PROTOCOL MDV3100-14 (C3431005)

A Multinational, Phase 3, Randomized, Double-Blind, Placebo-Controlled, Efficacy and Safety Study of Enzalutamide in Patients with Nonmetastatic Castration-Resistant Prostate Cancer

Statistical Analysis Plan

Version: 5.0

Date: 18-November-2019

[redacted]

TABLE OF CONTENTS

CCI [redacted]

List of Tables

List of Appendices

CCI [redacted]

LIST OF ABBREVIATIONS AND TERMS

Abbreviation	Definition
ALT	Alanine aminotransferase
AST	Aspartate aminotransferase
BPI-SF	Brief Pain Inventory Short Form
CT	Computed tomography
CTCAE	Common Terminology Criteria for Adverse Events
DMC	Data Monitoring Committee
ECOG	Eastern Cooperative Oncology Group
EORTC	European Organization for Research and Treatment of Cancer
EQ-5D-5L	European Quality of Life 5 Dimensions-5 Levels health questionnaire
FACT-P	Functional Assessment of Cancer Therapy-Prostate
ICR	Independent Central Radiology review
ITT	Intent-to-treat
LOCF	Last observation carried forward
MedDRA	Medical Dictionary for Regulatory Activities
MFS	Metastasis-free survival
MRI	Magnetic resonance imaging
NCI	National Cancer Institute
OS	Overall survival
PSA	Prostate-specific antigen
PT	Preferred term
QLQ-PR25	Quality of Life Questionnaire-Prostate 25 module
SAP	Statistical analysis plan
SOC	System organ class
TEAE	Treatment-emergent adverse event

1. INTRODUCTION

This statistical analysis plan (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 MDV3100-14 (C3431005): A Multinational, Phase 3, Randomized, Double-Blind, Placebo-Controlled, Efficacy and Safety Study of Enzalutamide in Patients with Nonmetastatic Castration-Resistant Prostate Cancer. This SAP is based on protocol version 5.0 dated 26 January 2018.

2. STUDY OVERVIEW

Protocol MDV3100-14 is a multinational, Phase 3, randomized, double-blind, placebo-controlled efficacy and safety study in patients with nonmetastatic castration-resistant prostate cancer. The primary objective is to determine the efficacy of enzalutamide (formerly MDV3100) (160 mg/day by mouth) compared with placebo as assessed by metastasis-free survival (MFS). Patients will maintain androgen deprivation during the study, either using a gonadotropin-releasing hormone (GnRH) agonist/antagonist or will have a history of bilateral orchiectomy.

A total of 1440 patients will be centrally randomized in a 2:1 ratio to enzalutamide or placebo treatment within each stratum of prostate-specific antigen (PSA) doubling time (< 6 months vs. ≥ 6 months) and prior or current use of bone targeting agents (yes vs. no).

PSA doubling time will be calculated by the sponsor, using the method of Pound et al, 1999.[1] 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:

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

The study consists of screening, double-blind treatment, safety follow-up, and long-term follow-up periods. Radiographic assessments will be performed as described in Appendix 1 approximately every 16 weeks. Patients will continue study drug until radiographic progression. Radiographic progression will be confirmed by independent central radiology review before radiographic imaging stops.

In addition to imaging, the following assessments of prostate cancer status will be made during the study: survival status, PSA values, use of new antineoplastic therapy, pain intensity and interference using the Brief Pain Inventory Short Form (BPI-SF), and quality of life as assessed by the Functional Assessment of Cancer Therapy-Prostate (FACT-P) questionnaire, European Quality of Life-5 Dimensions-5 Levels health questionnaire (EQ-5D-5L) and the European Organization for Research and Treatment of Cancer (EORTC) Quality of Life Questionnaire-Prostate 25 (QLQ-PR25) module.

Safety will be assessed by monitoring adverse events, clinical laboratory tests, physical examinations, and vital signs during the treatment period through 30 days after the last dose of study drug or initiation of a new antineoplastic treatment, whichever occurs first. An independent Data Monitoring Committee (DMC) will periodically monitor the safety data.

During long-term follow-up, patients will be monitored for survival status, new antineoplastic therapies for prostate cancer, opiate medications, skeletal-related events, and interventions due to locoregional metastasis (eg, radiation, transurethral resection of the prostate, nephrostomy tube placement).

As the study met the primary efficacy endpoint of metastasis-free survival and the established safety profile of enzalutamide was confirmed based on the data cutoff on 28 June 2017, all patients were unblinded. Eligible patients from the placebo group were offered enzalutamide at the discretion of the investigator. For patients who had not progressed radiographically, scans (CT (computed tomography)/MRI (magnetic resonance imaging) and bone scan) were to be performed per investigator discretion until patient has progressed radiographically. Continuation of treatment on the open-label period after radiographic progression were at the discretion of the investigator. Patients who did not participate in the open-label period or withdrew consent for further treatment would continue long-term follow-up assessments per protocol. Treatment with open-labeled enzalutamide were to be stopped upon disease progression when in the opinion of the Investigator, there was no added clinical benefit to continue treatment with enzalutamide.

Long-term follow-up data (including survival status, new antineoplastic therapies for prostate cancer, skeletal-related events, and interventions due to locoregional progression) were to be collected every 16 weeks.

3. STUDY OBJECTIVES

The primary study objective is to determine the efficacy of enzalutamide compared with placebo as assessed by MFS.

The secondary objectives are as follows:

- To evaluate the benefit of enzalutamide compared with placebo as measured by the following:
  - Time to PSA progression;
  - Time to first use of new antineoplastic therapy;
  - Overall Survival (OS);
  - Time to pain progression;
  - Time to first use of cytotoxic chemotherapy;
  - Chemotherapy-free disease-specific survival;

- Chemotherapy-free survival;
- PSA response rates;
- Quality of life as assessed by the Functional Assessment of Cancer Therapy-Prostate (FACT-P) questionnaire, European Quality of Life-5 Dimensions-5 Levels (EQ-5D-5L) health questionnaire and Quality of Life Questionnaire-Prostate 25 (QLQ-PR25) module.

- To evaluate safety.

No exploratory objectives are planned; exploratory analyses will be performed as deemed necessary to assess the benefit-risk profile of enzalutamide.

4. STUDY ENDPOINTS

4.1. Primary Endpoint

The study has a single primary efficacy endpoint of MFS assessed by blinded Independent Central Radiology Review (ICR).

4.2. Secondary Endpoints

Secondary endpoints are as follows:

- Efficacy

  Key secondary endpoints:

  - Time to PSA progression;
  - Time to first use of new antineoplastic therapy;
  - Overall Survival;

  Additional secondary endpoints:

  - Time to pain progression;
  - Time to first use of cytotoxic chemotherapy;
  - Chemotherapy-free disease-specific survival;
  - Chemotherapy-free survival;
  - PSA response;
  - Quality of Life as assessed by the FACT-P questionnaire, EQ-5D-5L health questionnaire, and QLQ-PR25 module.

- Safety.

4.3. Exploratory Efficacy Endpoints

There is no predefined exploratory endpoint.

5. SAMPLE SIZE CONSIDERATIONS

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

- 2:1 enzalutamide to placebo treatment allocation;
- An increasing nonuniform accrual of 0.25 patients per month per site with maximum accrual of 63 patients per month;
- For MFS a target hazard ratio of 0.72 at a 2-sided significance level of 0.05 with 90% power. Under an exponential model assumption, the target difference in Kaplan-Meier estimated median is 9 months (24 months versus 33 months). The median MFS of 24 months for the placebo group is based on published data.[2]

A total of 440 events provide 90% power to detect a target hazard ratio of 0.72 based on a 2-sided log-rank test and the overall significance level of 0.05. A sample size of approximately 1305 patients (870 enzalutamide and 435 placebo) will achieve 440 events. Approximately 10% of patients enrolled are expected to be lost to follow-up, will be found to have metastatic disease at study entry, or will have events censored due to required analytical methods, so approximately 1440 patients (960 enzalutamide and 480 placebo) will be enrolled. The time from date of first randomization until 440 MFS events are observed is estimated to be approximately 43 months.

Approximately 500 PSA progression events are expected at the time of the single MFS analysis. Based on a hazard ratio assumption of 0.60, this endpoint will have >95% power at a 2-sided significance level of 0.02 (see Section 9.7.1 for details on multiplicity adjustments on key secondary endpoints).

Approximately 360 new antineoplastic therapy events are expected at the time of the single MFS analysis. This endpoint will have 80% power to detect a hazard ratio of 0.70 at a 2-sided significance level of 0.02 (see Section 9.7.1 for details on multiplicity adjustments on key secondary endpoints).

The study is also powered for OS. Specifically, 590 death events will be required to have 85% power to detect a hazard ratio of 0.77 at a 2-sided significance level of 0.05. Under an exponential model assumption, the target difference in Kaplan-Meier estimated medians is 13.7 months (46 months vs. 59.7 months). If either time to PSA progression or time to new antineoplastic therapy endpoint fails to show significance, OS will be tested at a 2-sided significance level of 0.03. Under this scenario, 590 death events will provide ~79% power to detect a hazard ratio of 0.77 (see Section 9.7.1 for details on multiplicity adjustments on key secondary endpoints).

The sample size and power calculations were performed using the software package East 5.4 (Cytel Inc., Cambridge, MA).

6. ANALYSIS POPULATIONS

6.1. Efficacy Populations

The intent-to-treat (ITT) population is defined as all patients randomly assigned to study treatment and is based on randomized treatment assignment regardless of whether or not treatment was administered. Unless otherwise specified, all efficacy analyses will use the ITT population.

6.2. Safety Population

The safety population is defined as all randomized patients who receive at least 1 dose or partial dose of study drug (enzalutamide or placebo). Unless otherwise specified, all safety analyses will use the safety population according to the actual treatment received (not the treatment assigned).

7. DEFINITIONS, COMPUTATIONS, AND CONVENTIONS

7.1. Definitions and Computations

Study Day

Study day will be calculated in reference to the date of randomization (Study Day 1). For assessments conducted on or after the randomization date, study day is calculated as (assessment date – randomization date + 1). For assessments conducted before the randomization date, study day is calculated as (assessment date – randomization date). There will be no Study Day 0.

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

The date of the first dose of study drug is defined as the date a patient receives the first dose of study drug (enzalutamide or placebo). The date of the last dose of study drug is defined as the date a patient receives the last dose of study drug (enzalutamide or placebo).

Treatment Day

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

Treatment Period

The treatment period is defined as the period of time from the date and time of the first dose of study drug through the date of last dose.

Treatment-Emergent Period

The treatment-emergent period is defined as the period of time from the date and time of the first dose of study drug through the date of last dose + 30 days (or the day before initiation of a new antineoplastic treatment, whichever occurs first).

Treatment Duration

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

Treatment Duration (days) = Date of last dose of study drug – Date of first dose of study drug + 1.

Treatment Duration (months) = (Date of last dose of study drug – Date of first dose of study drug + 1)/30.4375.

Long-Term Follow-up Period

The long-term follow-up period will start the day after the last day of study drug treatment and will continue through the data cutoff date.

Baseline Value and Postbaseline Value

Unless otherwise specified, the baseline value is defined as the last measurement before the first administration (date and time) of study drug. If time is not available, then date only will be used. Postbaseline value is defined as a measurement taken after the first administration of study drug. Change from baseline is defined as (postbaseline value – baseline value). For laboratory assessments and vital signs, only date and time of study drug administration and measurement will be considered when calculating baseline value. If time is not available, then date only will be used.

New Antineoplastic Therapy, Cytotoxic Chemotherapy and Bone-Targeting Agent Classification

All treatment modalities study participants receive during the study are collected via the Concomitant Medications CRF page. As these classifications for concomitant medications are critical for study results, the study medical monitor will independently review each concomitant medication and determine whether they belong to one of these categories. The classification process will only consider the medication name (by preferred term), and will be performed blinded to treatment. The study medical monitor will sign the completed list of classifications, and the list will be documented in the study folder. No changes to this classification will be allowed after study unblinding.

7.2. 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 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 time-to-event right-censored 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 Medical Dictionary for Regulatory Activities (MedDRA);
- Prior therapies and concomitant medications will be coded using the World Health Organization (WHO) Drug Dictionary.

7.2.1. Rules for Missing Data

By-Visit Endpoints

All by-visit endpoints will be analyzed either as observed data. Missing data will not be imputed and only the observed records will be included.

Adverse Events and Concomitant Medications

When the onset/start date is a partial date, the date will be imputed to determine treatment emergence of adverse events and whether medication is prior or concomitant (or both). The following rules will be applied to impute onset/start date:

- When year and month are present:
  - If year and month are same as the first dose date, then use the first dose date;
  - If year and month are before the first dose date, then use the last day of the year and month;
  - If year and month are after the first dose date, then use the first day of the year and month.
- When only year is present:
  - If year is the same as the first dose date, then use the first dose date;
  - If year is before the first dose date, then use the last day of the year;
  - If year is after the first dose date, then use the first day of the year.
- If date, month, and year are all missing, do not impute.

7.2.2. 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 1 and Table 2.

Table 1: Visit Windows for Safety Assessments

Week	Start Day	Target Day	End Day
5	2	33	70
17	71	113	168
33+ Every 16 Weeks	(Week-1) × 7 – 55	(Week-1) × 7 + 1	(Week-1) × 7 + 56
Safety Follow-up	Last Dosing Date	Last Dosing Date + 30	Last Dosing Date + 30

If an assessment is located both in Safety Follow-up and a regular analysis window, it will only be assigned to Safety Follow-up Visit, and will be reported as such.

Table 2: Visit Windows for Efficacy Assessments

Week	Start Day	Target Day	End Day
17	2	113	168
33+ Every 16 Weeks	(Week-1) × 7 – 55	(Week-1) × 7 + 1	(Week-1) × 7 + 56

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.

8. TIMING OF ANALYSES

The single MFS analyses will be performed after approximately 440 MFS events occur. All secondary endpoints will be evaluated at this time. This will include the single analysis of time to PSA progression and time to first use of new antineoplastic therapy as well as the first interim analysis of overall survival. Approximately 135 death, 500 PSA progression, and 360 new antineoplastic therapy events are expected at the time of this analysis.

Two additional interim analyses and the final analysis of overall survival are planned after approximately 285, 440, and 596 deaths occur, respectively. Additional analyses of the primary or other secondary endpoints are not planned. A multiplicity adjusted inferential procedure will be used to maintain the family-wise 2-sided type I error rate at 0.05. Details on Type I error rate control methodology for the interim and final overall survival analyses are provided in Section 9.7.1 and Section 9.7.3. If an interim OS analysis is statistically significant, it will be reported as the final analysis and no subsequent analyses will be performed.

9. STATISTICAL METHODS

Efficacy analyses will be conducted using the ITT population. Efficacy data will be analyzed according to treatment assignment at randomization. Safety analyses will be conducted using the safety population according to the treatment actually received.

9.1. Patient Disposition

Patient disposition will be summarized by randomized treatment group as follows:

- Randomized patients by randomization stratification factors;
- Patients in the ITT, and safety populations and the reason for exclusion from 1 or more populations;
- Patients in the treatment and long-term follow-up periods;
- Patients discontinuing treatment and the reasons for discontinuation;
- Patients withdrawing from long-term follow-up and the reasons for withdrawal.

The following listings will be produced:

- Patients who received treatment different from the treatment randomized;
- Patients discontinuing from each study period after enrollment.

9.2. Protocol Deviations

Patients with major protocol deviations will be listed and summarized by treatment group as randomized. Categories of major deviations are as follows:

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

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.

9.3. Demographics and Baseline Characteristics

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

- Age, race, weight, and body mass index;
- Geographic region (North America, European Union, rest of world);
- Categorized PSA doubling time (< 6 months, ≥ 6 months);
- Baseline use of a bone targeting agent (yes, no);
- Eastern Cooperative Oncology Group (ECOG) performance status;
- Serum PSA;
- PSA doubling time (continuous);
- Quality of life (EQ-5D-5L health score and FACT-P global score);
- Selected laboratory parameters including: testosterone, absolute neutrophil count, platelet count, hemoglobin, total bilirubin, alanine aminotransferase (ALT), aspartate aminotransferase (AST);
- History of cardiovascular disease as defined in Appendix 2.

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

9.4. Disease Characteristics and Previous Therapies

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]);

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 therapy (yes/no) at baseline;
- History of surgical prostate cancer procedures (yes/no);
- Type of surgical procedures (prostatectomy, orchiectomy, transurethral resection of the prostate [TURP], cryoablation, and other).

9.5. Extent of Exposure and Compliance of Study Drug

Data from patients in the safety population will be used to summarize the extent of exposure, compliance, and dose modification of study drug by treatment group as treated.

Treatment duration, total number of capsules taken, and percent overall compliance rate will be calculated and summarized by treatment group as treated. Treatment duration will be calculated as (date of last dose of study drug – date of first dose of study drug + 1). For patients who are continuing study drug at the analysis cutoff date, the last known date of study drug taken will be used as date of last dose of study drug. Treatment duration will be summarized both as a continuous measure and a categorical measure ($\leq$ 3 months, 3 to $<$ 6 months, 6 to $<$ 12 months, $\geq$ 12 months). 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 subjects who have dose modifications (eg, dose reduction or dose withholding), 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%.

Patients with at least 1 dose modification (including dose reduction and dose withholding) and the reason for the dose modifications will be summarized by treatment group as treated.

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

9.6. Concomitant Medications

Concomitant medications taken during the study treatment period will be summarized for all patients in the safety population by treatment group as treated. Medications are considered concomitant if exposure occurs during the treatment-emergent period.

Subsequent therapies taken after study drug discontinuation will be summarized by treatment group and overall. In addition, subsequent antineoplastic therapies will be summarized separately by treatment group.

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

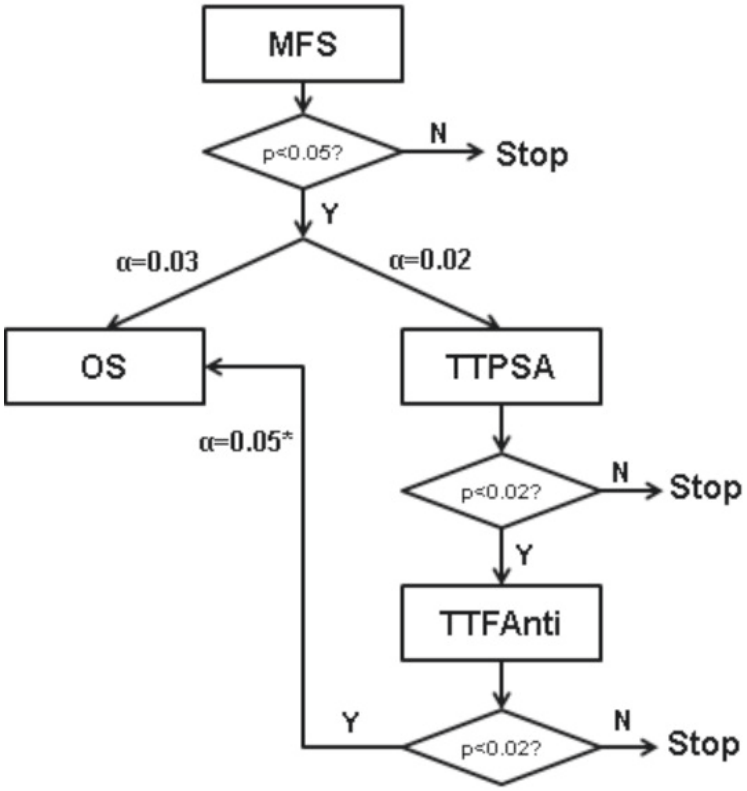
9.7. Efficacy Analyses

The single MFS analysis will be performed after approximately 440 MFS events occur. All secondary endpoints will be evaluated for efficacy at this time. This will include the single analysis of time to PSA progression and time to first use of new antineoplastic therapy as well as the first interim analysis of overall survival. Approximately 135 death, 500 PSA progression and 360 new antineoplastic therapy events are expected at the time of this analysis. Two additional interim analyses and the final analysis of overall survival are planned after approximately 285, 440, and 596 death events occur, respectively. No additional analyses of other efficacy endpoints are planned at the time of the additional interim and final analyses of overall survival. If an interim analysis of overall survival is statistically significant, it will be reported as the final analysis and no subsequent analyses will be performed.

9.7.1. Multiplicity Adjustment for Efficacy Analysis

Adjustment for multiplicity will be considered for MFS based on ICR assessment, time to PSA progression, time to first use of new antineoplastic therapy, and OS. All secondary endpoint analyses will be performed at the time of the single MFS analysis. To maintain the family-wise 2-sided type I error rate at 0.05, a parallel testing strategy between overall survival (with allocated type I error rate 0.03) and remaining key secondary endpoints (time to PSA progression and time to first use of new antineoplastic therapy with allocated type I error rate 0.02) will be performed. Testing strategy for primary and key secondary endpoints is summarized in Figure 1.

Figure 1. Testing Strategy for Primary and Key Secondary Endpoints



MFS, metastasis-free survival; OS, overall survival; TTPSA, time to prostate-specific antigen progression; TTFAnti, time to first use of new antineoplastic therapy.

** Overall survival will be tested at 0.05 only if both time to PSA progression and time to new antineoplastic therapy endpoints are significant.*

Details of primary and key secondary endpoints testing as a step-by-step approach will be as follows:

1. Compute p-value for MFS. If p-value < 0.05, declare statistical significance for MFS and go to step 2 below. Otherwise stop.

2. Compute p-value for time to PSA progression. If p-value < 0.02, declare statistical significance for time to PSA progression and go to step 3 below. Otherwise declare time to PSA progression failed to show significance, continue with step 4 below and test OS at significance level 0.03.

3. Compute p-value for time to first use of new antineoplastic therapy. If p-value < 0.02, declare statistical significance for time to first use of new antineoplastic therapy, continue with step 4 below and test OS at significance level 0.05. Otherwise, declare time to first use of new antineoplastic therapy failed to show significance, continue with step 4 below and test OS at significance level 0.03.

4. Compare OS with group sequential testing methodology based on 3 interim analyses and 1 final analysis (described in Section 9.7.3) at a significance level of 0.05 (if both time to PSA progression in step 2 and time to new antineoplastic therapy in step 3 are declared significant) or 0.03 (if either time to PSA progression in step 2 or time to new antineoplastic therapy in step 3 failed to show significance). The O'Brien-Fleming alpha spending function will be used to determine the stopping boundaries based on the number of events observed at the interim looks to control the overall 2-sided alpha at 0.05 or at 0.03.[3]

All other efficacy analyses (including sensitivity analyses for primary, and other secondary endpoints) and associated p-values will be deemed exploratory, for which no adjustment for multiplicity will be used.

The key efficacy analyses, order of testing, and multiplicity adjustment rules are summarized in Table 3.

Table 3: Key Efficacy Analyses and Multiplicity Adjustment

Order of Testing	Population	Declare Significant	Statistic	Strata
1. Metastasis-free survival [1]	ITT	$p < 0.05$	Stratified log-rank test [2]	PSA doubling time Prior or current bone targeting agents
	ITT, subgroup [2]	NA		
2. Time to PSA progression	ITT	$p < 0.02$	Stratified log-rank test	
3. Time to first use of new antineoplastic therapy	ITT	$p < 0.02$	Stratified log-rank test	
4. Overall survival [3]	ITT	$p < 0.05$ or $p < 0.03$ [4]	Stratified log-rank test	

ITT: intent-to-treat; NA: not applicable; PSA: prostate-specific antigen.

1. MFS is the primary analysis.
2. The MFS subgroup analysis will be performed with an unstratified log-rank test.
3. Three interim analyses and 1 final analysis are planned for OS, details of each analysis are provided in Section 9.7.3.
4. If both time to PSA progression and time to first use of new antineoplastic therapy are declared significant, OS will be tested at a 0.05 significance level; otherwise, OS will be tested at 0.03.

9.7.2. Primary Efficacy Analysis of Metastasis-Free Survival

MFS will be evaluated for the ITT population. MFS is defined as the time from randomization to the first date of radiographic progression (assessed by blinded ICR) or death on study (death within 112 days of treatment discontinuation without evidence of radiographic progression), whichever occurs first. Radiographic assessments from both

scheduled and unscheduled assessments will be used to determine events in the primary analysis. For patients not known to have had radiographic progression or death at the time of the analysis data cutoff, MFS time will be censored at the date of the last available scan before the analysis data cutoff date for the purposes of analysis. Patients who were randomized but later confirmed to have metastatic disease before randomization by ICR assessment or patients with no postbaseline tumor assessment information will be censored on the date of randomization. Patients who initiate antiandrogen receptor agents (eg, bicalutamide, flutamide, or nilutamide) without evidence of metastasis as per ICR review will not be censored for the primary MFS analysis. The details of the censoring rules are provided in Appendix 1.

Significance Level

A 2-sided alpha of 0.05 will be preserved for the primary endpoint of MFS and key secondary endpoints using the testing strategy outlined in Section 9.7.1. The primary endpoint of MFS and all secondary endpoints will be analyzed when approximately 440 MFS events based on ICR occur. See Section 5 for additional details.

Methodology

MFS will be compared between the 2 treatment groups using a stratified log-rank test. The strata will be PSA doubling time (< 6 months vs. ≥ 6 months) and prior or current use of a bone targeting agent (yes vs. no). Kaplan-Meier curves will be used to estimate the distribution of duration of MFS. The 50th percentile of Kaplan-Meier estimates will be used to estimate the median duration of MFS. A 2-sided 95% confidence interval will be provided for this estimate. In cases where the median is not reached as of the data cutoff, the 25th percentile and its 2-sided 95% confidence interval will be provided. The progression and censoring reason will also be summarized for each treatment arm.

The null and alternative hypotheses regarding MFS can be phrased in terms of the hazard ratio $\lambda_{ArmA}/\lambda_{ArmB}$, where λ_{ArmA} represents the hazard of death for group A (enzalutamide) and λ_{ArmB} represents the hazard of death for group B (placebo). A hazard ratio of $\neq 1$ indicates that the MFS is prolonged for patients randomized to group A (enzalutamide) compared with patients randomized to group B (placebo). The null and alternative hypotheses, respectively, can be written as follows:

$$H_0: \frac{\lambda_{ArmA}}{\lambda_{ArmB}} = 1 \qquad H_A: \frac{\lambda_{ArmA}}{\lambda_{ArmB}} \neq 1$$

The hazard ratio, $\lambda_{ArmA}/\lambda_{ArmB}$, will be estimated using a stratified Cox regression model with the same strata as above.

If the estimate of $\lambda_{ArmA}/\lambda_{ArmB} < 1$ and the results from the stratified log-rank test lead to the rejection of H_0 in favor of H_A, then it will be concluded that enzalutamide prolongs MFS compared to placebo.

To evaluate the follow-up time for MFS between the 2 treatment groups, the median follow-up time for MFS will be estimated according to the Kaplan-Meier estimate of potential follow-up, also known as "Reverse Kaplan-Meier".[4]

Subgroup Analyses of Metastasis-Free Survival

Subgroup analyses of MFS will be performed to determine whether the treatment effect is concordant among subgroups. The same methodology as in the analysis of MFS above will be used for each subgroup. All subgroup analysis will be performed using an unstratified log-rank test and will be performed using MFS assessed by ICR. The hazard ratios and their 95% confidence intervals will be displayed in a forest plot. The following variables will be used to define subgroups:

- PSA doubling time (< 6 months, ≥ 6 months);
- Baseline use of a bone targeting agent (yes, no);
- Baseline age category at baseline (at or below median, above median);
- ECOG performance status at baseline (0, 1);
- Geographic region (North America, European Union, and rest of the world);
- Total Gleason score at baseline (≤ 7, ≥ 8) at diagnosis;
- PSA value at baseline (at or below median, above median);
- Lactate dehydrogenase (LDH) value at baseline (at or below median, above median);
- Hemoglobin value at baseline (at or below median, above median).

Sensitivity Analyses of Metastasis-Free Survival

The following sensitivity analyses will be performed for MFS.

Modified MFS 1 - Including Progression After Alternative Therapy as an Event

A modified MFS event is defined as an MFS event with a modified censoring rule where progression after initiation of any prostate cancer treatment is also counted as an event. The stratified log-rank test will be used to compare the treatment groups. All methods from the primary efficacy analysis will be repeated. The details of the censoring rules are provided in Appendix 1.

Modified MFS 2 – Including Post Treatment Death as Event

A modified MFS 2 event is defined as an MFS event with a modified censoring rule where any death (including those occurred > 112 days after treatment discontinuation) is considered as progression. Modified MFS 2 is defined as the time from randomization to the first date of radiographic progression (assessed by independent central radiology review) or death,

whichever occurs first. The stratified log-rank test will be used to compare the treatment groups. All methods from the primary efficacy analysis will be repeated. The details of the censoring rules are provided in Appendix 1.

Impact of Antineoplastic Therapies

A modified MFS analysis will be performed to assess the sensitivity of MFS to antineoplastic therapy. In this analysis, patients receiving any antineoplastic therapy without radiographic evidence of metastasis will be censored at the last visit with adequate radiographic assessment prior to new antineoplastic therapy. A stratified log-rank test will be used to compare the treatment groups. The details of the censoring rules are provided in Appendix 1.

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. The stratified log-rank test will be used to analyze the MFS values. Furthermore, the concordance and discordance rate between the independent central radiology review and investigator assessment will be summarized using the radiographic metastasis status by the treatment groups.

Impact of Clinical Deterioration

In this sensitivity analysis, patients who discontinue study drug primarily due to clinical deterioration (defined as drop out due to adverse event, defined by investigator) prior to protocol-defined evidence of radiographic progression will be considered as having clinical deterioration, and the date of such an event will be set as the date of study drug discontinuation. The approach used for the primary analysis of MFS will be repeated after accounting for these events. The hazard ratio and its 95% confidence interval will be reported.

In addition to these sensitivity analyses, the time to tumor assessment visits will be summarized by treatment groups using the median and its 95% confidence interval.

9.7.3. Analyses of the Key Secondary Efficacy Endpoints

Time to PSA progression

PSA progression is defined according to Prostate Cancer Working Group 2 (PCWG2) guidelines. Time to PSA progression is defined as the time from randomization to the date of first PSA value demonstrating progression, which is subsequently confirmed. Patients without confirmed PSA progression at the time of analysis will be right censored at the date of last PSA assessment before the analysis data cutoff date for the purposes of analysis.

For patients with PSA declines at week 17, the PSA progression date is defined as the date that a $\geq$ 25% increase and an absolute increase of $\geq$ 2 µg/L (2 ng/mL) above the nadir is documented, which is confirmed by a second consecutive value obtained at least 3 weeks later.

For patients with no PSA declines at week 17, 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 baseline is documented, which is confirmed by a second consecutive value at least 3 weeks later.

As step 2 of the testing strategy described in Section 9.7.1, the ITT analysis for the time to PSA progression endpoint will be conducted with 0.02 alpha level (2-sided), if and only if, enzalutamide is assessed to be statistically superior to placebo for the primary endpoint of MFS. See Section 5 for additional details.

A stratified log-rank test will be used to compare the treatment groups with the same strata as in the time to MFS analysis described in Section 9.7.2.

Time to First Use of New Antineoplastic Therapy

Time to first use of new antineoplastic therapy is defined as the time from randomization to first use of new antineoplastic for prostate cancer. Patients not starting treatment with a new antineoplastic therapy at the time of analysis will be right censored at the date of last assessment before the analysis data cutoff date for the purposes of analysis.

As step 3 of the testing strategy described in Section 9.7.1, the ITT analysis for the time to first use of new antineoplastic therapy endpoint will be conducted with 0.02 alpha level (2-sided), if and only if, enzalutamide is assessed to be statistically superior to placebo for the primary endpoint of MFS and the secondary endpoint of time to PSA progression. See Section 5 for additional details.

A stratified log-rank test will be used to compare the treatment groups with the same strata as in the time to MFS analysis described in Section 9.7.2.

At the time of the third interim and at the final analysis of OS, this endpoint will be updated and provided as an exploratory analysis.

Overall Survival

OS is defined as the time from randomization to death from any cause. For patients who are alive at the time of the analysis data cutoff, OS time will be censored at the last date the patient was known to be alive or analysis data cutoff date, whichever is earlier. Patients with no postbaseline survival information will be censored on the date of randomization.

As step 4 of the testing strategy described in Section 9.7.1, the ITT analysis for the OS endpoint will be conducted with 0.05 (or 0.03 if either time to PSA progression or time to new antineoplastic therapy fails to show significance) alpha level (2-sided), if and only if, enzalutamide is assessed to be statistically superior to placebo for the primary endpoint of MFS. Three interim analyses and 1 final analysis for OS are planned. The data cutoff date will be applied to achieve the approximate number of deaths at each of these analyses.

Depending on the outcome of the time to first use of new antineoplastic therapy endpoint, the total type I error rate across the interim and final analyses will be controlled at 0.03 or 0.05 with the O'Brien-Fleming alpha spending function. The significance level will be fixed at

0.001 for the first interim analysis. For the other overall survival analyses, the significance levels will be recalculated based on the actual number of events at each analysis using the O'Brien-Fleming method, using the remaining type I error rate (0.029 or 0.049 depending on the outcome of the time to PSA progression and time to first use of new antineoplastic therapy endpoint). If an interim analysis of overall survival is statistically significant, it will be reported as the final analysis and no subsequent analyses will be performed. The approximate number of events and corresponding significance level at each analysis based on this methodology are provided in Table 4.

Table 4: Type I Error Spending for the Overall Survival Analyses

Analysis	Number of Death Events [1]	Significance Level	
		Error Rate: 0.03 [2]	Error Rate: 0.05 [3]
First interim	135	0.001	0.001
Second interim	285	0.001	0.002
Third interim	440	0.009	0.018
Final	596	0.026	0.044

[1] Approximate number of targeted events.

[2] Will be used if either time to PSA progression or time to first use of new antineoplastic therapy endpoint fails to show significance. The significance level will be fixed at 0.001 for the first interim analysis. For the other analyses, the significance levels will be recalculated based on the actual number of events at each analysis using the O'Brien-Fleming method.[3]

[3] Will be used if both time to PSA progression and time to first use of new antineoplastic therapy endpoint show significance. The significance level will be fixed at 0.001 for the first interim analysis. For the other analyses, the significance levels will be recalculated based on the actual number of events at each analysis using the O'Brien-Fleming method.[3]

The first interim analysis will be conducted at a 0.001 significance level at the time of the primary MFS analysis. A stratified log-rank test will be used to compare the 2 treatment groups. The number and percentage of death events in each treatment group will be summarized with Kaplan-Meier curves of overall survival with the hazard ratio and its 95% CI.

A second interim analysis is planned after approximately 285 death events are observed. A stratified log-rank test will be used to compare the 2 treatment groups at a 0.002 or 0.001 (depending on the outcome of time to PSA progression and time to new antineoplastic therapy tests) significance level (exact significance level will be calculated based on exact number of events using the O'Brien-Fleming alpha spending function). The number and percentage of death events in each treatment group will be summarized with Kaplan-Meier curves of overall survival with the hazard ratio and its 95% CI.

A third interim analysis is planned after approximately 440 death events are observed. A stratified log-rank test will be used to compare the 2 treatment groups at a 0.018 or 0.009 (depending on the outcome of time to PSA progression and time to new antineoplastic

therapy tests) significance level (exact significance level for this analysis will be calculated based on exact number of events using the O'Brien-Fleming alpha spending function). The number and percentage of death events in each treatment group will be summarized with Kaplan-Meier curves of overall survival with the hazard ratio and its 95% CI.

The final analysis is planned after approximately 596 death events are observed. A stratified log-rank test will be used to compare the 2 treatment groups at a 0.044 or 0.026 (depending on the outcome of time to PSA progression and time to new antineoplastic therapy tests) significance level (exact significance level for this analysis will be calculated using the O'Brien-Fleming alpha spending function). The number and percentage of death events in each treatment group will be summarized with Kaplan-Meier curves of overall survival with the hazard ratio and its 95% CI. If an interim OS analysis is statistically significant, it will be reported as the final analysis and no subsequent analyses will be performed.

At the time of the final analysis of OS, subgroup analyses will be conducted using the same analysis methods as for the subgroup analyses of MFS described in Section 9.7.2.

At the final analysis of OS, to adjust the treatment crossover effect, a sensitivity analysis of OS using Rank-Preserving Structural Failure Time Model (RPSFTM) method will be performed. Additionally, the inverse probability of treatment weighting (IPTW) method may also be performed as an exploratory analysis as appropriate.

9.7.4. Analysis of Other Secondary Endpoints

The following analyses will be performed in the ITT population. The analyses and resulting p-values will be deemed descriptive; adjustments for multiplicity are not planned.

Time to Pain Progression

Pain will be assessed using the score from the BPI-SF question 3: "Please rate your pain by marking the box beside the number that best describes your pain at its worst in the last 24 hours." Time to this event is defined as the time 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. Patients without observed pain progression at the time of analysis will be right censored at the date of last pain assessment for the purposes of analysis. A stratified log-rank test will be used on the ITT population to compare the 2 treatment groups with the same strata as in the time to MFS analysis described in Section 9.7.2.

Time to First Use of Cytotoxic Chemotherapy

Time to first use of cytotoxic chemotherapy is defined as the time from randomization to the first use of cytotoxic chemotherapy for prostate cancer. Patients not starting treatment with a cytotoxic chemotherapy for prostate cancer at the time of analysis will be right censored at the date of last assessment before the analysis data cutoff date for the purposes of analysis. A stratified log-rank test will be used on the ITT population to compare the 2 treatment groups with the same strata as in the time to MFS analysis described in Section 9.7.2.

At the time of the third interim and at the final analysis of OS, this endpoint will be updated and provided as an exploratory analysis.

Chemotherapy-Free Disease-Specific Survival

Chemotherapy-free disease-specific survival is defined as the time from randomization to first use of cytotoxic chemotherapy for prostate cancer or death due to prostate cancer as assessed by the investigator. Patients not starting treatment with a cytotoxic chemotherapy or not known to have died due to prostate cancer at the time of analysis will be right censored at the date of last assessment before the analysis data cutoff date for the purposes of analysis. A stratified log-rank test will be used to compare the treatment groups with the same strata as in the time to MFS analysis described in Section 9.7.2.

At the time of the third interim and at the final analysis of OS, this endpoint will be updated and provided as an exploratory analysis.

Chemotherapy-Free Survival

Chemotherapy-free survival is defined as the time from randomization to first use of cytotoxic chemotherapy for prostate cancer or death. Patients not starting treatment with a cytotoxic chemotherapy or not known to have died at the time of analysis will be censored at the date of last assessment before the analysis data cutoff date for the purposes of analysis. A stratified log-rank test will be used on the ITT population to compare the 2 treatment groups with the same strata as in the time to MFS analysis described in Section 9.7.2.

At the time of the third interim and at the final analysis of OS, this endpoint will be updated and provided as an exploratory analysis.

Analysis of Brief Pain Inventory (BPI-SF)

The 4 individual items used to evaluate pain intensity are ‘worst’, ‘least’, ’average’, and ‘now’ (current pain). Pain ratings range from 0 (zero) to 10 for each item. The mean pain severity score will be calculated using the 4 individual items. The 7 individual items used to evaluate pain interference include general activity, mood, walking ability, normal work, relations with other people, sleep, and enjoyment of life. Interference rating ranges from 0 (zero) to 10 for each item. BPI-SF pain interference will be scored as the mean of the 7 individual interference items. This mean can be used if more than 50% of the total items (4 of 7) have been completed at a given assessment. Descriptive statistics will be provided for the ITT population to summarize each individual item and the two domains scores (severity and interference) by treatment group. Missing data will be excluded from the analysis.

PSA Response

PSA response will be calculated at each visit as a decline from baseline in PSA (ng/mL) at the following maximal PSA response categories:

- Decline of ≥ 50% will be response;

- Decline of ≥ 90% will be response;
- Decline to undetectable level will be response.

Undetectable level is defined as below the limit of quantification of the centrally assessed PSA results. PSA response must be confirmed by a second consecutive value at least 3 weeks later. The percentage of patients in the ITT population with maximal PSA declines of at least 50%, 90%, and undetectable will each be compared between the 2 treatment groups using a stratified Cochran-Mantel-Haenszel mean score test with the same strata as in the time to MFS analysis described in Section 9.7.2. Missing PSA value will be assigned to be nonresponder.

Analysis of Quality of Life

The Functional Assessment of Cancer Therapy – Prostate (FACT-P) is a multidimensional, self-reported, quality-of-life instrument specifically designed for use with prostate cancer. It consists of 27 core items that assess patient function in 4 domains: Physical, Social/Family, Emotional, and Functional wellbeing, which is further supplemented by 12 site-specific items to assess for prostate cancer symptoms. Each item is rated on a 0 to 4 Likert-type scale. The FACT-P domain scores and global score will be calculated using the Manual of Functional Assessment of Chronic Illness Therapy (FACIT) Measurement System[5] and summarized for the ITT population along with their change from baseline by treatment group. Missing data will be excluded from the analysis.

The EQ-5D-5L questionnaire is a standardized instrument for use as a measure of health-related quality of life for men with prostate cancer. Patients will self-rate their current state of mobility, self-care, usual activities, pain/discomfort, and anxiety/depression on a 5-point categorical scale ranging from “no problem” to “extreme problem” within each dimension.

The EORTC QLQ-PR25, a module of the EORTC QLQ-30 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 5 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.

The number and percentage of patients with each response in each domain will be summarized descriptively by study visit for the ITT population for each treatment group. Missing data will be excluded from the analysis.

9.8. Safety Analyses

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

The treatment-emergent period will be defined as in Section 7.1. For incomplete date of last dose of study drug that are missing the day of the month, the 15th of the month will be assumed in determining the treatment-emergent period.

9.8.1. Adverse Events

The severity of all adverse events is to be evaluated by the investigator based on the National Cancer Institute (NCI) Common Terminology Criteria for Adverse Events (CTCAE) version 4.0 or later. All adverse events will be coded to preferred term (PT) and system organ class (SOC) using MedDRA 14.1 or higher.

Treatment-emergent adverse events (TEAEs) are adverse events that started or worsened in severity during the treatment-emergent period.

Patients with multiple occurrences of events for a given PT, SOC, or overall will only be counted once at the worst severity and strongest relationship to study drug for each PT, SOC and overall, respectively.

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 surveillance for each treatment group. Patients can have more than one occurrence of each event.

Tabular summaries including numbers and percentages of the following adverse events will be provided:

- TEAEs;
- Serious TEAEs;
- TEAEs by SOC and PT;
- TEAEs by SOC, PT, and maximum severity;
- TEAEs by SOC, PT, and strongest relationship to study drug;
- TEAEs by decreasing frequency of PT;
- TEAEs of grade 3 or higher;
- TEAEs of grade 3 or higher by SOC, PT, and relationship to study drug;
- TEAEs related to study drug (judged by the investigator) by SOC, PT, and severity;
- TEAEs leading to study drug discontinuation;
- TEAEs leading to study drug discontinuation by SOC, PT, and severity;
- TEAEs leading to dose reduction;

- TEAEs leading to dose reduction by SOC, PT, and severity;
- TEAEs leading to dose interruption;
- TEAEs leading to dose interruption by SOC, PT, and severity;
- Serious TEAEs by SOC, PT, and severity;
- Serious TEAEs by SOC, PT, and relationship to study drug;
- Serious TEAEs that caused study treatment discontinuation by SOC and PT;
- TEAEs leading to death by SOC and PT.

Additionally, Kaplan-Meier methods will be used to display the time to the first grade 3 or higher TEAEs by treatment group.

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 5.

Table 5: Subgroup and Supplemental Tabulations of Treatment-Emergent Adverse Events

Group Variable	Subgroups	Definition
Study Day Cut Points	≤ 60 days after initiation of study drug ≤ 180 days after initiation of study drug ≤ 365 days after initiation of study drug > 365 days after initiation of study drug	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 7.1.
Age (years)	< 65 ≥ 65 < 75 ≥ 75	-
Baseline Body Mass Index	≤ Median > Median	-
Geographic Region	North America European Union Rest of the World	Defined in Appendix 2.
History of Cardiovascular Disease	Yes No	Defined in Appendix 2.

9.8.1.1. Adverse 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 categorized treatment group. Additionally, for the following special adverse events, Kaplan-Meier methods will be used to display the time to the first occurrence of each of these events, if the number of events is deemed sufficiently large:

- Convulsions;
- Neutropenia;
- Cognitive and memory impairment disorders as identified by the higher-level group term for mental impairment;
- Hypertension;
- Hepatic impairment;
- Select cardiovascular events as defined in Appendix 2;
- Posterior reversible encephalopathy syndrome (PRES).

A detailed list of the derivation of each term is provided in Appendix 2.

9.8.2. Laboratory Assessments

Laboratory data in this study consist of hematology values, chemistry tests, and urine tests. Only data collected from the central laboratory 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 4.0. Parameters that have criteria available for both low and high values (e.g., 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 test elevations will be presented by categorized treatment group. Liver function test elevations are assessed by using postbaseline results in ALT, AST, total bilirubin, and alkaline phosphatase during the treatment-emergent period based upon the definitions presented in Table 6.

Table 6: Categories of Liver 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 with alkaline phosphatase < 2 times the upper limit of normal at concurrent visits will also be presented.

9.8.3. Vital Signs

Temperature, blood pressure (systolic and diastolic), and heart rate will be summarized at baseline and each subsequent scheduled assessment by treatment group as treated. Only data collected during the treatment-emergent period will be summarized. Baseline results are defined as the last vital sign results taken on or before the date and time of the first dose of study drug. Change from baseline will be calculated and presented for each parameter at all scheduled postbaseline assessment timepoints. The mean change from baseline and its

95% confidence interval for each treatment group and study visit will be plotted over time as well.

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 7.

Table 7: 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
	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.

10. REFERENCES

1. Pound CR, Partin AW, Eisenberger MA, Chan DW, Pearson JD. Walsh PC. Natural history of progression after PSA elevation following radical prostatectomy. JAMA 281(17):1591-7, 1999.

2. Nelson JB, Love W, Chin JL, et al. Phase 3, randomized, controlled trial of atrasentan in patients with nonmetastatic, hormone refractory prostate cancer. Cancer Nov 1;113(9):2478-87, 2008.

3. O'Brien PC, Fleming TR. A multiple testing procedure for clinical trials. Biometrics 35:549-56, 1979.

4. Schemper M, Smith TL. A note on quantifying follow-up in studies of failure time. Cont Clin Trials 17:343-6, 1996.

5. 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.

Appendix 1: Definition of Metastasis-Free Survival

Metastasis-free survival (MFS) is defined as time from randomization to radiographic progression as assessed by blinded independent central radiology review or death on study (death within 112 days of treatment discontinuation without evidence of radiographic progression), whichever occurs first.

The following table summarizes the protocol-specified rules for the radiographic evidence of progression (protocol version 4.0):

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 5 regions on a bone scan will not require confirmation with a second imaging modality.	Screening and every 16 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
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 censoring rules for the primary and sensitivity analyses of MFS are summarized in the following table:

Analysis	Censoring Rules	Date of Censoring
Primary analysis of MFS	Patients with no baseline or no postbaseline assessments	Date of randomization
	Patients who were randomized but confirmed metastatic at baseline by independent central radiology review	Date of randomization
	Patients who had no confirmed metastasis as per independent central radiology review or did not die prior to data cutoff date	Date of the last radiographic assessment prior to data cutoff date
	Patients who had no confirmed metastasis as per independent central radiology review but died after 112 days following last dose of study drug	Date of the last radiographic assessment prior to data cutoff date
	Patients who initiate cytotoxic chemotherapy, abiraterone acetate, or nonradioactive bone-targeting agents without evidence of metastasis as per independent central radiology review	Date of the last radiographic assessment prior to first use of any such therapy
	Patients who experience a skeletal-related event without evidence of metastasis as per independent central radiology review	Date of the last radiographic assessment prior to the earliest skeletal-related event
	Patients with radiation therapy performed for prostate cancer-related lesions without evidence of metastasis as per independent central radiology review	Date of the last radiographic assessment prior to the earliest use of radiation therapy
	Patients with 2 or more consecutive missed tumor assessment visits without evidence of metastasis as per independent central radiology review	Date of the last radiographic assessment prior to the missed visit date
MFS Based on Investigator Assessment	Same as MFS	Same as MFS
Impact of Antineoplastic Therapies	All censoring rules for the primary analysis of MFS	Same as the primary analysis of MFS
	Patients who initiate any antineoplastic therapy without evidence of metastasis as per independent central radiology review	Date of the last radiographic assessment prior to first antineoplastic therapy use
Modified MFS 1 - *Including Progression After Alternative Therapy as Event*	Patients with no baseline or no postbaseline assessments	Date of randomization
	Patients who had no confirmed metastasis as per independent central radiology review but died after 112 days following last dose of study drug	Date of the last radiographic assessment prior to data cutoff date
	Patients who experience a skeletal-related event without evidence of metastasis as per independent central radiology review	Date of the last radiographic assessment prior to the earliest skeletal-related event

Analysis	Censoring Rules	Date of Censoring
	Patients with 2 or more consecutive missed tumor assessment visits without evidence of metastasis as per independent central radiology review	Date of the last radiographic assessment prior to the missed visit date
Modified MFS 2 – Including Post Treatment Deaths as event	Patients with no baseline or no postbaseline assessments	Date of randomization
	Patients who were randomized but confirmed metastatic at baseline by independent central radiology review	Date of randomization
	Patients who had no confirmed metastasis as per independent central radiology review or did not die prior to data cutoff date	Date of the last radiographic assessment prior to data cutoff date
	Patients who initiate cytotoxic chemotherapy, abiraterone acetate, or nonradioactive bone-targeting agents without evidence of metastasis as per independent central radiology review	Date of the last radiographic assessment prior to first use of any such therapy
	Patients who experience a skeletal-related event without evidence of metastasis as per independent central radiology review	Date of the last radiographic assessment prior to the earliest skeletal-related event
	Patients with radiation therapy performed for prostate cancer-related lesions without evidence of metastasis as per independent central radiology review	Date of the last radiographic assessment prior to the earliest use of radiation therapy
	Patients with 2 or more consecutive missed tumor assessment visits without evidence of metastasis as per independent central radiology review	Date of the last radiographic assessment prior to the missed visit date

ITT, Intent-to-treat; MFS, metastasis-free survival.

Appendix 2: Study Specific Definitions

Cardiovascular disease will focus primarily on arterial thromboembolic disease processes with the following terms:

System Organ Class	Event
Cardiac	Coronary artery disease, acute coronary syndromes, ischemic cardiomyopathy, angina, and related/synonymous terms.
Nervous system	Transient ischemic attack, cerebral vascular accident, cerebral hemorrhage, and related/synonymous terms.
Surgical procedures	Coronary artery bypass grafting, angioplasty ± stent, peripheral revascularization procedures, and related/synonymous terms.
Vascular disorders	Arterial occlusive disease, arterial thromboembolic events, and related/synonymous terms.

Three major geographic regions are defined below:

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

EU, European Union.

Definitions for prespecified adverse events of special interest are listed below:

Event Grouping of Interest	Definition
Convulsions	Narrow SMQ of 'convulsions'
Neutropenia	PT: neutrophil count decreased, neutropenia, white blood cell count decreased, leukopenia, agranulocytosis, cyclic neutropenia, febrile neutropenia, neutropenic infection, neutropenic sepsis, neutrophil percentage decreased, band neutrophil count decreased, band neutrophil percentage decreased, and idiopathic neutropenia
Memory impairment	All PT in MedDRA High Level Group Term 'mental impairment disorders'
Hypertension	Narrow SMQ of 'hypertension'
Hepatic impairment	Narrow SMQ of 'hepatic failure, fibrosis and cirrhosis and other liver damage related conditions;' narrow SMQ of 'hepatitis, non-infectious;' and narrow SMQ of 'liver related investigations, signs and symptoms'
Major adverse cardiovascular events	Narrow SMQs of 'myocardial infarction,' 'haemorrhagic cerebrovascular conditions,' and 'ischemic cerebrovascular conditions,' and 'heart failure'
PRES	Narrow SMQ of 'noninfectious encephalopathy/delirium'

SMQ: standardized MedDRA query, PT: preferred term

PPD

[redacted]

CCI [redacted]

[redacted]

[redacted]

[redacted]