	
16.2.3.7	Urinalysis Laboratory Parameters	SAF	
16.2.3.8	Other Laboratory Parameters	SAF	
16.2.3.9	Vital Signs	SAF	

Title		Population	Programming notes
16.2.3.10	Orthostatic Blood Pressure and Pulse Rate	SAF-Orth	
16.2.3.11	Electrocardiogram (ECG) parameters	SAF	
16.2.3.12	Physical Examination	SAF	
16.2.3.13	Post-Void Residual (PVR) Volume	SAF	
16.2.3.14	Urodynamics Assessment	SAF-UDS	
16.2.3.15	Prostate Volume Measurement	SAF	