Protocol MDV3100-18

(C3431012)

A Phase 2, Open-Label, Single-Arm Study of 18F-Sodium Fluoride PET/CT Bone Imaging in Enzalutamide-Treated Chemotherapy-Naïve Patients with Bone-Metastatic Castration-Resistant Prostate Cancer

Statistical Analysis Plan (SAP)

Version: 1.0

Date: 24-SEPT-2018

[redacted]

TABLE OF CONTENTS

LIST OF TABLES 3
LIST OF FIGURES 4
APPENDICES 4
1. VERSION HISTORY 5
2. INTRODUCTION & STUDY DESIGN 5
2.1. Study Objectives 5
2.2. Study Design 6
3. ENDPOINTS AND BASELINE VARIABLES: DEFINITIONS AND CONVENTIONS 7
3.1. Primary Endpoint 8
3.2. Secondary Endpoints 9
CCI [redacted]
3.4. Baseline Variables 11
3.5. Safety Endpoints 12
4. ANALYSIS SETS 13
4.1. As Treated (AT) 13
5. GENERAL METHODOLOGY AND CONVENTIONS 13
5.1. Hypotheses and Decision Rules 13
5.2. General Methods 13
5.2.1. Pooling of Data by Study Site 13
5.2.2. Nominal Timepoints 13
5.2.3. Definition of Study Day 14
5.2.4. Unscheduled Assessments 14
5.2.5. Standard Derivations and Reporting Conventions 14
5.2.6. Analyses for Categorical Data 14
5.2.7. Analyses for Continuous Data 14
5.2.8. Analysis of Covariance 15
5.2.9. Analysis for Log-Transformed Data 15
CCI [redacted]
5.3. Methods to Manage Missing Data 16
6. ANALYSES AND SUMMARIES 17
6.1. Primary Endpoint 17

[redacted]

6.2. Secondary Endpoint(s) 17
CCI [redacted]
6.4. Subgroup Analyses....23
6.5. Baseline and Other Summaries and Analyses 23
6.5.1. Baseline Summaries....23
6.5.2. Protocol Deviations 24
6.5.3. Study Conduct and Patient Disposition 24
6.5.4. Prior and Current Therapies....24
6.5.5. Study Drug Exposure and Treatment Compliance 24
6.5.6. Concomitant Medications and Non-Drug Treatments 25
6.6. Safety Summaries and Analyses 26
6.6.1. Adverse Events 26
6.6.2. Adverse Events of Special Interest 27
6.6.3. Deaths 27
6.6.4. Laboratory Data 27
6.6.5. Vital Signs....29
6.6.6. Electrocardiogram....29
6.6.7. ECOG 29
7. INTERIM ANALYSES 29
8. REFERENCES 30
9. APPENDICES 31
10. ABBREVIATIONS 43

LIST OF TABLES

Table 1. Summary of Major Changes in SAP Amendments 5

LIST OF FIGURES

Figure 1. Study Schematic 7

APPENDICES

Appendix 1. Imputation of Missing/Partially Missing Dates 31
Appendix 2. Sample SAS Code 33
Appendix 3. Analysis Windows for Reporting 35
Appendix 4. Summary of Efficacy Analyses 36

1. VERSION HISTORY

This Statistical Analysis Plan (SAP) for study MDV3100-18 is based on the Protocol Amendment #2 dated 08-DECEMBER-2016.

Table 1. Summary of Major Changes in SAP Amendments

SAP Version	Change	Rationale
1.0	Not Applicable	Not Applicable

2. INTRODUCTION & STUDY DESIGN

This SAP provides detailed methodology for summary and statistical analyses of the data collected in study MDV3100-18. 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. Any deviations from this analysis plan will be described in the Clinical Study Report (CSR).

MDV3100-18 is a phase 2, open-label, single-arm trial of enzalutamide in adult male patients who are chemotherapy-naïve, and who have bone-metastatic castration-resistant prostate cancer.

In this study, approximately 20 adult male patients (age ≥18 years) will be enrolled across 3 sites in the United States. The primary objective is to evaluate F-NaF PET/CT imaging as a method for assessing treatment response in metastatic bone lesions either at time of disease progression, or after 2 years of treatment, whichever comes first.

The primary analysis will include all data up to a data cutoff date which will be determined when all patients have progressed or have otherwise been followed for two year or are no longer being followed due to withdrawal of consent, lost to follow-up, or death. All summaries and analyses will include all data pertaining to visits/assessments (both planned and unplanned) performed up to and including the data cutoff date.

2.1. Study Objectives

Primary Objectives:

- To evaluate F-NaF PET/CT imaging as a method for assessing treatment response in metastatic bone lesions either at time of disease progression, or after 2 years of treatment, whichever comes first.

Secondary Objectives:

- To evaluate the heterogeneity of response in metastatic bone lesions at the time the primary objective is assessed.

CCI

[redacted]

2.2. Study Design

This is a phase 2, open-label, single-arm trial of enzalutamide in adult male patients who are chemotherapy-naïve, and who have bone-metastatic castration-resistant prostate cancer.

In this study approximately 20 adult male patients will be enrolled across 3 sites in the United States. Eligible patients must have bone metastatic disease in at least 2 lesions with a PSA ≥2 ng/mL. Patients must not have received any prior chemotherapy previously.

Following screening, enrolled patients will have baseline 18F-NaF PET/CT imaging (NaF-1) and then begin open-label treatment with enzalutamide. PSA response will be determined at week 13 when a second 18F-NaF PET/CT imaging (NaF-2) is obtained. Patients with any increase in PSA at week 13 may discontinue treatment and proceed to safety follow-up unless the investigator considers continuation of study drug to be beneficial. For all continuing patients, PSA will be monitored every 8 weeks after week 13 and other study assessments will continue to be obtained. A third 18F-NaF PET/CT imaging (NaF-3) and a follow-up 99mTc-MDP bone scintigraphy will be performed when PSA progression (an increase of at least 25% AND an absolute increase of at least 2.0 ng/mL above nadir), radiographic progression (bone or soft tissue), or other clinically relevant progression (eg, significant bone pain or a skeletal-related event) is observed, or at 2 years without progression after initiation of treatment. No 99mTc-MDP bone scintigraphy will be required at the time the NaF-3 bone scan is obtained if bone disease progression was confirmed earlier on bone scintigraphy obtained at investigator discretion.

Total study duration will be approximately 36 months.

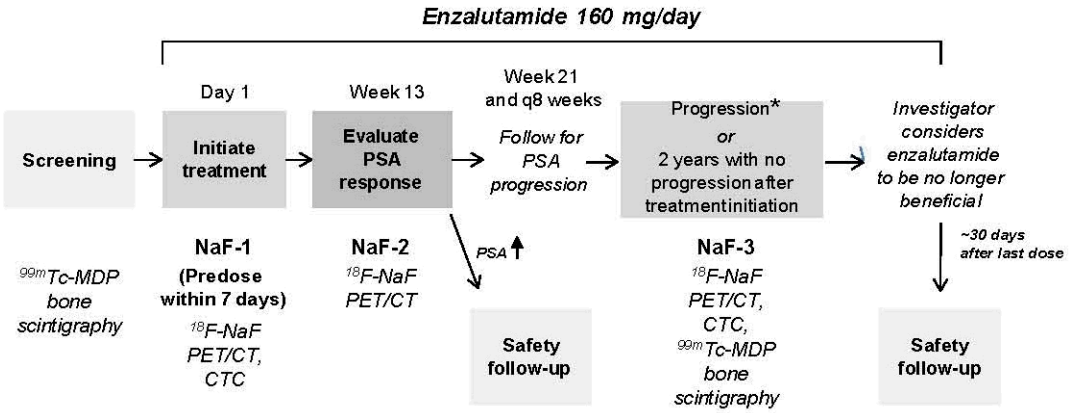
Enzalutamide will be administered one daily at a starting dose of 160mg/day. Dose modifications downwards are possible to either 120 or 80 mg/day, per protocol section 6.2.4.

[redacted]

Study periods include screening, treatment, and safety follow-up. Safety follow-up applies to patients who are off protocol therapy and continues to the earliest of 1) 30 days after last dose, 2) death, 3) lost to follow-up, 4) withdrawal of consent, or 5) enrollment into another anti-cancer clinical trial.

The study schematic is provided in Figure 1.

Figure 1. Study Schematic



3. ENDPOINTS AND BASELINE VARIABLES: DEFINITIONS AND CONVENTIONS

Standardized uptake value (SUV) is a unitless measure derived from the imaging PET/CT scans as a ratio of image-radioactivity concentration to the whole-body concentration. SUV and its various uptake measures described below will be analyzed at the central vendor and provided to the sponsor.

For a particular lesion (n), where n = 1, 2, 3, … N, a 3-dimensional volume of interest (VOI) will be analyzed. Within that VOI for that lesion, there are numerous voxels which compromise the volume. These voxels are 3-dimensional analogs of the 2-dimensional pixels. A particular voxel will be voxel (k), and there are K voxels compromising the entire VOI for that lesion, where k = 1, 2, 3, … K. Let SUV_k be the SUV for a voxel (k) within that particular lesion.

- **SUV(total)** is the total standardized uptake value. This measure assesses the total tumor burden.
    - First, each lesion (n) of interest produces its own SUV(total) by summing from k=1 to k=K of SUV_k.
    - Second, the sum from n=1 to n=N of each lesion's SUV(total) will produce the desired **SUV(total).**

- **SUV(mean)** is the mean standardized uptake value.
    - First, each lesion (n) of interest produces its own SUV(mean) by summing from k=1 to k=K of SUV_k and then dividing that sum by K.
    - Second, the mean from n=1 to n=N of each lesion's SUV(mean) will produce the desired **SUV(mean).**
- **SUV(max)** is the maximum standardized uptake value.
    - First, each lesion (n) of interest produces its own SUV(max) by taking the maximum value from k=1 to k=K of SUV_k.
    - Second, the maximum from n=1 to n=N of each lesion's SUV(max) will produce the desired **SUV(max).**
- **SUV(hetero)** is the heterogeneity measure of tumor activity across all lesions
    - First, each lesion (n) of interest produces its own SUV(mean) by summing from k=1 to k=K of SUV_k and then dividing that sum by K.
    - Second, that lesion's mean is subtracted from the **SUV(mean)** above, and that difference is then squared.
    - Third, those squared differences from n=1 to n=N are then summed, and then divided by N to produce the desired **SUV(hetero).**

3.1. Primary Endpoint

The single primary endpoint is the **proportion of patients with ≥1 responding bone lesion on the third scheduled F-NaF PET/CT scan** for this study. For clarity going forward, this scan will be referred to as NaF. Three such NaF scans are scheduled and are noted below:

Scan identifier	Scheduled to occur at
NaF-1	Screening
NaF-2	Week 13
NaF-3	Time of progression, or at 2 years after treatment start, whichever occurs first*

* If progression occurs before Week 13, then NaF-3 will constitute the scan at progression and support the primary analysis, even though NaF-2 hadn't occurred yet.

A lesion is considered responding if the change from baseline in total standardized uptake value (SUV(total)) is below the limit of agreement (LOA). The LOA will be estimated based upon test/retest analysis and provided by the vendor.

The NaF-3 will be triggered by the earliest of any of the following events:

- Prostate-specific antigen (PSA) progression:
    - Defined as percentage change from baseline ≥25% and nominal increase ≥2.0 ng/mL above the nadir (or above the baseline if no decline by week 13).
- Bone disease progression:
    - Defined as ≥2 new lesions assessed by bone scintigraphy.
- Soft tissue disease progression:
    - Defined by RECIST v1.1.
- Clinically relevant progression:
    - Defined by the investigator.
- Two years after start of therapy.

NaF scans will be interpreted by a central imaging vendor.

3.2. Secondary Endpoints

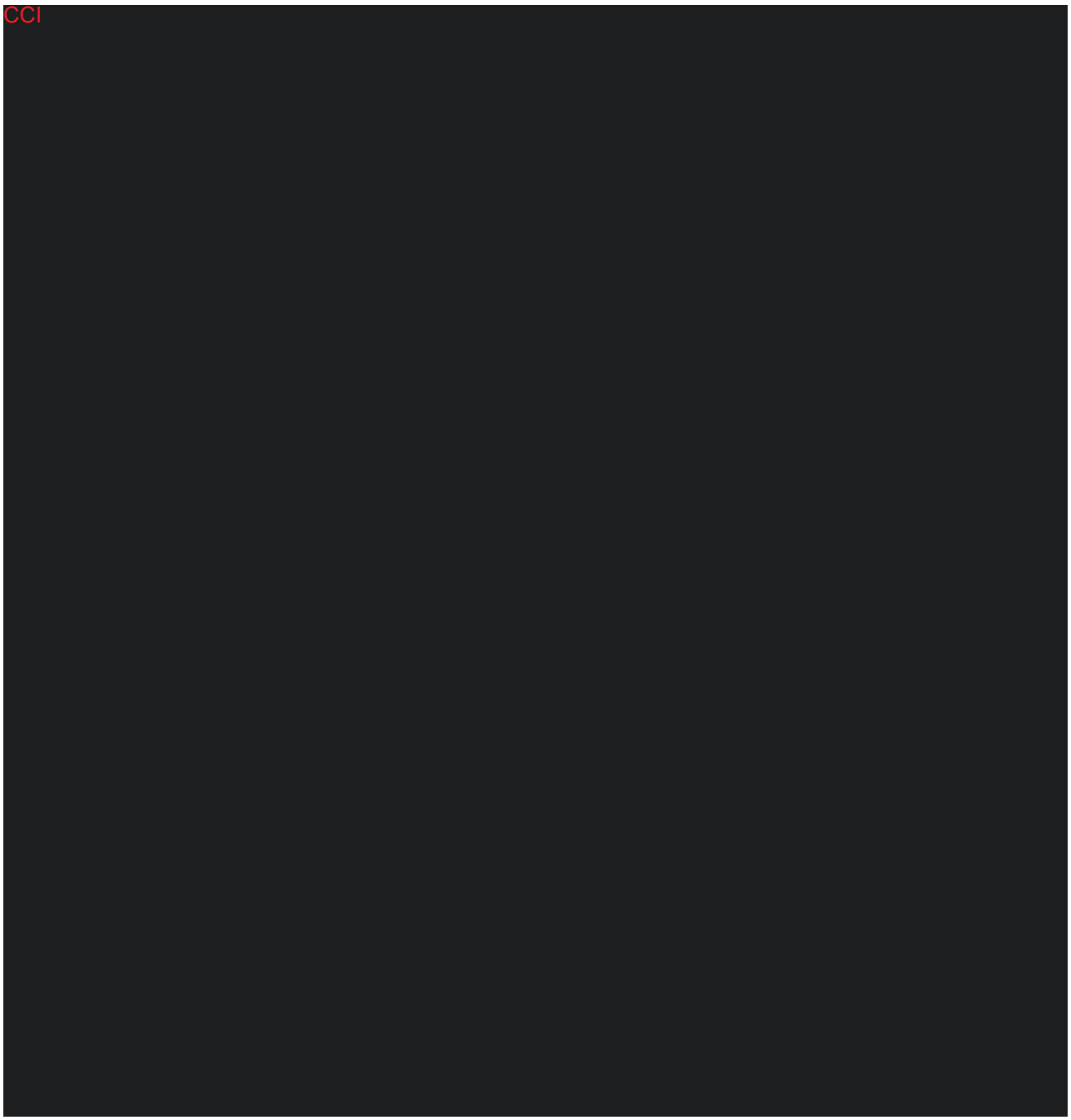
There is one secondary endpoint for this study, namely **the heterogeneity of response across all bone lesions at the time the primary endpoint is assessed, SUV(hetero)**. At the NaF-3, the observed SUV(hetero) will be assessed.

CCI

[redacted]

CCI

[redacted]



3.4. Baseline Variables

The date of first dose (start date) of study treatment is the earliest date of non-zero dosing of the study drug. The date of last dose of study treatment is the latest date of non-zero dosing of the study drug.

No windowing will be applied when defining baseline. Any deviations from the protocol specified window will be documented as protocol deviations. For efficacy analyses and baseline characteristics associated with tumor assessments, the last assessment prior to the date and time of first study drug will serve as the baseline assessment.

For safety, the last assessment performed on or prior to date of the first dose of study treatment will serve as the baseline assessment. If there are no observations meeting these criteria, then baseline is considered missing.

Post-baseline value is defined as a measurement taken after the first administration of study drug. Change from baseline is defined as (post baseline value – baseline value). Percent change from baseline is defined as [(post baseline value – baseline value)/(baseline value)] multiplied by 100%. Both 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.

3.5. Safety Endpoints

Safety endpoints will be summarized based on the on-treatment period unless otherwise specified.

On-treatment is defined as the period between the first dose of the study drug to the earliest of 30 days after the last dose of study drug treatment, the start of cytotoxic chemotherapy, the start of a lutamide (bicalutamide, nilutamide, or flutamide), or the start of a new investigational therapy. 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 firs dose of study treatment will be considered baseline assessments.

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

An adverse event is considered treatment emergent relative to a given treatment if:

- the event occurs for the first time during the on-treatment period and was not seen prior to the start of treatment, or
- the event was seen prior to the start of treatment but increased in severity during the on-treatment period.

Adverse events will be coded to preferred term and system organ class using the Medical Dictionary for Regulatory Activities (MedDRA) and classified by severity using the CTCAE, version 4.

Hematology and chemistry results will be programmatically graded according to the NCI CTCAE version 4 for relevant parameters. Parameters which cannot be graded will be summarized relative to the normal range (ie, normal range high or normal range low). Additional details are provided in Section 6.6.

4. ANALYSIS SETS

While the protocol notes an intent-to-treat population, the below As Treated population will be used instead since the study is single-arm and exploratory in nature. Certain restrictions will later be noted where baseline and certain post-baseline measurements are required for particular endpoints.

4.1. As Treated (AT)

The As Treated (AT) population is defined as all enrolled patients who receive any amount of study treatment.

5. GENERAL METHODOLOGY AND CONVENTIONS

There is no planned interim analysis. The final analysis will occur when all patients have progressed or have otherwise been followed for two year or are no longer being followed due to withdrawal of consent, lost to follow-up, or death. Since the study is phase 2 and exploratory, no adjustments for multiplicity will be made.

5.1. Hypotheses and Decision Rules

The primary purpose of this study is to assess the NaF imaging as a method for evaluating treatment response in bone lesions at the time of progression. The sample size of 20 patients was chosen for practical clinical considerations.

Due to the exploratory nature of this study, there is no formal hypothesis. Instead, estimation will be the primary aim.

5.2. General Methods

Confidence intervals will be done with 95% coverage. No adjustments to the confidence intervals are planned for any multiple testing.

5.2.1. Pooling of Data by Study Site

In order to provide overall estimates, data will be pooled across study sites. Furthermore, the small number of enrolled patients makes an analysis by study site not feasible. However, study site will be used as a factor in the ANCOVA models so as to improve precision of the estimates.

5.2.2. Nominal Timepoints

Visit windows will apply per Appendix 3. Analysis Windows for Reporting.

Note that NaF-1 and NaF-2 are scheduled for baseline and week 13, respectively. NaF-3, however, is more fluid and can occur depending on different disease progressions or 2-years of therapy, per Section 3.1. If NaF-3 occurs before the scheduled NaF-2, then that scan will be NaF-3, and the pre-planned NaF-2 will not exist.

For endpoints only collected once post-baseline at NaF-3 (eg, circulating tumor cells), the NaF-3 slotted value will be the measurement closest to the NaF-3 scan.

5.2.3. Definition of 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 +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.

5.2.4. Unscheduled Assessments

Unscheduled assessments will be used when deriving baseline and worst case on-treatment for safety and patient-reported outcome analyses (except where noted for baseline ECGs). Additionally, unscheduled assessments will be used for efficacy analyses (eg, defining date of PSA progression, date of last contact).

5.2.5. Standard Derivations and Reporting Conventions

The following conversion factors will be used to convert days into weeks, months or years: 1 week = 7 days, 1 month = 30.4375 days, 1 year = 365.25 days.

Percentages will be reported to one decimal place. The rounding will be performed to closest integer / first decimal using the common mid-point between the two consecutive values. Eg, 5.1 to 5.4 will be rounded to an integer of 5, and 5.5 to 5.9 will be rounded to an integer of 6.

5.2.6. 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.

If a binary proportion requires a confidence interval, it will be calculated with a 2-sided exact 95% CI using the Clopper-Pearson method (Clopper, C. J.; Pearson, E. S. (1934)). This confidence interval is an exact interval, not relying on any normal approximations. Sample SAS code can be found in Appendix 2. Sample SAS Code.

5.2.7. Analyses for Continuous Data

Descriptive statistics, including the number of observations (N), mean, standard deviation, median, minimum, 25th quantile, 75th quantile, and maximum values, will be provided for continuous variables.

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

5.2.8. Analysis of Covariance

A linear model for analysis of covariance (ANCOVA) will be used to analyze continuous data collected at baseline and at particular post-baseline visits. The dependent variable in the ANCOVA model will be the change from baseline at the specified post-baseline windowed visit. The model will include fixed effects terms for:

- Treatment (as a categorical variable);
- Baseline value (as a continuous variable);
- Study site (as a categorical variable).

The observed margins option (OM) will be implemented in assigning categorical weights when calculating the least squares means. There will be no imputation for missing data in the analyses. While there is only one treatment arm in this study, the inclusion of the fixed effect for treatment is needed in order to produce the desired least squares mean.

The model-adjusted mean change from baseline will be presented at the specified post-baseline visit using the least squares mean, its standard error, and its 95% confidence interval.

In the event of convergence problems, then the fixed effect term for study site will be removed and the model re-run.

Sample SAS code for the ANCOVA model is found in Appendix 2. Sample SAS Code.

5.2.9. Analysis for Log-Transformed Data

Endpoints that are markedly skewed can be analyzed with log-transformation. The natural log of the baseline and the natural log of post-baseline observed data will be used in the ANCOVA analysis from Section 5.2.8. Nominal change in the log-transformed data (log[post-baseline value] minus log[baseline value]) will be analyzed. The following will estimate a percent change from baseline and will be displayed in the tables.

Percent change for treatment = [exp (LSMean) – 1]*100%

Likewise, the confidence intervals for the percent change will be obtained similarly, where the above back-transformation is applied to the log-transformed confidence interval for the LSMean. If any values for analysis are zero, then a positive value of 0.0001 will be used instead; this will allow the computation of the natural log.

CCI [redacted]

CCI

[redacted]

5.3. Methods to Manage Missing Data

Missing data will not be imputed unless specified elsewhere in the SAP. Missing dates or partially missing dates will be imputed conservatively for adverse events and prior/concomitant medications/procedures. Specific rules for handling missing or partially missing dates for date of birth, adverse events, prior/concomitant medications/procedures, and diagnosis of cancer are provided in Appendix 1. Imputation of Missing/Partially Missing Dates. Any imputations will occur at the analysis dataset level and not the raw dataset level. Additionally, in all patient data listings imputed values will be presented and flagged as imputed.

Missing statistics, eg, when they cannot be calculated, should be presented as 'ND' for not done, 'NR' for not reached or 'NA' for not applicable. For example, if N=1, the measure of variability cannot be computed and should be presented as 'ND' or 'NA'. If too few events are observed in the KM analyses, then certain quartiles may not be estimable; hence NR should be presented.

Missing toxicity grades for adverse events:

> Prior to Study Treatment: If no toxicity grade is available or the grade is reported as unknown for an adverse event prior to the first study treatment, then Grade 1 will be assumed for purposes of defining a baseline grade for assessing if further occurrences are treatment emergent. However, if the patient experiences multiple episodes of the same AE prior to study treatment, then the maximum toxicity grade observed prior to study treatment will be utilized in assessing if further occurrences are treatment emergent.
>
> During Study Treatment: If no toxicity grade is available or the grade is reported as unknown for an adverse event during the study treatment, then the event will be considered treatment emergent unless a baseline event was reported as Grade 4.

In summaries which present maximum toxicity grade, the maximum of non-missing grades will be displayed. Missing grade will only be displayed for cases where a patient reported only one event and the grade is missing.

[redacted]

For incomplete day of last dose of study drug, the 15th of the month will be assumed in determining the treatment-emergent period.

6. ANALYSES AND SUMMARIES

All efficacy analyses will be based on the AT population. Any confidence intervals presented will be 2-sided 95% CI. No adjustments to the confidence intervals are planned for any multiple testing. The start day for time to event analyses will begin at the date of first study drug.

Safety analyses (adverse events, laboratory tests, etc.) will be based on the AT population.

A grid-summary of the efficacy analyses can be found in Appendix 4. Summary of Efficacy Analyses.

6.1. Primary Endpoint

The primary endpoint is the proportion of patients with ≥1 responding bone lesions at NaF-3. A lesion is considered responding if the change from baseline at NaF-3 in total standardized uptake value (SUV(total)) is below the limit of agreement (LOA).

The primary endpoint will be summarized with number and percent of patients with ≥1 responding bone lesions at NaF-3 along with the 95% CI using the Clopper-Pearson method.

This endpoint will be analyzed for all patients in the AT population who also have non-missing scans at both NaF-1 and NaF-3. Patients not having both NaF-1 and NaF-3 will be counted as non-evaluable (hence not included in the denominator of the point estimate).

6.2. Secondary Endpoint(s)

The secondary endpoint is the observed heterogeneity of response across all bone lesions, SUV(hetero). This endpoint will be assessed for each patient when their primary endpoint is assessed, and is based on the post-baseline scan NaF-3. This endpoint was analyzed at the central vendor and provided to the sponsor. Summary statistics will be used to describe the endpoint. Graphical displays utilizing histograms will also be produced.

This endpoint will be analyzed for all patients in the AT population who also have non-missing scans at both NaF-1 and NaF-3.

CCI [redacted]

CCI

[redacted]

CCI

[redacted]

CCI

[redacted]

CCI

[redacted]

CCI

[redacted]

CCI

[redacted]

6.4. Subgroup Analyses

Not applicable due to the small sample size and exploratory nature of the study.

6.5. Baseline and Other Summaries and Analyses

6.5.1. Baseline Summaries

The following demographic and baseline characteristics will be summarized for all patients in the AT population.

- Age, race, ethnicity, height (cm), weight (kg), and body mass index (kg/m^2).
- PSA (ng/mL) and Gleason scores.
    - From the medical history prostate cancer CRF page.
- Primary diagnosis.
    - Continuous summary of time from primary diagnosis to study enrollment (in years).
- ECOG status.
    - Tabulations with n (%) for the different categories.
- Brief Pain Inventory (Short Form).
    - Yes/No categories will be tabulated with n (%).
    - Categories scaled from 0 to 10, or from 0% to 100%, will be summarized as a continuous measure using descriptive statistics.
- Medical history.
    - Derived from the medical history non-prostate cancer CRF page.

Listings will be provided for these parameters.

[redacted]

6.5.2. Protocol Deviations

Protocol deviations will be compiled prior to database closure and will be summarized by category (n(%)) for the AT population. Categories will be assigned by the study clinician.

6.5.3. Study Conduct and Patient Disposition

Summaries will include the number and percentage of patients in AT population.

The number and percentage of patients who discontinue study drug and the reasons for discontinuation will be summarized for all patients in the AT population. Likewise the number and percentage of patients who discontinue the study itself will be summarized for all patients in the AT population.

6.5.4. Prior and Current Therapies

Prior therapies and current therapies will be summarized separately for patients in the AT population:

- Radiotherapy for prostate cancer.
- Surgical procedures for prostate cancer.

6.5.5. Study Drug Exposure and Treatment Compliance

Study Drug Exposure

Summary statistics will be provided for all patients in the AT population for:

- Treatment duration (months): For each patient, treatment duration is defined as (date of last dose – date of first dose +1)/30.4375. Treatment duration will be summarized both as a continuous measure and a categorical measure (<3 months, 3 to <6 months, 6 to <12 months, 12 to <24 months, 24 to <36 months, and ≥36 months).
- Cumulative dose (mg): For each patient, the cumulative dose is calculated based on the Study Drug Dosing CRF page. The daily total dose (mg) will be summed across the study.
- Average daily dose (mg/day): The average daily dose is defined as the cumulative dose divided by the dose exposure.
    - Dose exposure (days) is defined as the total number of actual days on study drug during the treatment period.

Treatment Compliance

- Percent overall compliance will be summarized both as a continuous measure and a categorical measure (<80%, ≥80%).

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%, as specified in section 10.5 of the protocol. Capsules not returned will be assumed to have been taken, as per section 6.3 of the protocol. The number of returned capsules is obtained from the Study Drug Accountability CRF page. A patient's expected number of capsules, if no dose modifications, will be calculated as [4 × (date of last dose – date of first dose of study drug + 1)]. Adjustments to the expected number of capsules are made if dose adjustments occurred:

- If a dose reduction to 120mg, then subtract one capsule from the above formula for every day on the 120mg dose.
- If a dose reduction to 80mg, then subtract two capsules from the above formula for every day on the 80mg dose.
- If a dose interruption, then subtract the appropriate number of capsules (4 or 3 or 2) from the above formula for every day of interruption.

Dose Interruptions and Reductions

Dose modifications include dose interruptions and/or reductions. A dose interruption will include a missing or a zero dose on the dosing page. A dose reduction is any dose that is lower than the previous dose. Dosage of enzalutamide may be reduced to either 120 mg/day, or 80 mg/day.

- Patients with at least 1 dose modification (interruption or reduction) and the reason for the dose modifications will be tabulated.
- Patients with at least 1 dose interruption will be tabulated.
- Patients with at least 1 dose reduction will be tabulated.

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

6.5.6. Concomitant Medications and Non-Drug Treatments

Concomitant medications taken during the study treatment period will be summarized for all patients in the AT population. Concomitant medications refer to all medications which started prior to first dose of study treatment and continued during the treatment-emergent period (earliest of 30 days after the last dose of study drug treatment, the start of cytotoxic chemotherapy, the start of a lutamide (bicalutamide, nilutamide, or flutamide), or the start of a new investigational therapy (whichever occurs first)) as well as those started during the on-treatment period.

Concomitant medications will be coded in the WHO Drug coding dictionary and will be tabulated by Anatomical Therapeutic Chemical (ATC) Classification level 2 and preferred term in descending order of frequency. In case of equal frequency regarding drug class (respectively drug name), alphabetical order will be used. A patient will be counted only

once within a given drug class and within a given drug name, even if he/she received the same medication at different times. Preferred Terms will be reported under each ATC class that it is included under within WHO Drug (no primary path is available in WHO Drug).

All concomitant medications and non-drug treatments recorded on the case report form will also be listed.

6.6. Safety Summaries and Analyses

All patients in the AT population will be used in the safety analyses.

6.6.1. Adverse Events

Adverse events will be coded using MedDRA, in accordance with the version(s) documented in the data management plan. Adverse events will be graded by the investigator using the CTCAE v4.0.

All analyses will be based on treatment emergent events unless otherwise specified. Treatment emergent is defined in Section 3.5. AEs not considered treatment emergent will be flagged in data listings. Only TEAEs with onset date to the earliest of 30 days after the last dose of study drug treatment, the start of cytotoxic chemotherapy, the start of a lutamide (bicalutamide, nilutamide, or flutamide), or the start of a new investigational therapy (whichever occurs first) will be included in the adverse event summaries.

Adverse event incidence will be summarized with frequency and percentage using MedDRA. Some summaries are by system organ class (SOC) and preferred term (PT), while others are only by PT (see below). In addition, adverse event incidence will also be summarized by severity and relationship to study drug. The denominator for the adverse event incidence will be based on the number of patients in the AT population unless otherwise specified.

Patients with multiple occurrences of events for a given preferred term (even if concurrent), system organ class, or overall will only be counted once at the worst severity and strongest relationship to study drug for each preferred term, system organ class, and overall, respectively. If relationship to study drug is missing, the adverse event will be counted as treatment-related. Adverse event listings will show missing relationship as missing.

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

- TEAEs;
- TEAEs by SOC and PT;
- TEAEs by PT and maximum severity;
- TEAEs (treatment related);
- TEAEs (treatment related) by SOC and PT;

- TEAEs (treatment related) by PT and maximum severity;
- TEAEs leading to study drug discontinuation by PT;
- TEAEs leading to temporary study drug interruptions by PT;
- TEAEs leading to dose interruptions or reductions by PT;
- Serious TEAEs;
- Serious TEAEs by SOC, PT and maximum severity.

6.6.2. Adverse Events of Special Interest

Not applicable.

6.6.3. Deaths

The frequency (number and percentage) of patients in the AT population who died and who also died within 30 days after last dose of study treatment as well as the primary reason for death will be tabulated.

Date and cause of death will be provided in individual patient data listing together with selected dosing information (date of first/last administration, dose).

6.6.4. Laboratory Data

Laboratory data in this study consist of hematology values and chemistry 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.

Laboratory results will be converted to International System of Units (SI) units which will be used for applying toxicity grades and for all summaries.

Quantitative data will be summarized using simple descriptive statistics (mean, standard deviation, median, quartiles, minimum, and maximum) of actual values and change from baseline for each nominal visit over time (ie, unscheduled assessments will be excluded). The total number of patients for change from baseline will include all patients in the treatment arm who have both a baseline and a value at the nominal visit.

As described in Section 3.4, baseline will defined as the last assessment performed on or prior to date of the first dose of study treatment. If there are multiple assessments that meet the baseline definition on the same day without the ability to determine which was truly last, then the worst grade will be assigned as the baseline grade.

Results collected as strict inequalities (eg, >10, <10) will be converted to numeric values by subtracting a factor of <0.001>. Expressions of the form" ≥" or "≤" will be converted to the end point. These numeric values will be evaluated for clinically significant abnormalities, but will not be included in calculations of summary statistics.

Additionally, laboratory results will be programmatically classified according to NCI-CTCAE version 4.0 grade. Non-numerical qualifiers will not be taken into consideration in the derivation of grade (eg, hypokalemia Grade 1 and Grade 2 are only distinguished by a non-numerical qualifier and therefore Grade 2 will not be derived). In summary statistics the number and percentage of patients corresponding to grades that only include non-quantitative criteria will be displayed as a blank or NA (not assessed) rather than 0. If there is any overlap between grade criteria (eg, CTCAE grading criteria for Creatinine Increased – a value can fall into one range based on comparison to ULN and another range based on comparison to baseline), the highest (worst) grade would be assigned to that record. Grade 5 is defined in the CTCAE criteria guidance as an event with an outcome of death. Since laboratory data does not collect an outcome, Grade 5 is not used when programmatically grading laboratory data. Grade 0 or Outside Toxicity Reference (OTR) is not defined specifically by in the CTCAE guidance. However, programmatically this is used as a category to represent those patients who did not meet any of the Grades 1 to 4 criteria. If the laboratory value is evaluable for CTCAE criteria grading (numeric value is present, valid units and ranges are present as required to allow conversion to standard units and grading), and does not qualify for any of the Grade 1-4 criteria for a given lab test, then the value is assigned as Grade 0 or OTR.

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 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. The change from baseline to post baseline value will be calculated for each laboratory parameter where available.

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 post baseline toxicity grade during the treatment-emergent period. The number and percentage of patients will be provided for each parameter.

For each laboratory parameter that is not able to be graded 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 post baseline 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, assessment date (day), laboratory value, and normal range flag (low, normal, and high). The baseline value will be flagged in this data listing.

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

Categories of Liver Function Test Elevations

Parameter	Elevation
AST	>5 x ULN
ALT	>5 x ULN
Total bilirubin	>5 x ULN

ALT, alanine aminotransferase; AST, aspartate aminotransferase; ULN, upper limit of normal

6.6.5. Vital Signs

Systolic and diastolic blood pressure (mmHg), height (cm), and weight (kg) at baseline and at all subsequent scheduled time points will be summarized. Changes from baseline value will be presented for all scheduled time points.

All recorded vital sign data will be listed.

6.6.6. Electrocardiogram

Echocardiogram findings that were collected during the screening period will be listed by patient.

6.6.7. ECOG

Descriptive summaries at baseline and at the safety follow-up will be provided. These include:

- Number and percentage of patients for each of the six categories.

The ECOG performance status scores will be listed.

7. INTERIM ANALYSES

Not applicable.

8. REFERENCES

1. Brookmeyer R, Crowley JJ. A confidence interval for median survival time. Biometrics. 1982 Mar;38 (1):29-41.

2. Clopper, C. J.; Pearson, E. S. (1934). "The use of confidence or fiducial limits illustrated in the case of the binomial". Biometrika. 26: 404–413.

3. Eisenhauer E, Therasse P, Bogaerts J, Schwartz LH, Sargent D, Ford R, Dancey J, Arbuck S, Gwyther S, Mooney M, Rubinstein L, Shankar L, Dodd L, Kaplan R, Lacombe D, Verweij J. New response evaluation criteria in solid tumors: Revised RECIST guideline version 1.1. Eur J Can 45: 228-47, 2009.

4. Kaplan EL, Meier P. Nonparametric estimation from incomplete observations. J Am Stat Assoc 1958; 53:457-481.

5. National Cancer Institute (NCI). Common Terminology Criteria For Adverse Events (CTCAE). http://evs.nci.nih.gov/ftp1/CTCAE/About.html

9. APPENDICES

Appendix 1. Imputation of Missing/Partially Missing Dates

Missing data will not be imputed unless otherwise specified. Any imputations will occur at the analysis dataset level and not the raw dataset level. Additionally, in all patient data listings imputed values will be presented and flagged as imputed.

For safety analyses, incomplete date of last dose of study drug such as missing the day, the 15th of the month will be used to impute the missing data. When imputing partial last dose dates, the last assessment date and death date will be taken into consideration. This imputation rule will be used to determine the treatment-emergent period.

Adverse Events

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

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

- If both month and day are missing and year = year of treatment start date, then set to treatment start date.
- If both month and day are missing and year ≠ year of treatment start date, then set to January 01.
- If day is missing and(month and year) = (month and year) of treatment start date, then set to treatment start date.
- If day is missing and (month and year) ≠ (month and year) of treatment start date, then set to first of the month.
- If start date is completely missing, set to treatment start date as long as adverse event end date is not before treatment start date.

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

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

Concomitant Medications

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

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

- If both month and day are missing, then set to January 01.
- If only day is missing, then set to the first of the month.

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

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

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

Diagnosis of Primary Cancer

If the diagnosis date of primary cancer is partially missing, the following rules will be applied to impute partial dates:

- If both month and day are missing and year = year of treatment start date, then set to treatment start date.
- If both month and day are missing and year ≠ year of treatment start date, then set to December 31.
- If day is missing and (month and year) = (month and year) of treatment start date, then set to treatment start date.
- If day is missing and (month and year) ≠ (month and year) of treatment start date, then set to the last day of the month.

Appendix 2. Sample SAS Code

Sample SAS code follows below. The code is meant for a general start. Further dialogue between the programmer and statistician is encouraged.

KM quartiles and plot

```
proc lifetest data=x1 alpha=0.05 conftype=loglog plot=s;
 time months*eventn(0);
 run;
```

The variable "months" is a continuous variable representing months of exposure. The variable "eventn" is a binary variable (0=censor, 1=event).

Hazard ratio

```
proc phreg data=x1;
 model months*eventn(0) = group / rl alpha=0.05;
 run;
```

The variable "months" is a continuous variable representing months of exposure. The variable "eventn" is a binary variable (0=censor, 1=event). The variable "group" is a binary variable for median cutpoint (below/above the median change from baseline).

Clopper-Pearson Confidence Interval

```
proc sort data=y1 out=y2;
 by descending response;

proc freq data=y2 order=data;
 tables response / binomial(exact) alpha=0.05;
 run;
```

The descending sort is included to ensure that the response modeled is 1 and not 0. The input dataset y1 has a one-record-per-patient format similar to:

PATIENT	RESPONSE
1001	0
1002	1
1003	0
…	…
…	…

ANCOVA Analysis

```
proc mixed data=z1;
 class treat studysite;
 model efchg = treat efbase studysite;
 lsmeans treat / cl alpha=0.05 om;
 run;
```

The variables "efchg" and "efbase" represent change from baseline and baseline values respectively. The OM option is used when the categorical term "studysite" is used in the model. If "studysite" is not included for convergence reasons, then the OM option can be removed.

While there is only one treatment arm in this study, the inclusion of the fixed effect for "treat" is needed in order to produce the desired least squares mean.

Appendix 3. Analysis Windows for Reporting

The below applies to certain windowed summaries and analyses (eg, descriptive statistics by visit, change from baseline by visit, etc.). Windows will be based on the patient's date of first dose of study medication; the first day of dosing is Day 1.

Adverse events, however, do not follow windows.

If there is more than one observation in a specific window, then the one closest to the target day will be selected for analysis. If there are two observations that are equidistant to the target day, then the first of these two observations will be selected for analysis. If there are two or more observations on the same selected day within the same analysis window, then the mean of these observations will be used.

Week	Target Day	Analysis Window
Week 5	29	Day 2 to Day 57
Week 13	85	Day 58 to Day 113
Week 21	141	Day 114 to Day 169
Week 29	197	Day 170 to Day 225
Week 37	253	Day 226 to Day 281
Week 45	309	Day 282 to Day 337
Week 53	365	Day 338 to Day 393
Week 61	421	Day 394 to Day 449
Week 69	477	Day 450 to Day 505
Week 77	533	Day 506 to Day 561
Week 85	589	Day 562 to Day 617
Week 93	645	Day 618 to Day 673
Week 101	701	Day 674 to Day 729
Week 109	757	Day 730 to Day 785
Week 117	813	Day 786 to Day 841
Week 125	869	Day 842 to Day 897
Week 133	925	Day 898 to Day 953
Week 141	981	Day 954 to Day 1009
Week 149	1037	Day 1010 to Day 1065
Week 157	1093	≥ Day 1066

Additional weeks if needed are derived by adding 8 weeks onto Week, and adding 56 days onto Target Day. Lower and upper bounds are also adjusted by adding 56 days.

Appendix 4. Summary of Efficacy Analyses

Endpoint	**Classification**	**Population**	**Descriptive** analysis	**Inferential** analysis
Proportion with responding bone lesions at NaF-3	Primary	AT*	Number and %	Point estimate and CI using Clopper-Pearson method
Heterogeneity of response at NaF-3	Secondary	AT*	Summary stats and histograms	

CCI

[redacted]

Endpoint	Classification	Population	Descriptive analysis	Inferential analysis
CCI				

Endpoint	Classification	Population	Descriptive analysis	Inferential analysis
CCI				

Endpoint	Classification	Population	Descriptive analysis	Inferential analysis
CCI				

Endpoint	Classification	Population	Descriptive analysis	Inferential analysis
CCI				

Endpoint	Classification	Population	Descriptive analysis	Inferential analysis
CCI				

Endpoint	Classification	Population	Descriptive analysis	Inferential analysis
CCI [redacted]				

Chg: change; bsl: baseline; Obs: observed, LSMean: least-squares mean

* And non-missing scans at NaF-1 and NaF-3.

CCI [redacted]

[redacted]

10. ABBREVIATIONS

AE	Adverse event
ANCOVA	Analysis of covariance
ALT	Alanine aminotransferase
AST	Aspartate aminotransferase
AT	As treated
BLQ	Below the level of quantification
CI	Confidence interval
CSR	Clinical study report
CTC	Circulating tumor cells
CTCAE	Common terminology criteria for adverse events
EpCAM	Epithelial cell adhesion molecule
GM	Geometric mean
HR	Hazard ratio
KM	Kaplan-Meier
LOA	Limit of agreement
LSMean	Least-squares mean
MedDRA	Medical dictionary for regulatory activities
NCI	National Cancer Institute
OM	Observed margins
PSA	Prostate specific antigen
PT	Preferred term
RECIST	Response evaluation criteria in solid tumors
SAP	Statistical analysis plan
SOC	System organ class
StD	Standard deviation
SUV	Standardized uptake value
TEAE	Treatment emergent adverse event
ULN	Upper limit of normal