Study Title:	HERO: A Multinational Phase 3 Randomized, Open-Label, Parallel Group Study to Evaluate the Safety and Efficacy of Relugolix in Men with Advanced Prostate Cancer
NCT Number:	03085095
Document Date:	Statistical Analysis Plan V3.0 / 07 November 2019

STATISTICAL ANALYSIS PLAN

Study Title:	HERO: A Multinational Phase 3 Randomized, Open-Label, Parallel Group Study to Evaluate the Safety and Efficacy of Relugolix in Men with Advanced Prostate Cancer
Investigational Product:	Relugolix
Protocol Number:	MVT-601-3201
Indication:	Advanced Prostate Cancer
Sponsor:	Myovant Sciences GmbH Viaduktstrasse 8 4051 Basel Switzerland
IND No.:	118736
Version/Effective Date:	V1.0 / 12 September 2019 V2.0 / 29 October 2019 V3.0 / 07 November 2019

STATISTICAL ANALYSIS PLAN APPROVAL SHEET

MVT-601-3201 (HERO): A Multinational Phase 3 Randomized, Open-label, Parallel Group Study to Evaluate the Safety and Efficacy of Relugolix in Men with Advanced Prostate Cancer

This statistical analysis plan has been approved by Myovant Sciences, Inc., agent for Myovant Sciences GmbH. The following signatures document this approval.

[redacted]

07 Nov 2019
Date

07 Nov 2019
Date

07 Nov 2019
Date

07 Nov 2019
Date

07 Nov 2019
Date

07-Nov-2019
Date

07 Nov 2019
Date

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVIATIONS

Term	Definition
3-M	3-month
ALT	alanine aminotransferase
AST	aspartate aminotransferase
ATC	Anatomical Therapeutic Chemical
$AUC_{0-\tau}$	area under the curve from time 0 to the end of the dosing interval
CI	confidence interval
C_{max}	maximum plasma concentration
CRF	case report form
CRFS	castration resistance free survival
CTCAE	Common Terminology Criteria for Adverse Events
DSMB	data and safety monitoring board
ECOG	Eastern Cooperative Oncology Group
eDISH	evaluation of drug induced serious hepatotoxicity
EMA	European Medicines Agency
EORTC	European Organisation of Research and Treatment of Cancer
EOT	end-of-treatment
EuroQol EQ-5D-5L	European Quality of Life 5-Dimension 5-Level Questionnaire
FSH	follicle-stimulating hormone
GnRH	gonadotropin-releasing hormone
ICF	informed consent form
IWRS	interactive web response systems
LH	luteinizing hormone
MedDRA	Medical Dictionary for Regulatory Activities
mITT	modified intent-to-treat (Population)
NCI CTCAE	National Cancer Institute's Common Terminology Criteria for Adverse Events
PD	pharmacodynamic
PMDA	Japanese Pharmaceutical and Medical Device Agency
PK	pharmacokinetics
PSA	prostate-specific antigen
Q12W	once every 12 weeks
QTcF	corrected QT interval Fridericia
SAP	statistical analysis plan
SD	standard deviation
t_{max}	time to maximum plasma concentration
ULN	upper limit of normal
VAS	visual analogue scale

Term	Definition
WHO	World Health Organization

1. INTRODUCTION

The purpose of this statistical analysis plan (SAP) is to describe the analyses planned for phase 3 study MVT-601-3201 (HERO), entitled "A Multinational Phase 3 Randomized, Open-Label, Parallel Group Study to Evaluate the Safety and Efficacy of Relugolix in Men with Advanced Prostate Cancer." This SAP is based on protocol amendment 3, dated 23 Oct 2018.

The study has two analyses, the primary and the final analyses. The primary analysis of the study will be performed when all patients randomized under protocol amendment 2 (targeting approximately 915 patients), dated 18 Jan 2018, have either completed 48 weeks of study treatment with the 30-day safety follow-up visit or discontinued early. The primary endpoint, secondary efficacy endpoints, and safety of relugolix in men with advanced prostate cancer will be evaluated.

The final analysis of the study will occur after approximately 390 patients with metastatic disease have been randomized to the study and have either completed 48 weeks of study treatment with the 30-day safety follow-up visit or discontinued early. As the enrollment in China is concurred with the enrollment of patients with metastatic disease under protocol amendment 3, Chinese patients with metastatic disease randomized prior to the completion of the enrollment of patients with metastatic disease will be included in the final analysis.

This SAP describes the statistical analysis plan for the primary and final analyses of the study. However, the SAP will be finalized prior to the database lock of the study for the primary analysis.

Any subsequent analyses conducted at later data cut-off dates (ie, analysis inclusive of all Chinese patients enrolled) will be described in separate SAPs.

1.1. Study Objectives and Endpoints

The study objectives and corresponding endpoints are listed in Table 1.

The endpoints in *italics* either are not listed in the protocol or have been modified. These endpoints are included in this SAP because they have been identified as important for assessment of treatment effect, on the basis of emerging literature review and clinical relevance to the study objectives.

For the primary efficacy endpoint, there are two evaluation criteria in the protocol to support different regulatory requirements for assessing benefit. The first criterion evaluates castration rate in approximately 610 patients randomized to relugolix. The second criterion evaluates the noninferiority of relugolix in approximately 610 randomized patients compared with leuprolide acetate in approximately 305 randomized patients.

Table 1: Study Objectives and Endpoints

Objective(s)	Endpoint(s)
Primary Efficacy	
To evaluate the ability of relugolix to achieve and maintain serum testosterone suppression to castrate levels of < 50 ng/dL (1.7 nmol/L) in men with androgen-sensitive advanced prostate cancer	The primary endpoint is the sustained castration rate, defined as the cumulative probability of testosterone suppression to < 50 ng/dL (1.7 nmol/L) while on study treatment from Week 5 Day 1 (study Day 29) through Week 49 Day 1 (study Day 337). • Evaluation Criterion 1: to determine whether the sustained castration rate (defined as the cumulative probability of testosterone suppression to < 50 ng/dL [1.7 nmol/L] while on study treatment from Week 5 Day 1 through Week 49 *Day 1*) for relugolix is ≥ 90%. The lower bound of the 95% confidence interval for the cumulative probability of sustained testosterone suppression in the relugolix treatment group will be calculated and must be at least 90% for this criterion to be met. • Evaluation Criterion 2: to establish the noninferiority of relugolix compared to leuprolide acetate every 3-month (3-M) depot injection as assessed by the cumulative probability of sustained testosterone suppression. The lower bound of the 95% confidence interval for the difference in the cumulative probability of sustained testosterone suppression between the two treatment groups will be calculated and must be greater than or equal to the noninferiority margin of -10% for this criterion to be met.

Key Secondary Efficacy (Alpha-Protected for Hierarchical Hypothesis Testing)	
To determine the time course and change in serum testosterone	Cumulative probability of testosterone suppression to < 50 ng/dL (1.7 nmol/L) prior to dosing on Week 1 Day 4
	Cumulative probability of testosterone suppression to < 50 ng/dL (1.7 nmol/L) prior to dosing on Week 3 Day 1
To evaluate the time course and magnitude of PSA reduction	Proportion of patients with PSA response (by Prostate Cancer Clinical Trials Working Group 3 [Scher et al, 2016] at Week 3 Day 1 followed with the confirmation at Week 5 Day 1
To determine the time course and change in serum testosterone	*Profound castration rate defined as cumulative probability of testosterone suppression to < 20 ng/dL (0.7 nmol/L) prior to dosing on Week 3 Day 1.*
To evaluate the effect of relugolix and leuprolide acetate on endocrine pharmacodynamic parameters.	FSH level at Week 25 Day 1
To describe the time course and magnitude of development of castration resistant prostate cancer [a]	Castration resistance free survival during the 48-week treatment in patients with metastatic prostate cancer
	Castration resistance free survival during the 48-week treatment in patients with or without metastatic prostate cancer
To evaluate testosterone recovery following discontinuation of study treatment	Cumulative probability of testosterone recovery to 280 ng/dL at the 90-day follow-up in approximately 150 patients who complete 48 weeks of treatment and who do not plan to start alternative androgen deprivation therapy within the following 12 weeks (or within 24 weeks following the last injection of leuprolide acetate 3-M depot)

Other Secondary Efficacy (Not for Hierarchical Hypothesis Testing)	
To determine the time course and change in serum testosterone	Sustained profound castration rate from Week 5 Day 1 through Week 49 Day 1 defined as the cumulative probability of testosterone suppression to < 20 ng/dL (0.7 nmol/L) while on treatment from Week 5 Day 1 through Week 49 Day 1; Sustained profound castration rate from Week 25 Day 1 through Week 49 Day 1 defined as the cumulative probability of testosterone suppression to < 20 ng/dL (0.7 nmol/L) while on treatment from Week 25 Day 1 through Week 49 Day 1.
To evaluate the time course and magnitude of PSA reduction during treatment	Proportion of patients with confirmed PSA response at Week 5 Day 1; *Proportion of patients with PSA concentration < 0.02 ng/mL (0.02 µg/L) at Week 25 visit.*
To describe the time course and magnitude of PSA progression and development of castration resistant prostate cancer during treatment	Time to PSA progression per Prostate Cancer Clinical Trials Working Group 3
To evaluate testosterone recovery following discontinuation of study treatment	Cumulative probability of testosterone recovery back to 50 ng/dL or back to baseline or 280 ng/dL at the 90-day follow-up in approximately 150 patients who complete 48 weeks of treatment and who do not plan to start alternative androgen deprivation therapy within the following 12 weeks (or within 24 weeks following the last injection of leuprolide acetate 3-M depot)
To evaluate the impact of treatment on quality of life using validated patient-report outcome instruments.	• Absolute values and changes from baseline in the scores of the EORTC-QLQ-C30 global health domain, and the EORTC-QLQ-PR25 sexual activity and hormonal-treatment-related symptom subdomains, at regular intervals during treatment, and as applicable during the Follow-up and/or End of Treatment visits; • Absolute values and changes from baseline of the remaining domains in the EORTC QLQC30 and EORTC QLQ-PR25, as well as the EuroQol EQ-5D-5L questionnaire, at regular intervals during treatment, and as applicable during the Follow-up visits.

Safety	
To evaluate the safety of relugolix 120 mg once daily in men with androgen-sensitive advanced prostate cancer.	Treatment-emergent adverse events, clinical laboratory tests, and vital sign measurements.
Pharmacodynamic	
To evaluate the effect of relugolix and leuprolide acetate on endocrine pharmacodynamic parameters.	Endocrine marker effects of relugolix and leuprolide acetate as measured as absolute values and change from baseline for: • LH at the Day 4, Week 5, Week 25, and Week 49 visits; • FSH at the Day 4, Week 5, Week 25, and Week 49 visits; • Dihydrotestosterone at the Week 5, Week 25, and Week 49 visits; • Sex hormone-binding globulin at Week 5, Week 25, and Week 49 visits.
Pharmacokinetics	
To collect relugolix plasma concentration data to further evaluate relugolix population pharmacokinetics and the relationship between relugolix exposure and serum testosterone.	Predose relugolix plasma concentrations
To characterize the relugolix plasma pharmacokinetic parameters in a subset of patients from China and Japan.	Single and repeat-dose plasma relugolix pharmacokinetic parameters such as C_{max}, $AUC_{0-\tau}$, and t_{max} in subsets of patients from China or Japan.
Exploratory	
To explore the overall survival	Overall survival defined as time from randomization to date of death prior to data cut-off date.
To explore the contribution of genetic variance on drug response.	The presence of polymorphisms in germline genes related to the hypothalamic-pituitary androgen pathway, prostate cancer risk, or to drug metabolizing enzymes and transporter proteins that might be implicated in the drug disposition, safety, or efficacy of relugolix.

Abbreviations: 3-M = every 3 months; $AUC_{0-\tau}$ = area under the curve from time 0 to the end of the dosing interval; C_{max} = maximum plasma concentration; EuroQol EQ-5D-5L = European Quality of Life 5-Dimension 5-Level Questionnaire; FSH = follicle-stimulating hormone; LH = luteinizing hormone; PSA = prostate-specific antigen; t_{max} = time to maximum plasma concentration.

[a] Castration resistance free survival will not be assessed in the primary analysis of the study. It will be evaluated at the time of approximately 390 patients with metastatic disease enrolled and completed the 48 weeks of study treatment.

2. STUDY DESIGN

2.1. Summary

The MVT-601-3201 (HERO) study is a phase 3, multinational, randomized, open-label, parallel-group study to evaluate the safety and efficacy of relugolix in patients with androgen-sensitive advanced prostate cancer who require at least 1 year of continuous androgen deprivation therapy. Oral relugolix or leuprolide acetate depot subcutaneous or intramuscular injection will be administered to patients with prostate cancer who require androgen deprivation therapy. The primary endpoint is sustained castration rate, defined as the cumulative probability of testosterone suppression to < 50 ng/dL (1.7 nmol/L) while on study treatment from Week 5 Day 1 (Study Day 29) through Week 49 Day 1 (Study Day 337).

Approximately 915 patients will be enrolled and randomized in a 2:1 ratio to the following two treatment groups:

- Oral relugolix 120 mg once daily, following a loading dose of 360 mg on Day 1;
- Leuprolide acetate 22.5 mg (or 11.25 mg in some Asian countries) every 3 months (3-M) depot subcutaneous or intramuscular injection.

Randomization will be stratified as follows:

- Geographic region: North and South America; Europe; Asia and Rest of World;
- Presence of metastatic disease: yes, no;
- Baseline age: ≤ 75 years old, > 75 years old.

After the initial 915 patients have been enrolled under protocol amendment 2, only patients with metastatic advanced prostate cancer will be enrolled under protocol amendment 3 to assess castration resistance free survival (ie, time to castration resistance) as a key secondary endpoint. A total of approximately 390 patients with metastatic disease as the high-risk subgroup for developing castration resistance, randomized in a 2:1 ratio to relugolix versus leuprolide acetate, are required to evaluate castration resistance free survival. In parallel, for approval in China 138 patients in China and Taiwan will be enrolled, with or without metastatic advanced prostate cancer, under protocol amendment 3. To support applications across various countries and regions, a total of approximately 1100 patients is expected to be enrolled in the study.

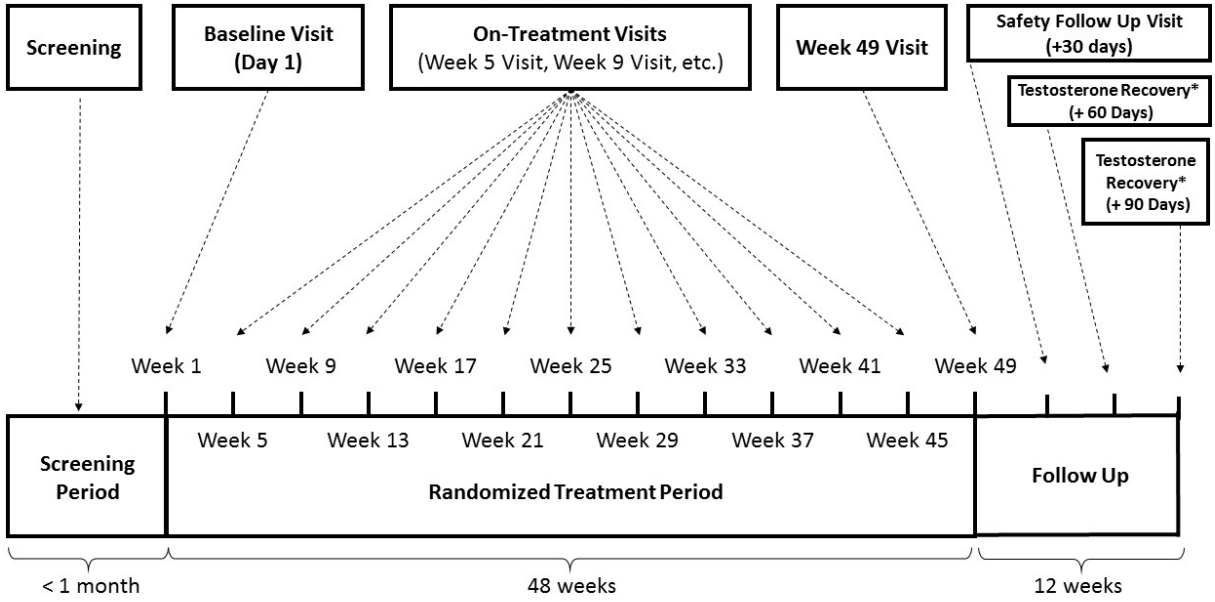
This study includes a screening period of up to 28 days, a treatment period of 48 weeks, and a follow-up period of up to 30 days. Approximately 100 patients randomized to relugolix and approximately 50 patients randomized to leuprolide acetate who completed 48 weeks of treatment and who do not plan to start alternative androgen deprivation therapy within the following 12 weeks will be followed for testosterone recovery at the 60- and 90-day follow-up visits.

Testosterone assayed using a liquid chromatography-tandem mass spectrometry method sensitive at least as low as 5 ng/dL (0.17 nmol/L) and prostate-specific antigen (PSA) will be assessed at Day 4, Week 3 Day 1 and Week 5 Day 1, then monthly afterwards. Additional serum endocrine evaluations and quality of life questionnaires will be collected throughout the study.

An external independent Data and Safety Monitoring Board (DSMB) was established to review periodic safety analyses. The roles and responsibilities of the independent DSMB are described in the DSMB charter. A separate SAP was created to specify the safety data analyses performed by an independent data coordinating center and reviewed by the DSMB.

A schematic of the study is presented in Figure 1.

Figure 1: Schematic of Study Design



Relugolix dosed daily (Baseline Day 1 – Week 48 Day 7)
Leuprolide Acetate dosed every 12 weeks (Baseline Day 1, Week 13 Day 1, Week 25 Day 1, and Week 37 Day 1)
*+60 Days and +90 Days Testosterone Recovery visits in subset of patients

2.2. Sample Size Justification

The following assumptions were used to determine the sample size for this study:

- True probabilities of sustained testosterone suppression are 94% and 96% for relugolix and leuprolide acetate, respectively;
- 2:1 randomization ratio (relugolix: leuprolide acetate);
- Dropout rate: 15%.

For Evaluation Criterion 1, 610 patients in the relugolix group will provide approximately 90% power to rule out a fixed probability of sustained testosterone suppression of < 90% at a two-sided type I error rate of 0.05.

For Evaluation Criterion 2, with a noninferiority margin of -10% and an overall two-sided type I error rate of 0.05, a total of approximately 915 patients (610 receiving relugolix, 305 receiving leuprolide acetate) will yield at least 99% power to declare the noninferiority of relugolix to leuprolide acetate. The 10% noninferiority margin for the comparison of relugolix versus

leuprorelin is based on regulatory precedence of pivotal assessment of gonadotropin-releasing hormone (GnRH) antagonist degarelix versus leuprorelin as well as studies of branded GnRH agonist generics.

The primary analysis will be performed separately for each evaluation criterion using data collected through 48 weeks after enrollment of approximately 915 patients.

Approximately 915 patients will be randomized in order to fulfill the regulatory requirements of all participating countries included in the primary efficacy endpoint of this study.

Under amendment 3 the study is also powered for the secondary endpoint of castration resistance free survival in the high-risk subgroup of patients with metastatic disease, although this is not part of the primary analysis. Approximately 107 confirmed castration-resistance events (ie, PSA progression while castrated or deaths due to any reasons) will need to be observed (or approximately 390 patients with metastatic disease will need to be enrolled) to detect a hazard ratio of 0.55 (relugolix vs. leuprolide acetate) with 85% power and a two-sided type I error of 5%, assuming a castration-resistant event–free rate of 60% at 48 weeks for the control arm, an 18-month enrollment period, 12 months of additional follow-up, and a 15% dropout rate.

With a total of approximately 1100 patients with or without metastatic disease randomized into the study through amendment 3, including approximately 138 Chinese patients, it is anticipated that approximately 149 confirmed castration-resistance events (PSA progression while castrated or deaths due to any cause) will be observed. Assuming an 18-month enrollment period, 12 months of additional follow-up, and a 10% dropout rate, the study will provide approximately 85% power to detect a hazard ratio of 0.6 (relugolix vs. leuprolide acetate) with a two-sided type I error of 5%.

3. PLANNED ANALYSES

3.1. Interim Analysis

There will be no interim efficacy analysis performed for this study.

Periodic safety data is reviewed by the DSMB. An independent data coordinating center will perform the periodic safety data analyses and provide the results of these analyses to the DSMB, as defined in the DSMB charter.

3.2. Primary Analysis

The primary analysis of the study will occur after approximately 915 patients have been randomized into the study, have had the opportunity to be evaluated for 48 weeks, and either have completed the 30-day safety follow-up visit or have discontinued early.

The statistical methods for the primary analysis of the study will be specified in this SAP.

There will be two dates of data analysis cut-off for the primary analysis. The first data analysis cut-off date will be the date when the last Week 49 Day 1 visit has occurred. The second analysis cut-off date will be the date of the last safety follow-up visit. Analyses from both data cut-off dates will be used as the basis for the clinical study report.

3.3. Final Analysis

The final analysis of the study will occur after approximately 390 patients with metastatic disease have been randomized into the study, have had the opportunity to be evaluated for 48 weeks of study treatment, and either have completed the 30-day safety follow-up visit or have discontinued early. The statistical methods for the final analysis of the study will be specified in this SAP.

3.4. Other Analyses

Other analyses, inclusive of all Chinese patients enrolled, are expected to be completed later and will be described separately.

4. GENERAL CONSIDERATIONS FOR DATA ANALYSES AND HANDLING OF MISSING DATA

4.1. Data Presentation Conventions

Leuprolide acetate injection with 22.5 mg or 11.25 mg (in Japan, Taiwan, and China) every 12 weeks (Q12W) will be presented in total in one column.

All statistical analyses will be conducted using SAS® Version 9.2 or higher.

Unless otherwise stated, a statistical test will be assessed at two-sided $\alpha = 0.05$ significance level and all confidence intervals (CIs) will be reported as two-sided.

Where appropriate, variables will be summarized descriptively by study visit. For the categorical variables, the count and proportions of each possible value will be tabulated by treatment group. For continuous variables, the number of subjects with non-missing values and the mean, median, standard deviation (SD), minimum, and maximum values will be tabulated.

Unless otherwise specified, the following conventions will be applied to all analyses:

- Mean and median values will be formatted to one more decimal place than the measured value. Standard deviation values will be formatted to two more decimal places than the measured value, and minimum and maximum values will be presented to the same number of decimal places as the measured value. If the measured value is large (eg, > 100), fewer decimal places may be displayed.
- Percentages will be rounded to one decimal place;
- p-values will be rounded to 4 decimal places. p-values < 0.0001 will be presented as "< 0.0001" and p-values > 0.9999 will be presented as "> 0.9999";
- 1 month = 30.4375 days. Month is calculated as (days/30.4375) rounded to 1 decimal place;
- 1 year = 365.25 days. Year is calculated as (days/365.25) rounded to 1 decimal place;
- Age = integer part of month (date of randomization – date of birth – X)/12, where X = 1 if day of randomization date < day of birth date and X = 0 otherwise. If only birth year is provided, 1 July of the birth year will be used to calculate age;
- 1 pound = 0.454 kg;
- 1 inch = 2.54 cm;
- Missing efficacy or safety data will not be imputed, unless otherwise specified;
- For laboratory results above or below sensitivity limits displayed as < or > a quantification threshold, 0.0000000001 will be subtracted or added, respectively, to the sensitivity limit to derive a numeric result for analyses;
- For safety analyses, calculation of percentages will be calculated on the basis of the number of patients in the analysis population in each treatment group;

- For by-visit observed data analyses, calculation of percentages will be calculated on the basis of the number of patients with non-missing data as the denominator, unless otherwise specified;
- For continuous endpoints, the summary statistics will include mean, standard deviation, median, and range (minimum and maximum);
- For time-to-event endpoints, the summary statistics will include median time to event-free survival, 25th and 75th percentiles and number of patients at risk at specified time points;
- For categorical endpoints, the summary statistics will include counts and percentages;
- Confidence intervals, when presented, will generally be constructed at the 95% level. For binomial variables, exact methods will be employed unless otherwise specified.

4.2. Analysis Populations

Three analysis populations are defined as below. Number and percent of patients meeting the definition of each analysis population will be summarized by treatment group.

4.2.1. Modified Intent-to-Treat Population

The modified intent-to-treat (mITT) population is defined as all randomized patients who have taken at least one dose of any study treatment. Unless otherwise specified, all analyses use the mITT population by treatment group as assigned by randomization (not by actual treatment received). The mITT population will be the primary population used for efficacy endpoints.

4.2.2. Per-Protocol Population

The per-protocol population will consist of those members of the mITT population who do not have important protocol deviations (see Section 5.3). The per-protocol population will not be analyzed if this population comprises > 95% or < 50% of the mITT population. This population will be used for sensitivity analysis of the mITT population for the primary efficacy endpoint. The per-protocol population and the associated subset of important protocol deviations will be identified prior to database lock.

4.2.3. Safety Population

The safety population is defined as all randomized patients who receive at least one dose of any study treatment. Unless otherwise specified, all safety analyses will use the safety population according to the actual treatment received (not the treatment assigned). Actual treatment received will be determined by the treatment received on Day 1.

4.3. Definitions, Computations, and Conventions

4.3.1. Date of First Dose and Date of Last Dose of Study Drug

The date of the first dose of study drug is defined as the date a patient receives the first dose of study drug. The date of the last dose of study drug is defined as the date a patient receives the last dose of study drug.

4.3.2. Treatment Day

Treatment day will be calculated in reference to the date of the first dose of study drug. Treatment Day 1 corresponds to the date a patient receives the first dose of study drug. For assessments conducted on or after the date of the first dose of study drug, treatment day will be calculated as (assessment date – date of first dose of study drug + 1). There will be no Treatment Day 0. Efficacy and safety analyses will be based on treatment day, unless otherwise specified.

4.3.3. Treatment Period

The treatment period is defined as the period of time from the date of the first dose of study drug through the date of last dose. The treatment period per protocol is 48 weeks (Treatment Day 336). For patients in the leuprolide acetate treatment arm, treatment period ends 12 weeks after the last leuprolide acetate injection.

4.3.4. Treatment Duration

Treatment duration is defined as the duration of time from the date of the first dose of study drug to the date of the last dose of study drug, calculated as follows:

For relugolix treatment arm, treatment duration is:

(Date of last dose of study drug – Date of first dose of study drug) + 1.

For leuprolide acetate treatment arm, treatment duration is:

(Date of last dose of study drug – Date of first dose of study drug) + 84.

4.3.5. Baseline Value and Postbaseline Value

Both date and time of study drug administration and measurement will be considered when calculating baseline value. Unless otherwise specified, the baseline value is defined as the last measurement on or before the first administration date (if time is not available) of study drug. Or the baseline value is defined as the last measurement before the first administration date and time of study drug. Postbaseline value is defined as any measurement taken after the first administration of study drug. Change from baseline is defined as (postbaseline value – baseline value). If time is not available, then date only will be used.

4.3.6. Visit Windows

Visit windows will be used to associate assessments with a scheduled visit and will be used only for summarizing data by visit. Visit windows for each assessment are defined as shown in Table 2. A separate set of visit windows for assessment of quality of life are defined in

Table 3. For both efficacy and safety assessments, the Treatment Day will be used to determine the associated visit window.

If more than one observation lies within the same visit window, the observation with the closest study day to the target study day will be used. If there are two observations equidistant to the scheduled target day, the earlier assessment will be used.

Table 2: Visit Windows

Visit	Start Day	Target Day	End Day
Week 1 Day 4[a]	1	4	8
Week 3 Day 1	9	15	21
Week 5 Day 1	22	29	43
Week 9 Day 1	44	57	71
Week 13 Day 1	72	85	99
Week 17 Day 1	100	113	127
Week 21 Day 1	128	141	155
Week 25 Day 1	156	169	183
Week 29 Day 1	184	197	211
Week 33 Day 1	212	225	239
Week 37 Day 1	240	253	267
Week 41 Day 1	268	281	295
Week 45 Day 1	296	309	323
Week 49 Day 1	324	337	365
Safety Follow-up [b]	Date of last dose + 2 days in relugolix arm Date of last dose + 8 days in leuprolide arm	Date of last dose + 30 days	Date of last dose + 45 (59 [d]) days
60-Day Testosterone Recovery [c]	Date of last dose + 46 days	Date of last dose + 60 days	Date of last dose + 75 days
90-Day Testosterone Recovery [c]	Date of last dose + 76 days	Date of last dose + 90 days	Date of last dose + 119 days

Date of last dose in leuprolide acetate ends 12 weeks after the last leuprolide acetate injection.
[a] Start day of Week 1 Day 4 only includes postbaseline assessments on or after the first dose.
[b] The safety follow-up visit window will be restricted to assessments prior to the date of initiation of another investigational agent or hormonal therapy or surgical intervention, whichever occurs first.
[c] Approximately 100 patients receiving relugolix and 50 patients receiving leuprolide acetate who complete the 48 weeks of study treatment and, in the opinion of the investigator, can remain off androgen deprivation therapy for 90 days will be included in the testosterone recovery 60- and 90-day follow-up visits.
[d] For patients not enrolled in the testosterone recovery follow-up, the end day of safety follow-up visit can be extended to Day 59 after the date of last dose.

Table 3: Visit Windows for Assessments in Quality of Life

Visit	Start Day	Target Day	End Day
Week 5 Day 1	9	29	57
Week 13 Day 1	58	85	127
Week 25 Day 1	128	169	211
Week 37 Day 1	212	253	295
Week 49 Day 1	296	337	365
Safety Follow-up [a]	Date of last dose + 2 days in relugolix arm Date of last dose + 8 days in leuprolide arm	Date of last dose + 30 days	Date of last dose + 45 (59 [c]) days
60-Day Testosterone Recovery [b]	Date of last dose + 46 days	Date of last dose + 60 days	Date of last dose + 75 days
90-Day Testosterone Recovery [b]	Date of last dose + 76 days	Date of last dose + 90 days	Date of last dose + 119 days

Date of last dose in leuprolide acetate ends 12 weeks after the last leuprolide acetate injection.
[a] The safety follow-up visit window will be restricted to assessments prior to the date of initiation of another investigational agent or hormonal therapy or surgical intervention, whichever occurs first.
[b] Approximately 100 patients receiving relugolix and 50 patients receiving leuprolide acetate who complete the 48 weeks of study treatment and, in the opinion of the investigator, can remain off androgen deprivation therapy for 90 days will be included in the testosterone recovery 60- and 90-day follow-up visits.
[c] For patients not enrolled in the testosterone recovery follow-up, the end day of safety follow-up visit can be extended to day 59 after the date of last dose.

4.4. General Rules for Missing Data

Handling of missing data for the efficacy endpoints is described in Section 7.

4.4.1. By -Visit Endpoints

By-visit endpoints will be analyzed using observed data, unless otherwise specified. For observed data analyses, missing data will not be imputed and only the observed records will be included.

4.4.2. Adverse Events

The imputation rule for the safety analyses will be used to address the issues with partial dates. The imputed dates will be used to determine the treatment-emergent period. For adverse events with a partial date, available date parts (year, month, and day) of the partial date will be compared with the corresponding date components of the start date and end date of the treatment-emergent period to determine if the event is treatment emergent. When in doubt, the adverse event will be considered treatment emergent by default. The following rules will be applied to impute partial dates for adverse events:

If start date of an adverse event is partially missing, impute as follows:

- If both Month and Day are missing and Year = Year of treatment start date, then set to treatment start date, as long as adverse event end date is not prior to treatment start date;
- If both Month and Day are missing and Year ≠ Year of treatment start date, then set to January 1;
- If Day is missing and Month and Year = Month and Year of treatment start date, then set to treatment start date, as long as adverse event end date is not prior to treatment start date;
- If Day is missing and Month and Year ≠ Month and Year of treatment start date, then set to first of the month;
- If start date is completely missing, set to treatment start date, as long as adverse event end date is not prior to treatment start date.

If end date of an adverse event is partially missing, impute as follows:

- If both Month and Day are missing, then set to December 31;
- If only Day is missing, then set to last day of the month;
- If end date is completely missing, do not impute.

4.4.3. Concomitant Medications

When the start date or end date of a medication is partially missing, the date will be imputed to determine whether the medication is prior or concomitant (or both). The following rules will be applied to impute partial dates for medications:

If start date of a medication is partially missing, impute as follows:

- If both Month and Day are missing, then set to January 1;
- If only Day is missing, then set to the first of the month;

If end date of a medication is partially missing, impute as follows:

- If both Month and Day are missing, then set to December 31;
- If only Day is missing, then set to last day of the month;

If start date or end date of a medication is completely missing, do not impute.

5. PATIENT POPULATION

5.1. Patient Disposition

The number of patients for each of the following categories will be summarized by treatment group:

- All randomized patients;
- Number and percentage of patients in the safety population;
- Number and percentage of patients who did not receive study treatment;
- Number and percentage of patients who completed 48 weeks of study treatment;
- Number and percentage of patients who prematurely discontinued study treatment and reasons for treatment discontinuation;
- Number and percentage of patients who participated in testosterone recovery follow-up.

Patient disposition will be summarized for all randomized patients.

5.2. Screen Failure

Reasons for screen failure will be summarized. Number and percentage of patients who did not pass screening will be based on the patients who signed the informed consent form but were not randomized.

5.3. Protocol Deviations

Protocol deviations will be categorized as important protocol deviation or protocol deviation according to the protocol deviation plan. Important protocol deviations will include but not be limited to the following:

- Patients who did not satisfy key entry criteria;
- Patients who met withdrawal criteria during the study but were not withdrawn;
- Patients who received the wrong treatment;
- Patients who received a prohibited concomitant medication that met criteria for an important protocol deviation.

Important protocol deviations will be summarized by deviation category for all patients in the mITT population. A by-patient listing of important protocol deviations will be provided.

Patient eligibility including inclusion criteria that are not met and exclusion criteria that are met at enrollment will be summarized for all patients in the mITT population.

A selected subset of the important protocol deviations that are likely to affect analysis of efficacy will be identified to define the per-protocol population prior to the database lock. This subset will include but will not be limited to the following import protocol deviations:

- Did not satisfy key entry criteria;

- Drug compliance < 75% over more than 1 consecutive month;
- Patient received prohibited concomitant medications that met criteria for important protocol deviations.

5.4. Demographics and Baseline Characteristics

Demographics and baseline characteristics for patients in the mITT population will be summarized by treatment group.

Demographics will include age, race, ethnicity and geographic region while baseline disease characteristics will include clinical disease stage, baseline PSA, presence of metastatic disease, and Eastern Cooperative Oncology Group (ECOG) status.

5.5. Medical History

Medical history will be coded using the Medical Dictionary for Regulatory Activities (MedDRA) and will be summarized by system organ class and preferred term.

In addition, summaries will be provided of medical and surgical treatment history specific to prostate cancer.

5.6. Prior and Concomitant Medications

Prior and concomitant medications taken during the study treatment period will be summarized for all patients in the mITT population by treatment group. Medications are considered concomitant if medications were used during the study treatment period. Medications will be coded using the World Health Organization (WHO) Drug Dictionary and summarized according to the Anatomical Therapeutic Chemical (ATC) Classification System and generic medication name.

Use of antiandrogens initiated during the first 4 weeks of study treatment (including bicalutamide, flutamide, nilutamide, and enzalutamide) will be summarized separately.

5.7. Prior and Current Procedures

Prior and current procedures, including radiotherapies, taken during the study treatment period will be summarized by treatment group for all patients in the mITT population. Procedures are considered current if they occur at or after the date of the first dose of study treatment.

6. STUDY DRUG EXPOSURE AND COMPLIANCE

Patients in the safety population will be summarized for extent of exposure and compliance of study drug by actual treatment received. Study drug exposure will be calculated based on the case report form (CRF) pages of "Relugolix Dosing Log" and "Leuprolide Acetate Administration." Compliance for relugolix will be based on the CRF of "Study Drug Accountability."

Summary statistics on extent of exposure and compliance will be provided as follows:

- Treatment duration in weeks;
- Cumulative dose administered (mg) for leuprolide and total tablets taken for relugolix;
- Dose intensity and relative dose intensity for leuprolide;
- Compliance for relugolix.

For leuprolide acetate arm:

Treatment duration in weeks = (the date of last dose – the date of first dose + 84) / 7.

Cumulative dose administered will be sum of all the doses administered.

Dose intensity = cumulative dose administered / ceiling function of (treatment duration in weeks/12).

Relative dose intensity = (dose intensity / 22.5 or 11.25 mg per treatment on Day 1) × 100.

For relugolix arm:

Treatment duration in weeks = (the date of last dose – the date of first dose + 1) / 7.

Relugolix compliance will be calculated as follows:

(total tablets taken / total tablets expected to be taken) × 100.

The total tablets taken will be calculated as:

total number of tablets dispensed - total number of tablets returned.

The total tablets expected to be taken will be calculated as:

3 tablets on Day 1 + 1 tablet per day × (treatment duration in Days -1).

In the case of unreturned bottles, tablets that were dispensed will be assumed to be taken. However, for patients who did not return for their last scheduled visit, tablets that were dispensed and not returned will be subtracted from both the total tablets taken and the total tablets expected to be taken in the calculation of relugolix compliance. In cases when the patient never returned any of the dispensed bottles, the total tablet taken, total tablets expected to be taken, and relugolix compliance will be set to "not able to calculate."

Summary statistics of relative dose intensity and study drug compliance (eg, mean, median, etc) will be presented along with a categorical summary (eg, < 80%, 80 to 100%, > 100%).

For this study, any dose of relugolix > 240 mg within a 24-hour window is an overdose, except for the 360-mg loading dose on Treatment Day 1. Events of overdose will be reported within 24 hours using a serious adverse event form, whether the overdose is associated with an adverse event or not. Patients with daily dose > 240 mg reported on the relugolix dosing log will be flagged in the listing of study drug administration.

7. EFFICACY ANALYSES

7.1. General Considerations

Unless otherwise specified, efficacy analyses will be conducted using the mITT population according to the randomized treatment assignment and will be stratified by the randomization stratification factors (geographic region, presence of metastatic disease, and baseline age). If the group of patients from any of the individual randomization stratification factors (eg, with presence of metastatic disease) comprises < 10% of the entire mITT population, this stratification factor will be collapsed for stratified analyses. In addition, if there were < 15 patients in one of the 12 strata, stratification factors of presence of metastatic disease and baseline age would be used in the stratified analysis for more robust strata-adjusted estimation of treatment effect. If there is a discrepancy > 5% between interactive web response systems (IWRS) and CRF on a stratification factor, the stratification factor reported in CRF will be used as primary analyses and that reported in IWRS will be used as sensitivity analyses, unless otherwise specified. In the case of a discrepancy > 5% between IWRS and CRF on the presence of metastatic disease stratification factor, patients with metastatic prostate cancer will be identified from CRF instead of from IWRS.

For efficacy endpoints related to testosterone test results, analytical testosterone results from the central bioanalytical laboratory using a liquid chromatography–tandem mass spectrometry method, with a lower limit of quantification of 5 ng/dL, will be used. Unless the endpoints are related to testosterone recovery, results from safety follow-up, 60-day and 90-day testosterone recovery (refer to Table 2) will be excluded from the efficacy analyses.

7.2. Multiplicity Adjustment

The primary and the key secondary efficacy analyses will be performed at an overall two-sided type I error of 0.05. A test will be deemed statistically significant if the two-sided p-value rounded to 4 decimal places is less than 0.05. If the result of the primary endpoint analysis meets the respective evaluation criterion of the primary endpoint, the key secondary endpoints will then be tested with a fixed-sequence testing procedure (as illustrated in Figure 2) to maintain the overall familywise error rate of 0.05 for the testing of primary and key secondary endpoints.

The key secondary endpoints with alpha protection for labelling purposes are identified and defined in Section 7.4.1.

All p-values (if provided) aside from the endpoints listed in the testing order are not adjusted in multiplicity, thus are at a nominal level of 0.05.

7.3. Primary Efficacy Endpoint

The study has two separate evaluation criteria for the primary efficacy endpoint to support different regulatory requirements for assessing benefit:

- Evaluation Criterion 1: to determine whether the sustained castration rate for relugolix is ≥ 90%. The lower bound of the 95% CI for the cumulative probability of sustained testosterone suppression in the relugolix treatment group will be calculated and must be at least 90% for this criterion to be met.

- Evaluation Criterion 2: to establish the noninferiority of relugolix to leuprolide acetate 3-M depot injection, as assessed by the cumulative probability of sustained testosterone suppression. The lower bound of the 95% CI for the difference in the cumulative probability of sustained testosterone suppression between the 2 treatment groups will be calculated and must be greater than or equal to the noninferiority margin of –10% for this criterion to be met.

Evaluation Criterion 1 is a regulatory requirement from the United States Food and Drug Administration (FDA). Evaluation Criterion 2 is required by European Medicines Agency (EMA) and Japanese Pharmaceutical and Medical Device Agency (PMDA). For the FDA, Evaluation Criterion 1 is the trial success criterion for the primary efficacy endpoint, while Evaluation Criterion 2 is the first to be tested in the order of ranked key secondary endpoints to assess noninferiority of relugolix compared with leuprolide acetate after Evaluation Criterion 1 is met.

7.3.1. Primary Efficacy Analyses

The primary hypotheses associated with two evaluation criteria for the primary endpoint in this study are:

1. Hypothesis 1, corresponding to Evaluation Criterion 1: the cumulative probability of testosterone suppression to < 50 ng/dL (1.7 nmol/L) for relugolix while on study treatment from Week 5 Day 1 through Week 49 Day 1 is ≥ 90%.

   Null hypothesis H_{01}: $\pi_R < 0.9$ versus Alternative hypothesis H_{a1}: $\pi_R \geq 0.9$

2. Hypothesis 2, corresponding to Evaluation Criterion 2: relugolix is noninferior to leuprolide acetate 3-M depot injection, as assessed by the cumulative probability of sustained testosterone suppression with a noninferiority margin of –10%.

   Null hypothesis H_{02}: $\pi_R - \pi_L < -10\%$ versus Alternative hypothesis H_{a2}: $\pi_R - \pi_L \geq -10\%$

   where π_R and π_L are the sustained castration rates for the relugolix and leuprolide acetate groups, respectively.

The primary analysis is the calculation of the Kaplan-Meier cumulative estimate of the proportion of patients who achieve and maintain castrate levels of testosterone (< 50 ng/dL) from Week 5 Day 1 through Week 49 Day 1.

The 95% CI for the Kaplan-Meier estimation is calculated using the exponential Greenwood formula via log-log transformation of the survival function.

Variance of treatment difference will be calculated using the formula

$$\hat{V}\left[\widehat{S_R}(t) - \widehat{S_L}(t)\right] = \hat{V}\left[\widehat{S_R}(t)\right] + \hat{V}[\widehat{S_L}(t)];$$

Where each of the variance of the Kaplan-Meier estimate will be calculated using the Greenwood's formula

$$\hat{V}\left[\hat{S}(t)\right] = \hat{S}(t)^2 \sum_{t_i \le t} \frac{d_i}{n_i(n_i - d_i)}$$

where n_i denotes the number of patients at risk at time t_i and d_i denotes the number of events observed at time t_i. 95% CI of treatment difference will be calculated via linear transformation of the difference in survival function with pooled variance, as follows:

$$[(\widehat{S_R}(t) - \widehat{S_L}(t)) - 1.96\sqrt{\hat{V}\left[\widehat{S_R}(t) - \widehat{S_L}(t)\right]}, (\widehat{S_R}(t) - \widehat{S_L}(t)) + 1.96\sqrt{\hat{V}\left[\widehat{S_R}(t) - \widehat{S_L}(t)\right]}\,].$$

Survival functions, starting at the Week 5 Day 1 (Day 29) to Week 49 Day 1 (Day 337), will be plotted and summarized by treatment group.

7.3.1.1. Definition of Testosterone Test Result at Week 5 Day 1

Serum concentration of testosterone at Week 5 Day 1 can be obtained at targeted Study Day 29, with a visit window from Day 22 to Day 43 inclusive (see Table 2). If more than one test result lies within the visit window, the result with the study day closest to Day 29 target date will be used. If there are two results equidistant to the scheduled target study day, the earlier assessment will be used as testosterone test result at Week 5 Day 1.

7.3.1.2. Kaplan-Meier Analysis and Censoring Rules

In general, patients with testosterone escape (any testosterone test result rising above the castration level [≥ 50 ng/dL]) from Week 5 Day 1 through Week 49 Day 1 will be considered as an event in the Kaplan-Meier analysis. The time from the date of the first dose to date of the first testosterone escape will be considered as the event time.

Patients who have not reached castration level at Week 5 Day 1 will be considered as having had an event at the target day of Week 5 Day 1.

Patients who discontinued from the study prior to Week 5 Day 1 will be censored at the target day of Week 5 Day 1. In addition, patients without a Week 5 Day 1 assessment will be considered to have had an event at the target day of Week 5 Day 1.

For patients reaching castrate levels at Week 5 Day 1, the following rules will be applied to the Kaplan-Meier analysis for estimation of sustained castration rate with consideration for missed visits. Time to event or censoring, whichever occurs first, will be used in Kaplan-Meier analysis.

a) Patients who had one or more consecutive missed visits (ie, a visit gap of > 42 days) and had a non-castrate assessment immediately after the missed visit(s) will be considered as having an escape at the target day of the earliest missed visit prior to the non-castrate assessment;

b) Patients with one missed visit who has a castrate assessment immediately before and after the missed visit will be assumed to be castrated at the missed visit;

c) Patients with two or more consecutive missed visits (ie, a visit gap of > 70 days) and who had a castrate assessment immediately before and after the missed visits will be censored at the last available testosterone assessment prior to the missed visits.

Note: A visit gap of 70 days for two missed visits was used to consider the duration between two expected visits (8 weeks), plus the visit windows allowed per protocol (7 days for each

expected visit with a total of 2 weeks as visit window for two visits). A visit gap of 42 days for one missed visit was used to consider the visit of every 4 weeks, and ± 7 days as the visit window, per protocol.

Otherwise, patients will be censored at the last available assessment prior to the follow-up visits, including patients who discontinue from the trial for reasons other than a non-castrate testosterone level.

If the above censoring rules are applied to a patient in multiple instances, the date of the earliest censoring for missed visits will be used as the date of censoring.

In addition, patients who had initiated therapies known to suppress testosterone (refer to Appendix Table 1) will be censored at the time of last testosterone assessment prior to the initiation of such therapies.

The time to event or censoring, whichever comes first, will be summarized by the Kaplan-Meier method. If the event time or censoring time is after Day 337, Day 337 will be used as the event time or censoring time.

For Hypothesis 1:

Estimation of the cumulative probability of testosterone suppression to < 50 ng/dL at Week 49 Day 1 (Day 337) for relugolix arm will be presented along with the 95% CI.

Similarly, estimation of the cumulative probability of testosterone suppression to < 50 ng/dL at Week 49 Day 1 (Day 337) for leuprolide acetate will be presented along with the 95% CI.

For Hypothesis 2:

Treatment difference in the cumulative probability of testosterone suppression to < 50 ng/dL at Week 49 Day 1 (Day 337) between relugolix and leuprolide acetate arms will be presented along with the 95% CI.

7.3.2. Sensitivity Analyses

To assess the robustness of the primary analyses for Evaluation Criteria 1 and 2, the following sensitivity analyses of the primary endpoint will be performed.

7.3.2.1. Sensitivity Analysis 1

Analyses of the primary endpoint will be repeated in the per-protocol population.

7.3.2.2. Sensitivity Analysis 2

In addition to the censoring rules under Section 7.3.1.2, patients who had received concomitant medications and herbal supplements that could possibly affect testosterone level during study treatment will be excluded from the analysis.

7.3.2.3. Sensitivity Analysis 3

Patients who had missed two or more consecutive visits after Week 5 Day 1 (censoring rule c in Section 7.3.1.2) or discontinued from the study early will be considered to have an event at the target day of the earliest missed visit.

7.3.2.4. Sensitivity Analysis 4

In order to assess the impact of delayed testosterone suppression to castration level, analyses of the primary endpoint will be repeated by considering that patients who have not reached castration level at Week 5 Day 1 are censored at Week 5 Day 1.

7.3.3. Subgroup Analyses

Subgroup analyses of primary efficacy endpoint will be performed to determine whether treatment effects are consistent across clinically important subgroups. Testosterone castration rates at Week 49 Day 1 (Day 337) for relugolix arm and their 95% CI will be displayed in a forest plot. Similarly, the difference in testosterone castration rates at Week 49 Day 1 between relugolix and leuprolide arms and their 95% CI will also be displayed in a forest plot. Subgroups will include but not be limited to the following:

- Geographic region (North and South America, Europe, Asia and Rest of World);
- Geographic region (North America, other regions);
- Geographic region (Asia [Japan, Korea, China, and Taiwan], other regions);
- Ethnicity (Hispanic or Latino, Not Hispanic or Latino);
- Presence of metastatic disease (yes, no) per IWRS;
- Presence of metastatic disease (yes, no) per CRF if > 5% discrepancy from IWRS;
- Age category (≤ 75 years old, > 75 years old);
- Race (white, Black/African American, Asian, or others);
- Clinical disease state presentations (biomedical or clinical relapse, newly diagnosed metastatic disease, advanced localized disease);
- Gleason score at study entry (< 8, ≥ 8);
- Baseline testosterone level (< 250 ng/dL, ≥ 250 ng/dL);
- Baseline PSA level (< 20 ng/mL, ≥ 20 ng/mL).

In addition, testosterone castration rates at Week 49 Day 1 (Day 337) for the two different dose levels (22.5 mg vs. 11.25 mg) in leuprolide arm and their 95% CI will be provided separately.

7.4. Secondary Efficacy Endpoints

7.4.1. Key Secondary Efficacy Endpoints with Alpha-Protection

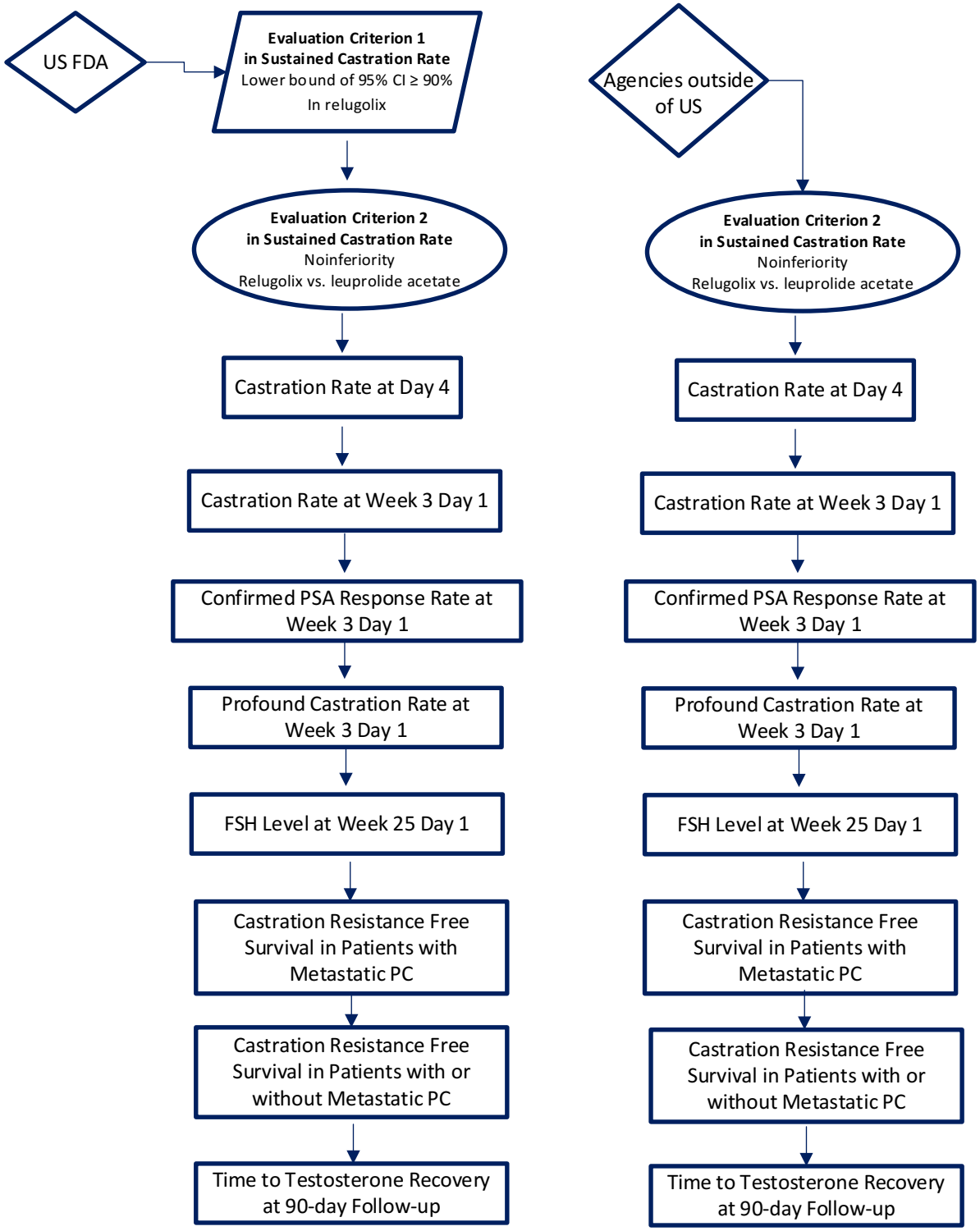
The key secondary endpoints will be tested in the following order (see Figure 2) for different regulatory agencies, with a fixed-sequence testing procedure to control the overall familywise error rate at a two-sided type I error rate of 0.05 across primary and key secondary endpoints.

For the testing order to be submitted to the FDA, Evaluation Criterion 2 for noninferiority of the primary efficacy endpoint will be assessed first in the order of hierarchical hypothesis testing for the key secondary endpoints after Evaluation Criterion 1 is met.

For the testing order to be submitted to EMA and PMDA, Evaluation Criterion 2 for noninferiority of the primary efficacy endpoint will be the primary efficacy assessment.

Timing of analyses for primary and key secondary endpoints is also included in Table 4. Castration resistance free survival and time to testosterone recovery back to 280 ng/dL at 90-day follow-up will be tested at the final analysis only if all the above endpoints reach statistical significance in the primary analysis. Endpoints in the higher order will be updated with descriptive statistics in the final analysis.

Figure 2: Diagram of Testing Order in Primary and Key Secondary Endpoints for Different Regulatory Agencies



Abbreviations: CI = confidence interval; EMA = European Medicines Agency; PC = prostate cancer; US FDA = United States Food and Drug Administration; PMDA = Japanese Pharmaceutical and Medical Device Agency.

Table 4: Testing Order and Timing of Analysis for Primary and Key Secondary Endpoints for Different Regulatory Agencies

Endpoints	Testing Order for the US FDA		Testing Order for EMA and PMDA	
	At Primary Analysis	At Final Analysis	At Primary Analysis	At Final Analysis
Sustained castration rate per Evaluation Criterion 1 (≥ 90% in relugolix)	**1**	**Update**	**NA**	**Update**
Sustained castration rate per Evaluation Criterion 2 (noninferiority of relugolix compared with leuprolide acetate)	**2**	**Update**	**1**	**Update**
Castration rate on Week 1 Day 4	**3**	**Update**	**2**	**Update**
Castration rate on Week 3 Day 1	**4**	**Update**	**3**	**Update**
Confirmed PSA response rate at Week 3 Day 1	**5**	**Update**	**4**	**Update**
Profound castration rate at Week 3 Day 1	**6**	**Update**	**5**	**Update**
FSH level at Week 25 Day 1	**7**	**Update**	**6**	**Update**
Castration resistance free survival during the 48-week treatment in patients with metastatic prostate cancer [a]	**NA**	**8**	**NA**	**7**
Castration resistance free survival during the 48-week treatment in patients with or without metastatic prostate cancer [a]	**NA**	**9**	**NA**	**8**
Time to testosterone recovery back to 280 ng/dL at the 90-day follow-up in patients participating in testosterone recovery follow-up [a]	**10** [b]	**NA**	**9** [b]	**NA**

Abbreviations: EMA = European Medicines Agency; FSA = follicle-stimulating hormone; NA = not applicable; PSA = prostate-specific antigen; US FDA = United States Food and Drug Administration.

[a] Castration resistance free survival and time to testosterone recovery back to 280 ng/dL at the 90-day follow-up will be tested at the final analysis only if all the above endpoints reach statistical significance in the primary analysis. Endpoints in the higher order will be updated with descriptive statistics in the final analysis.

[b] Analysis of time to testosterone recovery back to 280 ng/dL at the 90-day follow-up will be performed at the primary analysis. Testing order of time to testosterone recovery will be followed by castration resistance free survival in the final analysis.

7.4.2. Other Secondary Efficacy Endpoints

Other secondary efficacy endpoints include the following:

- Sustained profound castration rate from Week 5 Day 1 through Week 49 Day 1 defined as the cumulative probability of testosterone suppression to < 20 ng/dL (0.7 nmol/L) while on study treatment from Week 5 Day 1 through Week 49 Day 1;
- Sustained profound castration rate from Week 25 Day 1 through Week 49 Day 1 defined as the cumulative probability of testosterone suppression to < 20 ng/dL (0.7 nmol/L) while on study treatment from Week 25 Day 1 through Week 49 Day 1;
- Profound castration rate at Week 1 Day 4;
- Time to testosterone recovery back to > 50 ng/dL;
- Time to testosterone recovery back to > 280 ng/dL or baseline;
- PSA response rate at Week 3 Day 1 and Week 5 Day 1;
- Proportion of patients with PSA concentration < 0.02 ng/mL (0.02 µg/L) at Week 25 Day 1 visit;
- Time to PSA progression;
- FSH level over time (analysis details refer to Section 8).
- Absolute values and changes from baseline in the scores of each domain in EORTC-QLQ-C30 and EORTC-QLQ-PR25, at regular intervals during treatment, and during the Follow-Up and/or End of Treatment visits (analysis details refer to Section 7.6).

7.4.3. Analyses of Secondary Efficacy Endpoints

7.4.3.1. Castration Rate and Profound Castration Rate at Day 4 and Week 3 Day 1

Castration rate at Day 4 and Week 3 Day 1 (ie. Day 15) are in the key secondary endpoints and will be estimated by use of the Kaplan-Meier method based on time to first testosterone castration, which is defined as time from the date of first dose to the date of initial testosterone suppression to < 50 ng/dL (1.7 nmol/L).

The Kaplan-Meier method will be applied with the following,

- Patients with initial testosterone suppression to < 50 ng/dL (1.7 nmol/L) will be considered as an event at the time of initial suppression.
- Patients without testosterone suppression to < 50 ng/dL (1.7 nmol/L) will be censored at the last available assessment prior to the follow-up visits.

Cumulative incidence curve of time to first testosterone castration will be plotted and summarized by treatment group. Castration rates at Day 4 and Week 3 Day 1 based on the cumulative incidence curve of time to first testosterone castration will be estimated along with 95% CI.

Treatment difference in castration rates at Day 4 and Week 3 Day 1 between relugolix and leuprolide acetate will be estimated along with 95% CI using a similar method to that described

in Section 7.3.1. The stratified test statistics via log-log transformation of the difference in cumulative incidence curve at a fixed time point is constructed by (Klein 2007).

$$x^2 = \frac{\left\{\sum_{s=1}^{M} \log\left(-\log\left(\widehat{S_{R_s}}(t)\right)\right) - \log\left(-\log\left(\widehat{S_{L_s}}(t)\right)\right)\right\}^2}{\sum_{s=1}^{M}\left\{\frac{V(\widehat{S_{R_s}}(t))}{\left[\widehat{S_{R_s}}(t)\log\left(\widehat{S_{R_s}}(t)\right)\right]^2} + \frac{V(\widehat{S_{L_s}}(t))}{\left[\widehat{S_{L_s}}(t)\log\left(\widehat{S_{L_s}}(t)\right)\right]^2}\right\}}$$

Where s is the s^{th} stratum in the total of M strata and t = Day 4 and Week 3 Day 1 in this case. Each stratified test statistics at a fixed time under two samples follows a chi-square distribution with one-degree freedom. If the stratified test statistics cannot be calculated due to event having not occurred in certain stratum, unstratified test will be used instead. If unstratified test is still not estimable, the stratified version of the Mantel-Haenszel test using pooled Kaplan-Meier estimators in each stratum will be constructed as the following. This test statistics follows a chi-square distribution with one-degree freedom.

$$\chi^2_{MH} = \frac{\left\{\sum_{s=1}^{M} \frac{n_{R_s} n_{L_s}}{n_{R_s} + n_{L_s}} \left(\widehat{S_{R_s}}(t) - \widehat{S_{L_s}}(t)\right)\right\}^2}{\sum_{s=1}^{M} n_{R_s} n_{L_s} V(\widehat{S_{pooled_s}}(t))}$$

Profound castration rate at Day 4 and Week 3 Day 1 will be analyzed similarly as above with < 20 ng/dL (0.7 nmol/L) as testosterone suppression criterion. Profound castration rate at Week 3 Day 1 is one of the key secondary endpoints and will be compared between relugolix and leuprolide acetate using the stratified test statistics, as above.

Testosterone level and percentage change from baseline in testosterone level at each visit will be summarized by treatment group. Categorical summary of testosterone level in the first 4 weeks of treatment will be also summarized. Due to the different visit schedule in the original protocol versus protocol amendment 2, testosterone level in the first 4 weeks under the original protocol will be summarized additionally using the visit windows in Table 5. The same rule in Section 4.3.6 for handling multiple assessments in one visit will be used.

Table 5: Visit Windows in the First 4 Weeks under Original Protocol

Visit	Start Day	Target Day	End Day
Week 1 Day 4 [a]	1	4	5
Week 2 Day 1	6	8	11
Week 3 Day 1	12	15	18
Week 4 Day 1	19	22	25
Week 5 Day1	26	29	43

[a] Start day of Week 1 Day 4 only includes postbaseline assessments on or after the first dose.

7.4.3.2. Sustained Profound Castration Rate

Sustained profound castration rates are defined for two treatment periods:

- Sustained profound castration rate from Week 5 Day 1 through Week 49 Day 1 is defined as the cumulative probability of testosterone suppression to < 20 ng/dL (0.7 nmol/L) while on treatment from Week 5 Day 1 through Week 49 Day 1;
- Sustained profound castration rate from Week 25 Day 1 through Week 49 Day1 is defined as the cumulative probability of testosterone suppression to < 20 ng/dL (0.7 nmol/L) while on treatment from Week 25 Day 1 through Week 49 Day 1.

The Kaplan-Meier method described in Section 7.3.1.2 will be applied similarly by changing the castration level (< 50 ng/dL) to profound castration level (< 20 ng/dL). Sustained profound castration rate from Week 25 Day 1 through Week 49 Day 1 is described below.

Definition of Testosterone Test Result at Week 25 Day 1

Serum concentration of testosterone at Week 25 Day 1 can be obtained at Day 169 with a visit window from Day 156 to Day 183 inclusive. If more than one test result lies within the visit window, the result with the closest study day to Day 169 will be used as testosterone test result at Week 25 Day 1. If there are two results equidistant to the scheduled target day, the earlier assessment will be used as testosterone test result at Week 25 Day 1.

Kaplan-Meier Analysis

Any testosterone test result rising above the profound castration level (≥ 20 ng/dL) between Week 25 Day 1 and Week 49 Day 1 will be considered as an event in the Kaplan-Meier analysis.

Patients who have not reached profound castration level at Week 25 Day 1 will be considered as having had an event at the target day of Week 25 Day 1.

Patients who discontinued from the study prior to Week 25 Day 1 will be censored at the target day of Week 25 Day 1. In addition, patients without a Week 25 Day 1 assessment will be considered to have had an event at the target day of Week 25 Day 1.

For patients reaching profound castration levels at Week 25 Day 1, the following rules will be applied to the Kaplan-Meier analysis for estimation of sustained profound castration rate with consideration for missed visits. Time to event or censoring, whichever occurs first, will be used in Kaplan-Meier analysis.

a) Patients who had one or more consecutive missed visits (ie, a visit gap of > 42 days) and had a non-profound castrated assessment immediately after the missed visit(s) will be considered as having an escape at the target day of the earliest missed visit prior to the non-profound castrated assessment;

b) Patients with one missed visit who has a profound castrated assessment immediately before and after the missed visit will be assumed to be profound castrated at the missed visit;

c) Patients with two or more consecutive missed visits (ie, a visit gap of > 70 days) and who had a profound castrated assessment immediately before and after the missed visits will be censored at the last available testosterone assessment prior to the missed visits.

7.4.3.3. Prostate-Specific Antigen Response Rate at Week 3 Day 1 and Week 5 Day 1

Proportion of patients with PSA reduction from baseline at Week 3 Day 1 and Week 5 Day 1 will be summarized along with 95% exact CI. Patients without PSA assessment at Week 3 Day 1 and Week 5 Day 1 will be considered as nonresponders. The following two criteria will be considered for PSA reduction:

- > 50% reduction from baseline;
- > 90% reduction from baseline.

Confirmed PSA response at Week 3 Day 1 is one of the key secondary endpoints and defined as > 50% reduction in PSA from baseline at Week 3 Day 1 followed with confirmation at Week 5 Day 1. Treatment comparison between relugolix and leuprolide acetate will be performed using a stratified Cochran-Mantel-Haenszel test.

Change from baseline in PSA at Week 3 Day 1 and Week 5 Day 1 will be also summarized.

7.4.3.4. Undetectable Prostate-Specific Antigen Rate at Week 25 Day 1

Proportion of patients who have a PSA concentration < 0.02 ng/mL (0.02 µg/L) at Week 25 Day 1 will be summarized by treatment group along with 95% exact CI. Patients with missing PSA assessment at Week 25 Day 1 will be considered as nonresponders.

7.4.3.5. Time to Testosterone Recovery

Testosterone levels will be followed for up to 90 days in approximately 100 patients randomized to relugolix and approximately 50 patients randomized to leuprolide acetate who complete 48 weeks of treatment and who do not plan to start alternative androgen deprivation therapy within the following 12 weeks (or within 24 weeks following the last injection of leuprolide acetate). As the enrollment for the follow-up of testosterone recovery has been completed for the patients enrolled under protocol amendment 2, analysis of time to testosterone recovery will be performed in the primary analysis and there will be no updates in the final analysis.

Patients will be censored at the time of last follow-up visit or initiation of treatments known to suppress testosterone (refer to Appendix Table 1), whichever comes earlier.

Time to testosterone recovery will be summarized using the Kaplan-Meier method with the following 3 testosterone recovery levels:

- ≥ 50 ng/dL (1.7 nmol/L);
- Above baseline level or 280 ng/dL;
- Above 280 ng/dL (as one of the key secondary endpoints).

Cumulative incidence of time to testosterone recovery at the 30-day, 60-day and 90-day follow-up visits will be estimated by treatment group along with 95% CI.

The unstratified test via log-log transformation of the difference in cumulative incidence of time to testosterone recovery back to 280 ng/dL will be compared at the 90-day follow-up visit (unstratified test version as in Section 7.4.3.1).

7.4.3.6. Time to Prostate-Specific Antigen Progression

Prostate-specific antigen progression is defined per Prostate Cancer Clinical Trials Working Group 3 Guidelines (Scher et al, 2016) as the following:

1. With declining PSA from baseline: The first increase in PSA of ≥ 25% and ≥ 2 ng/mL above the nadir with confirmation by a consecutive PSA measurement at least 3 weeks later.

   The date of initial PSA progression will be considered as the date of event for the Kaplan-Meier analysis.

   The initial PSA progression can be confirmed with later PSA measurements as long as all PSA values in between from the initial PSA progression indicated PSA progression. For example, patient had initial PSA progression at Day 78, the second PSA progression at Day 92, and the third PSA progression at Day 106. The patient will be considered as having PSA progression at Day 78 with confirmation at Day 106.

2. Without declining PSA from baseline: PSA increase of 25% or greater and 2 ng/mL from baseline beyond Week 12 (> 84 days after the date of the first dose).

Patients without confirmed PSA progression will be censored at the last available assessment prior to the follow-up visits. As a sensitivity analysis, patients will be censored at the last available assessment prior to the follow-up visits or at the time of initiating any treatment that can affect or alter PSA levels (refer to Appendix Table 2), whichever comes earlier.

Time to PSA progression is defined as the time from the date of first dose to the date of the initial PSA progression which was subsequently confirmed and will be analyzed using the Kaplan-Meier method. Treatment comparison via hazard ratio will be performed using Cox proportional hazard model.

7.4.3.7. Castration Resistance Free Survival

Castration resistance free survival will not be analyzed for the primary analyses but will be analyzed at the final analyses when approximately 390 patients with metastatic disease have been randomized to the study and have had the opportunity to be evaluated for 48 weeks of study treatment and completed the 30-day safety follow-up or discontinued early.

Castration resistance free survival is in the key secondary endpoints and will be analyzed under metastatic patients only and mITT population. PSA results and deaths from follow-up visits will be excluded from the analysis. Patients with metastatic disease will be identified through presence of distant metastasis (M1) at study entry reported in the CRF. If the discrepancy between IWRS and CRF is > 5%, sensitivity analysis based on the presence of metastatic disease reported in the IWRS will be performed.

Castration resistance free survival is defined as the time from the date of first dose to the date of PSA progression while castrated or death due to any reason, whichever occurs earlier. A confirmed PSA progression prior to testosterone escape or the date of testosterone censoring will be considered as an event. Patients will be censored at the last PSA assessment prior to follow-up visits, the date of censoring for testosterone, or the date of testosterone escape, whichever comes earlier.

As a sensitivity analysis, patients will be censored at the last available assessment prior to the follow-up visits, the date of censoring for testosterone, the date of testosterone escape, or the time of initiating androgen receptor inhibitors (eg, enzalutamide), whichever comes earlier.

Castration resistance free survival will be analyzed using the Kaplan-Meier method. Treatment comparison with hazard ratio will be performed using Cox proportional hazard model and p-value will be provided using unstratified log-rank test on patients with metastatic disease and stratified log-rank test on mITT population. Proportionality in survival curves between two treatment groups will be assessed. CRFS rate at Week 49 Day 1 (Day 337) will be estimated along with 95% CI.

7.5. Exploratory Efficacy Endpoints and Analyses

7.5.1. Overall Survival

Overall survival is defined as time from randomization to date of death prior to data cut-off date. The Kaplan-Meier method will be used to describe survival distributions by treatment group. Patients will be censored at the last contact date prior to data cut-off date if patient is known to be alive prior to data cut-off date.

7.5.2. Pharmacogenomics Analyses

A separate SAP may be provided for pharmacogenomic analysis on the presence of polymorphisms in germline genes related to the hypothalamic-pituitary androgen pathway, prostate cancer risk, or to drug-metabolizing enzymes and transporter proteins that might be implicated in the drug disposition, safety, or efficacy of relugolix.

7.6. Patient Reported Outcomes

Patient reported outcome questionnaires (EORTC QLQ-C30, EORTC QLQ-PR25, and EuroQoL EQ-5D-5L) are completed by patients at Day 1, Week 5, Week 13, Week 25, Week 37, Week 49, and the 30-day safety follow-up visit. They are also completed in the 60-day and 90-day testosterone recovery follow up if patients are participating in the testosterone recovery follow up.

7.6.1. EORTC QLQ-C30 and QLQ-PR25

The scoring procedure for EORTC QLQ-C30 and QLQ-PR25 is shown in Table 6:

Table 6: Scoring Procedure

Domain	Number of Items	Item Range[a]	Item Numbers
EORTC QLQ-C30			
Global Health Status	2	6	29, 30
Functional Scales			
Physical functioning	5	3	1 to 5
Role functioning	2	3	6, 7

Emotional functioning	4	3	21 to 24
Cognitive functioning	2	3	20, 25
Social functioning	2	3	26, 27
Symptom Scales			
Fatigue	3	3	10, 12, 18
Nausea and vomiting	2	3	14, 15
Pain	2	3	9, 19
Dyspnoea	1	3	8
Insomnia	1	3	11
Appetite loss	1	3	13
Constipation	1	3	16
Diarrhoea	1	3	17
Financial difficulties	1	3	28
EORTC QLQ-PR25			
Symptom Scales			
Urinary symptoms	8	3	31 to 37, 39
Incontinence aid use	1	3	38
Bowel symptoms	4	3	40 to 43
Hormonal treatment-related symptoms	6	3	44 to 49
Functional Scales			
Sexual activity	2	3	50, 51
Sexual functioning	4	3	52 to 55

[a] Item range is the difference between the possible maximum and the minimum response to individual items (eg, range for items taking values from 1 to 4 is 3).

For all scales, the raw score is the average of component items:

$$\text{RawScore} = (I_1 + I_2 + \ldots + I_n)/n,$$

Where n is the number of responded items within the scale.

Missing Items

For multi-item scales, if at least half of the items from the scale have been answered, use all the items that were completed and apply the above raw score calculation. Otherwise, set the scale score to missing. For single-item scales, set the score to missing if the response of the item is missing.

Transformation of Scale Scores

Apply the below linear transformation to 0 – 100 to obtain the score:

Functional Scales:

$$\text{Score} = \left\{ 1 - \frac{RawScore-1}{range} \right\} \times 100$$

Symptom Scales/Global Health Status:

$$\text{Score} = \left\{ \frac{RawScore-1}{range} \right\} \times 100$$

A high scale score represents a higher response level. For example, a high score for a functional scale represents a high/healthy level of functioning while a high score for a symptom scale represents a high level of symptomatology/problems.

Analyses on Scale Scores

Scores and change from baseline scores from each scale will be summarized by each visit and treatment group.

For each scale, a mixed model using repeated measures will be performed with baseline scores, treatment, visit, randomization stratification factors (per EDC if the discrepancy between IWRS and CRF is > 5%), and treatment by visit interaction included as fixed effects. Change from baseline in each scale for each patient at each visit will be included as the dependent variable. An unstructured covariance matrix will be used in the mixed effect model. The least square means from the mixed model using repeated measures will be provided for each treatment group at each visit.

7.6.2. EuroQoL EQ-5D-5L

The EQ-5D-5L consists of the EQ-5D-5L descriptive system and the EQ Visual Analogue Scale (EQ VAS).

The descriptive system comprises 5 dimensions (mobility, self-care, usual activities, pain/discomfort, and anxiety/depression). Each dimension has 5 levels: no problems (1 as numerical score), slight problems (2 as numerical score), moderate problems (3 as numerical score), severe problems (4 as numerical score), and extreme problems (5 as numerical score). There should be only ONE response for each dimension. If more than one box is ticked for a single dimension, it should be treated as missing values. For each dimension, the count and proportion of improvement by level or at least one level will be tabulated by treatment group and each visit. The denominator for the proportion will be based on the number of patients who provided non-missing responses to the dimensions at baseline.

The EQ VAS records the patient's self-rated health on a 20-cm vertical visual analogue scale. It asks respondents to "mark an X on the scale to indicate how your health is TODAY" and then to "write the number you marked on the scale in the box below." If there is a discrepancy between where the respondent has placed the X and the number respondent has written in the box, administrator should use the number in the box. The non-missing EQ VAS score will be summarized, including 25th and 75th percentiles, at each visit by treatment group.

8. PHARMACODYNAMIC ANALYSES

Predose concentrations and change from baseline at each visit will be summarized in predose concentrations of luteinizing hormone (LH), follicle-stimulating hormone (FSH), dihydrotestosterone, and sex hormone-binding globulin. FSH concentration level at Week 25 Day 1 between relugolix and leuprolide acetate will be compared using the two-sample t-test.

9. PHARMACOKINETIC ANALYSES

Plasma concentrations in relugolix will be summarized and listed.

For patients from China and patients from Japan in the pharmacokinetic subset, pharmacokinetic parameters C_{max}, $AUC_{0-\tau,}$ and t_{max} from Day 1 and Week 2 Day 1 visit will be determined.

A separate SAP will be detailed for further population pharmacokinetic analyses.

10. SAFETY ANALYSES

Unless otherwise specified, safety analyses will be conducted using the safety population according to the treatment actually received.

10.1. Adverse Events

Per protocol, adverse event reporting period refers to adverse events planned to be collected from the time of the first dose of study drug through the safety follow up visit approximately 30 days after the end of treatment period, or the date of initiation of another investigational agent or hormonal therapy or surgical intervention following the end of treatment, whichever occurs first. Serious adverse events reported to the investigator after the safety reporting period should be reported to the sponsor if the investigator assesses the event as related to the study drug treatment.

The severity of all treatment-emergent adverse events will be evaluated by the investigator based on the National Cancer Institute's Common Terminology Criteria for Adverse Events (NCI CTCAE v4.03) and will be coded to preferred term and system organ class using the Medical Dictionary for Regulatory Activities (MedDRA) 22.0 or higher.

Treatment-emergent adverse event is defined as any adverse event that occur after administration of the first dose of study drug and prior to the end of the adverse event reporting period, except for drug related serious adverse events.

Adverse event summaries will be based on treatment emergent adverse events, unless otherwise specified. Any procedure-related adverse event reported during screening will be listed and flagged in listings by patient.

Tabular summaries with the number and percentage of patients with the following adverse events (but not limited to) will be provided:

- Overview of adverse events
- All adverse events
  - By system organ class and preferred term
  - By decreasing frequency of preferred term
  - By system organ class, preferred term, and maximum severity
  - Study drug-related per investigator by system organ class and preferred term
- Grade 3 or higher adverse events
  - By system organ class and preferred term
  - By decreasing frequency of preferred term
  - Study drug-related per investigator by system organ class and preferred term
- Adverse events leading to study drug withdrawn
  - By system organ class and preferred term
  - By decreasing frequency of preferred term

- Adverse events leading to study drug interruption
  - By system organ class and preferred term
  - By decreasing frequency of preferred term
- Adverse events resulting in fatal outcome
  - By decreasing frequency of preferred term
- Serious adverse events
  - By system organ class and preferred term
  - By decreasing frequency of preferred term
  - By system organ class, preferred term, and maximum severity
  - By system organ class, preferred term, and relationship to study drug
- Adverse events of clinical interest (alanine aminotransferase [ALT] or aspartate aminotransferase [AST] ≥ 3 × upper limit of normal [ULN])
  - By decreasing frequency of preferred term
  - By system organ class, preferred term, and maximum severity

In addition, adverse event categories defined in Table 7 will be summarized by decreasing frequency of preferred term.

Table 7: Constitution of Adverse Event Categories

Category	Search Criteria
Loss of bone mineral density	Osteoporosis/Osteopenia SMQ (broad) All PTs including the term "Fracture," excluding "Tooth fracture" and "Fracture of penis"
QTc prolongation	Torsades de Pointes/QT prolongation SMQ (broad)
Hepatic transaminase elevations	Drug-related hepatic disorders SMQ (narrow)
Carbohydrate and lipid metabolic effects	Dyslipidaemia SMQ (broad) Hyperglycemia/New onset diabetes mellitus SMQ (narrow)
Adverse Cardiovascular events	Major Adverse Cardiovascular Events: Myocardial infarction SMQ (broad) Central nervous system haemorrhages and cerebrovascular conditions SMQ (broad) Deaths due to all causes Ischemic heart disease Ischaemic heart disease SMQ (broad)

Vasomotor symptoms	The following 5 PTs will be included: 1. Hyperhidrosis 2. Feeling hot 3. Hot flush 4. Night sweats 5. Flushing
Mood Disorders	Depression and suicide/self-injury SMQ (broad)
Hypersensitivity	Hypersensitivity SMQ (narrow)

Abbreviations: PT = preferred term; SMQ = Standardised MedDRA Queries.

10.2. Laboratory Data

Laboratory parameters, including chemistry and hematology panels, specified as per protocol and collected from the central laboratory, will be tabulated and presented in by-patient listings. Urinalysis and hepatitis virus serological test results, if any, will be provided in by-patient listings only.

The NCI CTCAE grading scale with numeric component will be used to categorize toxicity grade for laboratory parameters (CTCAE v4.03). Parameters that have criteria available for both low and high values (eg, hypercalcemia for a high value of calcium and hypocalcemia for a low value of calcium) will be summarized for both criteria (low and high). Patients will only be counted once for each criterion. The same patient can be counted for both criteria if the patient has laboratory values meeting each criterion. Shift tables will be provided for each gradable parameter to summarize baseline toxicity grade versus worst postbaseline toxicity grade. For laboratory parameters that are not gradable by the CTCAE, a shift table based upon the normal range (low, normal, and high) will be provided for each parameter to summarize the baseline versus worst postbaseline results.

Boxplots of laboratory values over time will be plotted for key laboratory parameters. These laboratory parameters include but are not limited to hematology (hemoglobin, platelets, leukocytes, neutrophils), creatinine, glomerular filtration rate, and hepatic function panel (ALT, AST, ALP, and total bilirubin [TBL]).

The change from baseline to each postbaseline study visit will be presented by treatment group for each laboratory test.

The number and proportion of patients with liver test elevations will be presented by treatment group. Liver test elevations are assessed by using postbaseline results for ALT, AST, ALP, and TBL based on the definitions presented in Table 8.

Table 8: Categories of Liver Function Test Elevations

Laboratory Test	Category
ALT or AST	ALT or AST > ULN - < 3 × ULN ALT or AST ≥ 3 × to < 5 × ULN ALT or AST ≥ 5 × to < 8 × ULN ALT or AST ≥ 8 × to < 10 × ULN ALT or AST ≥ 10 × to < 20 × ULN ALT or AST ≥ 20 × ULN
Total bilirubin	Total bilirubin > 2 × ULN
ALT or AST and total bilirubin	ALT or AST ≥ 3 × ULN + total bilirubin > 2 × ULN
ALT or AST, total bilirubin, and ALP	ALT or AST ≥ 3 x ULN + total bilirubin > 2 × ULN + ALP < 2 × ULN

Abbreviations: ALP = alkaline phosphatase; ALT = alanine aminotransferase; AST = aspartate aminotransferase; ULN = upper limit of normal.

The number and percent of patients with either ALT or AST ≥ 3 times the ULN and TBL > 2 times ULN concurrently, defined as measurements occurred on the same day, will also be presented.

Worst postbaseline ALT and TBL values will be evaluated using the evaluation of drug induced serious hepatotoxicity (eDISH) concept.

Selected chemistry and hematology test results meeting pre-defined limits of change at any time and based on last observation on treatment will be also summarized per pre-defined threshold (refer to Appendix 2).

10.3. Other Safety Analyses

10.3.1. Electrocardiograms

Electrocardiogram assessment will be read by local reading. Overall interpretation will be presented in by-patient listing.

10.3.2. Visual Acuity

Visual acuity will be measured using the tumbling E chart with Visual Acuity Score. A 10-foot standard distance will be used and each correctly identified "E" will increase the Visual Acuity Score by one point. Visual Acuity Score at baseline and at each scheduled postbaseline assessment time point will be presented in by-patient listing. Visual acuity score ≤ 90 or decrease from baseline ≥ 10 points will be flagged in the by-patient listing.

10.3.3. Vital Signs

Blood pressure (systolic and diastolic), heart rate, and weight measurements will be summarized at baseline and each subsequent scheduled assessment by treatment group. Change from baseline will be calculated and presented for each parameter at all scheduled postbaseline assessment timepoints. All vital sign data will also be provided in by-patient listing.

Potentially clinically significant abnormalities in vital signs are defined in Table 9 and will be summarized by treatment group. Potentially clinically significant abnormalities will also be flagged in by-patient listing.

Table 9: Categories of Potentially Clinically Significant Abnormalities in Vital Signs

Vital Sign Parameter	Category
Systolic blood pressure	≥ 140 mmHg ≥ 180 mmHg
	≤ 90 mmHg
	Increase of ≥ 20 mmHg from baseline Decrease of ≥ 20 mmHg from baseline
Diastolic blood pressure	≥ 90 mmHg ≥ 105 mmHg
	≤ 50 mmHg
	Increase of ≥ 15 mmHg from baseline Decrease of ≥ 15 mmHg from baseline
Heart rate	≥ 120 bpm
	< 45 bpm
	Increase of ≥ 15 bpm from baseline Decrease of ≥ 15 bpm from baseline

REFERENCES

Klein J, Logan B, Harhoff M, Andersen PK. Analyzing Survival Curves at a Fixed Point in Time. Stat Med. 2007 Oct 30; 26 (24): 4505-19.

Scher H, Morris M, Stadler W, Higano C, Basch E, Fizazi K, et al. Trial Design and Objectives for Castration-Resistant Prostate Cancer: Updated Recommendations from the Prostate Cancer Clinical Trials Working Group 3. J Clin Onco. 2016 Apr 20; 34 (12):1402-18.

REVISION HISTORY

Table 10: Revision History

Version	Effective Date	Comments
1.0	12SEP2019	New document
2.0	29OCT2019	1. Naming change from "time to castration resistance" to "castration resistance free survival". 2. Updates on Kaplan-Meier analysis for the primary endpoint, sustained castration rate from Week 5 Day 1 to Week 49 Day 1, to align with FDA review comments in Sections 7.3.1.1, 7.3.1.2, 7.3.2.2 and 7.3.2.3. 3. Corrections in test statistics of difference in cumulative incidence curve at a fixed time point in Section 7.4.3.1. 4. Censoring rule b of analysis on sustained profound castration changed to align with sustained castration rate in Section 7.4.3.2.
3.0	06NOV2019	1. Aligned with FDA review comments, changed ≤ 50 ng/dL to < 50 ng/dL for castration level and changed ≤ 20 ng/dL to < 20 ng/dL for profound castration level. 2. To address FDA review comments, added "sustained profound castration rate from Week 5 Day 1 through Week 49 Day 1 as a secondary endpoint, in Table 1, Sections 7.4.2 and 7.4.3.2. 3. To address FDA review comments, similar censoring rules for the primary analysis of sustained castration rate applied to sustained profound castration rates in Section 7.4.3.2.

APPENDIX 1: LIST OF MEDICATIONS THAT HAVE IMPACT ON TESTOSTERONE OR PROSTATE-SPECIFIC ANTIGEN LEVEL

The following lists of medications that are likely to affect analysis of efficacy are identified and will be finalized prior to the database lock. Any changes in this appendix after the finalization of the SAP will not be considered as changes in the planned analysis.

Appendix Table 1: Medication List of Therapies Known to Suppress Testosterone

Generic Name	WHO Drug Dictionary Preferred Term
degarelix	DEGARELIX
	DEGARELIX ACETATE
goserelin	GOSERELIN
	GOSERELIN ACETATE
leuprolide	LEUPROLIDE
leuprorelin	LEUPRORELIN
	LEUPRORELIN ACETATE
abarelix	ABARELIX
histrelin	HISTRELIN
triptorelin	TRIPTORELIN
	TRIPTORELIN ACETATE
	TRIPTORELIN EMBONATE

Appendix Table 2: Medication List of Treatments that Can Affect or Alter Prostate-Specific Antigen

Generic Name	WHO Drug Dictionary Preferred Term
cabazitaxel	CABAZITAXEL
cisplatin	CISPLATIN
docetaxel	DOCETAXEL
doxorubicin	DOXORUBICIN
	DOXORUBICIN HYDROCHLORIDE
etoposide	ETOPOSIDE
fluorouracil	FLUOROURACIL
ifosfamide	IFOSFAMIDE
paclitaxel	PACLITAXEL
bicalutamide	BICALUTAMIDE
chlormadinone acetate	CHLORMADINONE ACETATE
enzalutamide	ENZALUTAMIDE
flutamide	FLUTAMIDE
nilutamide	NILUTAMIDE
epalutmide	EPALUTMIDE
estramustine	ESTRAMUSTINE
sipuleucel-T	SIPULEUCEL-T

APPENDIX 2: LIST OF PRE-DEFINED THRESHOLD IN SELECTED CHEMISTRY AND HEMATOLOGY TEST RESULTS

Chemistry Laboratory	
Liver Function	
ALT > ULN and < 3X ULN	Total BILI > ULN
ALT >= 3X ULN and < 5X ULN	Total BILI >2X ULN
ALT >= 5X ULN and < 10X ULN	
ALT >= 10X ULN and < 20X ULN	ALT or AST >= 3X ULN and Total BILI > 2X ULN
ALT >= 20X ULN	ALT or AST >= 3X ULN and Total BILI > 2X ULN and ALP < 2X ULN
	ALT > ULN and < 3X ULN
AST > ULN and < 3X ULN	
AST >= 3X ULN and < 5X ULN	GGT > ULN and < 3X ULN
AST >= 5X ULN and < 10X ULN	GGT >= 3X ULN and < 5X ULN
AST >= 10X ULN and < 20X ULN	GGT >= 5X ULN and < 10X ULN
AST >= 20X ULN	GGT >= 10X ULN and < 20X ULN
	GGT >= 20X ULN
ALT or AST > ULN and < 3X ULN	GGT > ULN and < 3X ULN
ALT or AST >= 3X ULN and < 5X ULN	
ALT or AST >= 5X ULN and < 10X ULN	
ALT or AST >= 10X ULN and < 20X ULN	
ALT or AST >= 20X ULN	
Renal Function	
CR > 1.5 mg/dL and > BL	GFR < 15 mL/min per 1.73 m^2
CR > 50% increase from BL	GFR >= 15 - < 30 mL/min per 1.73 m^2
	GFR >= 30 - < 60 mL/min per 1.73 m^2
	GFR >= 60 - < 90 mL/min per 1.73 m^2
	GFR >= 90 mL/min per 1.73 m^2

Chemistry Laboratory	
Metabolic Parameters	
Fasting Glucose	Highest Postbaseline Glucose
< 100 mg/dL at BL	Gluc >= 200 mg/dL
< 100 mg/dL at week 48	Gluc >= 200 mg/dL and >= 126 BL
>= 100 - < 126mg/dL at week 48	Gluc >= 500 mg/dL
>= 126 mg/dL at week 48	Gluc >= 500 mg/dL and >= 126 BL
>= 100 - < 126 mg/dL at BL	Total CHOL > 200 mg/dL and > BL
< 100 mg/dL at week 48	Total CHOL increase > 30 mg/dL from BL
>= 100 - < 126 mg/dL at week 48	
>= 126 mg/dL at week 48	HDL < LLN and < BL
	LDL > ULN and > BL
>= 126 mg/dL at BL	TRIG > ULN and > BL
< 100 mg/dL at week 48	
>= 100 - < 126 mg/dL at week 48	
>= 126 mg/dL at week 48	
Electrolytes and other Chemistry Parameters	
ALB < LLN and < BL	CK > 2X ULN and > BL
ALB > ULN and > BL	CK > 5X ULN and > BL
	CK > 10X ULN and > BL
ALP > 2X ULN and > BL	
ALP > 5X ULN and > BL	MG < LLN and < BL
ALP > 10X ULN and > BL	MG > ULN and > BL
CA < LLN and < BL	PHOS < LLN and < BL
CA > ULN and > BL	PHOS > ULN and > BL

K < LLN and < BL	NA < LLN and < BL
K > ULN and > BL	NA > ULN and > BL
Hematology Laboratory	
HCT < LLN and < BL	NEUT < LLN and < BL
HCT decrease >= 10 from BL	NEUT > ULN and > BL
HGB <= 10.5 g/dL and < BL	BASO < LLN and < BL
HGB decrease > 1 g/dL from BL	BASO > ULN and > BL
MCV < LLN and < BL	EOS < LLN and < BL
MCV > ULN and > BL	EOS > ULN and > BL
	EOS > 5% and > BL
WBC < LLN and < BL	
WBC > ULN and > BL	PLT < LLN and < BL
	PLT < 100 X 10^9/L and < BL
LYM < LLN and < BL	PLT > ULN and > BL
LYM > ULN and > BL	
	HbA1c >= 6.5 and > BL
MONO < LLN and < BL	HbA1c increase > 1.0 from BL
MONO > ULN and > BL	

BL = baseline; LLN = lower limit of normal; ULN=upper limit of normal.