STATISTICAL ANALYSIS PLAN

Final Version 1.00, dated 8-July 2020

A Randomized Phase IV Study Comparing Enzalutamide versus Flutamide in CRPC Patients Who Have Failed Combined Androgen Blockade Therapy with Bicalutamide plus ADT

ISN: 9785-MA-3051

Sponsor name
Astellas Pharma Inc. (API)/

2-5-1, Nihonbashi-Honcho, Chuo-ku, Tokyo

Table of Contents

I. LIST OF ABBREVIATIONS AND KEY TERMS

List of Abbreviations

Abbreviations	Description of abbreviations
AAT	Alternative Antiandrogen Therapy
CI	Confidence Interval
ADT	Androgen Deprivation Therapy
AE	Adverse Event
ALP	Alkaline Phosphatase
ALT	Alanine Aminotransferase
AST	Aspartate Aminotransferase
BFI	Brief Fatigue Inventory
BPI-SF	Brief Pain Inventory-Short Form
BMI	Body Mass Index
CAB	Combined Androgen Blockade
CR	Complete Response
CSR	Clinical Study Report
CT	Computed Tomography
CTCAE	Common Terminology Criteria for Adverse Events
ECG	Electrocardiogram
ECOG	Eastern Cooperative Oncology Group
EQ-5D-5L	European Quality of Life-5 Dimension-5 Level Instruments
FACT-P	Functional Assessment of Cancer Therapy–Prostate
GnRH	Gonadotropin Releasing Hormone
ICH	International Conference on Harmonization
ITT	Intent-to-Treat
LDH	Lactate Dehydrogenase
HIFU	High Intensity Focused Ultrasound
MedDRA	Medical Dictionary for Regulatory Activities
MFS	Metastasis-Free Survival
MRI	Magnetic Resonance Imaging
NE	Non-evaluable
ORR	Objective Response Rate
PCWG2	Prostate Cancer Clinical Trials Working Group 2
PD	Progressive Disease
PDAS	Pharmacodynamic Analysis Set
PK	Pharmacokinetics
PKAS	Pharmacokinetic Analysis Set
PR	Partial Response
PT	Preferred Term
PSA	Prostate-Specific Antigen
QOL	Quality of Life
RECIST	Response Evaluation Criteria in Solid Tumors
rPFS	Radiographic Progression-Free Survival
SAE	Serious Adverse Event
SAF	Safety Analysis Set
SAP	Statistical Analysis Plan
SOC	System Organ Class
SMQ	Standard MedDRA Query
TURP	Transurethral Resection of the Prostate

TTF1	Time to Treatment Failure 1
TTF2	Time to Treatment Failure 2
TTPP1	Time to PSA Progression 1
TTPP2	Time to PSA Progression 2
2nd TTPP	2nd Time to PSA Progression
ULN	Upper Limit of Normal
WHO-DD	World Health Organization Drug Dictionary

List of Key Terms

Terms	Definition of terms
Endpoint	A variable that pertains to the trial objectives
Variable	Any quantity that varies; any attribute, phenomenon or event that can have different qualitative or quantitative values.

1 INTRODUCTION

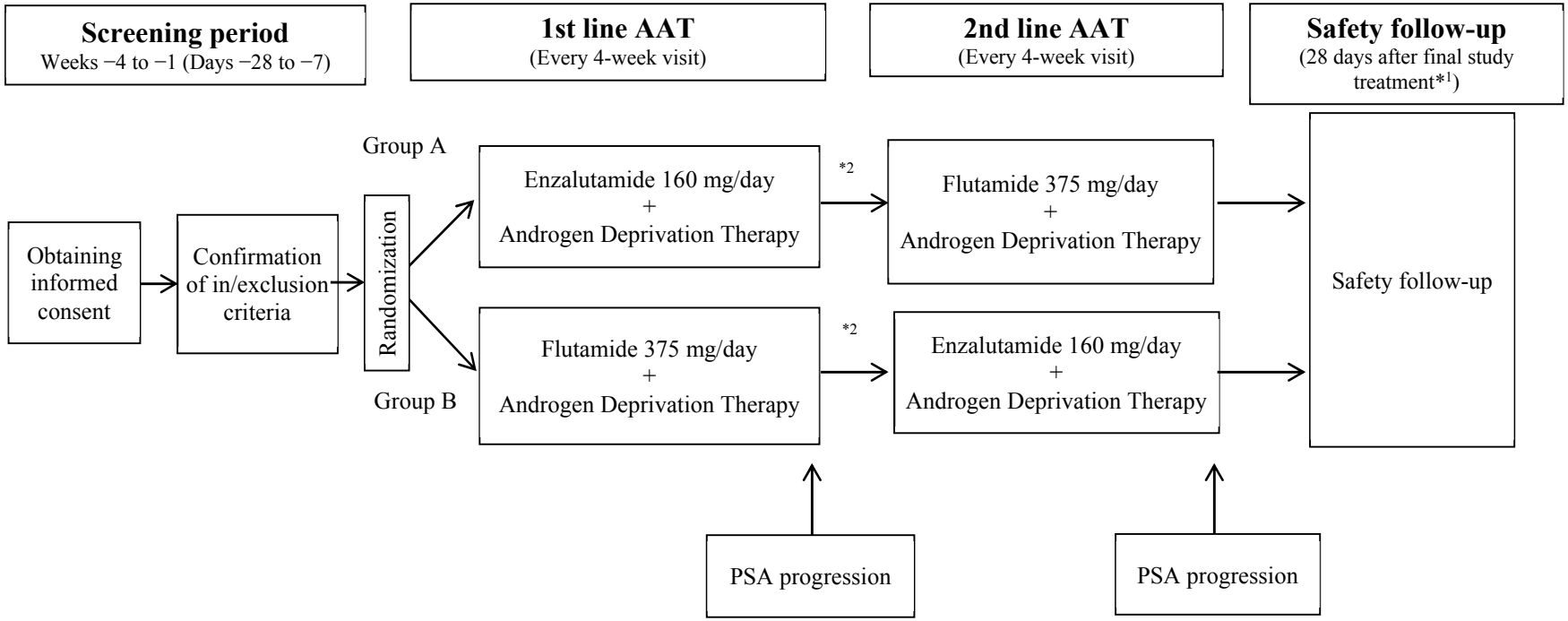
This Statistical Analysis Plan (SAP) contains a more technical and detailed elaboration of the principal features of the analysis described in the protocol, and includes detailed procedures for executing the statistical analysis of the primary and secondary endpoints and other data.

The SAP is finalized and signed prior to database hard lock.

This statistical analysis is prepared by [redacted] and approved by the responsible biostatistician of Astellas Pharma Inc.. Any changes from the analyses planned in the SAP will be justified in the Clinical Study Report (CSR).

2 FLOW CHART AND VISIT SCHEDULE

Flow Chart



*1: 28 days after final study treatment, or before starting other study drugs, or before starting other treatment for prostate cancer, whichever comes first

*2: Subjects who discontinued the 1st line AAT due to PSA progression will start the 2nd line AAT within 6 weeks after the date of examination that identified PSA progression. Subjects who discontinued the 1st line AAT due to other reasons will start the 2nd line AAT within 6 weeks from the day when the investigator or subinvestigator decided the discontinuation.

Table 1: Schedule of Assessments

1st line AAT

	Screening period	1st line AAT					Unscheduled[*a]
Visit Days	Days −28 to −7	1	29	57	85 and every subsequent 84 days	113 and every subsequent 28 days	–
Visit Weeks	−4 to −1	1	5	9	13 and every subsequent 12 weeks	17 and every subsequent 4 weeks	–
Allowed visit window (days)	–	–	± 7	± 7	± 7	± 7	–
Informed consent[*b]	X [*b]						
Subject background survey	X						
Confirmation of in/exclusion criteria	X	X					
Randomization/enrollment		X [*c]					
Vital signs[*d]	X	X	X	X	X	X	X
Height/body weight	X						
Laboratory tests[*e]	X	X	X	X	X	X	X
PSA	X	X			X	X	X
Abdominopelvic CT/MRI and bone scintigraphy (every 3 months)	X [*f]				X		X
Chest CT	X [*f]				X [*g]		X
ECOG Performance Status	X	X	X	X	X	X	X
FACT-P		X			X		
EQ-5D-5L		X			X		
Brief Pain Inventory-Short Form (BPI-SF)[*h]	X	X			X		
Brief Fatigue Inventory (BFI)	X	X			X		
Adverse event survey[*i]		X	X	X	X	X	X
Survey on previous/concomitant medication	X	X	X	X	X	X	X
Study drug prescription (every month)		X	X	X	X	X	

CT: Computed Tomography, ECOG: Eastern Cooperative Oncology Group, FACT-P: Functional Assessment of Cancer Therapy–Prostate

*a: Unscheduled visit may be made whenever adverse events should be assessed, or PSA progression and disease progression should be confirmed. Laboratory tests, PSA measurement, and diagnostic imaging will be performed when assessment is required.

*b: Informed consent will be obtained within 42 days before enrollment (Day 1) and before initiating the scheduled study procedures.

*c: Randomization should be performed on Day 1 after the confirmation of eligibility.

*d: Measure vital signs (blood pressure and pulse rate) at each visit.
*e: Laboratory tests include hematology, biochemistry, blood glucose test, and urinalysis.
*f: Imaging diagnosis performed before informed consent and within 42 days before randomization can be used as baseline data.
*g: Not required if no chest metastasis is identified in the screening chest CT.
*h: Instruct subjects to assess prostate cancer-related pains.
*i: Collect all adverse events in the period from the start of study treatment to the date of safety follow-up visit.

2nd line AAT

	2nd line AAT					Unscheduled[*1]	At the end or discontinuation[*2]	Safety follow-up
Visit Days	1	29	57	85 and every subsequent 84 days	113 and every subsequent 28 days	–	–	–
Visit Weeks	1	5	9	13 and every subsequent 12 weeks	17 and every subsequent 4 weeks	–	Final treatment	28 days after final treatment[*3]
Allowed visit window (days)	–	± 7	± 7	± 7	± 7	–	Within 7 days after final treatment[*11]	± 7
Enrollment	X						X	
Vital signs[*4]	X	X	X	X	X	X	X	X
Laboratory tests [*5]	X	X	X	X	X	X	X	X
PSA	X			X	X	X	X	X
Abdominopelvic CT/MRI and bone scintigraphy (every 3 months)				X [*6]		X	X [*12]	
Chest CT				X [*6,7]		X	X [*12]	
ECOG Performance Status	X	X	X	X	X	X	X	X
FACT-P	X [*8]			X [*8]			X [*12]	
EQ-5D-5L	X [*8]			X [*8]			X [*12]	
Brief Pain Inventory-Short Form (BPI-SF) [*9]	X [*8]			X [*8]			X [*12]	
Brief Fatigue Inventory (BFI)	X [*8]			X [*8]			X [*12]	
Adverse event survey[*10]	X	X	X	X	X	X	X	X
Previous/concomitant medication	X	X	X	X	X	X	X	X
Study drug prescription (every month)	X	X	X	X	X			

CT: Computed Tomography, ECOG: Eastern Cooperative Oncology Group, FACT-P: Functional Assessment of Cancer Therapy–Prostate

*1: Unscheduled visit may be made during the study period whenever adverse events should be assessed, or PSA progression and disease progression should be confirmed. Laboratory tests, PSA measurement, and diagnostic imaging will be performed when assessment is required. If the study is discontinued at an unscheduled visit, perform laboratory tests, PSA measurement, and diagnostic imaging wherever possible.

*2: For subjects who finished or discontinued the study with the 1st line AAT or 2nd line AAT, immediately perform examination/observation wherever possible.

*3: Conduct a visit for safety follow-up 28 days after the final study treatment, or before start of cytocidal chemotherapy, or before start of other new treatment for prostate cancer, whichever comes first.

*4: Measure vital signs (blood pressure and pulse rate) at each visit.

*5: Laboratory tests include hematology, biochemistry, blood glucose test, and urinalysis.

*6: Perform diagnostic imaging every 12 weeks after Day 1 of 1st line AAT, independently from switching medications.

*7: Not required if no chest metastasis is identified in screening chest CT.

*8: Start the evaluation after switching the drugs on Day 1 of 2nd line AAT, and implement every subsequent 12 weeks.

*9: Instruct subjects to assess prostate cancer-related pains.

*10: Collect all adverse events until the date of safety follow-up visit. If there is no safety follow-up visit, collect adverse events until 28 days after the final study treatment. If subjects do not show up on the safety follow-up visit, confirm adverse events by the telephone wherever possible.

*11: For subjects who discontinued the study treatment after the treatment interruption due to adverse events, adopt data within 7 days after the final study treatment as discontinuation data. If no data is available in the period within 7 days after the final study treatment, perform a discontinuation examination/observation immediately after the decision of discontinuation.

*12: Not required if it occurs within 28 days from the latest measurement.

3 STUDY OBJECTIVE(S) AND DESIGN

3.1 Study Objective(s)

3.1.1 Primary Objective

- To determine the benefit of enzalutamide + Androgen Deprivation Therapy (ADT) therapy as compared to flutamide + ADT therapy as assessed by time to prostate-specific antigen (PSA) progression with 1st line alternative antiandrogen therapy (AAT) (time to PSA progression 1 [TTPP1]).

3.1.2 Secondary Objectives

- To determine the order of AAT treatment as assessed by time to PSA progression with 1st line AAT + 2nd line AAT (time to PSA progression 2 [TTPP2]).
- To determine the benefit of enzalutamide + ADT therapy as compared to flutamide + ADT therapy as assessed by PSA response rate to 1st line AAT (proportion of subjects with a decrease by at least 50% or 90% from baseline irrespective of the timing).
- To determine the benefit of enzalutamide + ADT therapy as compared to flutamide + ADT therapy as assessed by PSA response rate to 1st line AAT at Week 13 (proportion of subjects with a decrease by at least 50% or 90% from baseline).
- To determine the benefit of enzalutamide + ADT therapy as compared to flutamide + ADT therapy as assessed by time to PSA decrease by 50% from baseline with 1st line AAT.
- To determine the benefit of enzalutamide + ADT therapy as compared to flutamide + ADT therapy as assessed by time to treatment failure of 1st line AAT (time to treatment failure 1 [TTF1]).
- To determine the order of AAT treatment as assessed by time to treatment failure of 2nd line AAT (time to treatment failure 2 [TTF2]).
- To determine the benefit of enzalutamide + ADT therapy as compared to flutamide + ADT therapy as assessed by radiographic progression-free survival (rPFS).

3.1.3 Safety Objective

- To determine the safety of enzalutamide + ADT therapy as compared to flutamide + ADT therapy.

3.1.4 Exploratory Objectives

- To evaluate time to PSA progression with 2nd line AAT (2nd time to PSA progression [2nd TTPP]).
- To evaluate PSA response rate to 2nd line AAT.
- To determine the benefit of enzalutamide + ADT therapy as compared to flutamide + ADT therapy as assessed by metastasis-free survival (MFS).

- To evaluate Quality of Life (QOL) on 1st line AAT and 2nd line AAT using Functional Assessment of Cancer Therapy–Prostate (FACT-P) and European Quality of Life 5-Dimension-5 Level instruments (EQ-5D-5L).
- To evaluate Brief Pain Inventory-Short Form (BPI-SF) with 1st line AAT and 2nd line AAT.
- To evaluate Brief Fatigue Inventory (BFI) with 1st line AAT and 2nd line AAT.
- To determine the benefit of enzalutamide + ADT therapy as compared to flutamide + ADT therapy as assessed by objective response rate (ORR) in soft tissue with 1st line AAT based on the best overall response in Response Evaluation Criteria in Solid Tumors (RECIST) guidelines.
- To evaluate the efficacy in subgroups according to disease-related patient demographics.

3.2 Study Design

This is a randomized, open-label, comparative study on enzalutamide and flutamide.

Patients with M0 or M1 castration-resistant prostatic neoplasm that relapsed during combined androgen blockade (CAB) therapy with bicalutamide will be randomized to Group A or B. Dynamic allocation will be performed with minimization with biased coin technique. The following stages (M0/N0, M0/N1, or M1) will be used as allocation factors.

- M0/N0: No distant metastasis, and no lymph node metastasis
- M0/N1: Without distant metastasis, but with metastasis in lymph nodes distal to the aortic bifurcation
- M1: With distant metastasis (including metastasis in lymph nodes proximal to the aortic bifurcation)

For subjects in Group A, enzalutamide is administered as the 1st line AAT. After the confirmation of PSA progression, medication is changed from enzalutamide to flutamide as the 2nd line AAT. For subjects in Group B, flutamide is administered as the 1st line AAT. After the confirmation of PSA progression, medication is changed from flutamide to enzalutamide as the 2nd line AAT. The treatment period with each drug should not be more than 2 years from the enrollment of the last subjects.

PSA progression will be defined according to the consensus guidelines of prostate cancer clinical trials working group 2 (PCWG2). For patients with PSA declines at Week 13, the PSA progression date is defined as the date that a $\geq$ 25% increase and an absolute increase of $\geq$ 2 ng/mL above the nadir is documented, which is confirmed by a second consecutive value obtained 3 or more weeks later. For patients with no PSA decline at Week 13, the PSA progression date is defined as the date that a $\geq$ 25% increase and an absolute increase of $\geq$ 2 ng/mL above the baseline is documented.

Data for the analysis of primary endpoint TTPP1 will be cut-off if not less than 135 events of PSA progression occurred in association with the 1st line AAT.

After the discontinuation of the 1st line AAT due to PSA progression or other reasons, the 2nd line AAT will be started. Subjects who discontinued the 1st line AAT due to PSA

progression will start the 2nd line AAT within 6 weeks after the date of examination that identified PSA progression. Subjects who discontinued the 1st line AAT due to other reasons will start the 2nd line AAT within 6 weeks from the day when the investigator or sub-investigator decided the discontinuation.

A safety follow-up visit should be performed wherever possible in 28 days from the final study treatment, or before starting other study drugs, or before starting other treatment for prostate cancer, whichever comes first. If subjects do not show up on the safety follow-up visit, confirm Adverse Events (AEs) by the telephone wherever possible.

3.3 Randomization

The subject registration center will perform dynamic allocation with minimization with biased coin technique according to the subject registration procedure in the ratio of 1:1 to the initial enzalutamide 160 mg/day group or the initial flutamide 125 mg × 3 times daily (375 mg/day) group. The facilities and the following stages (M0/N0, M0/N1, or M1) will be used as allocation factors.

- M0/N0: No distant metastasis, and no lymph node metastasis
- M0/N1: Without distant metastasis, but with metastasis in lymph nodes distal to the aortic bifurcation
- M1: With distant metastasis (including metastasis in lymph nodes proximal to the aortic bifurcation)

After the investigator or sub-investigator confirmed the eligibility of the subject on Day 1, the staff in charge at the study site will register the subject via the Subject Registration System, and confirm the study drugs that are randomly allocated.

4 SAMPLE SIZE

The number of evaluable subjects will be approximately 200 subjects (1st line AAT enzalutamide group: 100, 1st line AAT flutamide group: 100).

The sample size to evaluate the time to PSA progression with 1st line AAT (TTPP1), which is the primary endpoint of the study, was calculated in consideration of the following points:

- Median TTPP1 with enzalutamide: 10.5 months.
  [redacted]
- Median TTPP1 with flutamide: 6 months.
  Values reported in the literature: 5 months according to [Okihara et al., 2007], 6.25 months according to [Narimoto et al., 2010].
- The ratio of randomization of 1st line AAT enzalutamide and flutamide groups is 1:1.

- Subject enrollment period is 12 months, and the observation period is 24 months from the enrollment of the last subject.
- Type I error of the two-sided test is 0.05, and power is 90%.
- A log-rank test with the TTPP1 parameters shown above requires 135 events. The necessary number of subjects calculated based on the required number of events is 74 subjects per group.
  Required number of events and necessary number of subjects above are calculated with the formulas below based on [Schoenfeld, 1983]. The codes in the formulas are defined as follows.
  M_t: Median TTPP1 in enzalutamide group
  M_c: Median TTPP1 in flutamide group
  t_a: Enrollment period
  t_f: Observation period from the last subject enrollment
  $Z_{1-\alpha}$: 1-α% point of standard normal distribution
  $Z_{1-\beta}$: 1-β% point of standard normal distribution

  Hazard in enzalutamide group $(H_t) = \frac{\log_e 2}{M_t}$
  Hazard in flutamide group $(H_c) = \frac{\log_e 2}{M_c}$
  Hazard ratio for enzalutamide group/flutamide group $(HR) = \frac{H_t}{H_c}$
  Required number of events $(N_d) = \frac{4\times (Z_{1-\alpha}+Z_{1-\beta})^2}{(\log_e HR)^2}$ *The number after the decimal point will be rounded up to the nearest integer.
  Incidence of events of enzalutamide group
  $$(P_t) = 1 - \frac{1}{6}\left(\exp(-H_t \times t_f) + 4 \times \exp\left(-H_t \times (t_f + \frac{1}{2} \times t_a)\right) + \exp(-H_t \times (t_f + t_a))\right)$$
  Incidence of events of flutamide group
  $$(P_c) = 1 - \frac{1}{6}\left(\exp(-H_c \times t_f) + 4 \times \exp\left(-H_c \times (t_f + \frac{1}{2} \times t_a)\right) + \exp(-H_c \times (t_f + t_a))\right)$$
  Incidence of events of both groups $(P) = \frac{1}{2} \times (P_t + P_c)$
  Necessary number of subjects per group $(N) = \frac{1}{2} \times \frac{N_d}{P}$ * The number after the decimal point will be rounded up to the nearest integer.
- Considering the dropout rate of approximately 25%, 100 subjects are to be enrolled in each group.

Number of subjects	[redacted]	[redacted]
With PSA progression	[redacted]	[redacted]
Censored	[redacted]	[redacted]
25%	[redacted]	[redacted]
50%	[redacted]	[redacted]
75%	[redacted]	[redacted]

5 ANALYSIS SETS

In accordance with International Conference on Harmonization (ICH) recommendations in guidelines E3 and E9, the following analysis sets will be used for the analyses.

5.1 Intent-to-Treat (ITT)

Intent-to-Treat (ITT) population is defined as all subjects randomized in this study. The ITT population will be analyzed by randomized treatment group (treatment group based on randomization). Efficacy analyses will be done in the ITT population.

5.2 Safety Analysis Set (SAF)

For the statistical summary of the safety data, the safety analysis set (SAF) will be used. The SAF consists of all subjects who took at least one dose of study medication, and will be used for safety analyses.

5.3 Pharmacokinetic Analysis Set (PKAS)

The pharmacokinetic analysis set (PKAS) will not be established for this study.

5.4 Pharmacodynamic Analysis Set (PDAS)

The pharmacodynamic analysis set (PDAS) will not be established for this study.

6 ANALYSIS VARIABLES

6.1 Efficacy Endpoints

6.1.1 Primary Efficacy Endpoint(s)

Time to PSA Progression with 1st Line AAT (TTPP1)

PSA progression will be defined according to the consensus guidelines of PCWG2. For patients with PSA declines at Week 13, the PSA progression date is defined as the date that a

≥ 25% increase and an absolute increase of ≥ 2 ng/mL above the nadir is documented, which is confirmed by a second consecutive value obtained 3 or more weeks later. For patients with no PSA decline at Week 13, the PSA progression date is defined as the date that a ≥ 25% increase and an absolute increase of ≥ 2 ng/mL above the baseline is documented.
TTPP1 is defined as the period from the date of randomization to the date of PSA progression in the 1st line AAT period.

6.1.2 Secondary Efficacy Endpoints

Time to PSA progression with 1st line AAT + 2nd line AAT (TTPP2)
Time to PSA progression with 2nd line AAT is defined as the period from Day 1 of 2nd line AAT to the date of PSA progression with 2nd line AAT. TTPP2 will be the total of time to PSA progression with 1st line AAT and time to PSA progression with 2nd line AAT.

PSA response rate to 1st line AAT (decrease by at least 50% or 90% from baseline)
PSA response will be defined as PSA decreased by at least 50% or 90% from baseline when at least 3 weeks passed after the lowest PSA decreased by at least 50% or 90% from baseline in the 1st line AAT period after baseline. The proportion of the subjects who achieved PSA response will be evaluated.

PSA response rate to 1st line AAT at Week 13 (decrease by at least 50% or 90% from baseline)
PSA response will be defined as the lowest PSA at Week 13 decreased by at least 50% or 90% from baseline in the 1st line AAT period. The proportion of the subjects who achieved PSA response will be evaluated.

Time to PSA decrease by 50% from baseline with 1st line AAT
Time to PSA decrease by 50% with 1st line AAT is defined as the period from the date of randomization to the day when the decrease of PSA from baseline by 50% is first identified.

Time to treatment failure of 1st line AAT (TTF1)
Time to treatment failure is defined as the period to study drug discontinuation for any reason that includes disease progression, onset of AEs, subjects' request, or death. The following will be evaluated.
The period from randomization to discontinuation of 1st line AAT will be evaluated.

Time to treatment failure of 2nd line AAT (TTF2)
The period from randomization to discontinuation of 2nd line AAT will be evaluated.

rPFS

Imaging examination will be performed with chest CT, abdominopelvic CT or magnetic resonance imaging (MRI), and bone scintigraphy in the screening period, at study visit every 12 weeks after Week 1 (Day 1) of 1st line AAT irrespective of switching the drugs, and at

completion or discontinuation (chest CT will not be necessary at study visits after the screening period if chest metastasis is not observed in the image taken during the screening period). During the study period, imaging examination will be performed under the same condition for the same subject, and the image should be interpreted by the same reader at each study site wherever possible. Images taken within 42 days before randomization as part of daily consultation before informed consent may be used as data in the screening period. In case of disease progression suspected, imaging examination may be performed for confirmation at any time other than the scheduled examination.

The following will be evaluated with the images in the screening period as baseline.

rPFS is defined as the period from randomization to the time when radiographic disease progression is observed or death of any cause during the study period, whichever occurs earlier. With regard to radiographic disease progression, soft tissue lesion will be defined by RECIST guidelines version 1.1 (RECIST 1.1), bone lesion will be defined as the occurrence of at least 2 new bone lesions confirmed by bone scintigraphy in line with PCWG2 recommendation. Evaluation will be made in subjects with distant metastasis at baseline.

The following are procedures necessary for confirmation of radiographic disease progression.

Study visit where disease progression is observed	Criteria for disease progression	Criteria for confirmation of disease progression (necessity for imaging for confirmation and time of confirmation)	Criteria for imaging for confirmation of disease progression
Visits up to Week 13 of 1st line AAT	Bone lesion: ≥2 new bone lesions in comparison with baseline bone scintigraphy by PCWG2	Timing: ≥6 weeks after disease progression confirmed or at Week 25 Visit	≥2 new bone lesions in comparison with bone scintigraphy at Week 13 of 1st line AAT
	Soft tissue lesion: Disease progression detected by CT or MRI, and defined by RECIST 1.1	No need of imaging for confirmation	N/A
Visits after Week 13 of 1st line AAT	Bone lesion: ≥2 new bone lesions in comparison with bone scintigraphy at Week 13* of 1st line AAT	No need of imaging for confirmation	N/A
	Soft tissue lesion: Disease progression detected by CT or MRI, and defined by RECIST 1.1	No need of imaging for confirmation	N/A

*: Compare with baseline bone scintigraphy if data from Week 13 are missing.

6.1.3 Safety Endpoints

- Characteristics, frequency, and severity of AEs
- Safety-related laboratory tests (biochemistry and hematology)
- Vital signs (blood pressure and pulse rate)
- ECOG Performance Status

6.1.4 Exploratory Efficacy Endpoints

Time to PSA progression with 2nd line AAT (2nd TTPP)
2nd TTPP is defined as the period from the start date of 2nd line AAT (Day 1) to the date of PSA progression.

PSA response rate to 2nd line AAT (decrease by at least 50% or 90% from baseline)
PSA response will be defined as PSA decreased by at least 50% or 90% from baseline when at least 3 weeks passed after the lowest PSA decreased by at least 50% or 90% from baseline in the 2nd line AAT period after baseline. The proportion of the subjects who achieved PSA response will be evaluated.

MFS
MFS is defined as the period from randomization to new radiographic metastatic lesion or death of any cause during the study period, whichever occurs earlier. Evaluation will be made in subjects in whom distant metastasis is not confirmed at baseline.

QOL assessment with 1st line AAT and 2nd line AAT by FACT-P) and EQ-5D-5L
FACT-P

The assessment will be conducted at study visit at Week 1 (Day 1), Week 13, and every subsequent 12 weeks of 1st line AAT and 2nd line AAT, and at completion or discontinuation.

FACT-P is a multi-faceted QOL scale of self-assessment for patients with prostate cancer. FACT-P includes 27 major items to evaluate subjects' functions in 4 areas; that is, physical health, social/family, emotional, and functional well-being. Site-specific 12 items are added for assessment of prostate-related symptoms. Each item will be assessed with a 5-grade Likert scale. The subtotal in each area is a low rank score. The total of all items is the overall QOL score. The higher the QOL score, the higher is the QOL.
If there are missing items, subscale scores can be prorated. This is done by multiplying the sum of subscale by the number of items in the subscale, then dividing by the number of items actually answered. This can be done on the scoring guide or by using the formula below:
Prorated subscale score = [Sum of item scores] x [N of items in subscale]÷ [N of items answered]
When there are missing data, prorating by subscale in this way is acceptable as long as more than 50% of the items were answered (e.g., a minimum of 4 of 7 items, 4 of 6 items, etc). The total score is then calculated as the sum of the unweighted subscale scores. The FACT scale

is considered to be an acceptable indicator of patient quality of life as long as overall item response rate is greater than 80% (e.g., at least 22 of 27 FACT-G items completed). This is not to be confused with individual subscale item response rate, which allows a subscale score to be prorated for missing items if greater than 50% of items are answered.
For the Prostate Cancer-specific subscale, the procedure for scoring is the same as described above for the FACT-G. Again, over 50% of the items (e.g., 5 of 9 items, 7 of 12 items) must be completed in order to consider each subscale score valid. The total score consists of the sum of the FACT-G (the first 4 subscales common to almost all scales) plus the prostate-specific subscale.

EQ-5D-5L

EQ-5D-5L will be performed at study visit at Week 1 (Day 1), Week 13, and every subsequent 12 weeks of 1st line AAT and 2nd line AAT, and at completion or discontinuation.

EQ-5D-5L is a QOL scale for self-assessment that consists of 5 items related to health, such as mobility, self care, usual activities, pain/discomfort, and anxiety/depression. Each item will be evaluated in 5 levels from “no problem” to “extreme problem.” The last question is a visual analog scale to evaluate the present health status in the range from “best imaginable health status” to “worst imaginable health status.”

QOL utility index will be derived in accordance with the scoring method of the EQ-5D-5L Japanese version.

BPI-SF with 1st line AAT and 2nd line AAT

BPI-SF will be performed in the screening period, at study visit at Week 1 (Day 1), Week 13, and every subsequent 12 weeks of 1st line AAT and 2nd line AAT, and at completion or discontinuation.

BPI is a questionnaire slip verified as a self-assessment scale for subjects about the level of pain, effect of the pain on activities of daily living, and analgesic use. BPI used in this study is an abbreviated version (short form) that consists of questions in 9 items for which simple numerical evaluation scales from 0 to 10 are used. In order to evaluate prostate cancer-related pain in this study, the investigator or sub-investigator will instruct the subjects to describe prostate cancer-related pains.

The 4 individual items used to evaluate pain intensity are ‘worst’, ‘least’, ‘average’, and ‘now’ (current pain). Pain ratings range from 0 to 10 for each item. The mean pain severity score will be calculated using the 4 individual items. This mean can be used if all 4 items are completed at a given assessment. If any individual item is not completed, the mean cannot be used. Descriptive statistics will be provided to summarize each individual item and the mean score at baseline, and each visit for the ITT population.

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 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%, or 4 of 7, of the total items have been completed at a given assessment.

BFI with 1st line AAT and 2nd line AAT

BFI will be performed in the screening period, at study visit at Week 1 (Day 1), Week 13, and every subsequent 12 weeks of 1st line AAT and 2nd line AAT, and at completion or discontinuation.

BFI is a simple questionnaire slip that consists of questions in 10 items to evaluate the malaise (subjective symptoms characterized as debility that includes physical and mental wasting) of cancer patients.

The mean of the 9items (except for the first Yes/No question) are used as a global BFI score.

When there are missing data, global BFI score can be acceptable as long as more than 50% of the 9 items (except for the first Yes/No question) were answered (e.g., a minimum of 5 of 9 items). The global BFI score is then calculated as the mean of answered questions.

ORR in best overall soft tissue response of RECIST guidelines with 1st line AAT

According to RECIST 1.1, the investigator or sub-investigator will evaluate the overall response in target, non-target, and new lesions in soft tissue at each time point. At the end of the 1st line AAT period, the best overall response will be assessed. Of subjects with measurable soft tissue lesion (at least 1 target lesion according to RECIST 1.1) in baseline imaging, the proportion of subjects with objective response (complete or partial response) will be evaluated.

6.1.5 Other Efficacy Variables

PSA measurement

In the screening period, at study visit at Week 1 (Day 1), Week 13, Week 17, and every subsequent 4 weeks of 1st line AAT and 2nd line AAT, at completion or discontinuation, and at safety follow-up visit, blood sample for PSA will be collected, and PSA will be measured by the central laboratory designated by the Sponsor. The PSA measured on Week 1 (Day 1) of 1st line AAT and 2nd line AAT will be handled as baseline.

6.2 Safety Variables

Safety will be assessed by evaluation of the following variables:

- AEs (frequency, severity, seriousness, and relationship to study drug).
- Clinical laboratory variables (hematology, biochemistry including liver enzymes and total bilirubin, and urinalysis)
- Vital signs (systolic and diastolic blood pressure, pulse rate)
- ECOG Performance Status

AEs will be recorded from time of actually received the study drugs until the Safety Follow-up visit (28 days after the final treatment with the study drug, start date of cytocidal chemotherapy, or start date of new treatment for prostate cancer, whichever occurs first.). AEs will be assigned on treatment period (treatment arm) according to definition of treatment-emergent period in section 7.1.2. When assigning treatment arm cannot be judged due to starting date and/or end date of AE is missing, such AEs will be assigned on treatment depending on recorded CRF page. However, the AE that recorded as continues from 1st line AAT will be excluded from assignment of 2nd line AAT. TEAE is defined as an AEs observed within above treatment-emergent period or AEs that assigned the treatment arms. The severity of all AEs is to be evaluated by the investigator based on the National Cancer Institute's Common Terminology Criteria for Adverse Events (CTCAE) version 4.0.

A drug-related AE is defined as any AE with at least possible relationship to study treatment as assessed by the investigator or with missing assessment of the causal relationship.

6.3 Pharmacokinetic Variables

Not applicable.

6.4 Pharmacodynamic Variables

Not applicable.

6.5 Other Variables

- The duration of exposure
  For each subject and period, the length of time on treatment will be calculated in months, using the following formula.

  Treatment duration for 1st line (Months) = (Date of last dose study drug in 1st line-Date of first dose of study drug in 1st line+1)/30.4375

  Treatment duration for 2nd line (Months) = (Date of last dose study drug in 2nd line-Date of first dose of study drug in 2nd line+1)/30.4375

  Total treatment duration for 1st and 2nd line (Months) = Treatment duration for 1st line + Treatment duration for 2nd line

- Total number of study drugs taken

Total number of study drugs taken during the study is calculated as (number of study drugs dispensed at all study visits - number of study drugs indicated as having been returned and lost). Study drugs not returned will be assumed to have been taken.

Treatment compliance
Study drug accountability will be performed to document compliance with the dosing regimen. Subjects will be asked to bring back all remaining study drug at each study visit. Treatment compliance will be defined as the number of study drugs taken during the study divided by the expected number of study drugs, multiplied by 100%.

Treatment compliance for 1st or 2nd line (%)=

$$\frac{\text{[Total number of study drugs taken during 1st or 2nd line]}}{\text{[Expected number of study drugs during corresponding line]}} \times 100$$

Expected number of study drugs is calculated as sum for i (Xi x Yi x (assessment date of arbitrary visit i excluding first treatment of each line – date of visit (i-1) +1)) within each line, where i is visit sequence (=2nd ,3rd,...) which include unscheduled visit, X=1 and Y=4 for Enzalutamide plus ADT arm and X=3 and Y=1 for Flutamide plus ADT arm which are number of study drugs to be taken par day according to protocol. When prescribed dose amount is reduced, Y will be changed as reduced number of study drugs.

Example: Expected number of study drugs during corresponding line is calculated as =(1 x 4) x (2016/1/29-2016/1/1+1) + (1 x 2) x (2016/2/10-2016/1/29+1) for following records are reported;

For patient of enzalutamide plus ADT arm

Date of visit 1st: 2016/1/1 and 4 study drugs per day are prescribed

Date of visit 2nd: 2016/1/29 and 2 study drugs per day are prescribed (e.g. due to AE)

Date of visit 3rd (un schedule): 2016/2/10 and 2 study drugs per day are prescribed (e.g. due to AE)

- Previous and concomitant medication

  Treatment duration of bicalutamide (Months)=(End date of bicalutamide - Start date of bicalutamide + 1) / 30.4375

- Previous and concomitant medication

  Previous medication is defined as medication with at least one dose taken 4 weeks (28 days) before Day 1 to the day before initial study treatment.

Concomitant medication for 1st line is defined as medication with at least one dose taken between the date of first dose (inclusive) of study drug for 1st line AAT and the date of first dose (not inclusive) of study drug for 2st line AAT.

Concomitant medication for 2nd line is defined as medication with at least one dose taken between the date of first dose (inclusive) of study drug for 2nd line AAT and the date of safety visit (inclusive).

Concomitant medication through study medication is defined as medication with at least one dose taken between of first dose (inclusive) and the date of safety visit (inclusive).

- PSA doubling time (PSADT)

  PSA doubling time is defined from two consecutive PSA measurements (PSA_1, PSA_2) and measured dates (Day_1, Day_2) before study Day 1 as follow;

$$\text{PSA doubling time (month)} = \frac{\ln(2) * 12}{365.25 * \frac{\ln(PSA_2) - \ln(PSA_1)}{Day_2 - Day_1}}$$

7 STATISTICAL METHODOLOGY

7.1 General Considerations

7.1.1 General Principles

For continuous variables, descriptive statistics will include the number of subjects (n), mean, standard deviation, median, minimum and maximum. When needed, the use of other percentiles (e.g.10%, 25%, 75% and 90%) will be mentioned in the relevant section. Frequencies and percentages will be displayed for categorical data. A missing category will be added to categorical variables where data is missing.

Summaries based on ITT (e.g. disposition, baseline and efficacy data) will be presented by planned treatment group, unless specifically stated otherwise. Safety analysis and other summaries based on SAF will be presented by actual treatment received.

All statistical comparisons will be made using two sided tests at the α=0.05 significance level unless specifically stated otherwise. All null hypotheses will be of no treatment difference, all alternative hypotheses will be two-sided, unless specifically stated otherwise.

All data processing, summarization, and analyses will be performed using SAS® Version 9.4 or higher on Windows. Specifications for table, figures, and data listing formats can be found in the TLF specifications for this study.

The Age (Derived) variable from EDC will be used for analysis. If date of birth is incomplete, the variable Age from EDC will be used for analysis.

Year will be calculated as (days/365.25) rounded up to 1 significant digit.

Month will be calculated as (days/30.4375) rounded up to 1 significant digit.

Time to event or duration of event endpoints will be based on the actual date rather than Visit Day.

For the definition of subgroups of interest please refer to section 7.8.

7.1.2 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 prior to the randomization date, study day is calculated as (assessment date – randomization date). There will be no study day 0.

- Study Day for each line

Study day for each line will be calculated by same definition of "Study Day" replacing the reference date with date of randomization for 1st line AAT, and replacing with date of enrollment for 2nd line AAT respectively.

i.e. when assessment date be included between date of informed consent and date of enrollment for 2nd line AAT - 1 day: (assessment date – randomization date + 1). when assessment date is same or after date of enrollment for 2nd line AAT: (assessment date – date of enrollment for 2nd line AAT + 1)

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

The date of the first dose of study drug is defined as the date the subject received the first dose of any study drug (enzalutamide or flutamide, whichever comes first). The date of the last dose of study drug is defined as the date the subject received the last dose of study drug (enzalutamide or flutamide).

- Date of First Dose in each line and Date of Last Dose of Study Drug in each line

The date of the first dose of study drug in each line (1st or 2nd line) is defined as the date the subject received the first dose of study drug in each line (enzalutamide or flutamide). The date of the last dose of study drug in each line is defined as the date the subject received the last dose of study drug in each line (enzalutamide or flutamide).

- 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 the subject received 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. Study day 1 is not planed date of first dose of study drug, administration will

be started study day 2.

- Treatment-Emergent Period

The treatment-emergent period will be defined as the period of time from the date of the first dose of study drug to 28 days after the date of the last dose of study drug, or 1 day prior to the date of initiation of cytotoxic chemotherapy for prostate cancer or an investigational agent treating prostate cancer, whichever occurs first. And 1st line treatment-emergent period is defined as the period of time from the date of the first dose of study drug in 1st line to 28 days after the date of the last dose of study drug in 1st line, or 1 day prior to the date of initiation of cytotoxic chemotherapy for prostate cancer or an investigational agent treating prostate cancer, whichever occurs first. Regarding 2nd line treatment-emergent period is applied same rule for 2nd line date.

- Baseline Value and Post-Baseline Value

Unless otherwise specified, the baseline value is defined as the last non-missing measurement prior to the randomization of study drug. When analysis be provided separately by treatment lines, baseline value for 2nd line AAT is defined as the last non-missing value measurement prior to the enrollment for 2nd line AAT of study drug.
Post-baseline value is defined as a measurement taken after the randomization of study drug. For Post-baseline value for 2nd line is defined as a measurement taken after enrollment during of study for 2nd line AAT.
Baseline value is defined as the last non-missing measurement prior to randomization (or register of 2nd line AAT) and post-baseline value is defined as a measurement taken after randomization for subjects who did not receive study drug. Change from baseline is defined as (post-baseline value - baseline value). Both date and time of drug administration and measurement should be considered when calculating baseline value. If time is not available, then date only should be used.

- Treatment Group

Treatment group is defined as planned randomized treatment sequence which divide to initial enzalutamide 160 mg/day group or the initial flutamide 125 mg × 3 times daily (375 mg/day) group.

- Treatment Arm

Treatment arm is defined as administration drug group regardless 1st line or 2nd line treatment. That is enzalutamide plus ADT arm include subject who receive enzalutamide plus ADT at 1st line and 2nd line. Flutamide plus ADT include who receive flutamide plus ADT at 1st line and 2nd line.

- Treatment Line

Treatment line is defined as 1st line AAT and 2nd line AAT. When analysis by treatment line is provided, divided to data of 1st line AAT and 2nd line AAT, and calculate descriptive summarization or frequency etc.

7.2 Study Population

7.2.1 Disposition of Subjects

- Number of subjects with informed consent (overall only)
- Number of subjects randomized with discontinued before randomization (overall only)
- Number of subjects randomized, by treatment group and overall
- Number and percentage of subjects in each analysis set, by treatment group and overall
- Number and percentage of subjects in with transition to 2nd line AAT, by treatment group and overall
- Number and percentage of subjects who had discontinued study drug during 1st line AAT to end of 2nd line AAT, by treatment group and overall
- Number and percentage of subjects who had discontinued study drug by primary reason for treatment discontinuation during 1st line AAT to end of 2nd line AAT, by treatment group and overall
- Number and percentage of subjects who had discontinued study drug during 1st line AAT and not transition to 2nd line AAT, by treatment group and overall
- Number and percentage of subjects who had discontinued study drug by primary reason for treatment discontinuation during 1st line AAT, by treatment group and overall
- Number and percentage of subjects who had discontinued study drug by primary reason for treatment discontinuation during 2nd line AAT, by treatment group and overall
- Number and percentage of screening failure subjects for overall
- Number and percentage of subjects who participated in each study visit, by treatment group and overall for the ITT population

A listing of all randomized subjects discontinued from study treatment, by center and by treatment group, the specific reason for discontinuation, and the last dose day and date will be presented.

7.2.2 Demographic and Other Baseline Characteristics

Demographic and other baseline characteristics will be summarized by treatment group as randomized for all subjects in the ITT population. Additionally same summaries will be provided for subject in the SAF population.

Descriptive summary statistics will be presented for continuous variables of:

- Age
- Height
- Weight
- Body Mass Index (BMI)
- Baseline serum PSA value
- Baseline lactate dehydrogenase (LDH) value
- Baseline hemoglobin value
- Baseline alkaline phosphatase (ALP) value

- Baseline serum albumin
- Baseline creatinine value

Frequency tabulations will be presented for categorical variables of:
- Age category (grouped as <65, 65 to 74, 75 to 84, ≥85 years)
- Baseline ECOG performance status (0, 1, 2, 3/4)
- Baseline brief pain inventory question # 3 (0-1, 2-3, >3)
- History of prior cardiovascular disease (Yes, No)
- Disease stages at randomization (M0/N0, M0/N1, M1)
- Regional Lymph Nodes at randomization (Nx, N0, N1, Unknown)
- Distant Metastasis at randomization (Mx, M0, M1, Unknown)

7.2.3 Medical History

Summary statistics will be presented for the ITT population and SAF population for numeric variables of:
- Time (months) from initial diagnosis.
- Total Gleason score at initial diagnosis

Frequency tabulations will be presented for categorical variables of:
- Gleason score: primary score + secondary score (< 3 + 3, 3 + 3, 3 + 4, 4 + 3, 4 + 4, 4 + 5, 5 + 4, 5 + 5, other, and missing)
- Total Gleason score category at initial diagnosis (low [2-4], medium [5-7], high [8-10], Unknown or missing)
- Clinical Tumor Stage at initial diagnosis (Tx, T0, T1, T2, T3, T4, Unknown)
- Regional Lymph Nodes at initial diagnosis (Nx, N0, N1, Unknown)
- Distant Metastasis at initial diagnosis (Mx, M0, M1, Unknown)
- PSA progression at study entry (Yes, No)
- PSA doubling time (PSADT)
- Type of disease progression (PSA progression only, radiographic progression with/without PSA progression) at study entry
- Disease localization (bone only, soft tissue only, both bone and soft tissue, and none) at screening
- Target or nontarget soft tissue disease (i.e., at least one target or nontarget lesion per RECIST 1.1) at screening (Yes, No)
- Extent of disease at (bone, lymph node, visceral lung, visceral liver, visceral lung and/or liver, and other soft tissue) screening

In addition, ongoing medical history at randomization which not related to prostate cancer will be summarized by System Organ Class (SOC) and Preferred Term (PT). Medical history ongoing at randomization will be summarized.

7.2.4 Previous and Concomitant Medications

Medications will be coded with World Health Organization Drug Dictionary (WHO-DD)

(September 2016, HD Format B2).
Previous medication is defined as medication with at least one dose taken before initial study treatment. Number and percentage of subjects taking the previous medication by treatment arm for the ITT/SAF population.
Concomitant medication for 1st line is defined as medication with at least one dose taken between the date of first dose (inclusive) and the date of last dose of study drug for 1st line AAT and the date of first dose (not inclusive) of study drug for 2nd line AAT. Concomitant medication for 2nd line is defined as medication with at least one dose taken between the date of first dose (inclusive) and the date of last dose of study drug for 2nd line AAT and the date of safety visit (inclusive). As with previous medications, number and percentage of subjects taking the concomitant medication by treatment arm for the ITT/SAF population.

Number of subjects with at least one previous and concomitant medications will be summarized.

Data listings will be presented for previous and concomitant medications and therapies.

7.2.5 Previous Therapies for Prostate Cancer

Previous therapies for prostate cancer will be summarized by treatment group as randomized for the ITT population and SAF population as follows:

- Summary statistics for treatment duration of bicalutamide
- Number of subjects with gonadotropin releasing hormone (GnRH) agonist / antagonist and each drug
- Number of subjects with other drug therapies
- Number of subjects with radiation therapy

7.2.6 Prior Surgical Procedures for Prostate Cancer

Number and percentage of subjects with any prior surgical procedures for prostate cancer (prostatectomy, orchiectomy, transurethral resection of the prostate (TURP), cryoablation, pelvic lymph node dissection, high intensity focused ultrasound (HIFU), and other) will be tabulated by treatment group as randomized for the ITT population and SAF population.

7.3 Study Drugs

7.3.1 Exposure

Duration of exposure will be summarized in two ways for the ITT population and SAF.

- Descriptive statistics will be presented by treatment group and treatment line.
- Exposure time will be categorized according to the following categories by treatment group and treatment line:
    - less than 3 months
    - at least 3 months, less than 6 months
    - at least 6 months, less than 12 months
    - at least 12 months, less than 24 months
    - 24 months or more

Counts and percentages of subjects in each of these categories will be summarized for each treatment group.

Total treatment duration will be summarized by treatment group applying same rule as above.

7.3.2 Dose Modification

Subjects with at least one dose modification, including dose reduction and interruption, and the reason for the dose modification (adverse event or other) will be summarized by treatment group and treatment line for the ITT population and SAF. Also, subjects with at least one dose reduction and those reason will be summarized by treatment group and treatment line for the ITT population and SAF.

7.3.3 Treatment Compliance

Treatment compliance will be examined for subjects in the ITT population and SAF. For the details of treatment compliance calculation please refer to section 6.5.
Descriptive statistics will be presented for total number of study drugs taken per subject by treatment group and treatment line.
Percent compliance will be summarized in two ways:

- Descriptive statistics will be presented by treatment group and treatment line.
- Percent compliance will be categorized according to the following categories by treatment group and treatment line:
    - less than 20%
    - at least 20%, less than 40%
    - at least 40%, less than 60%
    - at least 60%, less than 80%
    - at least 80%, less or equal to 100%
    - greater than 100%, less than 120%
    - greater or equal to 120%

7.4 Analysis of Efficacy

All analysis of efficacy will be presented by planned treatment group for ITT population, unless specified otherwise. The data obtained at Day 1 visit will be used as the baseline for efficacy assessment except for rPFS for which the data at the Screening visit will be used as the baseline.

7.4.1 Analysis of Primary Endpoint

The primary efficacy analysis will be performed on the TTPP1 for the subjects in the ITT population. TTPP1 is the period from randomization to the first day when PSA progression is objectively identified. Convention of censoring is specified in table 7-1.

If a subject meets the criteria for more than one censoring rule, they will be censored with the earliest censoring date. The PSA progression date is the first date where progression definition is met, not confirmed.

Methodology

The benefit of enzalutamide as compared to flutamide will be assessed with the stratified log-rank test as the primary efficacy analysis. The stratification factor will be the disease stages (M0/N0, M0/N1, or M1) used as stratification factor for allocation.

The hypothesis for comparisons is as follows:

H0: The survival time distribution of enzalutamide group is the same as that with flutamide group.

H1: The survival time distribution of enzalutamide group is not the same as that with flutamide group.

The significance level is 0.05 (two-sided). ITT is used for the primary analysis.

Unstratified Cox's Proportional Hazards Model with treatment and disease stage as the covariate will be used to support the log rank test.

The benefit of enzalutamide compared to flutamide will be evaluated by hazard ratio (enzalutamide/flutamide) with its 95% confidence interval (CI) based on Cox Proportional Hazard Model.

Kaplan-Meier curves will be used to estimate the distribution of duration of TTPP1 and Stratified curve will be estimated by disease stage. The 50th percentile of Kaplan-Meier estimates will be used to estimate the median duration of TTPP1. A two-sided 95% CI will be provided for this estimate. The 25th and 75th percentiles, and the range (minimum, maximum) will be presented as well. The range will be determined including censored observations. Estimated Kaplan-Meier curve will be plotted.

7.4.2 Analysis of Secondary Endpoints

7.4.2.1 Time to PSA Progression with 1st Line AAT + 2nd Line AAT (TTPP2)

TTPP2 is defined as the period from randomization to the date of PSA progression with 2nd line AAT. The total of time to PSA progression with 1st line AAT and time to PSA progression with 2nd line AAT will be evaluated. Convention of censoring is specified in Table 7-1.

The distribution of the TTPP2 will be estimated using the Kaplan–Meier method. The median TTPP2 will be estimated using the 50 percentile value of the Kaplan–Meier curve. The two-sided 95% CI will be calculated based on the estimation.

7.4.2.2 PSA Response Rate to 1st Line AAT (Decrease by at Least 50% or 90% from Baseline)

The PSA response rate decreasing by at least 50% will be calculated by treatment groups of subjects with baseline PSA value and PSA measured at least once at least 3 weeks passed after the lowest PSA. For the comparison of the response rate between the enzalutamide and flutamide groups, the stratified Cochran–Mantel–Haenszel test will be used. The stratification factor will be the disease stages (M0/N0, M0/N1, or M1) used as stratification factor for allocation.

PSA decrease by at least 90% will also be compared at the 0.05 significance level (two-sided) between the enzalutamide and flutamide groups using the stratified Cochran–Mantel–Haenszel test.

Response rate, its Clopper-Pearson exact confidence interval and its count within stratification will be summarized for each treatment. Mantel-Haenszel common risk difference for response (enzalutamide - flutamide) and its 95% confidence interval based on the variance claimed by Sato (1989) will be estimated.

Water fall plot will be made to present the percent change of PSA from baseline at the lowest PSA value (except within 3 weeks from baseline) for patient who have at least one time point PSA reduction in the 1st line AAT period, and the highest PSA value (except within 3 weeks from baseline) for patient who have never had PSA reduction, respectively.

7.4.2.3 PSA Response Rate to 1st Line AAT at Week 13 (Decrease by at Least 50% or 90% from Baseline)

The PSA response rate decreasing by at least 50% (or by at least 90%) will be calculated by treatment groups of subjects with baseline PSA value with 1st line AAT and PSA measured at week 13. The PSA response rate at week 13 will be assessed by the same method as that with the PSA response rate to 1st line AAT. Water fall plot will be made to present the percent change of PSA from baseline at Week 13 PSA value for each patient.

7.4.2.4 Time to PSA Decrease by 50% from Baseline with 1st Line AAT

The time to PSA decrease by 50% with 1st line AAT is defined as the period from randomization to the day when the decrease of PSA from baseline by 50% is first identified. Convention of censoring is specified in Table 7-1.

The benefit of enzalutamide as compared to flutamide will be assessed with the stratified log-rank test as the primary efficacy analysis. The stratification factor will be the disease stages (M0/N0, M0/N1, or M1) used as stratification factor for allocation.

The distribution of the time to PSA decrease by 50% with 1st line AAT will be estimated using the Kaplan–Meier method. The median time to PSA decrease by 50% will be estimated using the 50 percentile value of the Kaplan–Meier curve. The two-sided 95% CI will be calculated based on the estimation.

7.4.2.5 Time to Treatment Failure of 1st Line AAT (TTF1)

TTF1 is defined as the period from randomization to the date of discontinuation of 1st line AAT. Convention of censoring is specified in Table7-1.

TTF1 will be assessed using the same method as that with time to PSA decrease by 50% from baseline with 1st Line AAT.

7.4.2.6 Time to Treatment Failure of 2nd Line AAT (TTF2)

TTF2 is defined as the period from randomization to the date of discontinuation of 2nd line AAT. Convention of censoring is specified in Table 7-1.

The distribution of the time to treatment failure of 2nd line AAT will be estimated using the Kaplan–Meier method. The median time to treatment failure of 2nd line AAT will be estimated using the 50 percentile value of the Kaplan–Meier curve. The two-sided 95% CI will be calculated based on the estimation.

7.4.2.7 Radiographic Progression-Free Survival (rPFS)

The rPFS is defined as the time from randomization to the first objective evidence of radiographic disease progression assessed by the blinded independent central review facility or on-study death, defined as death within 28 days after treatment discontinuation, whichever occurs first. rPFS will be assessed in the ITT population. Radiographic disease progression will include soft tissue disease progression and confirmed bone disease progression. The censoring rules are shown in Table 7-1. The analysis of rPFS will be presented by planned treatment group for subjects that having least one distant metastasis at baseline in ITT population.

rPFS will be assessed using the same method as that with TTPP1. i.e. Stratified log-rank test and unstratified cox regression will be applied to rPFS and estimate the distribution using Kaplan-Meier method.

7.4.3 Analysis of Exploratory Endpoints

7.4.3.1 Time to PSA Progression with 2nd Line AAT (2nd TTPP)

2nd TTPP is defined as the period from the start of 2nd line AAT to the first day when PSA progression is objectively identified. Convention of censoring is specified in Table 7-1.

The 2nd TTPP will be assessed by the same method as that with the Time to Treatment Failure of 2nd Line AAT (TTF2).

7.4.3.2 PSA Response Rate to 2nd line AAT (Decrease by at Least 50% or 90% from Baseline)

The PSA response rate to 2nd line AAT decreasing by at least 50% (or by at least 90%) will be calculated by treatment groups of subjects with baseline PSA value and PSA measured at least once at least 3 weeks passed after the lowest PSA. The PSA response rate to 2nd line AAT will be assessed by the same method as that with the PSA response rate to 1st line AAT.

Water fall plot will be made to present the percent change of PSA from baseline at the lowest PSA value (except within 3 weeks from baseline) for patient who have at least one time point PSA reduction in the 2nd line AAT period, and the highest PSA value (except within 3 weeks from baseline) for patient who have never had PSA reduction, respectively.

7.4.3.3 Metastasis-Free Survival (MFS)

MFS is defined as the period from randomization to new radiographic metastatic lesion or death of any cause during the study period, whichever occurs earlier. Evaluation will be made in subjects in whom distant metastasis is not confirmed at baseline. MFS will be

assessed by the same method as that with the Time to Treatment Failure of 2nd Line AAT (TTF2).

7.4.3.4 Functional Assessment of Cancer Therapy–Prostate (FACT-P) with 1st Line AAT and 2nd Line AAT

The FACT-P domain scores and total score and their change from baseline will be summarized by treatment group as randomized for the ITT population. These measurement values and changes from baseline will be plotted as mean +/- SD plot. The FACT-P domain score and total score will be calculated using the Manual of Functional Assessment of Chronic Illness Therapy (FACIT) Measurement System.

7.4.3.5 European Quality of Life-5 Dimension-5 Level Instruments (EQ-5D-5L) with 1st Line AAT and 2nd Line AAT

Following analysis will be applied for the ITT population. The subjects' current health state is assessed using a visual analogue scale form 0 (worst imaginable health state) to 100 (best imaginable health state). The current health state score will be summarized by visit and treatment group. QOL utility index will be derived according to section 6.1.4 will also be summarized by visit and treatment group. These measurement values will be plotted as mean +/- SD plot. Additionally, a within-subject change will be calculated per visit as the post-baseline measurement minus the baseline measurement and summarized by visit and treatment group in the same way.

7.4.3.6 Brief Pain Inventory-Short Form (BPI-SF) with 1st Line AAT and 2nd Line AAT

Descriptive statistics for Pain severity and Pain Interference and their change from baseline will be summarize by visit and treatment group for the ITT population. These measurement values and changes from baseline will be plotted as mean +/- SD plot.

7.4.3.7 Brief Fatigue Inventory (BFI) with 1st Line AAT and 2nd Line AAT

Descriptive statistics for global fatigue score and its change from baseline will be summarize by visit and treatment group for the ITT population. These measurement values and changes from baseline will be plotted as mean +/- SD plot.

Global fatigue score will be categorized according to the following categories by treatment group:

- Mild : global fatigue score be in range from 1 to 3.
- Moderate : global fatigue score be in range from 4 to 6.
- Severe : global fatigue score be in range from 7 to 10.

Counts and percentages of subjects in each of these categories will be summarized for each treatment group.

7.4.3.8 Objective Response Rate (ORR) in best Overall Soft Tissue Response of RECIST Guidelines with 1st Line AAT

For the comparison of ORR between the enzalutamide and flutamide groups, the stratified Cochran–Mantel–Haenszel test will be used. The stratification factor will be the disease stages (M0/N0, M0/N1, or M1) used as stratification factor for allocation.

ORR, its Clopper-Pearson exact confidence interval and its count within stratification will be summarized for each treatment. Mantel-Haenszel common risk difference for response (enzalutamide - /flutamide) and its 95% confidence interval based on the variance claimed by Sato (1989) will be estimated. And count and percentage of overall soft tissue response which is assessment using RECIST 1.1 will be tabulated. The summarized category is shown as follows:

- Complete response (CR)
- Partial response (PR)
- Progressive disease (PD)
- Stable disease (SD)
- Non-evaluable (NE)

Table 7-1 : Definition of time to event and censor

	TTPP1	**TTPP2**	**2nd TTPP**	**Time to PSA 50%/90% decline**	**TTF1**	**TTF2**	**rPFS**	**MFS**
Subjects without PSA value at baseline or in observation period	Considered as censoring at randomization	Considered as censoring at randomization	Excluded from analysis	Considered as censoring at randomization	Not considered as censoring *To be judged based on continuation of study treatment	Not considered as censoring *To be judged based on continuation of study treatment	Not considered as censoring *To be judged based on result of radiographic assessment	Not considered as censoring *To be judged based on result of radiographic assessment
Subjects without the result of radiographic assessment at baseline or in observation period	Not considered as censoring *To be judged based on result of PSA measurement	Not considered as censoring *To be judged based on result of PSA measurement	Not considered as censoring *To be judged based on result of PSA measurement	Not considered as censoring *To be judged based on result of PSA measurement	Not considered as censoring *To be judged based on continuation of study treatment	Not considered as censoring *To be judged based on continuation of study treatment	Considered as censoring at randomization	Considered as censoring at randomization
Subjects without PSA progression at final data cut-off (Including cases inconsistent with PSA progression defined by the protocol)	Subjects who continue the first study drug are considered as censoring at final PSA measurement	Considered as censoring at final PSA measurement	Subjects who continue the first study drug are excluded from analysis Subjects who have shifted to the second study drug are considered as censoring at final PSA measurement	Considered as censoring at final PSA measurement	Not considered as censoring *To be judged based on continuation of study treatment	Not considered as censoring *To be judged based on continuation of study treatment	Not considered as censoring *To be judged based on result of radiographic assessment	Not considered as censoring *To be judged based on result of radiographic assessment

Subjects without radiographic progression at final data cut-off (Including cases inconsistent with progression defined by the protocol)	Not considered as censoring *To be judged based on result of PSA measurement	Not considered as censoring *To be judged based on result of PSA measurement	Not considered as censoring *To be judged based on result of PSA measurement	Not considered as censoring *To be judged based on result of PSA measurement	Not considered as censoring *To be judged based on continuation of study treatment	Not considered as censoring *To be judged based on continuation of study treatment	Considered as censoring at final radiographic assessment	Considered as censoring at final radiographic assessment
Death during the study	Subjects who died during treatment with the first study drug are considered as censoring at final PSA measurement before death	Considered as censoring at final PSA measurement before death	Subjects who died during treatment with the first study drug are excluded from analysis Subjects who died during treatment with the second study drug are considered as censoring at final PSA measurement before death	Considered as censoring at final PSA measurement before death	The event is considered to occur at the time of death	The event is considered to occur at the time of death	The event is considered to occur at the time of death	The event is considered to occur at the time of death
Subjects who discontinued treatment during the study	Subjects who discontinue the first study drug are considered as censoring at final PSA measurement before discontinuation	Considered as censoring at final PSA measurement before discontinuation	Subjects who discontinue the first study drug are excluded from analysis Subjects who discontinue the second study drug are considered as censoring at final PSA measurement before discontinuation	Considered as censoring at final PSA measurement before discontinuation	The event is considered to occur at the time of discontinuation	The event is considered to occur at the time of discontinuation	Considered as censoring at final radiographic assessment before discontinuation	Considered as censoring at final radiographic assessment before discontinuation

Subjects who used prohibited concomitant medication or therapy during the study	Subjects who use one with the first study drug are considered as censoring at final PSA measurement before use	Considered as censoring at final PSA measurement before use	Subjects who use one with the first study drug are excluded from analysis Subjects who use one with the second study drug are considered as censoring at final PSA measurement before use	Considered as censoring at final PSA measurement before use	Not considered as censoring *To be judged based on continuation of study treatment	Not considered as censoring *To be judged based on continuation of study treatment	Considered as censoring at final radiographic assessment before use	Considered as censoring at final radiographic assessment before use
Subjects who miss PSA data at least twice consecutively at scheduled PSA measurement points	Considered as censoring at final PSA measurement before missing data occur	Considered as censoring at final PSA measurement before missing data occur	Subjects who continue the first study drug are excluded from analysis Subjects who have shifted to the second study drug are considered as censoring at final PSA measurement before missing data occur	Considered as censoring at final PSA measurement before missing data occur	Not considered as censoring *To be judged based on continuation of study treatment	Not considered as censoring *To be judged based on continuation of study treatment	Not considered as censoring *To be judged based on result of radiographic assessment	Not considered as censoring *To be judged based on result of radiographic assessment
Subjects who miss radiographic assessment at least twice consecutively at scheduled radiographic assessment points	Not considered as censoring *To be judged based on result of PSA measurement	Not considered as censoring *To be judged based on result of PSA measurement	Not considered as censoring *To be judged based on result of PSA measurement	Not considered as censoring *To be judged based on result of PSA measurement	Not considered as censoring *To be judged based on continuation of study treatment	Not considered as censoring *To be judged based on continuation of study treatment	Considered as censoring at final radiographic assessment before missing data occur	Considered as censoring at final radiographic assessment before missing data occur

7.4.4 Analysis of Other Variables

PSA measurement

Descriptive statistics for PSA measurement and its change from baseline will be summarized at by visit and treatment group. These measurement value and change from baseline will be plotted as mean +/- SD plot.

Swimmer's plot

Swimmer's plot will be made to present treatment flow and event occurrence for each patient. This plot will include time on and off of the 1st/2nd line AAT continuations as a bar chart, and the points at which drug related AEs (by CTCAE grades), PSA progression, radiographic progression, PSA response, and death.

7.5 Analysis of Safety

All analysis of safety will be presented by actual treatment group for SAF, unless specified otherwise. The last non-missing measurement prior to the first dose of study drug will be used as the baseline.

7.5.1 Adverse Events

The severity of all AEs is to be evaluated by the investigator based on the National Cancer Institute's CTCAE version 4.0. The coding dictionary for this study will be Medical Dictionary for Regulatory Activities (MedDRA) version 19.1. It will be used to summarize AEs by SOC and PT.

The safety analysis will be performed in the safety analysis set, and summarized by the time of assessment and by the treatment group that actually received the study drugs. The period of onset in the treatment period with the study drug is defined up to 28 days after the final treatment with the study drug, start date of cytocidal chemotherapy, or start date of new treatment for prostate cancer, whichever occurs first. Safety will be assessed by AEs, frequency of treatment discontinuation associated with AEs, and laboratory test value summary statistics. Descriptive statistics, and not inferential statistics, will be used.

The total number of AEs reported will be presented by treatment arm (i.e. enzalutamide v.s. flutamide) and overall total for the SAF. The same analysis will be provided dividing treatment lines. AEs which have occurred during 1st line AAT and 2nd line AAT will be included to analysis. In this 'event level' analyses, all AEs will be counted separately. The following categories will be presented:

- Number of AEs,
- Number of AEs leading to study drug discontinuation,
- Number of drug-related AEs,
- Number of drug-related AEs leading to study drug discontinuation,
- Number of grade 3 or higher AEs,
- Number of grade 3 or higher drug-related AEs,
- Number of serious adverse events (SAEs),

- Number of drug-related SAEs,
- Number of SAEs leading to study drug discontinuation,
- Number of drug-related SAEs leading to study drug discontinuation,
- Number of SAEs leading to death,
- Number of drug-related SAEs leading to death.

The number and percentage of subjects experiencing AEs will be presented by treatment arm and overall total for the SAF. The same analysis will be provided dividing treatment lines. AEs which have occurred during 1st line AAT and 2nd line AAT will be included to analysis. In this 'subject level' analyses, multiple AEs within the same category for a particular subject will only be counted once and the denominator for the percentage calculation will be the number of subjects in the SAF. The following categories will be presented:

- Number and percentage of subjects with AEs,
- Number and percentage of subjects with AEs leading to study drug discontinuation,
- Number and percentage of subjects with drug-related AEs,
- Number and percentage of subjects with drug-related AEs leading to study drug discontinuation,
- Number and percentage of subjects with grade 3 or higher AEs,
- Number of grade 3 or higher drug-related AEs,
- Number and percentage of subjects with SAEs,
- Number and percentage of subjects with drug-related SAEs,
- Number and percentage of subjects with SAEs leading to study drug discontinuation,
- Number and percentage of subjects with drug-related SAEs leading to study drug discontinuation,
- Number and percentage of subjects with SAEs leading to death,
- Number and percentage of subjects with drug-related SAEs leading to death.

The number and percentage of subjects experiencing AEs will be tabulated by SOC, PT, treatment arm and overall total for the SAF. The same analysis will be provided dividing treatment lines. AEs which have occurred during 1st line AAT and 2nd line AAT will be included to analysis. For this and similar analyses, an AE will only be counted once for each subject if the same AE is experienced on more than one occasion. Tables will be sorted by descending overall total of SOC and PT frequency. The following summaries will be produced:

- AEs,
- AEs leading to study drug discontinuation,
- drug-related AEs,
- drug-related AEs leading to study drug discontinuation,
- grade 3 or higher AEs,
- SAEs,

- drug-related SAEs,
- SAEs leading to study drug discontinuation,
- drug-related SAEs leading to study drug discontinuation,
- SAEs leading to death,
- drug-related SAEs leading to death,
- TEAEs excluding serious adverse events that equal to or exceed a threshold of 2.0% in any treatment group, and
- Common TEAEs that equal to or exceed a threshold of 2.0% in any treatment group.

The number and percentage of subjects experiencing AEs will be tabulated by SOC, PT, maximum severity, treatment arm and overall total for the SAF. The same analysis will be provided dividing treatment lines. AEs which have occurred during 1st line AAT and 2nd line AAT will be included to analysis. For this and similar analyses, a subject reporting the same adverse event more than once is counted once at the maximum severity when calculating incidence. Tables will be sorted by descending overall total of SOC and PT frequency. The following summaries will be produced:

- AEs,
- drug-related AEs,
- SAEs,

The listing for subjects those who meet the special situations such as overdose, drug abuse, inadvertent or accidental exposure, or medication error will be prepared.

7.5.2 Clinical Laboratory Evaluation

Laboratory data consist of hematology, chemistry, and urinalysis laboratory tests. The National Cancer Institute's CTCAE version 4.0, will be used to categorize toxicity grade for the laboratory parameters. Normal ranges will be implemented to identify values that are outside the normal range.

Quantitative clinical laboratory variables from the hematology, chemistry, and urinalysis panels will be summarized using mean, standard deviation, minimum, maximum and median for each treatment group at each visit. These measurement values will be plotted as mean +/- SD plot. Additionally, a within-subject change will be calculated per visit as the post-baseline measurement minus the baseline measurement and summarized in the same way.

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 by treatment line, separately. The number and percentage of subjects with at least one occurrence of grade 3 or grade 4 laboratory values in the treatment-emergent period will be summarized for each parameter and treatment group. The number and percentage of subjects with toxicity grade increase of two or more above baseline will also be summarized for each parameter and treatment group.

Each laboratory result will be classified as low (L), normal (N), or high (H) at each visit according to the laboratory supplied reference ranges. A high level overview shift table will be presented per parameter for each treatment group. Subjects who are normal at baseline and become low/high/low or high at any post-baseline time point will be summarized in the tables.

7.5.2.1 Liver function tests

The following potentially clinically significant criteria for liver tests – defined as ALP, Alanine Transaminase (ALT), total bilirubin, Aspartate Transaminase (AST) and their combination are defined. The subject's highest value during the investigational period will be used.

Parameter	Criteria
ALT	> 3xULN
	> 5xULN
AST	> 3xULN
	> 5xULN
Total Bilirubin	> 2xULN
ALP	> 1.5xULN
ALT and/or AST AND Total Bilirubin(*)	(ALT and/or AST > 3xULN) and total bilirubin > 2xULN

(*) Combination of values measured within same sample

The number and percentage of subjects with potentially clinically significant values in liver enzyme and total bilirubin tests during the investigational period will be presented by treatment arm. The same analysis will be provided dividing treatment lines.

7.5.3 Vital Signs

Vital signs (systolic blood pressure, diastolic blood pressure, and pulse rate) will be summarized using mean, standard deviation, minimum, maximum and median by treatment group and visit. These measurement values will be plotted as mean +/- SD plot. Additionally, a within-subject change will be calculated per visit as the post-baseline measurement minus the baseline measurement and summarized by treatment group and visit in the same way.

7.5.4 ECOG performance status

The number and percentage ECOG performance status will be presented by treatment arm. The same analysis will be provided dividing treatment lines.

7.5.5 Electrocardiograms (ECGs)

Not applicable.

7.5.6 Pregnancies

Not applicable.

7.5.7 Other Safety-Related Observations

Not applicable.

.

7.6 Analysis of PK

Not applicable.

7.7 Analysis of PD

Not applicable.

7.8 Subgroups of Interest

Primary efficacy endpoint, TTPP2, PSA Response Rate to 1st Line AAT, PSA Response Rate to 1st Line AAT at Week 13 and rPFS, analyses will be repeated by the subgroups defined on the basis of the categorized variables listed below:

- Baseline ECOG performance status (0 vs ≥1)
- Age category (<75 and ≥75 years), (at or below median vs above median)
- Total Gleason score (≤7 and ≥8) at diagnosis
- Type of progression (PSA progression only vs radiographic progression with or without PSA progression) at study entry
- Baseline PSA value (at or below median vs above median)
- PSA value at initial diagnosis (at or below median vs above median)
- Baseline LDH value (at or below median vs above median)
- Baseline hemoglobin value (at or below median vs above median)
- Baseline ALP value (at or below median vs above median)
- Treatment duration of bicalutamide (at or below 12 months vs above 12 months), (at or below median vs above median)
- PSA nadir value on CAB therapy (≤4ng/mL vs >4ng/mL), (≤0.2ng/mL vs >0.2ng/mL)
- PSA response rate to 1st line AAT at Week 13 (below 50% vs at or above 50% reduction from baseline), (below 90% vs at or above 90% reduction from baseline) *except for the analysis of PSA Response Rate to 1st Line AAT at Week 13.
- Baseline NLR value (at or below median vs above median)
- Baseline testosterone value (at or below median vs above median)
- Treatment duration of first line hormonal therapy (at or below 12 months vs above 12 months), (at or below median vs above median)
- Disease stages (M0/N0 vs M0/N1 vs M1)
- PSA doubling time (at or below median vs above median)

Also, selected safety variables (treatment emergent adverse events) will be summarized by the treatment group for the subgroups defined on the basis of the categorized variables listed below:

- Baseline ECOG performance status (0 vs ≥1)
- Age category (<75 and ≥75 years), (at or below median vs above median)

7.9 Other Analyses

A list of Flutamide-related and non-serious adverse event will be derived.

7.10 Interim Analysis (and Early Discontinuation of the Clinical Study)

Not applicable.

7.11 Handling of Missing Data, Outliers, Visit Windows, and Other Information

Refer to the data specification document in which more details are provided.

7.11.1 Missing Data

As a general principle, no imputation of missing data will be done. Exceptions are the start and stop dates of concomitant medication and radiotherapy for prostate cancer. The imputed dates will be used to determining whether a medication or radiotherapy is prior to/on or after the first dose of study drug. The following rules will be applied to impute partial dates for medications and radiotherapy for prostate cancer:
If start date of a medication/radiotherapy is partially missing, impute as follows:

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

If end date of a medication/radiotherapy 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 and/or end date of a medication/radiotherapy is completely missing, do not impute.

Listings will present the actual partial dates; imputed dates will not be shown.

7.11.2 Outliers

All values will be included in the analyses.

7.11.3 Visit Windows

[Acceptable Time Ranges of Efficacy Tests]
(1) PSA

1st line AAT

Study Visit	Reference Date	Acceptable Range
Screening visit	−28 days to −7 days	28 days to 7 days before randomization
Week 1 (Day 1)	1	-
Week 13 (Day 85)	85	Reference date ± 7 days
Week 17 (Day 113)	113	Reference date ± 7 days
Every subsequent 4 weeks	Every subsequent 28 days	Reference date ± 7 days

2nd line AAT

Study Visit	Reference Date	Acceptable Range
Week 1 (Day 1)	1	-
Week 13 (Day 85)	85	Reference date ± 7 days
Week 17 (Day 113)	113	Reference date ± 7 days
Every subsequent 4 weeks	Every subsequent 28 days	Reference date ± 7 days
At completion or discontinuation	Day of completion or discontinuation visit	Reference date + 7 days
Safety follow-up	Final study treatment + 28 days	Reference date ± 7 days

(2) Abdominopelvic CT/MRI, Bone scintigraphy, Chest CT

Study Visit	Reference Date	Acceptable Range
Screening visit[*1]	−28 days to −7 days	28 days to 7 days before randomization
Week 13 (Day 85)	85	Reference date ± 14 days
Every subsequent 12 weeks[*2]	Every subsequent 84 days	Reference date ± 14 days
At completion or discontinuation[*3]	Day of completion or discontinuation visit	Reference date + 14 days

*1: Images taken within 42 days before randomization as part of daily consultation before informed consent may be used as data in the screening period.

*2: Perform every 12 weeks, irrespective of switching medications.

*3: Not required if it occurs within 28 days from the latest measurement.

(3) FACT-P[*1], EQ-5D-5L[*1], BPI-SF, BFI

1st line AAT

Study Visit	Reference Date	Acceptable Range
Screening visit	−28 days to −7 days	28 days to 7 days before randomization
Week 1 (Day 1)	1	-
Week 13 (Day 85)	85	78 days to 92 days
Every subsequent 12 weeks	Every subsequent 84 days	Reference date ± 7 days

2nd line AAT

Study Visit	Reference Date	Acceptable Range
Week 1 (Day 1)	1	-
Week 13 (Day 85)	85	78 days to 92 days
Every subsequent 12 weeks	Every subsequent 84 days	Reference date ± 7 days
At completion or discontinuation*2	Day of completion or discontinuation visit	Reference date + 7 days

*1: Do not perform on screening visit.

*2: Not required if it occurs within 28 days from the latest measurement.

[Acceptable Time Ranges of Safety Tests]

Laboratory test (biochemistry, hematology, and urinalysis), Vital signs (blood pressure and pulse rate), ECOG Performance Status

1st line AAT

Study Visit	Reference Date	Acceptable Range
Screening visit	−28 days to −7 days	28 days to 7 days before randomization
Week 1 (Day 1)	1	-
Week 5 (Day 29)	29	Reference date ± 7 days
Week 9 (Day 57)	57	Reference date ± 7 days
Week 13 (Day 85)	85	Reference date ± 7 days
Every subsequent 4 weeks	Every subsequent 28 days	Reference date ± 7 days

2nd line AAT

Study Visit	Reference Date	Acceptable Range
Week 1 (Day 1)	1	-
Week 5 (Day 29)	29	Reference date ± 7 days
Week 9 (Day 57)	57	Reference date ± 7 days
Week 13 (Day 85)	85	Reference date ± 7 days
Every subsequent 4 weeks	Every subsequent 28 days	Reference date ± 7 days
At completion or discontinuation	Day of completion or discontinuation visit	Reference date + 7 days
Safety follow-up	Final study treatment + 28 days	Reference date ± 7 days

The value which assessment day is the closest to the defined target day within these windows is used. If two values are equally close, the earlier is used in the analysis. Note that a value which is assessed prior to Day 1 of the 1st line AAT or 2nd line AAT will be available as baseline assessment once it is investigated and approved by the review board. Also, a value which is assessed as Day 1 of the 2nd line AAT will be utilized as end/discontinuation of 1st line AAT when a subject start 2nd line AAT without confirmation of PSA progression or because of other than PSA progression.

8 DOCUMENT REVISION HISTORY

Version	**Date**	**Changes**	**Comment/rationale for change**
1.00	8-July-2020	NA	Document finalized

9 REFERENCES

- Ikeda S, Shiroiwa T, Igarashi A, Noto S, Fukuda T, Saito S, Shimozuma K. Developing a Japanese version of the EQ-5D-5L value set. J. Natl. Inst. Public Health, 64(1):47-55, 2015.

- Sato, T. On the Variance Estimator of the Mantel-Haenszel Risk Difference. Biometrics, 45, 1323–4, letter to the editor, 1989.

10 APPENDICES

10.1 Appendix 1: Key Contributors and Approvers

List of Key Contributors and Approvers

Key Contributors

The following contributed to or reviewed this Statistical Analysis Plan as relevant to their indicated discipline or role.

Primary author (s)

[redacted]

Contributors and Reviewers

[redacted]

Author and Approver Signatories

Prepared by:	[redacted]	Date:	[redacted]
Approved by:		Date:	
		Date:	
Reviewed by:		Date:	