Official Protocol Title:	Phase II Trial of Pembrolizumab (MK-3475) in Subjects with Metastatic Castration-Resistant Prostate Cancer (mCRPC) (KEYNOTE-199)
NCT number:	NCT02787005
Document Date:	25-SEP-2018

Supplemental Statistical Analysis Plan (sSAP)

TABLE OF CONTENTS

Table of Contents 1
Listing of Tables 3
1 INTRODUCTION 4
2 SUMMARY OF CHANGES 4
3 ANALYTICAL AND METHODOLOGICAL DETAILS 4
3.1 Statistical Analysis Plan Summary 4
3.2 Responsibility for Analyses/In-House Blinding 5
3.3 Hypotheses/Estimation 6
3.4 Analysis Endpoints 6
3.4.1 Efficacy/Immunogenicity/Pharmacokinetics Endpoints 6
3.4.2 Safety Endpoints 8
3.5 Analysis Populations 8
3.5.1 Efficacy Analysis Populations 8
3.5.2 Safety Analysis Populations 8
3.5.3 Biomarker Discovery Population 9
3.6 Statistical Methods 9
3.6.1 Statistical Methods for Efficacy Analyses 9
3.6.2 Statistical Methods for Safety Analyses 12
3.6.3 Statistical Methods for Biomarker 12
3.6.4 Summaries of Baseline Characteristics, Demographics, and Other Analyses 13
3.7 Interim Analyses 13
3.8 Multiplicity 13
3.9 Sample Size and Power Calculations 13
3.10 Subgroup Analyses and Effect of Baseline Factors 14
3.11 Compliance (Medication Adherence) 15
3.12 Extent of Exposure 15
3.13 Statistical considerations for Patient-Reported Outcomes (PRO) 15
3.13.1 Analysis Populations 16
3.13.2 Statistical Methods 16
3.13.3 Compliance Summary 17
4 REFERENCES 18
5 APPENDICES 19

5.1 FACT-P (version 4)* 19
5.2 FACT-P Scoring Guidelines 22
5.3 FACT-P Score Category 24

LISTING OF TABLES

Table 1 Censoring Rules for DOR 9
Table 2 Censoring Rules for Analyses of rPFS 10
Table 3 Analysis Strategy for Key Efficacy Variables 10
Table 4 Confidence Intervals for Different Observed Response Rates and Sample Size 14
Table 5 PRO Data Collection Schedule 16
Table 6 Mapping Relative Day to Analysis Visit 16

1 INTRODUCTION

This supplemental SAP (sSAP) is a companion document to the protocol. In addition to the information presented in the protocol SAP which provides the principal features of confirmatory analyses for this trial, this supplemental SAP provides additional statistical analysis details/data derivations and documents modifications or additions to the analysis plan that are not "principal" in nature and result from information that was not available at the time of protocol finalization. The PRO analysis plan will also be included in this sSAP. There will be a separate PK analysis plan as well as biomarker analysis plan.

2 SUMMARY OF CHANGES

Section 3.6.1	Update censoring rule for rPFS and DOR per company TA standard
Section 3.10	Update subgroup analysis

3 ANALYTICAL AND METHODOLOGICAL DETAILS

3.1 Statistical Analysis Plan Summary

Key elements of the statistical analysis plan are summarized below; the comprehensive plan is provided in Sections 3.2-3.13.

Study Design Overview	Phase II Trial of Pembrolizumab (MK-3475) in Subjects with mCRPC (KEYNOTE-199)
Treatment Assignment	This is an open-label study. All subjects will receive pembrolizumab 200 mg administered intravenously (IV) every 3 weeks (Q3W).
Analysis Populations	Efficacy: The All Subjects as Treated (ASaT) population will serve as the primary population for the analysis of efficacy data in this study. The ASaT population consists of all allocated subjects who received at least one dose of study treatment. -The analysis population for ORR consists of all subjects with measurable disease at baseline in ASaT. Safety: All Subjects as Treated (ASaT)
Primary Endpoint(s)	Cohorts 1, 2 and 4: Objective response rate (ORR) per RECIST 1.1 assessed by central imaging vendor
Secondary Endpoints	Duration of response (DOR) per RECIST 1.1 and PCWG3-modified RECIST 1.1, Disease control rate (DCR), radiographic progression-free (rPFS) where progressive disease (PD) in bone metastases will be determined by radionuclide bone scan using PCWG3 criteria and PD for all other tumors will be determined using by RECIST 1.1 by central imaging vendor, PSA response rate, time to PSA progression, overall survival (OS) Cohorts 4 and 5 only: duration of PSA response, time to initiation of cytotoxic chemotherapy, time to new-anticancer therapy, time to first skeletal-related event

Statistical Methods for Key Efficacy/Immunogenicity/ Pharmacokinetic Analyses	-For ORR: the point estimate and 95% CI will be calculated using an exact method based on binomial distribution (Clopper-Pearson method)
Statistical Methods for Key Safety Analyses	Counts and percentages of AEs will be provided.
Interim Analyses	One interim analysis will be performed in this study. The interim analysis is summarized below. Details are provided in Section 3.7. • IA: o Timing: To be performed about 27 weeks after the first 100 subject with RECIST 1.1-measurable disease (Cohorts 1 and 2) and all subjects in Cohort 3 have been enrolled. o Scope of analyses: ORR, DOR and safety will be evaluated
Multiplicity	There will be no hypothesis testing performed in this study. For the primary and secondary objectives in this Phase II study, no adjustment to control for Type I error is planned.
Sample Size and Power	The planned sample size is approximately 370 including: Cohorts 1-3: subjects with mCRPC previously treated with docetaxel-based chemotherapy; Cohort 1 includes subjects with PD-L1 positive, RECIST 1.1-measurable disease and Cohort 2 includes subjects with PD-L1 negative, RECIST 1.1-measurable disease (n≈200 in Cohorts 1 and 2 combined); Cohort 3 includes subjects with bone-metastases, RECIST 1.1 non-measurable (n≈50) Cohorts 4 and 5: subjects who are on pre-chemotherapy enzalutamide treatment and failing or showing signs of failing enzalutamide treatment Cohort 4: subjects with RECISIT 1.1-measurable disease (n≈80) Cohort 5: subjects bone metastases only or bone-predominant disease (n≈40)

3.2 Responsibility for Analyses/In-House Blinding

The statistical analysis of the data obtained from this study will be the responsibility of the Clinical Biostatistics department of the SPONSOR.

This is an open-label trial; therefore, the Sponsor, investigators, and subjects will know the treatment administered.

Association between biomarker status and clinical response/resistance will be explored. The subjects-level biomarker results will be masked in the database to the study team, including clinical, statistical, statistical programming, and data management personnel throughout the study period until database lock for the primary study report. Access to the biomarker subject-level results will be limited to an unblinded Sponsor clinical scientist, unblinded Sponsor data management personnel, unblinded Sponsor statistician, and unblinded Sponsor statistical programmer who will be responsible for biomarker data review and analysis. Analyses related to biomarker scores will be conducted by a sponsor unblinded statistician. Unblinded analyses related to biomarker scores will be limited to biomarker score vs. outcome at the interim analysis and conducted by a sponsor unblinded biomarker statistician.

3.3 Hypotheses/Estimation

This trial is being conducted as an open-label study. There is no randomization in the study. Objectives and associated estimations of the study are stated in Protocol Section 3.

3.4 Analysis Endpoints

3.4.1 Efficacy/Immunogenicity/Pharmacokinetics Endpoints

Primary

- **Objective Response Rate (ORR) – per RECIST 1.1 assessed by central imaging vendor**

Proportion of subjects in the analysis population who have complete response (CR) or partial response (PR) where responses are determined by RECIST 1.1 assessed by central imaging vendor.

Secondary

- **Duration of Response (DOR) – per RECIST 1.1 assessed by central imaging vendor**

For subjects who demonstrated CR or PR, DOR is defined as the time from first documented evidence of CR or PR until progressive disease (PD) assessed by central imaging using RECIST 1.1 or death due to any cause, whichever occurs first.

- **Duration of Response (DOR) – per PCWG3-modified RECIST 1.1 assessed by central imaging vendor**

For subjects who demonstrated CR or PR, DOR is defined as the time from first documented evidence of CR or PR until progressive disease (PD) assessed by central imaging where PD will be determined by radionuclide bone scan using RECIST 1.1/PCWG3 criteria and PD for all other tumors will be determined using RECIST 1.1 or death due to any cause, whichever occurs first.

- **Disease Control Rate (DCR) – per PCWG3-modified RECIST 1.1 assessed by central imaging vendor**

Proportion of subjects in the analysis population who have CR or PR or stable disease (SD) or Non-CR/Non-PD for at least 6 months, by central imaging vendor where PD in bone-only tumors will be determined by radionuclide bone scan using PCWG3 criteria and PD for all other tumors will be determined using RECIST 1.1.

- **PSA Response Rate**

Proportion of subjects in the analysis population who have PSA response defined as at least 50% (≥50%) decline from baseline measured twice at least 3 weeks apart, i.e. the number of subjects with PSA response divided by the total number of subjects with PSA measurement at baseline. Response must be confirmed by a second PSA value 3 or more weeks later.

- **Time to PSA Progression**

Time to PSA progression is defined as the time from first day of study treatment to the date of PSA progression. Subjects without PSA progression will be censored at the last PSA assessment date. For subjects who have a decline in PSA, the PSA progression date is defined as the date that an increase of 25% or more and an absolute increase of 2ng/mL or more from the nadir is documented.

For subjects who have no decline in PSA, the PSA progression date is defined as the date that an increase of 25% or more and an absolute increase of 2ng/mL or more, at least 12 weeks from the baseline is documented. If the PSA value at progression is less than the baseline, then PSA progression must be confirmed by a second value 3 or more weeks later.

.

- **Radiographic progression-free survival (rPFS) – per PCWG3-modified RECIST 1.1 assessed by central imaging vendor**

Progression-free-survival (rPFS) is defined as the time from first day of study treatment to the documented disease progression by central imaging vendor where PD in bone-only tumors will be determined by radionuclide bone scan using PCWG3 criteria and PD for all other tumors will be determined using RECIST 1.1 or death due to any cause, whichever occurs first.

- **Overall Survival (OS)**

Overall survival (OS) is defined as the time from first day of study treatment to the time of death.

- **Duration of PSA response (Cohorts 4 and 5 only)**

Duration of PSA response is defined as the time from PSA response, when the PSA value first declines by at least 50% of the baseline (must be confirmed by a second value), to the date of PSA progression at which there is an increase of 25% or more from the nadir PSA, provided the absolute increase from the nadir PSA is at least 2ng/mL.

- **Time to initiation of cytotoxic chemotherapy (Cohorts 4 and 5 only)**

Time to initiation of cytotoxic chemotherapy is defined as the time from first day of study treatment to the time of initiation of cytotoxic chemotherapy for prostate cancer.

- **Time to new-anticancer therapy (Cohorts 4 and 5 only)**

Time to new-anticancer therapy is defined as the time from first day of study treatment to the time of new-anticancer therapy for prostate cancer

- **Time to first skeletal-related event (Cohorts 4 and 5 only)**

Time to initiation of first skeletal-related event is defined as the time from first day of study treatment to the first skeletal-related event, which is defined as radiation therapy or surgery to bone, pathologic bone fracture, spinal cord compression, or change or antineoplastic therapy to treat bone pain.

See Section 3.6.1 for censoring rules.

Exploratory

Exploratory endpoints of this study include patient-reported outcomes (PROs), ORR by immune-related RECIST 1.1 (irRECIST) and DOR, DCR and rPFS by PCWG3-modified immune-related RECIST (irRECIST).

3.4.2 Safety Endpoints

The primary safety objective of this trial is to characterize the safety and tolerability of pembrolizumab in subjects with mCRPC. The primary safety analysis will be based on subjects who experienced toxicities as defined by CTCAE criteria. Safety will be assessed by quantifying the toxicities and grades experienced by subjects who have received pembrolizumab, including serious adverse events (SAEs) and events of clinical interest (ECIs).

Safety will be assessed by reported adverse experiences using CTCAE, Version 4.0. The attribution to study treatment, time-of-onset, duration of the event, its resolution, and any concomitant medications administered will be recorded. AEs will be analyzed including but not limited to all AEs, SAEs, fatal AEs, and laboratory changes

3.5 Analysis Populations

Subjects will be entered into the trial when they are assigned an allocation number by the IVRS/IWRS.

3.5.1 Efficacy Analysis Populations

The All Subjects as Treated (ASaT) population will serve as the primary population for the analysis of efficacy data in this study. The ASaT population consists of all allocated subjects who received at least one dose of study treatment.

The analysis of DCR, PSA response rate, time to PSA progression, time to initiation of cytotoxic chemotherapy, time to new-anticancer therapy, time to first skeletal-related event, rPFS, and OS will be based on the ASaT population.

The analysis of ORR will be based on all subjects with measurable disease at baseline in the ASaT population. The analysis of DOR will be based on all responders with measurable disease at baseline in the ASaT population.

The analysis of duration of PSA response will be based on all PSA responders with at least 50% decline from baseline measured twice at least 3 weeks apart.

For the PSA response, the primary analysis will comprise all subjects with measurable PSA at baseline; the secondary analysis will comprise all subjects with elevated PSA at baseline as a sensitivity analysis.

3.5.2 Safety Analysis Populations

The All Subjects as Treated (ASaT) population will be used for the analysis of safety data in this study. The ASaT population consists of all allocated subjects who received at least one dose of study treatment.

At least one laboratory or vital sign measurement obtained subsequent to at least one dose of study treatment is required for inclusion in the analysis of each specific parameter. To assess change from baseline, a baseline measurement is also required.

3.5.3 Biomarker Discovery Population

The biomarker discovery population consists of the first 100 subjects with RECIST1.1 -measurable disease in Cohorts 1 and 2 who are clinically evaluable, which is defined as any subjects who received at least one dose of study drug and had at least 3 post-baseline scans, or discontinued due to radiographic/clinical progression or death before reaching week 27.

The biomarker discovery population will be used in this study for the evaluation of the association between biomarker score and clinical efficacy and subsequently to determine a biomarker cut-point (if applicable and necessary). The biomarker discovery population will be excluded from any subsequent validation analyses for any biomarker selected population, if applicable. These subjects will still be included in the general efficacy analyses for all cohorts combined and by each cohort irrespective of biomarker status.

More details will be provided in a separate biomarker analysis plan.

3.6 Statistical Methods

3.6.1 Statistical Methods for Efficacy Analyses

For ORR, DCR and PSA response rate, the point estimate and 95% confidence interval will be provided using the exact binomial method proposed by Clopper and Pearson (1934) [3]. Subjects in the analysis population without response data will be counted as non-responders.

For rPFS, DOR, time to PSA progression, duration of PSA response, time to initiation of cytotoxic chemotherapy, time to new-anticancer therapy, time to first skeletal-related event, and OS, Kaplan-Meier (KM) curves and median estimates from the KM curves will be provided as appropriate. Censoring rules for DOR are summarized in Table 1.

Table 1 Censoring Rules for DOR

Situation	Date of Progression or Censoring	Outcome
No progression nor death, no new anti-cancer therapy initiated	Last adequate disease assessment	Censor (non-event)
No progression nor death, new anti-cancer therapy initiated	Last adequate disease assessment before new anti-cancer therapy initiated	Censor (non-event)
Death or progression immediately after ≥ 2 consecutive missed disease assessments or after new anti-cancer therapy, if any	Earlier date of last adequate disease assessment prior to ≥ 2 missed adequate disease assessments and new anti-cancer therapy, if any	Censor (non-event)
Death or progression after ≤ 1 missed disease assessments and before new anti-cancer therapy, if any	PD or death	End of response (Event)
A missed disease assessment includes any assessment that is not obtained or is considered inadequate for evaluation of response.		

Censoring rules for rPFS are summarized in Table 2 .

Table 2 Censoring Rules for Analyses of rPFS

Situation	Rule
PD or death documented after ≤ 1 missed disease assessment, and before new anti-cancer therapy, if any	Progressed at date of documented PD or death
PD or death documented immediately after ≥ 2 consecutive missed disease assessments or after new anti-cancer therapy, if any	Censored at last disease assessment prior to the earlier date of ≥ 2 consecutive missed disease assessment and new anti-cancer therapy, if any
No PD and no death; and new anticancer treatment is not initiated	Censored at last disease assessment
No PD and no death; new anticancer treatment is initiated	Censored at last disease assessment before new anticancer treatment

Table 3 summarizes the key efficacy analyses.

Table 3 Analysis Strategy for Key Efficacy Variables

.

Endpoint/Variable (Description, Time Point)	Statistical Method	Analysis Population	Missing Data Approach
ORR per RECIST 1.1 assessed by central imaging vendor	Exact method based on binomial distribution (Clopper-Pearson method)	Subjects with measurable disease at baseline in ASaT in all below populations • Cohorts 1 and 2 • Cohort 1 • Cohort 2 • Cohort 4	Subjects with missing data are considered non-responders

DOR per RECSIT 1.1 and per PCWG3-modified RECIST 1.1 assessed by central imaging vendor	Summary statistics using Kaplan-Meier method	Responders with measurable disease at baseline in ASaT in all below populations • Cohorts 1 and 2 • Cohort 1 • Cohort 2 • Cohort 4	Details are provided in a separate table
DCR per PCWG3-modified RECIST 1.1 assessed by central imaging vendor	Exact method based on binomial distribution (Clopper-Pearson method)	ASaT • Cohorts 1+2+3 • Cohorts 1+2 • Cohorts 4+5 • Each Cohort	Subjects with missing data are considered disease not under control
PSA Response Rate[§]	Exact method based on binomial distribution (Clopper-Pearson method)	ASaT • Cohorts 1+2+3 • Cohorts 1+2 • Cohorts 4+5 • Each Cohort	Subjects with missing data are considered non-responders
Time to PSA Progression[§]	Summary statistic using Kaplan-Meier method	ASaT • Cohorts 1+2+3 • Cohorts 1+2 • Cohorts 4+5 • Each Cohort	Censored at last assessment date
Radiographic PFS per PCWG3-modified RECIST 1.1 assessed by central imaging vendor	Summary statistics using Kaplan-Meier method	ASaT • Cohorts 1+2+3 • Cohorts 1+2 • Cohorts 4+5 • Each Cohort	Censored at last assessment date
OS	Summary statistics using Kaplan-Meier method	ASaT • Cohorts 1+2+3 • Cohorts 1+2 • Cohorts 4+5 • Each Cohort	Censored at last date known alive
Duration of PSA response	Summary statistics using Kaplan-Meier method	PSA responders in • Cohorts 4 and 5 • Cohort 4 • Cohort 5	Censored at last assessment date if no new anti-cancer therapy initiated or last adequate disease assessment before new anti-cancer therapy initiated

Time to initiation of cytotoxic chemotherapy	Summary statistics using Kaplan-Meier method	ASaT in • Cohorts 4 and 5 • Cohort 4 • Cohort 5	Censored at last date known alive
Time to new-anticancer therapy	Summary statistics using Kaplan-Meier method	ASaT in • Cohorts 4 and 5 • Cohort 4 • Cohort 5	Censored at last date known alive
Time to first skeletal-related event	Summary statistics using Kaplan-Meier method	ASaT in • Cohorts 4 and 5 • Cohort 4 • Cohort 5	Censored at last date known alive
[§]: subjects without PSA at baseline will be excluded.			

3.6.2 Statistical Methods for Safety Analyses

Safety will be summarized for the overall safety population for Cohorts 1, 2, and 3 in the post-chemotherapy population and overall safety for Cohorts 4 and 5 in the pre-chemotherapy population. If there is no difference among individual patient populations, an overall analysis will be performed.
Safety and tolerability will be assessed by clinical review of all relevant parameters including adverse experiences (AEs), laboratory tests, and vital signs. Counts and percentages of AE will be provided.
Summary statistics (counts, percentages, means, standard deviations, etc.) will be provided for the safety endpoints as appropriate. The summary statistics of count and percentage will be provided for the incidence rate of any AE, any serious AE, any Grade 3-5 AE, any drug-related AE, any serious and drug-related AE, any Grade 3-5 and drug-related AE, dose interrupted due to AE, discontinuation due to AE, any immune-related AE (irAE), death and specific AEs of interest.
The summary statistics of mean and standard deviation will be provided for change from Baseline Results (Labs and Vital Signs) within each treatment cycle. The summary statistics of count and percentage will be provided for laboratory worsening from Baseline in terms of CTCAE grades.

3.6.3 Statistical Methods for Biomarker

The evaluation of a general positive association between biomarker score and ORR/DCR will be investigated via standard logistic regression as well as generalized additive models. The profiles of PPV (positive predictive value), NPV (negative predictive value), and the percentage of subjects above a given cut-off, along with intervals quantifying the uncertainty in those profiles, will be estimated as a function of potential cut-offs. A biomarker cut-off that maintains high NPV (e.g., near or above 90%) while achieving meaningful enrichment of response and largely capturing subjects showing durable clinical benefit is sought. Receiver operating characteristic curve analysis will also be used to understand the sensitivity and specificity profile and examine cut-offs that might be suggested based on the ROC curve and their appropriateness with regard to PPV and NPV.

3.6.4 Summaries of Baseline Characteristics, Demographics, and Other Analyses

The number and percentage of subjects screened, enrolled, the primary reasons for screening failure, and the primary reason for discontinuation will be displayed. Demographic variables, baseline characteristics, primary and secondary diagnoses, and prior and concomitant therapies will be summarized either by descriptive statistics or categorical tables.

3.7 Interim Analyses

Efficacy and safety data will be monitored over time in this single-arm, open-label, multi-cohort study.

One interim analysis will be conducted in the first 100 subjects with RECIST 1.1-measurable disease by RECIST 1.1 (Cohort 1 and Cohort 2) and all subjects in Cohort 3 at about 27 weeks after the completion of enrollment of all these subjects. Other analyses at the interim analysis will include DOR and safety.

A biomarker analysis is planned at the time of the interim analysis where a subset of subjects (the biomarker discovery population, about 100 subjects from Cohorts 1 and 2, see Section 3.5.3 for details) will be used to explore the association between the biomarker score and clinical efficacy/resistance and also to determine the biomarker cut-point. More subjects may be allocated to the biomarker discovery population, if necessary.

The final analysis of ORR in the population with measurable disease by RECIST 1.1 (Cohort 1, Cohort 2, and Cohort 4) will be performed after all subjects in Cohorts 1 to 5 have been enrolled and all subjects have accumulated mature data and had adequate follow -up. All other analyses, including safety, DOR, DCR, PSA response, time to PSA progression, rPFS, and OS will be analyzed at this time.

3.8 Multiplicity

There will be no hypothesis testing performed in this study. For the primary and secondary objectives in this Phase II study, no adjustment to control for Type I error is planned.

3.9 Sample Size and Power Calculations

The primary objective of this study is to estimate ORR in Cohorts 1 and 2 combined, Cohort 1, Cohort 2, and Cohort 4. The planned sample size is approximately 370 with about 200 subjects in Cohorts 1 and 2, 50 subjects in Cohort 3, 80 subjects in Cohort 4, and 40 subjects in Cohort 5. With a sample size of 200, 100, and 80, the confidence intervals for an observed response rate of 10%, 15%, 20%, 25% and 30% are listed below for 95% confidence level. In addition, about 50 subjects with bone-predominant disease at baseline for Cohort 3 will be enrolled.

Table 4 Confidence Intervals for Different Observed Response Rates and Sample Size

Sample Size	Number of Subjects with a response	Observed Response Rate	95% CI	95% CI Width
200	20	10%	(6.2%, 15.0%)	8.8%
200	30	15%	(10.4%, 20.7%)	10.4%
200	40	20%	(14.7%, 26.2%)	11.5%
200	50	25%	(19.2%, 31.6%)	12.4%
200	60	30%	(23.7%, 36.9%)	13.1%
100	10	10%	(4.9%, 17.6%)	12.7%
100	15	15%	(8.6%, 23.5%)	14.9%
100	20	20%	(12.7%, 29.2%)	16.5%
100	25	25%	(16.9%, 34.7%)	17.8%
100	30	30%	(21.2%, 40.0%)	18.7%
80	8	10%	(4.4%, 18.8 %)	14.3%
80	12	15%	(8.0%, 24.7 %)	16.7%
80	16	20%	(11.9%, 30.4 %)	18.6%
80	20	25%	(16.0%, 35.9 %)	19.9%
80	24	30%	(20.3%, 41.3 %)	21.0%

3.10 Subgroup Analyses and Effect of Baseline Factors

For the primary objective ORR and DOR, subgroup analyses to be performed include but may not be limited to following subgroups based on

- Age (<65 years, ⩾65 years)
- Race (White, Other)
- PD-L1 expression level (positive, negative),
- Geographic region (North America, EU region, Rest of the World)
- ECOG performance status (0-1, 2) at baseline
- Visceral Disease based on both target and non-target lesions at screening (Yes with Liver, Yes without liver , No)
- Number of prior chemotherapy regimens (1, >1)
- Prior targeted endocrine therapy: abiraterone only vs. enzalutamide only vs. abiraterone plus enzalutamide vs. other
- Total Gleason score (≤ 7 and ≥ 8) at diagnosis
- Baseline PSA value (at or below median versus above median);

  Categories of subgroup may be subject to change according to actual data. Categories with few than 5 subjects might not be considered or might be combined into other categories.

3.11 Compliance (Medication Adherence)

Drug accountability data for trial treatment will be collected during the study. Any deviation from protocol-directed administration will be reported.

3.12 Extent of Exposure

Extent of Exposure for a subject is defined as number of cycles in which the subject receives the study medication infusion. Summary statistics will be provided on Extent of Exposure for ASaT population.

3.13 Statistical considerations for Patient-Reported Outcomes (PRO)

The patient-reported outcomes are exploratory objectives in this study and no formal hypotheses were formulated. The objective is to evaluate changes in health-related quality of life assessment from baseline using FACT-P and to characterize utilities using EuroQoL-5D.

PRO Endpoints:

- The mean score change from baseline in FACT-P total score, Prostate Cancer Subscale (PCS) Score, PCS pain-related score at Week 27
- The mean score change from baseline in EQ-5D VAS score at Week 27
- The number and proportions of deterioration/stable/improvement from baseline in FACT-P score category at Week 27

For multi-item scale(s), the analysis will focus on the subscale score rather than each single item. Results at Week 36 and 45 weeks will also be reported if there is data from enough subjects at these time points.

Scoring Algorithm:

EQ-5D Scoring: EQ-5D utility score should be calculated based on the European algorithm [4]. The five health state dimensions in this instrument include the following: mobility, self-care, usual activities, pain/discomfort, and anxiety/depression. Each dimension is rated on a three-point scale from 1 (extreme problem) to 3 (no problem). The eEQ-5D also includes a graded (0 to 100) vertical visual analog scale on which the subject rates his or her general state of health at the time of the assessment.

FACT-P Scoring: FACT-P (version 4.0, Appendix 5.1) consists of FACT-G (general) which contains a 27-item self-report questionnaire measuring general HRQoL in four domains (physical, social, emotional, and functional well-being) and 12 prostate cancer-specific items. FACT-P score will be calculated based on the Manual of the Functional Assessment of Chronic Illness Therapy (FACIT) Measurement System, version 4.0 (November 1997) [1] and Measuring quality of life in men with prostate cancer using the functional assessments of cancer therapy prostate instrument [5].

The schedule for PRO data collection:

Table 5 provides the schedule for PRO data collection

Table 5 PRO Data Collection Schedule

Trial Period:	Treatment Cycles [a]									End of Treatment	Post-Treatment
Treatment Cycle/Title:	1	2	3	4	7	10	13	16	19	Discontinuation	Safety Follow-up
Week	0	3	6	9	18	27	36	45	54	X	X
EuroQol EQ-5D; FACT-P[b]	X	X	X	X	X	X	X	X	X	X	X

a: Each cycle is 3 weeks

b: Electronic patient reported outcomes (ePROs) should ideally be administered prior to all study procedures (EQ 5D first, followed by FACT-P). All ePROs are to be administered pre-infusion at Cycles 1, 2, 3, and 4 and every 3 cycles thereafter (Cycle 7, 10, 13, 16, and 19), at treatment discontinuation, and again 30 days after last dose of trial treatment. If the subject does not complete the ePROs at a scheduled time point, the MISS_MODE form must be completed to capture the reason the assessment was not performed.

The general rule of mapping relative day to analysis visit is provided in Table 6.

Table 6 Mapping Relative Day to Analysis Visit

Week	0	3	6	9	18	27	36	45	54
Target Day	1	21	42	63	126	189	252	315	378
Range	<=1	2-31	32-52	53-94	95-157	158-220	221-283	284-346	347-

At each scheduled visit, FACT-P and EQ-5D will be collected. If there are multiple PRO collections within any of the stated time windows, the closest collection to the target day will be used.

3.13.1 Analysis Populations

The primary analysis approach for the pre-specified exploratory PRO endpoints will be based on a quality of life related full analysis set (FAS) population. This population consists of all allocated subjects who have received at least one dose of study medication, and have completed at least one PRO assessment.

3.13.2 Statistical Methods

Descriptive statistics (mean and 95% CI) of change from baseline will be presented for EQ-5D VAS score, FACT-P prostate cancer subscale score, and FACT-P total score on observed data alone and using cLDA(constrained longitudinal data analysis) model.

The cLDA model is specified as follows:

$$E(Y_{it}) = \gamma_0 + \alpha Z_t, i = 1, 2, \ldots, t = ,0,1,2, \ldots,$$

where Y_{it} is the PRO score for subject i, at visit t, $\gamma_0$is the baseline mean, α is the response effect and Z_t is the status at visit t.

Summary statistics (mean and SE) for FACT-P total scores by scheduled visit will be plotted.

3.13.2.1 Analysis of the Proportions of Deterioration/Stable/Improvement

Predetermined minimally important difference (MID) is used to classify FACT-P PRO score as "improved" (increase from baseline ≥ MID), "stable", or "deteriorated" (decrease from baseline ≥ MID) from baseline as this magnitude of change is perceived by patients as being clinically significant [2]. Details are provided in Appendix 5.3.

The number and proportion of patients who "improved", "stable", or "deteriorated" from baseline will be summarized.

3.13.3 Compliance Summary

Completion and compliance FACT-P and EQ-5D by visit and by Cohort will be described based on PRO FAS population. Numbers and percentages of complete and missing data at each visit will be summarized.

Completion rate in the FAS population is defined as the percentage of number of subjects who complete at least one item over the number of subjects in the FAS population at each time points.

$$Completion\ Rate = \frac{Number\ of\ Subjects\ who\ Complete\ at\ least\ one\ Item}{Number\ of\ Allocated\ Subjects}$$

The completion rate is expected to shrink in the later visit during study period due to the subjects who discontinued early. Therefore, another measurement, Compliance Rate, defined as the percentage of observed visit over number of eligible subjects who are expected to complete the PRO assessment (not including the subjects missing by design (such as death, discontinuation, translation not available) will be employed as the support for completion rate.

$$Compliance\ Rate = \frac{Number\ of\ Subjects\ who\ Complete\ at\ least\ one\ Item}{Number\ of\ Eligible\ Subjects\ who\ are\ Expected\ to\ Complete}$$

Reason for non-completion will be summarized.

4 REFERENCES

1. Cella D: Manual of the Functional Assessment of Chronic Illness Therapy (FACIT) Measurement System. Center on Outcomes, Research and Education (CORE), Evanston Northwestern Healthcare and Northwestern University, Evanston IL, Version 4 1997

2. Cella D, Nichol MB, Eton D, Nelson JB, Mulani P. Estimating clinically meaningful changes for the Functional Assessment of Cancer Therapy–Prostate: results from a clinical trial of patients with metastatic hormone-refractory prostate cancer. Value Health 2009;12(1):124–9.

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

4. EQ-5D-3L User Guide, Oct 2013.

5. Esper P, Mo F, Chodak G, et al. Measuring quality of life in men with prostate cancer using the functional assessments of cancer therapy prostate instrument. Urology. 1997;50 (6) :920-8

5 APPENDICES

5.1 FACT-P (version 4)*

FACT-P (version 4) - page 1

Below is a list of statements that other people with your illness have said are important. **Please circle or mark one number per line to indicate your response as it applies to the past 7 days.**

	PHYSICAL WELL-BEING	**Not at all**	**A little bit**	**Some-what**	**Quite a bit**	**Very much**
GP1	I have a lack of energy	0	1	2	3	4
GP2	I have nausea	0	1	2	3	4
GP3	Because of my physical condition, I have trouble meeting the needs of my family	0	1	2	3	4
GP4	I have pain	0	1	2	3	4
GP5	I am bothered by side effects of treatment	0	1	2	3	4
GP6	I feel ill	0	1	2	3	4
GP7	I am forced to spend time in bed	0	1	2	3	4

	SOCIAL/FAMILY WELL-BEING	**Not at all**	**A little bit**	**Some-what**	**Quite a bit**	**Very much**
GS1	I feel close to my friends	0	1	2	3	4
GS2	I get emotional support from my family	0	1	2	3	4
GS3	I get support from my friends	0	1	2	3	4
GS4	My family has accepted my illness	0	1	2	3	4
GS5	I am satisfied with family communication about my illness	0	1	2	3	4
GS6	I feel close to my partner (or the person who is my main support)	0	1	2	3	4
Q1	*Regardless of your current level of sexual activity, please answer the following question. If you prefer not to answer it, please mark this box* ☐ *and go to the next section.*					
GS7	I am satisfied with my sex life	0	1	2	3	4

FACT-P (version 4) - page 2

Please circle or mark one number per line to indicate your response as it applies to the past 7 days.

	EMOTIONAL WELL-BEING	**Not at all**	**A little bit**	**Some-what**	**Quite a bit**	**Very much**
GE1	I feel sad	0	1	2	3	4
GE2	I am satisfied with how I am coping with my illness	0	1	2	3	4
GE3	I am losing hope in the fight against my illness	0	1	2	3	4
GE4	I feel nervous	0	1	2	3	4
GE5	I worry about dying	0	1	2	3	4
GE6	I worry that my condition will get worse	0	1	2	3	4

	FUNCTIONAL WELL-BEING	**Not at all**	**A little bit**	**Some-what**	**Quite a bit**	**Very much**
GF1	I am able to work (include work at home)	0	1	2	3	4
GF2	My work (include work at home) is fulfilling	0	1	2	3	4
GF3	I am able to enjoy life	0	1	2	3	4
GF4	I have accepted my illness	0	1	2	3	4
GF5	I am sleeping well	0	1	2	3	4
GF6	I am enjoying the things I usually do for fun	0	1	2	3	4
GF7	I am content with the quality of my life right now	0	1	2	3	4

*: from www.facit.org

FACT-P (version 4) - page 3

Please circle or mark one number per line to indicate your response as it applies to the past 7 days.

	ADDITIONAL CONCERNS	Not at all	A little bit	Some-what	Quite a bit	Very much
C2	I am losing weight	0	1	2	3	4
C6	I have a good appetite	0	1	2	3	4
P1	I have aches and pains that bother me	0	1	2	3	4
P2	I have certain parts of my body where I experience pain	0	1	2	3	4
P3	My pain keeps me from doing things I want to do	0	1	2	3	4
P4	I am satisfied with my present comfort level	0	1	2	3	4
P5	I am able to feel like a man	0	1	2	3	4
P6	I have trouble moving my bowels	0	1	2	3	4
P7	I have difficulty urinating	0	1	2	3	4
BL2	I urinate more frequently than usual	0	1	2	3	4
P8	My problems with urinating limit my activities	0	1	2	3	4
BL5	I am able to have and maintain an erection	0	1	2	3	4

5.2 FACT-P Scoring Guidelines

FACT-P Scoring Guidelines (Version 4) – Page 1

Instructions:*
1. Record answers in "item response" column. If missing, mark with an X
2. Perform reversals as indicated, and sum individual items to obtain a score.
3. Multiply the sum of the item scores by the number of items in the subscale, then divide by the number of items answered. This produces the subscale score.
4. Add subscale scores to derive total scores (TOI, FACT-G & FACT-P).
5. **The higher the score, the better the QOL.**

Subscale	Item Code	Reverse item?		Item response	Item Score
PHYSICAL WELL-BEING (PWB)	GP1	4	-	_____	=_____
	GP2	4	-	_____	=_____
	GP3	4	-	_____	=_____
	GP4	4	-	_____	=_____
Score range: 0-28	GP5	4	-	_____	=_____
	GP6	4	-	_____	=_____
	GP7	4	-	_____	=_____

Sum individual item scores: _____
Multiply by 7: _____
Divide by number of items answered: _____=**PWB subscale sco**

SOCIAL/FAMILY WELL-BEING (SWB)	GS1	0	+	_____	=_____
	GS2	0	+	_____	=_____
	GS3	0	+	_____	=_____
	GS4	0	+	_____	=_____
Score range: 0-28	GS5	0	+	_____	=_____
	GS6	0	+	_____	=_____
	GS7	0	+	_____	=_____

Sum individual item scores: _____
Multiply by 7: _____
Divide by number of items answered: _____=**SWB subscale scor**

EMOTIONAL WELL-BEING (EWB)	GE1	4	-	_____	=_____
	GE2	0	+	_____	=_____
	GE3	4	-	_____	=_____
	GE4	4	-	_____	=_____
Score range: 0-24	GE5	4	-	_____	=_____
	GE6	4	-	_____	=_____

Sum individual item scores: _____
Multiply by 6: _____
Divide by number of items answered: _____=**EWB subscale scor**

FUNCTIONAL WELL-BEING (FWB)	GF1	0	+	_____	=_____
	GF2	0	+	_____	=_____
	GF3	0	+	_____	=_____
	GF4	0	+	_____	=_____
Score range: 0-28	GF5	0	+	_____	=_____
	GF6	0	+	_____	=_____
	GF7	0	+	_____	=_____

Sum individual item scores: _____
Multiply by 7: _____
Divide by number of items answered: _____=**FWB subscale scor**

FACT-P Scoring Guidelines (Version 4) – Page 2

Subscale	Item Code	Reverse item?		Item response	Item Score
PROSTATE	C2	4	-	_____	=_____
CANCER	C6	0	+	_____	=_____
SUBSCALE	P1	4	-	_____	=_____
(PCS)	P2	4	-	_____	=_____
	P3	4	-	_____	=_____
Score range: 0-48	P4	0	+	_____	=_____
	P5	0	+	_____	=_____
	P6	4	-	_____	=_____
	P7	4	-	_____	=_____
	BL2	4	-	_____	=_____
	P8	4	-	_____	=_____
	BL5	0	+	_____	=_____

Sum individual item scores: _____
Multiply by 12: _____
Divide by number of items answered: _____ =**PC Subscale score**

To derive a FACT-P Trial Outcome Index (TOI):
Score range: 0-104

_____ + _____ + _____ = _____ =**FACT-P TOI**
(**PWB** score) (**FWB** score) (**PCS** score)

To Derive a FACT-G total score:
Score range: 0-108

_____ + _____ + _____ + _____ = _____ =**FACT-G Total sco**
(**PWB** score) (**SWB** score) (**EWB** score) (**FWB** score)

To Derive a FACT-P total score:
Score range: 0-156

_____ + _____ + _____ + _____ + _____ = _____ =**FACT-P Total sco**
(**PWB** score) (**SWB** score) (**EWB** score) (**FWB** score) (**PCS** score)

*For guidelines on handling missing data and scoring options, please refer to the Administration and Scoring Guidelines in the manual or on-line at www.facit.org.

5.3 FACT-P Score Category

	No. of Items	Possible Score Range	Established MID[a] Range	MID used in this study
FACT-P Total Score	39	0-156	6-10	10
Physical well-being	7	0-28	2-3	3
Social/family well-being	7	0-28	2-3	3
Emotional well-being	6	0-24	2-3	3
Functional well-being	7	0-28	2-3	3
FACT-G scale	27	1-108	5-9	9
Prostate cancer subscale	12	0-48	2-3	3
a: MID: minimum important difference				