A PHASE 3, RANDOMIZED, EFFICACY AND SAFETY STUDY OF ENZALUTAMIDE IN COMBINATION WITH LEUPROLIDE, ENZALUTAMIDE MONOTHERAPY, AND LEUPROLIDE ALONE IN MEN WITH HIGH-RISK NONMETASTATIC PROSTATE CANCER PROGRESSING AFTER DEFINITIVE THERAPY

Statistical Analysis Plan
(SAP)

Version: 4.0

Date: 2-June-2023

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

1. VERSION HISTORY

This Statistical Analysis Plan (SAP) for study MDV3100-13 (C3431004) is based on the Protocol Amendment #5 version 6.0 dated 18-May-2023.

Table 1. Summary of Major Changes in SAP Amendments

SAP Version	Change	Rationale
1		Original
1.0*	Adjustment of endpoints, testing hierarchy, and sample size calculation	• Align with Protocol Amendment #2 • MFS between enzalutamide monotherapy versus placebo plus leuprolide was moved to key secondary endpoints. • Time to prostate-specific antigen (PSA) progression and time to first use of new antineoplastic therapy and time to distant metastasis were added as key secondary endpoints. • OS, proportion of patients per group who remain treatment-free 2 years after suspension of study drug treatment at week 37 due to undetectable PSA, and time to castration resistance were moved to other secondary endpoints. • Two endpoints (prostate cancer-specific survival and time to metastasis) were removed from other secondary endpoints; and one additional endpoint (time to symptomatic progression) was added as other secondary endpoint. • The associated testing hierarchy was included. • Overall sample size reduced from 1860 patients to 1050 in the 3 treatment arms. • Number of MFS events in the blinded arms reduced from 480 to 217, target HR changed from 0.76 to 0.65, and power increased to 90% from 85%.

Table 1. Summary of Major Changes in SAP Amendments

SAP Version	Change	Rationale
2		• An interim analysis for efficacy/futility at 70% of the required number of Metastasis-Free Survival (MFS) events was included. • Progression-free survival on first subsequent therapy (PFS2) was included as an exploratory endpoint. • Sensitivity analyses to account for informative censoring due to patient discontinuation prior to the development of radiographically detectable metastatic disease (ie, event of the primary end point) included. • Analysis of time to prostate-specific antigen (PSA) progression was revised to align with protocol. • Overall survival (OS) was moved to a key secondary endpoint. • Time to distant metastasis was moved to another secondary endpoint. • Interim analyses and final analysis of overall survival were included with the associated testing procedure.
3.0	Section 5.1 Hypotheses and Decision Rules	Based on the lower target HR of 0.58, the number of MFS events for the primary endpoint analysis was revised in protocol amendment #4 from 336 to 197 MFS events across all 3 arms, at which time approximately 142 rather than 237 MFS events are expected to occur in the 2 blinded treatment groups combined for the primary hypothesis.
	Section 5.1 Hypotheses and Decision Rules and Section 6.2.4 Overall Survival	Interim analysis for MFS efficacy/futility was removed in protocol amendment #4. Following removal of the interim analysis for MFS, the OS design was revised from 3-look to 2-look group sequential design in protocol amendment #4.

Table 1. Summary of Major Changes in SAP Amendments

SAP Version	Change	Rationale
	Section 5.1.1 Hypotheses and Sample Size Determination	Patient number for sample size determination as 1050 participants was clarified in protocol amendment #4. 1068 participants is the actual enrollment number.
	Section 6.9 Sensitivity Analyses for Potential COVID-19 Pandemic Related Issues	Sensitivity analyses for potential coronavirus disease of 2019 (COVID-19) pandemic-related issues were added.
	Section 5.2.4 Adequate Tumor Assessment	Clarification of adequate tumor assessment.
	Section 6.5.1 Progression-Free Suvival on First Subsequent Therapy (PFS2) (Table 9)	PFS2 censoring rule was clarified based on the data collection on PD2 (the first progressive disease after initiation of first subsequent antineoplastic therapy).
	Section 6.7.1 Baseline Summaries	Clarification on some baseline and disease characteristics
	Section 2.1 Study Objectives, Section 3.1 Primary Endpoint(s), Section 3.2 Secondary Endpoint(s) and Section 3.3 Exploraty Endpoint(s)	Change to tabulation format for objectives and endpoints to enhance clarity.
	Section 3.5 Safety Endpoints	Clarification of modified treatment period.
	Section 6.1.1.3 Senstivity/Robustness Analyses	MFS sensitivity 5 was specified for additional scenario(s) of patient discontinuation prior to the development of radiographically detectable metastatic disease.
	Section 6.4 Patient Reported Outcomes -Pain Section 3.2.2.9 and 3.2.2.10 PRO – Quality of Life	Patient reported outcome (PRO) analyses were specified to reflect a supplemental SAP specific to PRO data.
	Terms “anticancer therapy”	Clarification Change to terms of antineoplastic therapy.
	Section 6.8.1 Adverse Events (Table 10)	Clarification of study day cut points in subgroup and supplemental tabulations of TEAEs.

Table 1. Summary of Major Changes in SAP Amendments

SAP Version	Change	Rationale
	Section 6.8.4 Event of Special Interest and Appendix 2	Clarification of adverse events of special interest.
3.1	Section 5.2.7 Visit window for Table 4, Table 5	Clarification of visit window usage scopes
	Section 5.2.5 PSA doubling time calculation	Clarification of PSA doubling time calculation
	Section 5.2.5 Modified treatment duration	Clarification in definition of modified treatment duration
	Section 6.1.1.1 Table 7	Clarification of the censoring rules applied to MFS events by either radiographic progression or death
	Section 6.4.1	Clarification of time to clinically relevant pain progression
	Section 6.4.2	Clarification of time to deterioration in global FACT-P score
	Section 6.6 and 6.7.1 subgroup and baseline characteristics	Clarification of subgroup and baseline characteristics of prior radiation therapy. The prior prostatectomy and radiation therapy is added in baseline characteristics. The current surgical procedure is removed from the baseline disease characteristics.
	Appendix 2	Update of adverse events of special interests per MedDRA version update
4	Section 5.1.2 and 5.1.3	Clarification of the significant level for OS analysis based on the planned testing procedure and the actual results from tests for the primary endpoint and key secondary endpoints at the time of final analysis for primary endpoint Clarification of the updated results at the time of final analysis of OS from endpoints including time to first use of new antineoplastic therapy, time to first symptomatic skeletal event and PFS2 to be

Table 1. Summary of Major Changes in SAP Amendments

SAP Version	Change	Rationale
		summarized descriptively and provided as exploratory analyses
	Section 6.2.3, 6.3.8 and 6.5.1	Clarification of these associated endpoints to be summarized descriptively and provided as exploratory analyses at the time of final analysis of OS
	Section 6.2.4	Clarification of OS subgroup analyses Prespecification of the additional analyses to adjust for the potential confounding factor of post-study systemic anticancer therapy for the OS analysis if appropriate
	Section 6.8	Clarification of safety summary and evaluation to align with protocol amendment #5

[*] SAP (dated 25 October 2018) was Pfizer version 1.0 after study transition from Medivation.

2. INTRODUCTION

This SAP contains definitions of analysis populations and endpoints, outlines the timing of statistical analyses, and provides a comprehensive description of statistical analyses to be implemented to assess the clinical efficacy and safety of Protocol MDV3100-13 (C3431004): A Phase 3, Randomized, Efficacy and Safety Study of Enzalutamide in Combination With Leuprolide, Enzalutamide Monotherapy, and Leuprolide Alone in Men With High-Risk Nonmetastatic Prostate Cancer Progressing After Definitive Therapy. This document may modify the plans outlined in the protocol; however, any major modifications of the primary endpoint definition or its analysis will also be reflected in a protocol amendment.

The primary analysis will include all data up to a clinical cut-off date which is determined by the number of events achieved for Metastasis-Free Survival (MFS) by Blinded Independent Central Review (BICR). The cut-off date is determined once a data extract (before database lock) is available which indicates that the required number of events for MFS by BICR have occurred.

Due to data cleaning activities, the final number of events may deviate from the planned number. The data cut-off date will not be adjusted retrospectively in this case. All summaries and analyses will include all data pertaining to visits/assessments performed up to and including the data cutoff date.

Any deviations from this analysis plan will be described in the Clinical Study Report (CSR).

2.1. Study Objectives

Analysis	Objective	Endpoint
Primary Objectives and Endpoints		
Efficacy	• To evaluate efficacy as measured by MFS of enzalutamide (160 mg/day) in combination with leuprolide versus placebo in combination with leuprolide for the treatment of men with high-risk nonmetastatic prostate cancer that is progressing after radical prostatectomy or radiotherapy or both.	• MFS of enzalutamide (160 mg/day) in combination with leuprolide versus placebo plus leuprolide
Key Secondary Objectives and Endpoints		
Efficacy	To evaluate efficacy of enzalutamide as measured by alpha-protected key secondary endpoints with respect to the following: • MFS between enzalutide monotherapy versus placebo plus leuprolide;	• MFS for participants receiving enzalutamide monotherapy versus placebo plus leuprolide
	• PSA progression;	• Time to PSA progression
	• Time to first use of new antineoplastic therapy;	• Time to first use of new antineoplastic therapy
	• Overall survival (OS).	• OS
Other Secondary Objectives and Endpoints		
Efficacy	To evaluate efficacy of enzalutamide as measured by other secondary endpoints as follows: • Time to distant metastasis;	• Time to distant metastasis
	• Proportion of treatment-free patients at 2 years after treatment suspension;	• Proportion of patients per group who remain treatment-free 2 years after suspension of study treatment at week 37 due to undetectable PSA
	• Proportion of patients with undetectable PSA at 2 years after treatment suspension.	• Proportion of patients per group with undetectable PSA 2 years after suspension of study treatment at week 37 due to undetectable PSA

Analysis	Objective	Endpoint
	• Undetectable PSA at week 36	• Proportion of patients per group with undetectable PSA at 36 weeks on study drug
	• Time to resumption of any hormal therapy following suspension of treatment at week 37	• Time to resumption of any hormonal therapy following suspension at week 37 due to undetectable PSA
	• Time to castration resistance	• Time to castration resistance
	• Time to symptomatic progression	• Time to symptomatic progression
	• Time to first SSE	• Time to first symptomatic skeletal event
PRO	• Time to clinically relevant pain progression defined as a 2-point increase from baseline on the Brief Pain Inventory (Short form) question 3 • Time to a 10-point decline (deterioration) in global FACT-P score. • Pain by BPI-SF and QoL by FACT-P, EQ-5D-5L and QLQ-PR25 questionnaires	• Time to clinically relevant pain progression defined as a 2-point increase from baseline on the Brief Pain Inventory (Short form) question 3.Time to a 10-point decline (deterioration) in global FACT-P score. • Pain by BPI-SF and QoL by FACT-P, EQ-5D-5L and QLQ-PR25 questionnaires • Additional PRO endpoints can be found in a PRO-specific SAP
Safety	To evaluate safety during on-treatment period, and modified treatment period, or re-initiation treatment period as specified	• Adverse Events as characterized by type, frequency, severity, seriousness, and relationship to study drug • Laboratory and vital signs abnormalities as characterized by type, frequency and severity
Exploratory Objectives and Endpoints		
Efficacy	To evaluate efficacy of enzalutamide as measured by exploratory endpoint as follows: • Progression-free survival on first subsequent therapy (PFS2)	• PFS2

2.2. Study Design

This is an international, Phase 3, randomized study of enzalutamide plus leuprolide, enzalutamide monotherapy, and placebo plus leuprolide in approximately 1050 men with high-risk nonmetastatic prostate cancer progressing after radical prostatectomy or radiotherapy or both. Treatment with enzalutamide monotherapy will be open label. Treatment with enzalutamide and placebo will be double-blind in combination with open-label leuprolide.

No prior cytotoxic chemotherapy or androgen deprivation therapy (with exceptions) is allowed. The primary efficacy endpoint is MFS in the combination treatment arms. The

comparison of MFS between enzalutamide monotherapy versus placebo plus leuprolide is a key secondary endpoint.

Central randomization (1:1:1) will be used to assign patients to one of the following study treatments:

- Enzalutamide plus leuprolide;
- Enzalutamide monotherapy;
- Placebo plus leuprolide.

Randomization will be stratified by screening PSA (≤10 ng/mL versus >10 ng/mL), PSA doubling time (≤3 months versus >3 to ≤9 months), and prior hormonal therapy (yes versus no).

Study treatments will be administered as follows, depending on treatment assignment:

- Enzalutamide will be administered as four 40-mg soft gelatin capsules by mouth once daily (160 mg/day) with or without food;
- Placebo capsules, identical in appearance to enzalutamide capsules, will be administered in the same manner as enzalutamide;
- Leuprolide acetate (leuprorelin acetate), 22.5 mg will be given as a single intramuscular or subcutaneous injection once every 12 weeks (for a minimum of 3 doses, providing 36 weeks of treatment).
- Patients will continue with the assigned study treatment until radiographic progression confirmed by BICR, unacceptable toxicity, gross noncompliance, administration of prohibited concomitant therapy, death, or withdrawal of consent.

At week 37, study treatment will be suspended for patients whose PSA values are undetectable (<0.2 ng/mL) at week 36 as determined by the central laboratory; PSA and testosterone will be measured every 3 months thereafter by the central laboratory. As the PSA values will not be provided to study sites or patients, study sites will be notified whether a patient's PSA value meets the specified threshold to determine whether or not to continue or suspend study treatment at week 37. Study treatment may be suspended only once (at week 37) due to undetectable PSA and will be reinitiated if subsequent central laboratory PSA values increase to ≥2.0 ng/mL for patients with prior prostatectomy or ≥5.0 ng/mL for patients without prostatectomy. Study sites will be notified if the PSA value meets the threshold specified for reinitiation of study treatment. Patients with detectable PSA values at week 36 will continue treatment without suspension until permanent treatment discontinuation criteria are met.

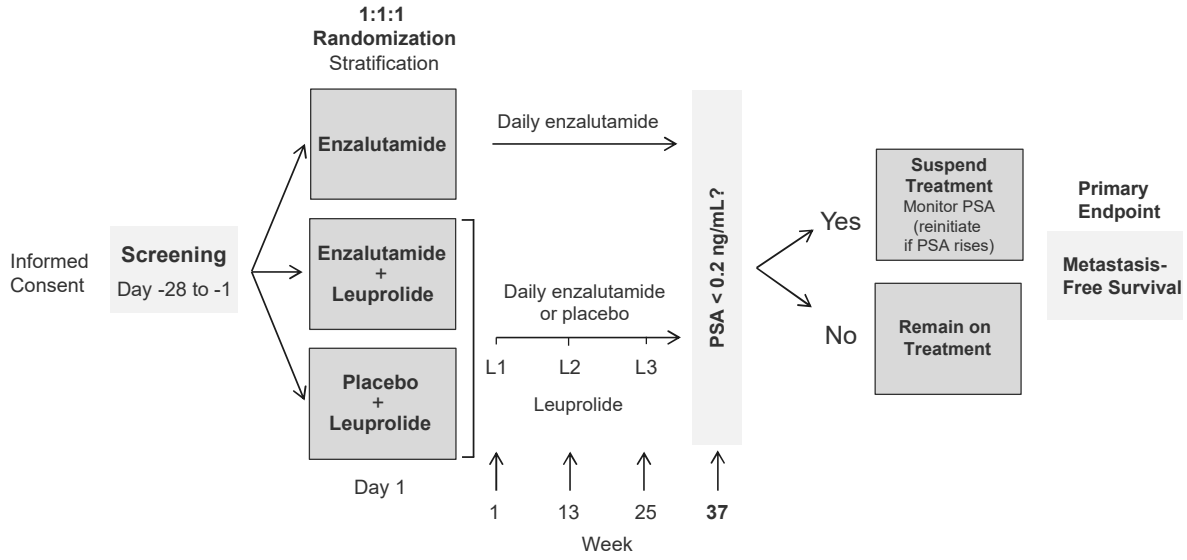
All patients will have a safety follow-up after permanent discontinuation of randomized study treatment. Safety follow-up should occur approximately 30 days after the last dose of randomized study treatment. Long-term follow-up will occur after safety follow-up (every

12 weeks based on the 12-week visit schedule determined at randomization) which may include imaging for patients who have not yet had confirmed radiographic progression.

2.3. Study Schematic

The study schematic is provided in Figure 1.

Figure 1. Study Schematic



Primary Assessment: Radiographic imaging approximately every 6 months

3. ENDPOINTS AND BASELINE VARIABLES: DEFINITIONS AND CONVENTIONS

3.1. Primary Endpoint(s)

Analysis	Endpoint	Additional notes
Primary Endpoint		
Efficacy	• The primary efficacy endpoint is MFS of enzalutamide plus leuprolide versus placebo plus leuprolide using BICR assessment of radiographic progression for soft tissue disease as defined by Response Evaluation Criteria in Solid Tumors (RECIST) v1.1 and for bone disease by the appearance of 1 or more metastatic lesions on bone scan.	See below

MFS is defined as the duration of time in months between randomization and the earliest objective evidence of radiographic progression by central imaging or death without radiographic progression, whichever occurs first. Refer to Section 6.1.1 for details on censoring rules.

3.2. Secondary Endpoint(s)

3.2.1. Key Secondary Endpoints

Analysis	Endpoint	Additional notes
Key Secondary Endpoints		

Efficacy	• MFS between enzalutamide monotherapy versus placebo plus leuprolide will be defined as above for the combination comparison.
	• The following 3 endpoints will be assessed as key secondary endpoints for both the enzalutamide plus leuprolide versus placebo plus leuprolide comparison as well as enzalutamide monotherapy versus placebo plus leuprolide comparison (see Section 5.1.3 for details on the testing strategy for key secondary endpoints).
	  • Time to PSA progression
	  • Time to first use of new antineoplastic therapy
	  • OS
		See Section 6.2

3.2.1.1. Metastasis-free survival between enzalutamide monotherapy versus placebo plus leuprolide

Metastasis-free survival between enzalutamide monotherapy versus placebo plus leuprolide will be defined as above in Section 3.1 for the combination comparison.

3.2.1.2. Time to PSA Progression

Time to PSA progression is defined as the time in months from randomization to the date of the first PSA value demonstrating progression, while patients are on study treatment, which is subsequently confirmed at least 3 weeks later.

For patients who do not have treatment suspended due to undetectable PSA at week 36: those with PSA decline at week 25, the PSA progression date is defined as the date that a ≥25% increase and an absolute increase of ≥2 µg/L (2 ng/mL) above the nadir (or baseline for patients with no PSA decline by week 25), which is subsequently confirmed at least 3 weeks later.

For patients who have treatment suspended due to undetectable PSA at week 36: following treatment suspension and reinitiation of study drug, the PSA progression date is defined as the date that a ≥25% increase and an absolute increase of ≥2 µg/L (2 ng/mL) above the nadir (re-baselined after week 36), which is confirmed by a second consecutive value obtained at a subsequent study visit at least 3 weeks later. PSA re-baseline value is defined as the last PSA assessment prior to or on the date of reinitiation. The nadir for these patients, is defined as minimum PSA value at the PSA re-baseline value visit and subsequent visits. Refer to Section 6.2.2 for details on censoring rules.

3.2.1.3. Time to First Use of New Antineoplastic Therapy

Time to first use of new antineoplastic therapy is defined as the time in months from randomization to first use of new antineoplastic therapy for prostate cancer. This will include medications used specifically for prostate cancer treatment including cytotoxic chemotherapy, abiraterone acetate, hormonal agents, prostate cancer vaccines, or nonradioactive bone-targeting agents and systemic radiopharmaceuticals for prostate cancer. Medications will be recorded on the 'Long Term Follow-up Antineoplastic Therapies' CRF page. Patients not starting treatment with a new antineoplastic therapy at the time of analysis will be censored at the date of last assessment before the analysis data cutoff date.

3.2.1.4. Overall Survival

Overall survival is defined as the time in months between randomization and death due to any cause. Refer to Section 6.2.4 for details on censoring rules.

3.2.2. Other Secondary Endpoints

Analysis	Endpoint	Additional notes
Other Secondary Endpoints		
Efficacy	The 2 comparisons (Enzalutamide in combination with leuprolide versus placebo in combination with leuprolide, Enzalutamide monotherapy versus placebo in combination with leuprolide) will be carried out for all other secondary endpoints: • Time to distant metastasis • Proportion of patients per group who remain treatment-free 2 years after suspension of study treatment at week 37 due to undetectable PSA • Proportion of patients per group with undetectable PSA 2 years after suspension of study treatment at week 37 due to undetectable PSA • Proportion of patients per group with undetectable PSA at 36 weeks on study drug • Time to resumption of any hormonal therapy following suspension at week 37 due to undetectable PSA • Time to castration resistance • Time to symptomatic progression • Time to first symptomatic skeletal event	See Section 6.3
PRO	• Time to clinically relevant pain progression defined as time in months from randomization to the time a 2-point increase from baseline on the BPI-SF question 3. • Time to functional status decline (deterioration) defined as the time in months from randomization to the time of a 10-point decline (deterioration) in global FACTP score from baseline. • Pain by BPI-SF and QoL by FACT-P, EQ-5D-5L and QLQ-PR25 questionnaires • Additional PRO endpoints can be found in a separate SAP specific to PROs	See Section 6.4

3.2.2.1. Time to distant metastasis

The time to distant metastasis is defined as the time in months from randomization to the earliest objective evidence of distant soft tissue metastases or metastatic bone disease by BICR. Soft tissue disease including lymph nodes above the aortic bifurcation and outside the pelvis and any bone metastases will be counted as distant metastases. Refer to Section 6.3.1 for details on censoring rules.

3.2.2.2. Proportion of patients who remain treatment-free 2 years after suspension of study treatment at week 37 due to undetectable PSA

The proportion of patients who remain treatment-free 2 years after suspension of study treatment at week 37 due to undetectable PSA is defined as number of patients remaining treatment-free 2 years after suspension of study treatment at week 37 due to undetectable PSA divided by the total number of patients with suspension of study drugs at week 37 due to undetectable PSA in each treatment group.

3.2.2.3. Proportion of patients per group with undetectable PSA 2 years after suspension of study treatments at week 37 due to undetectable PSA

The proportion of patients per group with undetectable PSA 2 years after suspension of study treatments at week 37 due to undetectable PSA is defined as number of patients with undetectable PSA 2 years after suspension of study treatments at week 37 due to undetectable PSA divided by the total number of patients with suspension of study drugs at week 37 due to undetectable PSA in each treatment groups.

3.2.2.4. Proportion of patients per group with undetectable PSA at 36 weeks on study drug

The proportion of patients per group with undetectable PSA at 36 weeks on study drug is defined asby the number of patients with undetectable PSA at 36 weeks divided by the total number of patients with PSA value at week 36 in each treatment group.

3.2.2.5. Time to resumption of any hormonal therapy following suspension at week 37 due to undetectable PSA

The time to resumption of any hormonal therapy following suspension at week 37 due to undetectable PSA is defined as the time in months between the date of treatment suspension at week 37 due to undetectable PSA and the date that hormonal therapy is restarted. Patients not starting hormonal therapy at the time of analysis will be censored at the date of last assessment before the analysis data cutoff date.

3.2.2.6. Time to castration resistance

Time to castration resistance applies only to patients receiving leuprolide treatment and is defined as the time in months from randomization to the first occurrence of radiographic disease progression by BICR, PSA progression (as defined above in Section 3.2.1) or symptomatic skeletal event whichever occurs first with castrate levels of testosterone (<50 ng/dL). The latest testosterone value measured prior to or at the date of radiographic disease progression by BICR, PSA progression or a symptomatic skeletal event is used to

determine if the event is a castration resistance event. Refer to Section 6.3.6 for details on censoring rules.

3.2.2.7. Time to symptomatic progression

Time to symptomatic progression is defined as the time in months from randomization to development of a skeletal-related event, worsening of disease-related symptoms requiring initiation of a new antineoplastic therapy, or development of adverse events and clinically significant signs and/or symptoms due to loco-regional tumor progression requiring opiate use, surgical intervention or radiation therapy, whichever occurs first. Refer to Section 6.3.7 for details on censoring rules.

3.2.2.8. Time to first symptomatic skeletal event

Time to first symptomatic skeletal event is defined as the time in months from randomization to use of radiation therapy (external beam radiation therapy or radionuclides) or surgery to bone for prostate cancer, findings of clinically apparent pathologic bone fracture or of spinal cord compression, or new use of opiate and/or systemic antineoplastic therapy due to bone pain collected in the SSE CRF, whichever occurs first. Because skeletal events are expected to occur after radiographic progression, the analysis of time to first symptomatic skeletal event will be performed with the final overall survival analysis. Refer to Section 6.3.8 for details on censoring rules.

3.2.2.9. Patient-Reported Outcomes (PRO)

The assessment of pain progression will be conducted using the BPI-SF (Cleeland 2006, Cleeland 2009) that has been subject to extensive validation processes. This questionnaire uses a self-reported scale assessing level of pain, its effect on activities of daily living, and analgesic medication use.

This study will use the short form containing 9 main questions related to pain and analgesic medication use. The primary question (paraphrased) is “On a scale of 0 to 10, please rate your pain at its worst in the last 24 hours.”

Time to clinically relevant pain progression is defined as the time in months from randomization to onset of pain progression, where pain progression is defined as a 2-point or more increase from baseline in the question 3 score.
Additional pain analyses with further details about questionnaires and associated derived variables can be found in a separate SAP specific to PROs.

Patient-Reported Outcomes (PRO) - Quality of Life Quality of life will be assessed by FACT-PEQ5D-5L and QLQ-PR25 questionnaires.

The FACT-P (Cella et al. 1993, Esper et al. 1997, 1999) is a multidimensional, self-reported, quality of life instrument specifically designed for use with prostate cancer patients. It consists of 27 core items that assess patient function in 4 domains: Physical, Social/Famil y, Emotional, and Functional wellbeing, which is further supplemented by 12 site-specific items to assess prostate cancer related symptoms. Each item is rated on a 0 to 4 Likert-type scale,

and then combined to produce subscale scores for each domain, as well as a global quality of life score with higher scores representing better quality of life.

Time to functional status deterioration will be assessed and is defined as the time in months from randomization to the time of a 10-point decline (deterioration) in global FACT-P score. The FACT-P domain scores and global score will be calculated using the Manual of Functional Assessment of Chronic Illness Therapy (FACIT) Measurement System (Cella, 1997).

The EQ-5D-5L questionnaire (Herdman et al. 2011, Janssen et al. 2013) is a standardized instrument for use as a measure of health-related quality and general health status. With the EQ-5D-5L questionnaire, patients will self-rate their current state of mobility, selfcare, , pain/discomfort, and anxiety/depression on a 5point categorical scale ranging from "no problem" to "extreme problem" within each dimension and EQ-VAS.

The European Organization for Research and Treatment of Cancer (EORTC) QLQ-PR25 (Van Andel et al. 2008) , a module of the EORTC QLQ30 questionnaire, was developed to assess the quality of life of patients with prostate cancer. Patients will self-rate their current state of pain as it relates to urination, ease and frequency of urination, and bowel and other problems during the past week. Patients will also answer 5 questions about weight loss/gain and sexual interest and 4 questions about sexual activity during the past 4 weeks. Patients will choose 1 of 4 possible responses that record level of intensity (not at all, a little, quite a bit, very much) within each dimension.

Additional QoL analyses with further details about questionnaires and associated derived variables can be found in a separate SAP specific to PROs.

3.3. Exploratory Endpoint(s)

Analysis	Endpoint	Additional notes
Exploratory Endpoints		
Efficacy	• PFS2 will be assessed for the 2 comparisons (Enzalutamide in combination with leuprolide versus placebo in combination with leuprolide, Enzalutamide monotherapy versus placebo in combination with leuprolide).	See Section 3.3.1

3.3.1. Progression-free survival on first subsequent therapy

Progression-free survival on first subsequent therapy (PFS2) is defined as the time in months from date of randomization to date of investigator-determined disease progression (PSA progression, progression on imaging, or clinical progression) or death due to any cause, whichever occurs first, while the patient was receiving first subsequent therapy for prostate cancer. For censoring rules see Section 6.5.

3.4. Baseline Variables

3.4.1. Study Drug, Study Treatment, and Baseline Definitions

The date of first dose (start date) of study treatment in the combination arms is the earliest date of non-zero dosing for each of the study drugs (enzalutamide/placebo or leuprolide).

The date of first dose (start date) of study treatment in the enzalutamide monotherapy arm is the earliest date of non-zero dosing of enzalutamide.

The date of last dose of study treatment in the combination arms is the latest date of non-zero dosing for each of the study drugs.

The date of last dose of study treatment in the enzalutamide monotherapy arm is the latest date of non-zero dosing of enzalutamide.

Definition of Baseline:

No windowing will be applied when defining baseline. For example, the protocol requires safety assessments to be performed within 28 days prior to randomization; however, values outside this window will not be excluded when determining baseline assessments. Any deviations from the protocol specified window will be documented as protocol deviations.

Definition of baseline for efficacy and baseline analyses

The last measurement prior to randomization will serve as the baseline measurement for efficacy and baseline analyses.

Definition of baseline for PRO

Per protocol, the first PRO assessment is planned to occur prior to dosing on Day 1; therefore, the last measurement on or prior to the day of first dose of study treatment (on or prior to randomization for patients randomized but not dosed) will be used as the baseline measurement.

Definition of baseline for safety analyses

The last available assessment on or prior to the date of first dose of study treatment (or prior to randomization for patients randomized but not dosed) is defined as the 'baseline' value or 'baseline' assessment for safety analyses. If an assessment is planned to be performed prior to the first dose of study treatment in the protocol and the assessment is performed on the same day as the first dose of study treatment, it will be assumed that it was performed prior to study treatment administration, if assessment time is not collected or is missing. If assessment times are collected, the observed time point will be used to determine pre-dose on study Day 1 for baseline calculation (this only applies for safety). Unscheduled assessments will be used in the determination of baseline.

3.4.2. Stratification

Randomization is stratified by the following, as recorded in the Interactive Response Technology (IRT):

- Screening PSA ≤10 ng/mL versus >10 ng/mL,
- PSA doubling time ≤3 months versus >3 to ≤9 months,
- Prior hormonal therapy versus no prior hormonal therapy

Unless otherwise specified, stratified analyses will utilize strata as defined in the randomization system.

3.5. Safety Endpoints

Safety will be assessed through adverse event (AE) reporting, laboratory and vital sign assessments. AEs will be graded by the investigator according to the Common Terminology Criteria for Adverse Events (CTCAE) version 4.03 and coded using the Medical Dictionary for Regulatory Activities (MedDRA) version in use at the time of database release.

Safety endpoints will be summarized based on the on-treatment period, modified treatment period, or re-initiation treatment period as specified.

On-Treatment Period

The on-treatment period is defined as the time from the date of first dose of study treatment through a minimum of 30 days after last dose of study treatment, or the start day of new antineoplastic drug therapy – 1 day, where the last dose of study treatment is the final dose after any suspension period due to undetectable PSA. Adverse events occurring on the same day as the first dose of study treatment will be considered to have occurred during the on-treatment period. All other assessments which occur on the same day as the first dose of study treatment will be considered baseline assessments (see Section 3.5).

Safety data collected outside the on-treatment period as described above will be listed but not summarized.

Modified Treatment Period

For patients whose treatment is suspended due to undetectable PSA at week 36 the modified period is defined as the period of time of study drug exposure starting from the date of the first dose of study drug through 30 days after last dose prior to the treatment suspension PLUS the time period starting from date of first dose at study drug reinitiation through a minimum of 30 days after last dose of study treatment, or the start day of new antineoplastic drug therapy – 1 day. If date of first dose at study drug reinitiation is earlier than 30 days after last dose prior to the treatment suspension, then the modified treatment period is the same as that on-treatment period.

For all other patients the modified treatment period is the same as the on-treatment period.

Reinitiation Treatment Period

The reinitiation treatment period is defined only for patients who suspend treatment due to undetectable PSA at week 36 and subsequently reinitiate treatment and is defined as the time

period starting from date of first dose at study drug reinitiation through minimum 30 days after last dose of study treatment, or start day of new antineoplastic drug therapy – 1 day.

Treatment-Emergent Adverse Events

Treatment-emergent adverse events (TEAEs) are those events with onset dates occurring during the on-treatment period for the first time

Modified Treatment-Emergent Adverse Events

Modified treatment-emergent adverse events (mTEAEs) are adverse events occurred during the modified treatment period (ie, events occurring or worsening during the treatment suspension period after 30 days of last dose prior to treatment suspension (and prior to treatment suspension if applicable) are excluded).

Reinitiation Treatment-Emergent Adverse Events

Reinitiation treatment-emergent adverse events (rTEAEs) are adverse events occurred for the first time during the dosing period after suspension and reinitiation of study treatment as defined by the reinitiation treatment period.

3.5.1. Adverse Events of Special Interest

The TEAEs of special interests will be defined based on a list of MedDRA Preferred Terms provided by Pfizer Pharmacolvigillance. A final list will be provided to programming prior to database release.

3.5.2. Labs

Hematology and chemistry result will be programmatically graded according to the National Cancer Institute's (NCI) CTCAE version 4.03 for relevant parameters.

4. ANALYSIS SETS

Data for all subjects will be assessed to determine if participants meet the criteria for inclusion in each analysis population prior to unblinding and releasing the database and classifications will be documented per standard operating procedures.

4.1. Intent-to-Treat

The intent-to-treat (ITT) population is defined as all patients randomly assigned to study treatment. The ITT population (or a subset thereof) will be used for all efficacy analyses and PRO analyses unless otherwise specified, and will be analyzed based on randomized treatment assignment, regardless of whether or not study treatment was administered.

4.2. Safety Analysis Set

The safety analysis set will include all randomized patients who receive at least 1 dose of any study drug. Unless otherwise specified, all safety analyses will use the safety population according to the actual treatment received (not the treatment assigned).

4.3. Evaluable Intent-to-Treat (eITT) Set

The evaluable ITT (eITT) set is defined as all ITT patients who have confirmed nonmetastatic disease at baseline by independent central radiology review. The eITT population will be used for certain efficacy analyses.

5. GENERAL METHODOLOGY AND CONVENTIONS

5.1. Hypotheses and Decision Rules

5.1.1. Hypotheses and Sample Size Determination

The study is designed to test the following primary hypothesis:

$$H_{01}: \frac{\lambda_{TrtA}}{\lambda_{TrtC}} = 1, \text{ versus } H_{a1}: \frac{\lambda_{TrtA}}{\lambda_{TrtC}} \neq 1, \text{ where}$$

Trt A: Enzalutamide plus leuprolide;

Trt C: Placebo plus leuprolide.

The sample size and event-driven readout was chosen to adequately power the testing of the primary endpoint of MFS.

The comparison of MFS for enzalutamide monotherapy versus placebo in combination with leuprolide is a key secondary endpoint.

The following assumptions were used in determining the sample size calculation for the MFS endpoint:

- Overall 2-sided Type I error rate: 0.05
- Patient accrual rate: An adaptive enrollment rate has been modified into 2 arms to fit into EAST v6.4 software from Cytel
- Randomization: 1:1:1
- Median MFS for the control group: 55 months

An observed 142 MFS events in the 2 blinded treatment groups will provide approximately 90% power to detect a target hazard ratio of 0.58 using a 2-sided log-rank test with a 0.05 level of significance. This target hazard ratio corresponds to a difference of approximately 40 months in median MFS assuming an exponential distribution for MFS and a constant hazard rate for each group. For the key secondary hypothesis of MFS for the monotherapy arm, the target effect size, and expected number of MFS events will be the same as the primary hypothesis in the combination arm. As a 2-sided alpha of 0.03 will be utilized for the monotherapy comparison, the power for this analysis will be 86% with 142 MFS events observed. At the time of the final analysis, at least 197 MFS events total are expected for the 3 treatment groups. The study will require approximately 1050 patients (350 in each group) to achieve the 197 MFS events across the 3 treatment arms. This sample

size calculation accounts for a 5% loss to follow-up by the end of 4 years for all 3 treatment groups.

An actual enrollment of 1068 patients will also allow for an assessment for the key secondary endpoint of OS.

The significance level associated with the test of OS, in blinded treatment arm and monotherapy arm comparisons, will depend on the outcome of the key secondary endpoints (see Figure 2). If all key secondary endpoints for the blinded treatment arms are statistically significant then an α of 0.02 will be contributed to the comparison of OS according to the gatekeeping procedure. Similarly, if all key secondary endpoints for the monotherapy treatment arm are statistically significant then an α of 0.03 will be contributed to the comparison of OS. With 191 deaths, the power to detect a hazard ratio of 0.67 is 79.0% (if all key secondary endpoints are significant for both the blinded arm and monotherapy arm comparisons), 72.3% (if the key secondary endpoints are significant for the monotherapy arm only) or 67.0% (if the key secondary endpoints are significant for the blinded arms only) using a 2-sided log-rank test at a significance level of 0.05, 0.03 or 0.02, respectively, and a 2-look group sequential design with a Haybittle-Peto efficacy boundary (Haybittle et al. 1971, Peto et al. 1976).

5.1.2. Decision Rules

There will be no interim analysis for the primary endpoint (MFS). The interim and final analyses for the key secondary endpoint OS will be performed after the target number of events have occurred in the 3 treatment arms. A maximum of 2 distinct analysis cut-offs are planned according to the numbers of events described below:

- Final MFS and OS interim analyses at the time when 197 MFS events have occurred for the 3 treatment groups;
- Final OS analysis at the time when 271 deaths have occurred for the 3 treatment groups.

Table 2 shows the efficacy boundary associated with the final analysis of the primary endpoint of MFS between enzalutamide plus leuprolide and placebo plus leuprolide, according to independent central radiology review.

Table 2. MFS Based on Independent Review (2 Blinded Treatment Arms) - Efficacy Boundary

Analysis	Analysis Cut-Off Trigger	Number of MFS Events (2 Blinded Arms)[a]	Fraction of Required MFS Events	p-value (z-value) for Efficacy
Final MFS	197 MFS events in 3 arms	142	100%	≤0.05 (-1.9600)

MFS = metastasis-free survival

[a] Number of events expected for MFS in blinded treatment arms assuming a hazard ratio of 0.58.

Table 3 shows the analysis triggers for OS as well the associated Haybittle-Peto efficacy boundaries. The significance level for the analyses of this key secondary endpoint is determined by the gatekeeping procedure.

Table 3. OS (2 Blinded Treatment Arms) - Efficacy Boundaries

Analysis	Analysis Cut-Off Trigger	Number of OS Events (2 Blinded Arms)[a]	Fraction of Required OS Events	p-value (z-value) for Efficacy[b]
IA OS	197 MFS events in 3 arms	82	43%	≤0.0001 (-3.89059)
Final OS	271 OS events in 3 arms	191	100%	≤0.04999 (-1.96001)

OS = overall survival; IA = interim analysis; MFS metastatic-free survival.

a. Number of events expected for OS in blinded treatment arms assuming a hazard ratio of 0.67.

b. The p-values and z-values noted for OS are those associated with the scenario where all the key secondary endpoints for the blinded treatment arms and the monotherapy arm are statistically significant (α=0.05).

At the time of final analysis for MFS, a statistically significant difference for MFS between the enzalutamide plus leuprolide and placebo plus leuprolide arms (in favor of enzalutamide plus leuprolide) would trigger subsequent analyses of the remainder of the endpoints in Figure 2, including the key secondary endpoint of OS.

At the time of the final analysis for MFS, OS data will not be mature. The final analysis of overall survival will be performed at the time of the specified target numbers of deaths have occurred for the 3 treatment groups in Table 3. Long-term follow-up data (survival status and skeletal related events and new prostate cancer therapies, and progression on first subsequent therapy) will be collected every 12 weeks until the final OS analysis.

Based on the planned testing procedure in Section 5.1.3 and the actual results from tests for the primary endpoint and key secondary endpoints at the time of final analysis for primary endpoint, the analysis of OS for the combination arm will be tested at the alpha level of 0.05. The analysis of OS for the monotherapy comparison will be tested hierarchically as planned at the same alpha level as the combination arm comparison if significant (Figure 2 and Table 3).

In the final OS analysis, the significance level will be adjusted for the interim OS analysis following the specified Haybittle-Peto efficacy boundaries (Table 3). The analysis of OS in the aforementioned final analyses will be based on an ITT approach.

At the time of final analysis of OS, updated results from following secondary endpoints including time to first use of new antineoplastic therapy, time to first symptomatic skeletal event and progression free survival on first subsequent therapy will be summarized

descriptively and provided as exploratory analysis. No other efficacy endpoints (only OS) will be formally tested.

5.1.3. Multiplicity Adjustment for Efficacy Analysis

Alpha protected efficacy analyses will include tests for the primary endpoint of MFS for enzalutamide plus leuprolide versus placebo plus leuprolide, and all 3 key secondary efficacy endpoints (time to PSA progression, time to first antineoplastic therapy, and overall survival) for the combination comparisons. Additionally MFS, time to PSA progression, time to first antineoplastic therapy, and overall survival will be tested for enzalutamide monotherapy versus placebo plus leuprolide.

If the test for the primary endpoint (MFS in the combination arms) is significant at the full 2-sided alpha level of 0.05, the key secondary endpoints for the combination arms will be tested at a 2-sided alpha of 0.02 utilizing a hierarchical approach to preserve the family-wise Type I error rate. The remaining 0.03 alpha will be allocated to compare MFS as well as other key secondary endpoints for enzalutamide monotherapy versus placebo plus leuprolide. The efficacy analyses and the multiplicity adjustment rules are summarized in Figure 2.

The details of the approach and testing flowchart are as follows:

Step 1. Test MFS for enzalutamide plus leuprolide versus placebo plus leuprolide: Compute 2-sided p-value for MFS. If and only if p-value <0.05, declare statistical significance for MFS for the enzalutamide plus leuprolide versus placebo plus leuprolide treatment arm, then go to step 2 and step 3 below (as illustrated in testing flowchart Figure 2, the testing path is split into two branches). Otherwise stop.

Step 2. Testing of the 2 key secondary endpoints for enzalutamide plus leuprolide versus placebo plus leuprolide: Allocate a 0.02 alpha for testing the 3 key secondary endpoints using a hierarchical approach to preserve the family-wise Type I error rate. Proceed as follows:

If time to PSA progression is significant at the 2-sided alpha level of 0.02, declare statistical significance for this endpoint. Then the alpha of 0.02 is carried forward to test for the next endpoint of the time to first use of antineoplastic therapy. Otherwise stop.

If the time to first use of antineoplastic therapy endpoint is significant at the 2-sided alpha level of 0.02, declare statistical significance for this endpoint, then the alpha is carried forward to test the third endpoint of overall survival. Otherwise stop.

Step 3. Test MFS for enzalutamide monotherapy versus placebo plus leuprolide: Compute the 2-sided p-value for MFS. If and only if p-value <0.03, declare statistical significance for MFS for this treatment arm, then go to step 4 below. Otherwise stop.

Step 4. Testing of the 2 key secondary endpoints for enzalutamide monotherapy versus placebo plus leuprolide:

If time to PSA progression is significant at the 2-sided alpha level of 0.03, declare statistical significance for this endpoint. Then the alpha of 0.03 is carried forward to test for the next endpoint of the time to first use of antineoplastic therapy. Otherwise stop.

If the time to first use of antineoplastic therapy endpoint is significant at the 2-sided alpha level of 0.03, declare statistical significance for this endpoint, then the alpha level is carried forward to test the third endpoint of overall survival. Otherwise stop.

Step 5. Testing of overall survival for enzalutamide plus leuprolide versus placebo plus leuprolide

If all key secondary comparisons for the combination arm comparisons are significant at the alpha level of 0.02, and for the monotherapy comparisons at the alpha level of 0.03, then OS for the combination arm will be tested at the alpha level of 0.05. Otherwise, only the alpha from the significant set of key secondary comparisons will be carried forward.

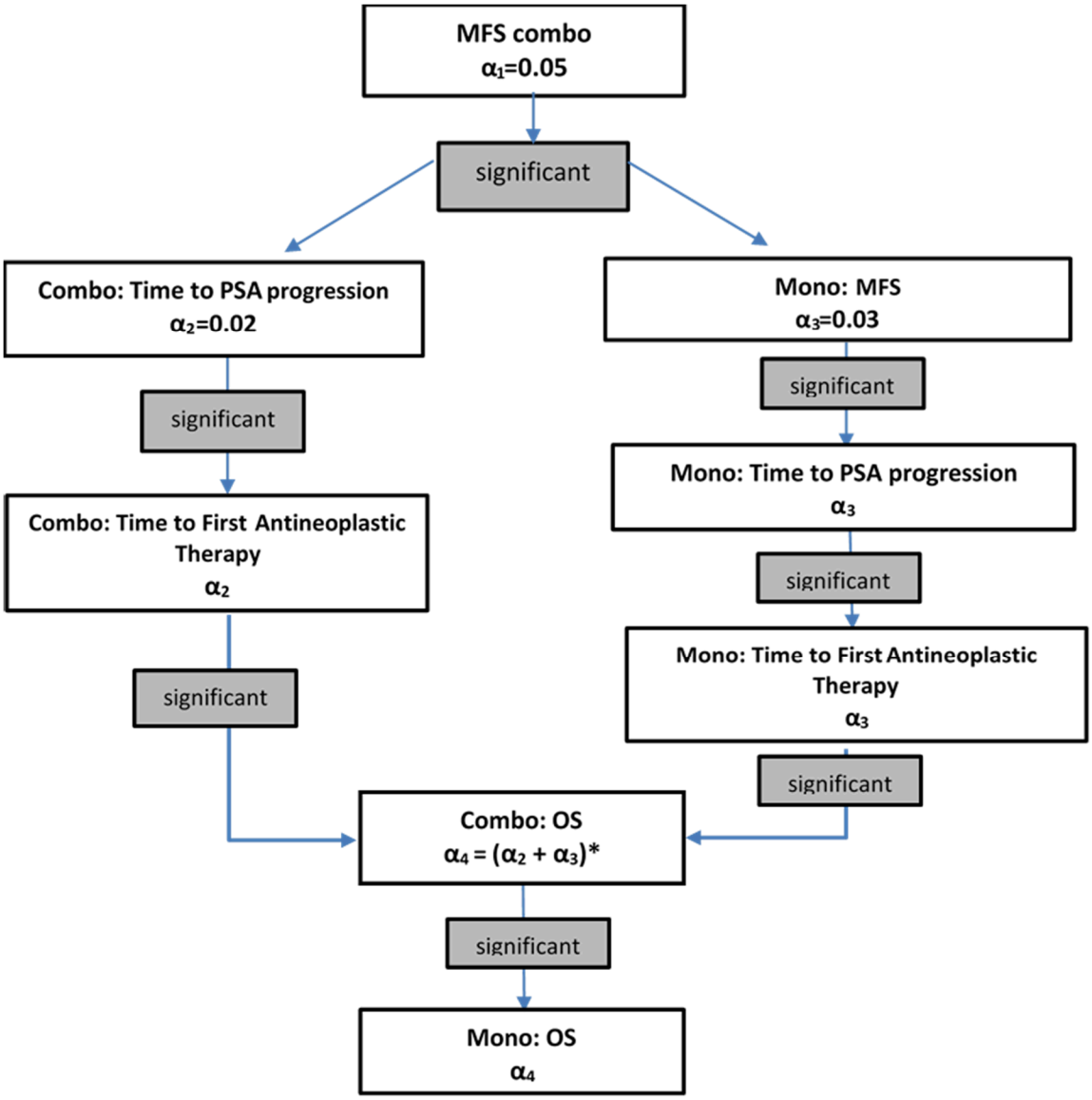
Testing of overall survival for enzalutamide monotherapy versus placebo plus leuprolide

- OS for the monotherapy comparison will be tested in a hierarchical fashion at the same alpha level as the combination arm comparison if significant.

All other efficacy analyses and associated p-values will be deemed exploratory, for which no adjustment for multiplicity will be made.

At the time of final analysis of OS, no other efficacy endpoints (only OS) will be formally tested.

Figure 2. Key Efficacy Analyses and Multiplicity Adjustment



* $\alpha_4 = \alpha_2 + \alpha_3 = 0.05$ if all Combo and Mono comparisons are significant
$= \alpha_2 = 0.02$ if Combo comparisons are significant and Mono are not
$= \alpha_3 = 0.03$ if Mono comparisons are significant and Combo are not

5.2. General Methods

5.2.1. Pooling of Centers

In order to provide overall estimates of treatment effects, data will be pooled across centers. The 'center' factor will not be considered in statistical models or for subgroup analyses due to the high number of participating centers in contrast to the anticipated small number of patients randomized at each center.

5.2.2. Presentation of Continuous and Qualitative Variables

Continuous variables will be summarized using descriptive statistics ie, number of non-missing values and number of missing values [ie, n (missing)], mean, median, standard deviation (SD), minimum, maximum and first and third quartile (Q1 and Q3).

Qualitative variables will be summarized by frequency counts and percentages. Unless otherwise specified, the calculation of proportions will include the missing category. Therefore counts of missing observations will be included in the denominator and presented as a separate category.

In case the analysis refers only to certain visits, percentages will be based on the number of patients still present in the study at that visit, unless otherwise specified.

5.2.3. Date of Last Contact

The date of last contact will be derived for patients not known to have died at the data cutoff date using the latest complete date (ie, imputed dates will not be used in the derivation) among the following:

- All patient assessment dates (eg, laboratory blood draws), vital signs, physical exam, performance status, (tumor assessments);
- Start and stop dates of concomitant therapies including non-drug treatments or procedures;
- Completion dates for PRO Questionnaires;
- Start and end dates of antineoplastic therapies administered after study treatment discontinuation including systemic therapy, radiation, and surgeries;
- AE start and end dates;
- Last date of contact collected on the 'Survival Follow-up' Case Report Form (CRF) (do not use date of survival follow-up assessment unless status is 'alive');
- Study treatment start and end dates;
- Randomization date; and

- Date of discontinuation on disposition CRF pages (do not use if reason for discontinuation is lost to follow-up or death).

Only dates associated with actual examinations of the patient will be used in the derivation. Dates associated with a technical operation unrelated to patient status such as the date a blood sample was processed or dates data were entered into the CRF will not be used. Assessment dates after the data cutoff date will not be applied to derive the last contact date.

5.2.4. Adequate Tumor Assessment

For soft tissues assessments by BICR, no assessments of CR, PR, or stable disease will be made given that the patient population is non-metastatic at study entry. The assessment options will be no disease (ND), no progression (NN), progressive disease (PD), or not evaluable (NE).

For bone assessment by BICR the assessment options will be No Bone Metastases (NBM), Bone metastases - unconfirmed (BPDu), Bone metastases (BPD), or NE.

For analyses based on BICR, a tumor assessment is defined as adequate when both bone and soft tissue assessments are evaluable (there is no NE in either bone or soft tissue) and the difference between bone and soft assessments dates is less than 25 weeks (of note, radiographic assessments are at 24 week ± 5 days intervals in study schedule of activities). If either bone or soft tissue assessment is NE, the assessment would be considered as "inadequate".

For analyses based on investigator assessments, a tumor assessment is defined as adequate if both a bone scan and a soft tissue assessment are performed and the difference between bone and soft assessments dates is less than 25 weeks.

However, an assessment indicating PD from either a bone scan or a soft tissue assessment will be considered as adequate for declaring PD for analyses based on BICR and investigator assessment.

5.2.5. Definitions and Computations

Study Day

The study day for assessments occurring on or after the first dose of study treatment (eg, adverse event onset, tumor measurement) will be calculated as:

Study day = Date of the assessment/event - start date of study treatment/randomization date + 1.

The study day for assessments occurring prior to the first dose of study treatment (eg, baseline characteristics, medical history) will be negative and calculated as:

Study day = Date of the assessment/event – start date of study treatment.

The study day will be displayed in all relevant data listings.

Treatment Duration

The following durations will be calculated (of note, the data cutoff date is considered in calculation when last dose date and antineoplastic therapy date are not available):

Treatment Duration = Date of last dose of study drug – Date of first dose of study drug + 1

Treatment Suspension Duration (only for patients who suspend due to undetectable PSA at week 36) = Last Date of suspension without drug (ie, First dose after suspension-1 or Date of data cutoff, whichever occurs first) -First date of suspension +1. Treatment Suspension Duration will be missing for patients who do no suspend treatment due to undetectable PSA at week 36.

Modified Treatment Duration = Treatment duration – Treatment Suspension Duration + 1 for patients who suspend due to undetectable PSA at week 36. The Modified Treatment Duration will equal the Treatment duration for patients who do not suspend treatment due to undetectable PSA at week 36.

Treatment Duration After Reinitiation (only for patients who suspend due to undetectable PSA at week 36 and subsequently reinitiate treatment) = Date of last dose of study drug – First dose after suspension + 1. Treatment Duration after Reinitiation will be missing for patients who do no suspend treatment due to undetectable PSA at week 36 as well as patients who suspend but do not reinitiate treatment.

Long-Term Follow-up Period

Long-term follow up begins after safety follow-up. Visits repeat every 12 weeks (±7 days) based on the 12-week visit schedule determined at randomization.

PSA Doubling Time Calculation

Patients' PSA doubling time will be calculated by the sponsor prior to randomization. All PSA values used for calculation will be examined during randomization stage and captured on the CRF page. To calculate PSA doubling time:

a. All PSA values used in the calculation should be ≥0.20 ng/ml and follow a rising trend;

b. All PSA values obtained over a maximum period of 12 months should be included in the calculation;

c. Minimum requirements for the calculation are 3 PSA values obtained over 3 months with a minimum of 4 weeks between measurements;

d. PSA values from the same local laboratory are preferred;

e. All PSAs should be obtained following primary treatment for prostate cancer.

PSA doubling time calculation will be performed by using the method of Pound et al (Pound, 1999). A linear regression model will be used with log-transformed PSA values as the response and the duration of time from the first reference PSA value as the explanatory

variable. The slope from the fitted regression line will be used to calculate PSA doubling time in months as follows:

$$\text{PSA Doubling Time (months)} = \frac{\text{LN}(2)*12}{(365.25*\text{Slope})}.$$

Of note, LN stands for natural logarithm function.

5.2.6. Conventions

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

- 1 year = 365.25 days. Year is calculated as (days/365.25) rounded up to 1 significant digit;
- 1 month = 30.4375 days. Month is calculated as (days/30.4375) rounded up to 1 significant digit;
- Age will be calculated as the integer part of (date of randomization – date of birth + 1)/365.25;
- 1 pound = 0.454 kg;
- 1 inch = 2.54 cm;
- Time-to-event (TTE) or duration of event endpoints will be based on the actual date rather than the associated visit date;
- Missing efficacy or safety data will not be imputed unless otherwise specified;
- For laboratory results collected as < or > a numeric value, 0.0000000001 will be subtracted or added, respectively, to the value;
- For safety analyses, percentages will be calculated based on the number of patients in the analysis population in each treatment group;
- For by-visit observed data analyses, percentages will be calculated based on the number of patients with nonmissing data as the denominator unless otherwise specified;
- For TTE data, the summary statistics and descriptions will include Kaplan-Meier plots and/or life tables;
- For other continuous endpoints, the summary statistics will include mean, standard deviation, median, and range (minimum and maximum);
- For categorical endpoints, the summary statistics will include counts and percentages;
- Medical history will be coded using the MedDRA version in use at the time of database lock;

- Prior therapies and concomitant medications will be coded using the World Health Organization (WHO) Drug Dictionary.

- The terms "patients", "participant", and "subjects" may appear in this document and are considered interchangeable."

5.2.7. 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 safety and efficacy assessments will be defined as shown in Table 4 and Table 5. Visit windows for efficacy assessments are created ±14 from the target day to capture the efficacy endpoints precisely.

Table 4. Visit Windows for Safety Assessments

Week	Start Day	Target Day	End Day
13	2	85	126
25	127	169	210
37	211	253	294
49	295	337	378
49+ Every 12 Weeks	(Week-1) × 7 – 41	(Week-1) × 7 + 1	(Week-1) × 7 + 42

Note: At week 41 (±3 days), there will be a telephone contact to all patients to collect adverse event information only.
Note: For PRO assessments, visit windows are defined in Table 4, because the PRO schedules is the same as safety assessments.

Table 5. Visit Windows for Efficacy Assessments

Week	Start Day	Target Day	End Day
25	155	169	183
49	323	337	351
49+ Every 24 Weeks	(Week-1) × 7 – 13	(Week-1) × 7 + 1	(Week-1) × 7 + 15

Note: the visit window is applied to the analyses by visit, such as summary tables by nominal visit, but not to the time to event endpoints analyses.

If more than 1 assessment is within a given visit window, the assessment closest to the target date will be used. If 2 assessments are equally close to the target day, the earlier assessment will be used. If there are more than 1 assessment on the same date, the assessment with the earliest collection time will be used.

At week 36 (day 246±5), there will only be PSA collection. Adjusted visit windows for PSA collection at weeks 36 and 37 will be (days 241, 246, 249) and (days 250, 253, 258) respectively.

For PRO assessments, the visit windows will be defined as shown in Table 4

5.2.8. Unscheduled Assessments

Unless otherwise specified, unscheduled assessments will not be displayed in summary tables by nominal visit/timepoint. Unscheduled assessments will be used when deriving baseline and worst case on-treatment for safety and PRO analyses. Additionally, unscheduled assessments will be used for efficacy analyses (eg, defining date of progression/censoring, date of last contact).

5.2.9. Analyses for Continuous Data

Descriptive statistics, including the number of observations (N), mean, standard deviation, median, minimum, and maximum values, will be provided for continuous variables.

In case the analysis refers only to certain visits, percentages will be based on the number of patients with a non-missing assessment at that visit, unless otherwise specified.

5.2.10. Analyses for Longitudinal Data

A mixed-model repeated measures (MMRM) approach will be used to analyze continuous data collected over time, mainly PRO data. Further details on methodology is available in a supplemental SAP specific to PRO data.

5.2.11. Analyses for Categorical Data

The number and percentage of patients in each category will be provided for categorical variables. Unless otherwise specified, the calculation of proportions will include the missing category. Therefore counts of missing observations will be included in the denominator and presented as a separate category.

In case the analysis refers only to certain visits, percentages will be based on the number of patients with an assessment at that visit, unless otherwise specified.

Some analyses will report the odds ratio and its 95% Confidence Interval (CI). The CI will first be constructed on the log-scale using the approximate variance of log (odds ratio) as (1/a) + (1/b) + (1/c) + (1/d) where a, b, c, d are the cells from a 2-by-2 table. Afterwards the log (odds ratio) and its CI will be exponentiated back to original scale.

5.2.12. Analyses for Time-to-Event Data

A 2-sided stratified log-rank test will be used for comparing treatments unless otherwise specified. Hazard ratios and the associated 95% 2-sided confidence intervals are estimated by stratified Cox proportional hazards models, including treatment as a covariate. Stratified analyses will utilize strata as defined in the randomization system unless otherwise specified. Unstratified analyses may be used in sensitivity analyses, subgroup analyses and supportive analyses, if applicable.

TTE endpoints will be summarized using the Kaplan-Meier method and estimated survival curves will be displayed graphically when appropriate. Graphs will describe the number of patients at risk over time. The median, quartiles, and probabilities of an event at particular points in time will be estimated by the Kaplan-Meier method. Confidence intervals for

medians and quartiles are based on the Brookmeyer-Crowley method. Confidence intervals for the estimated probability of event at a particular time point will be generated using the Greenwood's formula.

5.3. Methods to Manage Missing Data

Unless otherwise specified, all data will be evaluated as observed, and no imputation method for missing values will be used.

In all patient data listings, imputed values will be presented and flagged.

Missing statistics, eg, when they cannot be calculated, should be presented as 'ND' or 'NA'. For example, if N=1, the measure of variability (SD) cannot be computed and should be presented as 'ND' or 'NA'.

5.3.1. Missing Dates

For purposes of data listings, dates will reflect only the information provided by the investigator on the CRF.

If start dates for adverse events or concomitant medications are completely missing, a worst case approach will be taken whereby the events will be considered treatment emergent and the medications will be considered concomitant. If only partial information are available (eg, only a month and year or only a year) and the partial information provide sufficient information to indicate the dates are prior to the start of study treatment (eg, month/year less than month/year of first dose) then these will be considered to have started prior to treatment; if the partial information show information to indicate the date might be equal to the start of study treatment (eg, month and/or year equal to month and/or year of first dose) then the start of study treatment will be imputed; otherwise a similar worst case approach will apply and these will be considered to have started after treatment.

For the start dates for adverse events for reinitiation period, the similar approach as above will apply except the reinitiation of study treatment will be used instead of the start of study treatment.

5.3.1.1. Date of Last Dose of Study Treatment

No imputation will be done for first dose date. Date of last dose of study treatment, if unknown or partially unknown, will be imputed as follows:

- If the last date of study treatment is completely missing and there is no End of Treatment (EOT) CRF page and no death date, the patient should be considered to be ongoing and use the data cutoff date for the analysis as the last dosing date; or

- If the last date of study treatment is completely or partially missing and there is either an EOT CRF page or a death date available (on or prior to the data cutoff date), then impute this date as the last dose date:

  = 31DECYYYY, if only Year is available and Year < Year of min (EOT date, death date),

  = Last day of the month, if both Year and Month are available and Year = Year of min (EOT date, death date) and Month < the month of min (EOT date, death date), or

  = min (EOT date, death date), for all other cases.

5.3.1.2. Missing or Partial Death Dates

Missing or partial death dates will be imputed based on the last contact date:

- If the entire date is missing it will be imputed as the day after the date of last contact; or
- If the day or month is missing, death will be imputed to the maximum of the full (non-imputed) day after the date of last contact and the following:
  - Missing day: First day of the month and year of death, or
  - Missing day and month: January first of the year of death.

5.3.1.3. Date of Start of New Antineoplastic Therapy

Incomplete dates for new antineoplastic therapy will be imputed as follows and will be used to determine censoring dates for efficacy analyses:

- The end date of new antineoplastic therapy will be included in the imputations for start date of new antineoplastic therapy. If the end date of new antineoplastictherapy is
  - completely missing then it will be ignored in the imputations below
  - partially missing with only year (YYYY) available then the imputations below will consider 31DECYYYY as the end date of the new antineoplastictherapy
  - partially missing with only month and year available then the imputations below will consider the last day of the month for MMMYYYY as the end date of the new antineoplastic therapy

For patients who have not discontinued study treatment at the analysis cut-off date, last dose of study treatment is set to the analysis cut-off date in the imputations below.

If the start date of new antineoplastic therapy is completely or partially missing then the imputed start date of new antineoplastic therapy is derived as follows:

- Start date of new antineoplastic therapy is completely missing

- Imputed start date = min [max(PD date + 1, last dose of study treatment + 1), end date of new antineoplastic therapy]

- Only year (YYYY) for start of antineoplastic therapy is available

  IF YYYY < Year of min [max(PD date + 1, last dose of study treatment + 1), end date of new antineoplastic therapy] THEN imputed start date = 31DECYYYY;

  ELSE IF YYYY = Year of min [max(PD date + 1, last dose of study treatment + 1), end date of new antineoplastic therapy]

  THEN imputed start date = min [max(PD date + 1, last dose of study treatment + 1), end date of new antineoplastic therapy]

  ELSE IF YYYY > Year of min [max(PD date + 1, last dose of study treatment + 1), end date of new antineoplastic therapy]

  THEN imputed start date = 01JANYYYY

- Both Year (YYYY) and Month (MMM) for start of antineoplastic therapy are available

  IF

  YYYY = Year of min [max(PD date + 1, last dose of study treatment + 1), end date of new antineoplastic therapy], AND

  MMM < Month of min [max(PD date + 1 day, last dose of study treatment + 1 day), end date of new antineoplastic therapy]

  THEN

  imputed start date = DAY (Last day of MMM) MMM YYYY ;

  ELSE IF

  YYYY = Year of min [max(PD date + 1, last dose of study treatment + 1), end date of new antineoplastic therapy], AND

  MMM = Month of min [max(PD date + 1 day, last dose of study treatment + 1 day), end date of new antineoplastic therapy]

  THEN

  imputed start date = min [max(PD date + 1 day, last dose of study treatment + 1 day), end date of new antineoplastic therapy]);

ELSE IF

YYYY = Year of min [max(PD date + 1, last dose of study treatment + 1), end date of new antineoplastic therapy], AND

MMM > Month of min [max(PD date + 1 day, last dose of study treatment + 1 day), end date of new antineoplastic therapy]

THEN

imputed start date = 01 MMM YYYY;

ELSE IF

YYYY < Year of min [max(PD date + 1, last dose of study treatment + 1), end date of new antineoplastic therapy]

THEN

imputed start date = DAY (Last day of MMM) MMM YYYY;

ELSE IF

YYYY > Year of min [max(PD date + 1, last dose of study treatment + 1), end date of new antineoplastic therapy]

THEN

imputed start date = 01 MMM YYYY.

5.3.1.4. AE Stop Date:

Ongoing events will have the AE Stop Date set to one of the following values:

- If the day or month is missing, AE Stop Date will be imputed the same way as that in Section 5.3.1.5.
- Date of Death, if imputed date is after date of death.
- If the date is completely missing, no imputation will be performed.

Adverse Events are deemed similar if they have the same verbatim term.

5.3.1.5. Other Missing or Partial Dates

Imputation methods for other partial dates as follows:

- If the day of the month is missing for a start date used in a calculation, the first day of the month will be used to replace the missing date.

- If both the day and month are missing, the first day of the year is used.
- For stop dates, the last day of the month, or last day of the year is used if the day or day and month are missing, respectively.
- If the date is completely missing, no imputation will be performed.

6. ANALYSES AND SUMMARIES

6.1. Primary Endpoint(s)

6.1.1. Metastasis-Free Survival (MFS) Between enzalutamide in combination with leuprolide versus placebo in combination with leuprolide

6.1.1.1. Primary Analysis

The primary efficacy analysis will compare MFS based on BICR assessment between enzalutamide in combination with leuprolide versus placebo in combination with leuprolide using a 2-sided stratified log-rank test as described in Section 5.1. The primary population for analysis will be the ITT population. Strata will be based on those specified in the randomization system.

MFS is defined as the time in months between randomization and the earliest objective evidence of radiographic progression as determined by BICR, or death due to any cause without evidence of radiographic progression, whichever occurs first. Radiographic assessment of soft tissue disease will be according to RECIST v1.1 on CT scan or on MRI. Additional bone lesions will be assessed separately using radionuclide bone scans.

Table 6 summarizes the protocol-specified rules for the radiographic evidence of progression.

Table 6. Protocol-Specified Rules for the Radiographic Evidence of Progression

Tissue Type	Method of Assessment	Schedule	Comment
Bone	A whole-body radionuclide bone scan will consist of 5 regions including skull, thorax, spine, pelvis, and extremities. Radiographic progression for bone disease is defined as the appearance of 1 or more metastatic lesion on bone scan. When bone lesions are found in a single region, confirmation with a second imaging modality (plain film, CT, or MRI) will be required. Appearance of metastatic lesions in 2 or more of the	Screening and every 24 weeks thereafter (earlier if progression is clinically suspected) until radiographic progression is confirmed by independent central radiology review	All study films should be read locally at the study site and submitted to the central imaging unit for independent central radiology review

Table 6. Protocol-Specified Rules for the Radiographic Evidence of Progression

Tissue Type	Method of Assessment	Schedule	Comment
	5 regions on a bone scan will not require confirmation with a second imaging modality.		
Soft Tissue	Assessment of soft tissue disease will be done by CT or MRI. Radiographic progression for soft tissue disease is defined by RECIST 1.1.		

CT = computed tomography; MRI = magnetic resonance imaging; RECIST 1.1 = Response Evaluation Criteria in Solid Tumors version 1.1.

The treatment effect will be estimated using a Cox's Proportional Hazard model stratified by the randomization strata to calculate the hazard ratio. Each stratum will define a separate baseline hazard function, ie, for the i-th stratum the hazard function is expressed as: $h(i;t) = h(i,0;t) \exp(x\beta)$, where $h(i,0;t)$ defines the baseline hazard function for the i-th stratum and x defines the treatment arm (0 = control arm, 1 = experimental arm) and β is the unknown regression parameter.

Ties will be handled by using the Exact option in SAS.

In addition, the 95% CIs for the hazard ratio will be reported at the final analyses for MFS.

Kaplan-Meier estimates (product-limit estimates) will be presented by treatment arm together with a summary of associated statistics including the median MFS time with 2-sided 95% CIs. The CIs for the median will be calculated according to Brookmeyer and Crowley (1982) and the CIs for the survival function estimates at the time points defined above will be derived using the log-log transformation according to Kalbfleisch and Prentice (2011) with back transformation to a CI on the untransformed scale. The estimate of the standard error will be computed using Greenwood's formula.

Frequency (number and percentage) of patients with each event type (PD or death) and censoring reasons will be presented by treatment arm.

For patients not known to have had radiographic progression and who have not died at the time of the analysis data cutoff, MFS time will be censored at the date of the last adequate assessment (see Section 5.2.4) on or before the analysis data cutoff date. For patients who were randomized but later confirmed to have metastatic disease at enrollment or who had no adequate postbaseline tumor assessment, information will be censored on the date of randomization. Patients who initiate antineoplastic therapy such as cytotoxic chemotherapy, abiraterone acetate, hormonal agents, prostate cancer vaccines, nonradioactive bone-targeting agents and systemic radiopharmaceuticals for prostate cancer, or any antineoplastic therapy

without evidence of metastasis will be censored at the last adequate assessment prior to initiation of such therapy. (Note: if progression occurs on the same date as the start of such therapy the progression will count as an event.) Additionally patients with documented progression or death after 2 or more missing tumor assessment will be censored at the last adequate assessment prior to the missing assessments. In this study, tumor assessments are assessed every 24 weeks. Therefore, time without adequate assessment is defined as 49 weeks (allowing for a +/- 1 week window for assessments). The censoring rules for the primary and sensitivity analyses of MFS are summarized in Table 7.

Table 7. The Censoring Rules for the Primary and Sensitivity Analyses of MFS

Analysis	Censoring Rules	Date of Censoring
Primary analysis of MFS	Patients with no baseline or no post baseline assessments who have not died within 49 weeks after randomization	Date of randomization
	Patients who were randomized but confirmed metastatic at baseline	Date of randomization
	Patients who had no confirmed metastasis and did not die prior to data cutoff date	Date of the last adequate radiographic tumor assessment prior to data cutoff date
	Patients who initiate antineoplastic therapy such as cytotoxic chemotherapy, abiraterone acetate, hormonal agents, prostate cancer vaccines, nonradioactive bone-targeting agents and systemic radiopharmaceuticals for prostate cancer, or any antineoplastic therapy without evidence of metastasis	Date of the last adequate radiographic tumor assessment prior to first use of any such therapy
	Patients with radiation therapy performed for prostate cancer-related lesions without evidence of metastasis	Date of the last adequate radiographic tumor assessment prior to the earliest use of radiation therapy
	Patients with evidence of metastasis or death after 2 or more consecutive missed tumor assessment visits	Date of the last adequate radiographic tumor assessment prior to the first missed visit date

MFS = Metastatic Free Survival; of note, the censoring rules are applied to MFS events by either radiographic progression, or death due to any cause without evidence of radiographic progression; of note, antineoplastic therapies according to the above search criteria undergo medical review to confirm.

6.1.1.2. Time of Follow-Up for MFS

A Kaplan-Meier summary table for MFS follow-up duration will also be generated to assess the follow-up time in the treatment arms by reversing the MFS censoring and event indicators.

6.1.1.3. Sensitivity/Robustness Analyses

The following sensitivity analyses will be performed for MFS.

Sensitivity 1: **Including Events Regardless of Initiation of Antineoplastic Therapies**

Censoring rules will follow those in the primary MFS analysis except that events occurring for the first time after the initiation of antineoplastic therapy will not be censored and be considered as events. A 2-sided stratified log-rank test (same as the primary analysis) will be used to compare the treatment groups.

Sensitivity 2: **MFS on eITT Population**

MFS for the eITT population will also be analyzed as a sensitivity analysis. The definition of MFS and censoring rules will be consistent with primary analysis. A 2-sided stratified log-rank test will be used to compare the treatment groups. All methods from the primary efficacy analysis will be repeated.

Sensitivity 3: **MFS Based on Investigator Assessment**

MFS as assessed by the investigator will also be analyzed as a sensitivity analysis. The definition of MFS and censoring rule will be consistent with primary analysis. A 2-sided stratified log-rank test will be used to analyze the MFS values. Furthermore, the concordance and discordance rates between the independent central radiology review and investigator assessment will be summarized using the metastasis status by the treatment groups.

Table 8 outlines the possible outcomes by investigator and BICR (Amit et al. 2011).

Table 8. Possible Outcomes for Investigator vs BICR

		BICR	
Investigator		**PD**	**No PD**
	PD	a = a1 + a2 + a3	b
	No PD	c	d

a1 = number of agreements on timing and occurrence of PD; a2 = number of times agreement on PD event but investigator declares PD event later than BICR; a3 = number of times agreement on PD event but investigator declares PD event earlier than BICR; PD = progressive disease.

N = a+b+c+d representing the number of patients in the FA population.

The timing agreement of progression has to be defined as a window of ± 7 days.

The following measure of discordance/concordance will be calculated for each treatment arm:

- Total Event Discrepancy Rate: (b+c) / N
- Early Discrepancy Rate (EDR): (a3+b) / (a+b)
- Late Discrepancy Rate (LDR): (a2+c) / (a2+a3+b+c)
- Overall Discrepancy Rate: (a2+a3+b+c) / N
- Overall Concordance Rate: (a+d)/N

The EDR represents the positive predictive value of investigator assessment and quantifies the frequency with which the investigator declares progression earlier than BICR within each arm as a proportion of the total number of investigator assessed PDs.

The LDR quantifies the frequency with which the investigator declares progression later than BICR as a proportion of the total number of discrepancies within the arm.

Discordance metrics are calculated for each treatment arm. For each metric, the difference in discordance between the experimental and control arms is used to evaluate potential bias. If the discordance is similar across the treatment arms then this suggests the absence of evaluation bias favoring a particular arm. A negative differential discordance for EDR and/or a positive differential discordance for LDR may be indicative of investigator evaluation bias in favor of the experimental arm (Amit et al, 2011).

Sensitivity 4: Impact of Clinical Progression

In this sensitivity analysis, patients who discontinue study drug primarily due to clinical deterioration prior to protocol-defined evidence of radiographic progression will be considered as having clinical progression. For this analysis, MFS is defined as the duration of time between randomization and the earliest objective evidence of metastatic disease, date of study drug discontinuation for clinical progression, or death, or evidence of clinical progression, whichever occurs first. The censoring rules used for the primary analysis will be utilized. The hazard ratio and its 95% confidence interval will be reported.

Sensitivity 5: Impact of Censoring Due to Discontinuation Prior to Radiographic Progression for Patients Notified of PSA Progression or Progression by Positron Emission Tomography (PET) Imaging of Prostate-specific Membrane Antigen (PSMA)

For the censoring of MFS for the patients who reached PSA progression and discontinued study treatment prior to the development of radiographically detectable metastatic disease, and the potential for this to be informative censoring, a reference-based imputation method based on Bayes Gibbs sampling as outlined by Lu, Li, and Koch (Lu et al. 2015) will be implemented to assess the impact of the above censoring. If applicable, the inverse probability of censoring weighting (IPCW) method by Robins and Finkelstein may be used to adjust for the above censoring. For the censoring of MFS for the patients who discontinued

study treatment due to progression by PSMA-PET scan prior to the development of radiographically detectable metastatic disease, if applicable, the similar methods described above may be used to adjust for the censoring.

In order to assess the impact of patients initiating novel androgen inhibitors (such as enzalutamide, apalutamide, darolutamide and abiraterone) prior to the development of radiographically detectable metastatic disease, if applicable, the following sensitivity analyses may be performed: the Rank-Preserving Structural Failure Time Model (RPSFTM) (Robins & Tsiatis 1991), IPCW method (Robins & Finkelstein, 2000) and the two-stage method (Latimer & Abrams 2014).

6.2. Key Secondary Endpoint(s)

6.2.1. MFS Between Enzalutamide Monotherapy Versus Placebo Plus Leuprolide

MFS between enzalutamide monotherapy versus placebo plus leuprolide will be defined as above for primary analysis of the combination comparison. Analysis of this endpoint will be performed using the 2-sided stratified log-rank test to compare the 2 treatment groups with the same strata described in Section 6.1.1. A Cox proportional hazards model will be used to evaluate the MFS analysis to calculate the HR and its 95% CI.

The same sensitivity analyses as those specified for the primary comparison between enzalutamide in combination with leuprolide versus placebo in combination with leuprolide in Section 6.1.1 will be implemented.

6.2.2. Time to PSA Progression

Only results from PSA samples taken before the initiation of any new prostate cancer therapy and after the start of study drug will be considered.

PSA progression is defined in Section 3.2.1 as the date that a ≥25% increase and an absolute increase of ≥2 µg/L (2 ng/mL) above the nadir (or baseline for patients with no PSA decline by week 25) that is confirmed by a second consecutive value at least 3 weeks later. The date of PSA progression is the first date the PSA progression is observed. For patients who have suspended treatment at week 37 and later reinitiated treatment, baseline will be defined as the last PSA assessment prior to or on the date of reinitiation. The date of PSA progression is the first date the PSA progression is observed.

PSA progression is only defined during active study treatment; therefore, patients meeting PSA progression during the suspension period will be censored unless the PSA progression criteria are subsequently met following treatment reinitiation. Time to PSA progression will be censored on the date of the last PSA sample taken. Patients with PSA progression after 2 or more consecutive missed PSA assessments (ie, time interval >6 months or 182 days between 2 consecutive PSA samples) will be censored on the date of last PSA assessment prior to the missed assessments. In patients with no baseline PSA and patients with no post-baseline PSA results, time to PSA progression will be censored on the date of randomization.

Time to PSA progression will be compared between treatment groups using a 2-sided stratified log-rank test.

6.2.3. Time to First Use of New Antineoplastic Therapy

New antineoplastic therapy includes medications used specifically for prostate cancer treatment including hormonal treatments, immunotherapy, chemotherapy and investigative agents.

Time to first use of new antineoplastic therapy will be compared between treatment groups using a 2-sided stratified log rank test. In patients with no new antineoplastic therapy initiated for prostate cancer after randomization, time to start of new antineoplastic therapy will be censored on the last visit date or the date of randomization, whichever occurs last.

At the time of final analysis of OS, this endpoint will be summarized descriptively and provided as exploratory analysis.

6.2.4. Overall Survival

OS is defined as the time in months between randomization and death due to any cause. One interim analysis and a final analysis of overall survival will be performed. The first interim will be performed at the time of the final analysis of the MFS primary endpoint. The overall survival will be compared between treatment groups using a 2-sided stratified log rank test. Patients without an event date will be censored at the date of the last contact.

At the time of final analysis of OS, subgroup analyses will be conducted using the same analysis methods as for the subgroup analyses of MFS described in Section 6.6.

At the final analysis of OS, an additional confounding factor for the OS analysis may come from patients who choose to access commercial enzalutamide, combination of enzalutamide and leuprolide and other commercially available treatments for the patients. The sponsors may explore the effect of crossover and post-study systemic anticancer therapy using rank-preserving structural failure time model (RPSFTM), and the inverse probability of censoring weighting (IPCW) method if appropriate. Justifications, limitations and assumptions of RPSFTM (Robins & Tsiatis 1991) and IPCW (Robins & Finkelstein, 2000) methods will be adequately evaluated.

6.3. Other Secondary Efficacy Endpoint(s)

The following efficacy analyses will be performed for the ITT population. The resulting 2-sided p-values will be considered descriptive and no adjustment will be made for multiplicity. Treatment group comparisons will be between the combination arms of enzalutamide plus leuprolide versus placebo plus leuprolide and between enzalutamide monotherapy therapy versus placebo plus leuprolide unless otherwise specified.

6.3.1. Time to Distant Metastasis

The time to distant metastasis is defined as the time in months from randomization to the earliest objective evidence of distant soft tissue metastases or metastatic bone disease by BICR. Soft tissue disease including lymph nodes above the aortic bifurcation and outside the pelvis and any bone metastases will be counted as distant metastases.

The time to distant metastasis will be compared between treatment groups using a 2-sided stratified log-rank test.

For patients not known to have distant metastases at the time of the analysis data cutoff, time to distant metastasis will be censored at the date of the last adequate assessment (see Section 5.2.4) on or before the analysis data cutoff date. Patients who were randomized but later confirmed to have metastatic disease at enrollment or patients with no postbaseline tumor assessment information will be censored on the date of randomization. Patients who initiate cytotoxic chemotherapy, abiraterone acetate, hormonal agents, or nonradioactive bone-targeting agents and systemic radiopharmaceuticals for prostate cancer, or any antineoplastic therapy without evidence of metastasis will be censored at the last adequate assessment prior to initiation of such therapy. (Note: if distant metastasis occurs on the same date as the start of such therapy the distant metastasis will count as an event.) Additionally patients with distant metastasis after 2 or more missing tumor assessment will be censored at the last adequate assessment prior to the missing assessments. In this study tumor assessments are assessed every 24 weeks. Therefore time without adequate assessment is defined as 49 weeks (allowing for a +/- 1 week window for assessments).

6.3.2. Proportion of Patients Who Remain Treatment-Free at 2 Years After Suspension of Study Treatment

The proportion of patients per group who remain treatment-free 2 years after suspension of study treatment at week 37 due to undetectable PSA will be compared between treatment groups using the stratified Cochran-Mantel-Haenszel test. Two-sided 95% CIs per group will be reported using the Clopper-Pearson method. The differences in proportions between treatment groups and 95% CI using Wald method will be provided.

6.3.3. Proportion of Patients With Undetectable PSA at 2 Years

The proportion of patients per group with undetectable PSA 2 years after suspension of study treatment at week 37 due to undetectable PSA will be compared between treatment groups using the stratified Cochran-Mantel-Haenszel test. Two-sided 95% CIs per group will be reported using the Clopper-Pearson method. The differences in proportions between treatment groups and 95% CI using Wald method will be provided.

6.3.4. Proportion of Patients With Undetectable PSA at 36 Weeks

The proportion of patients per group with undetectable PSA at 36 weeks will be compared between treatment groups using the stratified Cochran-Mantel-Haenszel test. Two-sided 95% CIs per group will be reported using the Clopper-Pearson method. The differences in proportions between treatment groups and 95% CI using Wald method will be provided.

6.3.5. Time to Resumption of any Hormonal Therapy

The time to resumption of any hormonal therapy is defined as the time in months between the date of treatment suspension at week 37 due to undetectable PSA and the date that hormonal therapy is restarted. The time to resumption of any hormonal therapy will be compared between treatment groups using a 2-sided log-rank test. Patients without observed resumption of any hormonal therapy at the time of analysis will be censored at the date of

last visit. A review of medication prior to unblinded will be performed to identify normal therapies.

6.3.6. Time to Castration Resistance

Time to castration resistance applies only to patients receiving leuprolide treatment and is defined as the time in months from randomization to the first occurrence of radiographic disease progression by BICR, PSA progression (as defined in Section 3.2.1) or symptomatic skeletal event whichever occurs first with castrate levels of testosterone (<50 ng/dL). The latest testosterone value measured prior to or at the date of radiographic disease progression by BICR, PSA progression or a symptomatic skeletal event is used to determine if the event is a castration resistance event.

Time to castration resistance is defined as the time in months from randomization to the first castration-resistant event (radiographic disease progression, PSA progression or symptomatic skeletal event), whichever occurs first. In patients with no documented castration resistance event, the time to castration resistance will be censored on the latest date from: the date of last radiographic assessment prior to the start of any new prostate cancer therapy, the last PSA sample taken prior to the start of any new prostate cancer therapy and prior to 2 or more consecutive missed PSA assessments (if applicable), and the last visit date performed. Patients with no baseline radiographic assessment, patients with no post baseline radiographic assessments, patients with all post-baseline radiographic assessments documented as “Not Evaluable”, patients with no baseline PSA, and in patients with no postbaseline PSA results, the time to castration resistance will be censored on the date of randomization.

6.3.7. Time to Symptomatic Progression

Time to symptomatic progression is defined as the time in months from randomization to development of a skeletal-related event, worsening of disease-related symptoms requiring initiation of a new systemic antineoplastic therapy, or development of adverse events and clinically significant signs and/or symptoms due to loco-regional tumor progression requiring opiate use, surgical intervention or radiation therapy, whichever occurs first. Patients without observed symptomatic progression at the time of analysis will be censored at the last visit date.

6.3.8. Time to First Symptomatic Skeletal Event (SSE)

Time to first symptomatic skeletal event is defined as the time in months from randomization to use of radiation therapy (external beam radiation therapy or radionuclides) or surgery to bone for prostate cancer, findings of clinically apparent pathologic bone fracture or of spinal cord compression, or new use of opiate and/or systemic antineoplastic therapy due to bone pain collected in the SSE CRF, whichever occurs first. Because skeletal events are expected to occur after radiographic progression, the analysis of time to first symptomatic skeletal event will be performed with the final overall survival analysis.

The time to first symptomatic skeletal event will be compared between treatment groups using a 2-sided stratified log-rank test. In patients with no SSE by the time of the data

analysis cut-off date, the time to first SSE will be censored on the last visit date or the date of randomization, whichever occurs last.

At the time of final analysis of OS, this endpoint will be updated and summarized descriptively and provided as exploratory analysis.

6.4. Patient Reported Outcomes

The following analyses will be performed for the ITT population.

6.4.1. Pain

The assessment of pain progression will be conducted using the BBPI-SF questionnaire. For each treatment arm and at each time point, the number and percentage of patients who completed the questionnaire will be summarized.

Time to clinically relevant pain progression is defined as the time in months from randomization to onset of pain progression BPI-SF, where pain progression is defined as a 2-point or more increase from baseline in the question 3 score. Patients without observed pain progression at the time of analysis will be censored at the date of last pain assessment on or before the analysis data cutoff date. Patients who were randomized, but with no baseline or no post baseline scores, or with a score that is too high (>8 in BPI-Item 3) at baseline to be able to show further deterioration, for any of the three previous scenarios, will be censored on the date of randomization.

The assessment of pain progression will be compared between the 2 treatment groups using a 2-sided stratified logrank test.

BPI-SF domain scores and change from baseline will be summarized by treatment arm as randomized. Pain Severity score, will be calculated by averaging the ratings of the four pain severity items (items 3, 4, 5 and 6). All four severity items must have been completed for the mean to be calculated. BPI-SF Pain Interference score is calculated as the average of the seven pain interference items (items 9A-9G). The pain interference score will be calculated if more than 50% of the total items, or four out of seven, have been completed on a given administration.

Subjects with missing baseline scores are not assessable for baseline description or change from baseline. Patients with no baseline scores or no post baseline scores are not assessable for baseline description or change from baseline summaries.

Additional PRO analysis of pain will be described in a supplemental SAP specific to PRO data.

6.4.2. Quality of Life

Quality of life will be assessed using FACT-P, EQ-5D-5L, and QLQ-PR25 questionnaires.

For each treatment arm and at each time point, the number and percentage of patients who completed the FACT-P, EQ-5D-5L, and QLQ-PR25 questionnaires will be summarized. An instrument is considered complete if at least one scale can be computed.

Time to functional status decline (deterioration) by the FACT-P questionnaire is defined as the time in months from randomization until the date of occurrence of a 10-point decrease from baseline in FACT-P total score. Patients without functional status deterioration at the time of analysis will be censored at the date of last FACT-P assessment on or before the analysis data cutoff date. Patients with a score that is too low (<10 in FACT-P total score) at baseline to be able to show further deterioration, or patients with no baseline score, or no post-baseline score, for any of the three previous scenarios, the time to deterioration will be censored at date of randomization.
The time to decline (deterioration) in global FACT-P score will be compared between treatment groups using a stratified 2-sided log-rank test with the same strata as in the time to MFS analysis described in Section 6.1.1.

The FACT-P domain scores and their change from baseline will be summarized by treatment arm as randomized for the ITT population. For handling missing items, if half or more questions within scale are answered then a score will be calculated for that scale. Otherwise the patient score for that scale will be missing. The FACT-P total score is the sum of all 5 subscale scores. The total score will be calculated only if the overall item response rate is greater than 80% (i.e., a minimum of 32 of 39 items currently scored in the FACT-P have been answered), and no subscale scores are missing. FACT-P data will be summarized descriptively by study visit.

For EQ-5D-5L, proportion of patients in each category, EQ-VAS score, and change from baseline will be summarized by treatment arm in the ITT population. On the EQ-5D-5L, questions not answered will be considered as missing items, and will neither be imputed nor utilized. EQ-5D-5L data will be summarized descriptively by study visit.

The QLQ-PR25 domain scores and their change from baseline will be summarized by treatment arm as randomized for the ITT population. The QLQ-PR25 domain scores and total score will be calculated using the EORTC QLQ-C30 Scoring Manual (Fayers 2001). If less than half of the constituent items on the QLQPR25 have been answered for a multi-item subscale, that subscale will be considered missing. Single-item subscales will be considered missing if the constituent item is incomplete. QLQ-PR25 data will be summarized descriptively by study visit.

Subjects with missing baseline scores are not assessable for baseline description or change from baseline.

Additional PRO analyses will be described in a supplemental SAP specific to PRO data.

6.5. Exploratory Endpoint(s)

6.5.1. Progression-Free Survival on First Subsequent Therapy (PFS2)

PFS2 is defined as the time in months from the date of randomization to the first occurrence of investigator-determined disease progression (PSA progression, progression on imaging, or clinical progression) or death due to any cause, whichever occurred first, while the patient was receiving first subsequent therapy for prostate cancer.

PFS2 (months) = [date of event or censoring – start date +1]/30.4375

A patient will be considered to have an event if

1. the patient had objective PD on or prior to the start of the first subsequent antineoplastic therapy, AND had investigator-determined disease progression (PSA progression, progression on imaging, or clinical progression) after the start of the first subsequent antineoplastic therapy)

OR

2. the patient died.

If there are multiple therapies within a regimen with different start dates, the earliest start date will be utilized. PFS2 will be analyzed using a 2-sided stratified log-rank test. The HR and 95% CI will be provided. Kaplan-Meier estimates will be presented by treatment arm with the median and 95% CIs.

The censoring and event date options to be considered for PFS2, each corresponding censoring reason and its hierarchy are presented in Table 9. Frequency (number and percentage) of patients with an event (PSA progression, progression on imaging, clinical progression after next line treatment or death) and censoring reasons will be presented by treatment arm.

Table 9. Outcome, Event Dates, Censoring Reasons and Hierarchy for PFS2 Analyses

Scenario	Date of Event/ Censoring	Event/ Censoring Reason/ Censoring Hierarchy
(No PD[a]) and (no death)	Date of lastadequate tumor assessment[b] documenting no PD	Censored/ No PD/ 1
(No PD[a]) and death	Date of death	Event (Death)
(PD[a] date > NTX[c] start date) and (no death)	Start date of NTX[c]	Censored/

Table 9. Outcome, Event Dates, Censoring Reasons and Hierarchy for PFS2 Analyses

Scenario	Date of Event/ Censoring	Event/ Censoring Reason/ Censoring Hierarchy
		Start of new antineoplastic therapy before PD/ 2
(PD[a] date > NTX[c] start date) and death	Date of death	Event (Death)
(PD[a] date ≤ NTX[c] start date) and (PD2[d] date is non-missing)	Date of PD2[d]	Event (PD after start of first subsequent antineoplastic therapy)
(PD[a] date ≤ NTX[c] start date) and (no PD2[d]) and (no death)		
• If [withdrawal of consent in the end of study page]	Withdrawal of consent date in the end of study page	Censored/ Withdrawal of consent/ 3
• Else if [lost to follow-up in any disposition page]	Last contact date	Censored/ Lost to follow-up/ 4
• Else if no prior conditions are met	Last contact date	Censored/ Ongoing without PFS2 event/ 5
(PD[a] and no NTX[c]) and (no death)	Last contact date	Censored/ Ongoing without PFS2[e] event/ 6
(PD[a] date ≤ NTX[c] start date) and (no PD2[d]) and death	Date of death	Event (Death)
(PD[a] and no NTX[c]) and death	Date of death	Event (Death)

a. PD is the first objective progressive disease by investigator assessment, without considering any censoring rules
b. If there are no adequate post-baseline assessments, then the censoring date is the date of randomization. If patient has initiated the first subsequent antineoplastic therapy, the last adequate post-baseline assessment on or prior to start date of the first subsequent antineoplastic therapywill be considered.
c. NTX is the first subsequent antineoplastic therapy collected in “Long Term Follow-up Antineoplastic Therapies” CRF page .
d. PD2 is the first progressive disease (PSA progression, progression on imaging, or clinical progression), after initiation of first subsequent antineoplastic therapy by investigator assessment without considering any censoring rules.

Table 9. Outcome, Event Dates, Censoring Reasons and Hierarchy for PFS2 Analyses

Scenario	Date of Event/ Censoring	Event/ Censoring Reason/ Censoring Hierarchy

e. PFS2 is the Progression-Free Survival on First Subsequent Therapy

The PFS2 time or censoring time and the reasons for censoring will also be presented in a patient listing.

At the time of final analysis of OS, this endpoint will be summarized descriptively and provided as exploratory analysis.

6.6. Subgroup Analyses

Subgroup analyses of the MFS endpoint for both enzalutamide in combination with leuprolide versus placebo in combination with leuprolide and enzalutamide monotherapy versus placebo in combination with leuprolide will be performed to determine whether the treatment effect is concordant among subgroups. If the subgroups are too small (<10 events), the analyses would not be provided. The unstratified hazard ratios and their 95% confidence intervals will be displayed in a forest plot. The following variables defined by baseline subject characteristics will be used to define subgroups:

- PSA doubling time (≤3 months, >3 to ≤6 months, >6 to ≤9 months);
- Baseline use of a bone targeting agent (yes, no);
- Baseline age category at baseline (<65, ≥65);
- Race (African-American or Black, White, Other);
- Body mass index calculated from height and weight (at or below median, above median);
- ECOG performance status at baseline (0, 1);
- Geographic region (North America, Europe, and rest of the world);
- Total Gleason score at baseline (≤7, ≥8) at diagnosis;
- Prior neoadjuvant/adjuvant hormonal therapy (yes, no);
- Prior radiation therapy (yes, no);
- Prior prostatectomy (yes, no);

- History of cardiovascular disease (yes, no);
- PSA value at baseline (≤10 ng/mL, >10 ng/mL);

6.7. Baseline and Other Summaries and Analyses

6.7.1. Baseline Summaries

Demographic and Baseline Characteristics

The following demographic and baseline characteristics will be summarized by treatment group as randomized for all patients in the ITT population and by geographic region:

- Age, race, weight, and body mass index;
- Geographic region (North America, Europe, rest of world) (Appendix 1);
- Categorized screening PSA doubling time (≤3 months ,>3 to ≤6 months, and >6 to ≤9 months), and continuous screening PSA doubling time;
- Prior neoadjuvant/adjuvant hormonal therapy (yes, no) and/or short course of hormonal therapy (yes, no);
- Prior prostatectomy (yes, no);
- Prior radiation therapy (yes, no);
- Prior prostatectomy and radiation therapy (yes, no);
- Eastern Cooperative Oncology Group (ECOG) performance status;
- Baseline serum PSA (≤10 ng/mL, >10 ng/mL), and continuous baseline serum PSA (ng/mL);
- History of cardiovascular disease (yes, no).

Listings will be provided for these parameters for all randomized patients.

Disease Characteristics

The following medical history and disease characteristics will be summarized by treatment group as randomized for all patients in the ITT population:

- Time (months) from initial pathological diagnosis or first treatment for prostate cancer to randomization, whichever is earlier;
- Total Gleason score category at diagnosis (Low [2–4], Medium [5–7], and High [8-10]);
- Clinical Stage (T1, T2, T3);

- Pathologic tumorurgical stage;
- Lymph node stage;
- Pelvic soft tissue lesion(s);

Previous Therapies

Previous therapies for prostate cancer will be summarized by treatment group as randomized for all patients in the ITT population as follows:

- Number of prior prostate cancer therapies;
- Number of prior hormonal therapies;
- Prior nonhormonal therapy use (yes/no);
- Use of bisphosphonate or denosumab for non-prostate cancer therapy (yes/no) at baseline;
- History of surgical prostate cancer procedures (yes/no).

6.7.2. Study Conduct and Patient Disposition

Study conduct will be summarized by randomized treatment group as follows:

- Randomized patients by randomization stratification factors;
- Reason for exclusion from 1 or more populations;

Patients Disposition will be summarized as follows:

- Patients discontinuing treatment and the reasons for discontinuation (as documented on the case report form (CRF), including death, lost to follow-up, protocol violation, withdrew consent to continue treatment, disease progression, adverse event, or other);
- Patients withdrawing from long-term follow-up and the reasons for withdrawal.

The number of patients in each of the specified populations (ITT, eITT, safety) will be provided. Additionally, the number and percentage of patients on treatment and in follow-up will be provided for the safety population.

The following listings will be produced:

- Patients who received treatment different from the treatment randomized;
- Patients discontinuing from each study period after enrollment (study center, treatment group, reason, dose of enzalutamide, and the duration of treatment).

6.7.3. Protocol Deviations

Patients with major protocol deviations will be listed and summarized by treatment group as randomized. Categories of major deviations are as follows but not limited to:

- Eligibility criteria were not met
- Developed criteria for discontinuation of study drug but did not discontinue study treatment
- Received excluded concomitant medication
- Received the wrong treatment or incorrect dose
- Did not sign informed consent before study-specific procedures were performed

A detailed list of all major protocol deviations will be determined before the study is unblinded and a listing of all major deviations will be provided.

6.7.4. Study Treatment Exposure

Data from patients in the safety population will be used to summarize exposure and compliance.

Enzalutamide

The summary of treatment exposure for enzalutamide will include the following information:

- Treatment duration (months)
- Treatment suspension duration (months)
- Modified treatment duration (months)
- Duration of treatment after reinitiation (months)
- Cumulative dose (mg)
- Average daily dose (mg/day).

Treatment duration, treatment suspension duration, modified treatment duration and duration of treatment after reinitiation (see Section 5.2.5) will be summarized both as a continuous measures and as categorical measures (≤6 months, 6 to <12 months, 12 to <24, 24 to <36, 36 to <48, 48 to <60, 60 to <72, and ≥72 where appropriate).

The cumulative dose is the sum of the actual dose levels that the patient received (ie, total dose administered).

Average daily dose is the cumulative dose divided by the modified treatment duration.

The total number of capsules taken and percent overall compliance will also be summarized. Number of capsules taken will be calculated based on the number of capsules dispensed at all study visits minus the number of capsules indicated as returned. The total cumulative dose in milligrams (mg) will be calculated as (40 mg × number of capsules taken), as each capsule is equivalent to 40 mg of the study drug or placebo.

Percent overall compliance rate will be defined as the number of capsules taken during the study divided by the expected number of capsules, multiplied by 100. Capsules from bottles not returned will be subtracted from the number of capsules taken and the expected number of capsules in calculating percent overall compliance rate. Each patient will be taking 4 capsules each day while on study treatment. A patient's expected number of capsules will be calculated as [4 × (date of last dose prior to data cutoff date – date of first dose of study drug + 1)]. However, for patients who have dose modifications (eg, dose reduction or dose withholding, or drug suspension due to undetectable PSA), the expected number of capsules will be calculated according to the dose modification. Percent overall compliance will be summarized both as a continuous measure and a categorical measure in increments of 20%.

A data listing will be provided to present dose administration, modifications, and the derived compliance variables.

Dose Reduction and Modification

A dose reduction is defined as a non-zero dose that is less than the prior dose.

The number and percentage of patients with at least 1 dose reduction as well as a breakdown of dose reductions (1 / 2 / 3 / ≥4) will be summarized by treatment arm.

Reasons for dose reductions will also be summarized. Patients can contribute to more than 1 reason if multiple dose reductions occurred for different reasons, but will only be counted once per reason. Percentages will be calculated based on the total number of patients in safety analysis set.

An interruption is defined as a 0 mg dose administered on 1 or more days (not including the treatment suspension period due to undetectable PSA at week 36). The number and percentage of patients with at least 1 dose interruption will be summarized by treatment arm.

Reasons for dose interruption will also be summarized. Patients can contribute to more than 1 reason if multiple dose interruptions occurred for different reasons, but will only be counted once per reason. Percentages will be calculated based on the total number of patients in safety analysis set.

Leuprolide

Number of injections before and after the suspension period by treatment arms will be summarized. A summary of the number of missed injections will also be provided.

A data listing will be provided to present the number of treatments, dose administration, route of treatment (subcutaneous or intramuscular injection) and modifications for leuprolide.

6.7.5. Concomitant Medications

Concomitant medications taken during the study treatment period will be summarized for all patients in the safety population by treatment group. Medications are considered concomitant if exposure occurs during the on-treatment period. Medications will be summarized by WHO drug dictionary anatomical therapeutic chemical (ATC) classification system and generic medication name.

In addition, subsequent antineoplastic therapies will be summarized separately by treatment group. Medications will be summarized by WHO drug dictionary ATC classification system and generic medication name.

All medications recorded on the case report form will be listed.

6.8. Safety Summaries and Analyses

All patients in the safety population will be used in the safety analyses. Safety analyses will be summarized by treatment group.

The on-treatment period will be defined as in Section 5.2.5. For incomplete date of last dose of study drug and incomplete start date of the first new antineoplastic treatment that are missing the day of the month, please refer to the missing data Section 5.3.1.

At the time of analyses after protocol amendment #5, safety data for the patients treated in protocol defined open-label (OL) stage will be summarized by double blinded stage and OL stage separately.

At the time of analyses after protocol amendment #5, based on the protocol defined data collection, safety will be evaluated by the frequency of serious adverse events, frequency and severity of adverse events, frequency of study drug discontinuation due to adverse events, and frequency of new clinically significant changes in clinical laboratory values and vital signs.

6.8.1. Adverse Events

The severity of all adverse events is to be evaluated by the investigator based on the NCI CTCAE version 4.03. All adverse events will be coded to preferred term (PT) and system organ class (SOC) using MedDRA version in use at the time of database release.

TEAEs will be summarized by PT and/or SOC, toxicity grade, and relationship to study treatment. An event will be considered treatment related if the investigator considered the event related to one or both of study drugs given in combination. The treatment related summaries will be separated for one or both study drugs when given in combination.

Separate summaries will be provided for serious TEAEs regardless of causality and serious TEAE related to study treatment.

Separate summaries will be provided for TEAEs (regardless of relationship to study treatment):

- Leading to permanent discontinuation from enzalutamide,
- Leading to permanent discontinuation form leuprolide,
- Leading to enzalutamide dose interruptions,
- Leading to leuprolide dose interruptions,
- Leading to enzalutamide dose reductions, and
- Leading to leuprolide dose reductions.

mTEAEs and rTEAEs defined as in Section 3.5 will be summarized by toxicity grade and relationship to study treatment. Separate summaries will also be provided for:

- Serious mTEAE regardless of causality,
- Serious mTEAE related to study treatment,
- Serious rTEAE regardless of causality, and
- Serious rTEAE related to study treatment.

Patients with multiple occurrences of events for a given PT, SOC will only be counted once.

Adverse event time-adjusted rates will be calculated as the number of occurrences of event divided by the number of patient-years of treatment-emergent (for the corresponding treatment emergent period)surveillance for each treatment group. Patients can have more than 1 occurrence of each event.

Additionally, Kaplan-Meier methods will be used to display the time to the first Grade 3 or higher TEAE by treatment group. Patients who are not known to have had Grade 3 or higher TEAE by the analysis cutoff date are censored at the date last known to be Grade 3 or higher TEAE-free. Kaplan-Meier methods will also be used to display the time to the first SAE by treatment group. Patients who are not known to have had SAE by the analysis cutoff date are censored at the date last known to be SAE-free.

Subgroup and supplemental tabulations of TEAEs by categorized treatment group, Grade 3 and higher TEAEs by categorized treatment group, and serious TEAEs by categorized treatment group will be created as shown in Table 10.

Table 10. Subgroup and Supplemental Tabulations of Treatment-Emergent Adverse Events (TEAEs)

Group Variable	Subgroups
Study Day Cut Points	0-30 days after initiation of study drug 1-37 weeks days after initiation of study drug >37 weeks after initiation of study drug, Denominator for the above categories will be those patients that are on treatment at the beginning of the period. Adverse events will be assigned to day categories based on the treatment day of the start date (or worsening date) of the adverse event. Study days after initiation of study drug use the study day computations presented in Section 5.2.5.
Age (years)	<65 ≥65 <75 ≥75 <85 ≥85
Baseline Body Mass Index	BMI Underweight: BMI is less than 18.5. Normal weight: BMI is 18.5 to 24.9. Overweight: BMI is 25 to 29.9. Obese: BMI is 30 or more
Geographic Region	North America Europe Rest of the World
History of Cardiovascular Disease	Yes No

6.8.2. Deaths

The frequency (number and percentage) of patients in the safety analysis set who died and who died within 30 days after last dose of study treatment as well as the primary reason for death, will be tabulated based on information from 'Survival Follow-Up' CRFs, by treatment arm.

6.8.3. Laboratory Data

Laboratory data in this study consist of hematology values and chemistry tests. Only data collected from the central laboratory at baseline and during the treatment-emergent period will be summarized. Laboratory data collected outside the treatment-emergent period will only be listed in the data listings.

Normal ranges will be implemented to identify values that are outside the normal ranges and create the NCI toxicity grade using the CTCAE version 4.03. Parameters that have criteria available for both low and high values (eg, hypercalcaemia for a high value of calcium and hypocalcaemia 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. For each laboratory parameter, the baseline laboratory value is defined as the last laboratory value

collected on or before the date and time of the first dose of study drug. The change from baseline to postbaseline value will be calculated for each laboratory parameter.

For laboratory parameters that are gradable by the CTCAE, a shift table will be provided for each parameter to summarize baseline toxicity grade versus worst postbaseline toxicity grade during the treatment-emergent period. The number and percentage of patients with at least 1 occurrence of Grade 3 or Grade 4 laboratory values in the treatment-emergent period will be summarized for each parameter and treatment group.

For each laboratory parameter that is not gradable by the CTCAE, a shift table based on the normal range (low, normal, and high) will be provided to summarize the baseline result versus both the lowest and the highest postbaseline result during the treatment-emergent period.

For patients with Grade 3 and/or Grade 4 laboratory values in the treatment-emergent period, a by-patient data listing will be presented to display data including visit label (eg, Week 17), assessment date (day), laboratory value, normal range flag (low, normal, and high), and change from baseline value. The baseline value will be flagged in this data listing.

The number and proportion of patients with liver function test elevations will be presented by categorized treatment group. Liver function test elevations are assessed by using postbaseline results in ALT, AST, and total bilirubin during the treatment-emergent period based upon the definitions presented in Table 11.

Table 11. Categories of Liver Function Test Elevations

Laboratory Test	Category
ALT, AST	Postbaseline result ≥3 × upper limit of normal
	Postbaseline result ≥3 × upper limit of normal and worse than baseline
	Postbaseline result ≥5 × upper limit of normal
	Postbaseline result ≥10 × upper limit of normal
	Postbaseline result ≥20 × upper limit of normal
ALT or AST	Postbaseline result ≥3 × upper limit of normal
Total Bilirubin	Postbaseline result ≥2 × upper limit of normal
Alkaline Phosphatase	Postbaseline result ≥1.5 × upper limit of normal

ALT = alanine aminotransferase; AST = aspartate aminotransferase.

The number and percent of patients with either ALT or AST ≥3 times the upper limit of normal and total bilirubin ≥2 times the upper limit of normal at concurrent and nonconcurrent visits will also be presented.

The number and percent of patients with either ALT or AST ≥3 times the upper limit of normal and total bilirubin ≥2 times the upper limit of normal Alkaline Phosphatase < 2 times the upper limit of normal at concurrent visits will also be presented.

An e-DISH scatter plot of maximum ALT vs maximum total bilirubin on study will also be presented. Of note, the protocol specified criteria of potential drug-induced liver injury (Hy's law) is applied in the evaluation of potential drug-induced liver injury cases.

6.8.4. Events of Special Interest

The following TEAEs of special interests will be summarized for the overall safety population and by SOC and PT for each treatment group.

- Falls
- Fatigue
- Fracture
- Secondary primary malignancies
- Convulsions (seizure)
- Neutrophil count decreased Loss of consciousness
- Cognitive and memory impairment
- Hypertension
- Hepatic disorder
- Ischemic heart disease (IHD)
- Other selected cardiovascular events
- Posterior reversible encephalopathy syndrome (PRES)
- Renal disorder
- Thrombocytopenia
- Musculoskeletal Events
- Severe cutaneous adverse reaction (SCAR)
- Angioedema
- Rash

Additionally, for the following special adverse events, Kaplan-Meier methods will be used to display the time to the first occurrence of each event.

- Secondary primary malignancies

- Hypertension
- Selected cardiovascular events

The TEAEs of special interests will be defined based on a list of MedDRA Preferred Terms confirmed by Pfizer Pharmacovigillance. A final list will be provided to programming prior to database release.

Appendix 2 provides the definition of AEs of special interests.

The treatment emergent hypertension incidence will be summarized by baseline charatristics including baseline history of dyslipidemia, diabetes mellitus, cardiovascular, hypertension, BMI and age subgroup. The associated list for dyslipidemia and hyperglycemia standardized MedDRA queries (SMQs) will be provided to programming prior to database release.

6.8.5. Vital Signs

Temperature, blood pressure (systolic and diastolic), and heart rate will be summarized at baseline and each subsequent scheduled assessment for the safety population. Change from baseline will be calculated and presented for each parameter at all scheduled postbaseline assessment timepoints.

The number and proportion of patients experiencing potentially clinically significant abnormalities during the treatment-emergent period will be summarized by treatment group. The definitions of potentially clinically significant abnormalities are shown in Table 12.

Table 12. Potentially Clinically Significant Abnormalities in Vital Signs

Parameter	Criteria for Potentially Clinically Significant Abnormalities
Systolic Blood Pressure	Absolute result >180 mm Hg and increase from baseline >40 mm Hg
	Absolute result <90 mm Hg and decrease from baseline >30 mm Hg
	Final visit or 2 consecutive visits results ≥10 mm Hg change from baseline
	Final visit or 2 consecutive visits results ≥15 mm Hg change from baseline
	Final visit or 2 consecutive visits results ≥20 mm Hg change from baseline
	Final visit or most extreme result ≥140 mm Hg
	Final visit or most extreme result ≥180 mm Hg
	Final visit or most extreme result ≥140 mm Hg and ≥20 mm Hg change from baseline
	Final visit or most extreme result ≥180 mm Hg and ≥20 mm Hg change from baseline
Diastolic Blood Pressure	Absolute result >105 mm Hg and increase from baseline >30 mm Hg
	Absolute result <50 mm Hg and decrease from baseline >20 mm Hg
	Final visit or 2 consecutive visits results ≥5 mm Hg change from baseline
	Final visit or 2 consecutive visits results ≥10 mm Hg change from baseline

Table 12. Potentially Clinically Significant Abnormalities in Vital Signs

Parameter	Criteria for Potentially Clinically Significant Abnormalities
	Final visit or 2 consecutive visits results ≥15 mm Hg change from baseline
	Final visit or most extreme result ≥ 90 mm Hg
	Final visit or most extreme result ≥105 mm Hg
	Final visit or most extreme result ≥90 mm Hg and ≥15 mm Hg change from baseline
	Final visit or most extreme result ≥105 mm Hg and ≥15 mm Hg change from baseline
Heart Rate	Absolute result >120 bpm and increase from baseline >30 bpm
	Absolute result <50 bpm and decrease from baseline >20 bpm

bpm = beats per minute; mm Hg = millimeters of mercury.

6.8.6. Electrocardiogram

Electrocardiograms are only performed at baseline; therefore, findings will be listed.

6.8.7. Physical Examination

Abnormal findings in physical examination will be listed.

6.9. Sensitivity Analyses for Potential COVID-19 Pandemic Related Issues

The sensitivity analyses in this section are planned to mitigate COVID-19 pandemic potential impact if applicable.

The assessments of potential impact on efficacy and safety analyses in Sections 6.1, 6.2, 6.3, 6.5 and 6.8, include the listings/summaries for AEs associated with COVID-19, discontinuation from treatment and study associated with COVID-19, protocol deviations associated with COVID-19.

Additionally, the following will be evaluated for the potential impact of COVID-19 and summarized by treatment arm if appropriate:

- Deaths due to COVID-19
- Missing and/or delayed assessment for radiographic/PSA progression due to COVID-19 (summary of the cases with unconfirmed radiographic progression or PSA progression, or cases with long gap of missing 2 consecutive scans before disease progression between treatment arms by the reasons due to COVID-19 or not)
- Changing ePRO data collection method to alternative phone interview.

Based on the above assessments, additional sensitivity analyses could be considered as follows:

- For the TTE endpoints in Sections 6.1, 6.2 and 6.3:
  - Censor deaths due to COVID-19
  - For imaging-based TTE endpoints, use alternative definition (for example, no censoring related to 2 or more consecutive missed/delayed tumor assessment visits due to COVID-19)
- For the analysis of quality of life endpoints in Section 6.5:
  - Consider the alternative collection method as additional covariate in modeling

6.9.1. COVID-19 Anchor Date

If additional analyses are needed to assess the impact of COVID-19 on the study population and study data, an anchor date will be used as a start date for COVID-19 pandemic based on Pfizer guidance and standard operating procedure (SOP):

For global pandemic reference date: use the date the World Health Organization designated COVID-19 as a global pandemic - March 11, 2020

When producing data summaries intended to show the potential impacts of COVID-19 on the study, data may be presented as “before” and “during,” where the anchor date is included in the “during” group.

A different anchor date may be used for purposes of regulatory submission should the regulatory authority requests.

7. REFERENCES

1. Amit O, Mannino F, Stone AM, et al. Blinded independent central review of progression in cancer clinical trials: results from a meta-analysis. European Journal of Cancer 2011;47:1772-1778.

2. Brookmeyer R, Crowley JJ. A confidence interval for the median survival time. Biometrics 1982;38:29-41.

3. Cella D. Manual of the functional assessment of chronic illness therapy (FACIT) measurement system. Center on Outcomes, Research and Education (CORE), Evanston Northwestern Healthcare and Northwestern University, Evanston IL, Version 4, 1997.

4. Cella DF, Tulsky DS, Gray G, Sarafian B, Linn E, Bonomi A, Silberman M, Yellen SB, Winicour P, Brannon J. The Functional Assessment of Cancer Therapy scale: development and validation of the general measure. J Clin Oncol. 1993;11:570–579.

5. Cleeland CS 2006. The measurement of pain from metastatic bone disease: Capturing the patient's experience. Clinical cancer research 12(20 Pt 2), 6236s–6242s.

6. Cleeland CS 2009. The Brief Pain Inventory User Guide https://www.mdanderson.org/documents/Departments-and-Divisions/Symptom-Research/BPI_UserGuide.pdf

7. Esper P, Mo F, Chodak G, Sinner M, Cella D, Pienta KJ. Measuring quality of life in men with prostate cancer using the functional assessment of cancer therapy-prostate instrument. Urology. 1997;50:920–928.

8. Esper P, Hampton JN, Smith DC, Pienta KJ. Quality-of-life evaluation in patients receiving treatment for advanced prostate cancer. Oncol Nurs Forum. 1999;26:107–112.

9. Van Andel G, Bottomley A, Aaronson NK, et al. An international field study of the EORTC QLQ-PR25: A questionnaire for assessing the health-related quality of life of patients with prostate cancer, European Journal of Cancer, Volume 44, Issue 16, 2008, Pages 2418-2424.

10. Haybittle J. Repeated assessment of results in clinical trials of cancer treatment. Brit J Radiol 1971;44:793–797.

11. Herdman M, Gudex C, Badia X, et al. Qual Life Res 2011 Dec;20(10):1727-1736.

12. Janssen MF, Pickard AS, Swinburn P, et al. Busschbach J Qual Life Res 2013 Sep;22(7):1717-1727.

13. Kalbfleisch JD, Prentice RL. Statistical analysis of failure time data, 2nd Edition. Hoboken, Wiley Interscience; 2011.

14. Latimer NR, Abrams KR. NICE DSU Technical Support Document 16: Adjusting survival time estimates in the presence of treatment switching. 2014.

15. Lu K, Li D, and Koch G. Comparison between 2 controlled multiple imputation methods for sensitivity analyses of time-to-event data with possibly informative censoring. Statistics in Biopharmaceutical Research, 2015;Vol. 7, 199-213.

16. Peto R, Pike MC, Armitage P, et al. Design and analysis of randomized clinical trials requiring prolonged observation of each patient. I. Introduction and design. Br J Cancer 1976;34(6):585–612.

17. Pound CR, Partin AW, Eisenberger MA, et al. Natural history of progression after PSA elevation following radical prostatectomy. JAMA 1999;281(17):1591-7.

18. Robins JM, Tsiatis AA. Correcting for non-compliance in randomized trials using rank preserving structural failure time models. Commun Stat Theory Methods 1991;20(8):2609-31.

19. Robins JM, Finkelstein DM. Correcting for noncompliance and dependent censoring in an AIDS clinical trial with inverse probability of censoring weighted (IPCW) log rank tests. Biometrics 2000;56(3):779-788.

8. LIST OF ABBREVIATIONS

Abbreviation	Term
AE	adverse event
ALT	alanine aminotransferase
AST	aspartate aminotransferase
ATC	anatomic therapeutic chemical
BICR	Blinded Independent Central Review
BPD	Bone metastases
BPDu	Bone metastases - unconfirmed
BPI-SF	Brief Pain Inventory (Short form)
CI	confidence interval
COVID-19	coronavirus disease of 2019
CRF	case report form
CSR	clinical study report
CTCAE	Common Terminology Criteria for Adverse Events
ECOG	Eastern Cooperative Oncology Group
EDR	early discrepancy rate
eITT	evaluable intent-to-treat
EQ-5D-5L	European Quality of Life 5-Dimensions 5-Level Health Questionnaire
EORTC	The European Organization for Research and Treatment of Cancer
EOS	end of study
EOT	end of treatment
FACIT	functional assessment of chronic illness therapy
FACT-P	functional assessment of cancer therapy-prostate
HLGT	high-level group term
HLT	high-level term
HR	Hazard ratio
IA	interim analysis
IHD	ischemic heart disease
INV	Investigator
IPCW	inverse probability of censoring weighting
IRT	interactive response technology
ITT	intent-to-treat
LDR	late discrepancy rate
LN	natural logarithm
MedDRA	Medical Dictionary for Regulatory Activities
MFS	metastatic free survival
MMRM	Mixed-Model Repeated Measures
mTEAE	modified treatment emergent adverse events
NCI	National Cancer Institute
NA	not applicable
NBM	no bone metastatses
ND	no disease; no bone metastases (as applicable)

Abbreviation	Term
NE	not evaluable
NN	no progression
OS	overall survival
PET	positron emission tomography
PD	progressive disease
PDu	bone metastasis – unconfirmed
PFS	progression-free survival
PFS2	Progression-free survival on first subsequent therapy
PRO	patient-reported outcome
PRES	posterior reversible encephalopathy syndrome
PSA	prostate-specific antigen
PSMA	prostate-specific membrane antigen
PT	preferred term
Q1	first quartile
Q3	third quartile
QLQ-PR25	Quality of Life Questionnaire-Prostate 25
QoL	Quality of Life
RECIST	Response Evaluation Criteria in Solid Tumors
REML	restricted maximum likelihood
RPSFTM	Rank-Preserving Structural Failure Time Model
rTEAE	reinitiation treatment emergent adverse events
SAP	statistical analysis plan
SCAR	Severe cutaneous adverse reaction
SD	standard deviation
SMQ	Standardized MedDRA Query
SOC	system organ class
SOP	standard operating procedure
SSE	symptomatic skeletal event
TEAE	treatment emergent adverse events
TNM	tumour, node, metastasis
TTE	Time-to-event
TURP	transurethral resection of the prostate
WHO	World Health Organization

9. APPENDICES

Appendix 1. Definition of Geographic Regions

Three major geographic regions are defined below:

Region	Countries
North America	United States, Canada
Europe	All European countries and Russia
Rest of the world	All countries not included in North America and Europe

Appendix 2. Definition of Adverse events of special interests

Event Grouping of Interest	Definition
Convulsions (seizure)	Narrow SMQ of "Convulsions".
Hypertension	Narrow SMQ of "Hypertension".
Neutrophil count decreased	Preferred terms of "Neutrophil count decreased", "Neutropenia", "Agranulocytosis", "Granulocyte count decreased", "Granulocytopenia", "Febrile neutropenia", "Neutrophil percentage decreased", "Band neutrophil count decreased", and "Band neutrophil percentage decreased", "Neutropenic sepsis", "Neutropenic infection", "Neutrophil count abnormal"
Cognitive and memory impairment	All preferred terms in MedDRA high-level group term (HLGT) "mental impairment disorders"
Ischemic heart disease (IHD)	Narrow SMQs of "myocardial infarction" and "other ischemic heart disease"
Other selected cardiovascular events	Narrow SMQs of "haemorrhagic central nervous system vascular conditions", "ischaemic central nervous system vascular conditions" and "cardiac failure"
Posterior reversible encephalopathy syndrome (PRES)	Preferred term of posterior reversible encephalopathy syndrome
Fatigue	Preferred terms of fatigue, asthenia
Renal disorder	Broad SMQ of "Acute Renal Failure"
Second primary malignancies*	Narrow SMQs of "Malignant or unspecified tumours" customized to exclude preferred terms of "Congenital fibrosarcoma" , "Congenital malignant neoplasm" , "Congenital retinoblastoma" , ", "Metastases to…" , "Metastasis" , "Metastatic neoplasm" , "Prostate cancer…" , "Carcinoid tumour of the prostatec, and "Neoplasm prostate" AND (inclusive of) Narrow SMQ of "Myelodysplastic syndrome" AND (inclusive of) All preferred terms under HLT of "Myeloproliferative disorders (excl leukaemias)" Note: Non-melanoma skin cancers are excluded (preferred terms of "Basal cell carcinoma" , "Basosquamous carcinoma" , "Basosquamous carcinoma of skin" , "Keratoacanthoma" , "Skin cancer" , "Skin cancer metastatic" , "Squamous cell carcinoma" , "Squamous cell carcinoma of skin" , "Lip squamous cell carcinoma", "Bowen's disease")
Fall	Preferred term of "Fall"
Fracture	All preferred terms under the MedDRA HLGT: "Fractures" , "Bone and Joint Injuries"
Loss of consciousness	Preferred terms of "Loss of consciousness" , "Syncope" , and "Presyncope"
Thrombocytopenia	Preferred terms of Thrombocytopenia, Platelet count decreased.

Event Grouping of Interest	Definition
Musculoskeletal events	Preferred terms of Back pain, Arthralgia, Myalgia, Musculoskeletal pain, Pain in extremity, Musculoskeletal stiffness, Muscular weakness, Muscle spasms
Severe cutaneous adverse reactions (SCAR)	Narrow SMQ of "Severe cutaneous adverse reactions"
Angioedema	Narrow SMQ of "Angioedema"
Rash	Preferred terms of "Butterfly rash", "Exfoliative rash", "Eyelid rash", "Genital rash", "Heliotrope rash", "Mucocutaneous rash", "Nodular rash", "Paraneoplastic rash", "Penile rash", "Perineal rash", "Rash", "Rash erythematous", "Rash follicular", "Rash macular", "Rash maculo-papular", "Rash maculovesicular", "Rash morbilliform", "Rash neonatal", "Rash papular", "Rash papulosquamous", "Rash pruritic", "Rash pustular", "Rash rubelliform", "Rash scarlatiniform", "Rash vesicular", "Septic rash", "Systemic lupus erythematosus rash", "Vasculitic rash", "Viral rash", "Vulvovaginal rash"
Hepatic disorder	Narrow SMQs of "hepatic failure, fibrosis and cirrhosis and other liver damage related conditions", "hepatitis, non-infectious" and "liver related investigations, signs and symptoms"
SMQ: standardized MedDRA query, PRES: posterior reversible encephalopathy syndrome, SCAR: severe cutaneous adverse reactions, PT: preferred term, HLT: high-level term, HLGT: high-level group term. Broad SMQ of "Acute Renal Failure" includes either broad or narrow SMQ of "Acute Renal Failure". *AE terms identified as second primary malignancies according to the above search criteria undergo medical review to confirm.	