Statistical Analysis Plan

Amendment 2

Modra Pharmaceuticals

M18MDP

"A multicentre Phase IIb trial to evaluate the efficacy and tolerability of ModraDoc006/r in subjects with metastatic Castration Resistant Prostate Cancer (mCRPC), suitable for treatment with a taxane"

Covance Study ID: 000000182507

Document Version: Version 3.0 Final

Document Date: 07Oct2021

Covance Inc. CDCS
Clinical Development Commercialization Services

Table of Contents

Reviewers

The following reviews of the SAP were conducted:

Name and Title	Role	Version Last Reviewed	Company/ Organization
Deborah Morgan, BSc MSc	Project Manager	Internal Draft 0.1	Covance
Pierre Gobbens	Senior Principal Statistical Programmer	Internal Draft 0.2	Covance
Ronald L Nolen	Medical Writer	Internal Draft 0.2	Covance
Markus Roters	Associate Director, Biostatistics (Peer Review)	Internal Draft 1.1	Covance
Inga Jansonaite/Maksimovice	Senior Statistical Programmer	Draft 2.3	Labcorp
Marianne Keessen	Project Director	Draft 2.3	Modra Pharmaceuticals

Glossary of Abbreviations

Abbreviation	Term
AE	Adverse event
ALAT	Alanine Amino Transferase
ASAT	Aspartate Amino Transferase
ATC	Anatomical Therapeutic Chemical
BOR	Best Overall Response
BSA	Body Surface Area
BUN	Blood Urea Nitrogen
CI	Confidence Interval
COVID-19	Coronavirus Disease of 2019
CR	Complete Response
mCRPC	Metastatic Castration Resistant Prostate Cancer
CT	Computed Tomography
DCR	Disease Control Rate
DOR	Duration of response
ECG	Electrocardiogram
ECOG	Eastern Cooperative Oncology Group
EQ	EuroQol
eCRF	Electronic Case report form
EOT	End of Treatment
FACT	Functional Assessment of Cancer Therapy
FAS	Full Analysis Set
FWB	Functional well-being
GGT	Gamma-Glutamyl Transferase
HRQoL	Health Related Quality of Life
HR	Hazard Ratio
ICH	International Council on Harmonization
IV (or i.v.)	intravenous
LD	Lactate Dehydrogenase
MDRD	Modification of Diet in Renal Disease
MedDRA	Medical Dictionary for Regulatory Activities
Mg	Magnesium
MRI	Magnetic Resonance Imaging
NCI-CTCAE	National Cancer Institute – Common Terminology Criteria for Adverse Events
ORR	Objective Response Rate
PC	Prostate Cancer
PCS	Prostate cancer subscale
PCWG3	Prostate Cancer Clinical Trials Working Group 3
PD	Progressive Disease
PFS	Progression Free Survival
PP	Per Protocol
PR	Partial Response
PSA	Prostate Specific Antigen
PT	Preferred Term
PWB	Physical well-being
Q1	First quartile
Q3	Third quartile

RBC	Red Blood Cell
Q 21 days	once every 21 days (Q3W)
RECIST	Response Evaluation Criteria in Solid Tumors
r (in PFS)	radiographic
/r (in ModraDoc006/r)	ritonavir
RR	Response Rate
SAE	Serious Adverse Event
SAF	Safety Population
SD	Stable Disease (also Standard Deviation)
SOC	System Organ Class
TTF	Time to Treatment Failure
TTP	Time to progression
TFLs	Tables, Figures and Listings
TSQM	Treatment Satisfaction Questionnaire for Medication
VAS	Visual Analogue Scale
WBC	White Blood Cell
WHO	World Health Organization

1. Source Documents

The Statistical Analysis Plan was written based on the following documentation:

Document	Date	Version
Protocol	31Mar2020	Version 3.0
eCRF	13Apr2020	DEV 02.009

2. Protocol Details

2.1 Study Objectives

The primary objective is to determine the efficacy of ModraDoc006/r, as measured by radiographic Progression Free Survival (rPFS), compared to standard treatment with intravenous (i.v.) docetaxel in subjects with metastatic Castration Resistant Prostate Cancer (mCRPC).

The secondary objectives are as follows:

- To evaluate the efficacy of ModraDoc006/r, as measured by Prostate Cancer Clinical Trials Working Group 3 (PCWG3)-modified Response Evaluation Criteria in Solid Tumors (RECIST) v1.1 criteria of objective response rate (ORR), disease control rate (DCR) and duration of response (DOR) compared to standard treatment with i.v. docetaxel in subjects with mCRPC.
- To evaluate the clinical outcome in terms of rPFS at 6 months and time to progression (TTP) of ModraDoc006/r compared to i.v. docetaxel
- To evaluate the outcome in terms of PSA tumor marker evaluation for PSA response, PSA-PFS and time to PSA progression of ModraDoc006/r compared to i.v. docetaxel
- To compare the time to the first skeletal-related event between ModraDoc006/r and i.v. docetaxel
- To determine the safety and tolerability of ModraDoc006/r compared to i.v. docetaxel
- To compare subject's Health Related Quality of Life (HRQoL) response of ModraDoc006/r and docetaxel i.v.

2.2 Overall Study Design

This is a multicenter open label 1:1 randomized phase Iib trial to determine the efficacy and tolerability of ModraDoc006/r (i.e., ModraDoc006 in combination with ritonavir therapy) versus standard i.v. docetaxel in mCRPC subjects (**Figure 1**).

Cohort 1 will receive docetaxel at 75 mg/m^2 given i.v. as a one-hour infusion on Day 1 every 21 days plus 5 mg oral prednisone twice daily.

Cohort 2 will receive ModraDoc006/r 20 mg oral docetaxel in combination with 200 mg ritonavir in the morning and 20 mg oral docetaxel in combination with 100 mg ritonavir in the evening (7-12 hours after the morning dose), on Day 1, 8 and 15 of a 21-day cycle, plus 5 mg oral prednisone twice daily.

Of note, patients enrolled before and under protocol amendment 1 (Protocol V2.0) who are dosed with 30 mg of ModraDoc006 combined with 200 mg ritonavir in the morning and 20 mg of ModraDoc006 combined with 100 mg ritonavir in the evening, will, at the investigator's discretion, be able to continue at this dose.

Patients can be replaced for following reason:

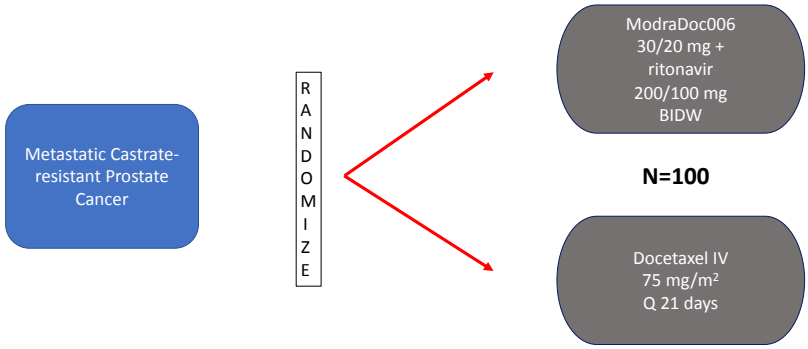
- Patients who fail to receive any administration of docetaxel i.v./ oral ModraDoc006/r
- Patients who fail to undergo any efficacy assessment after baseline measurement

As stated in protocol, patients who discontinue due to toxicity related to study drug will not be replaced.

Treatment in both cohorts will continue until disease progression, unacceptable toxicity, or discontinuation for any other reason. The end of the trial is defined as the time point when all subjects have discontinued trial treatment and have been given follow-up for safety measurements according to the trial assessment schedule. Ten months after last evaluable patient enrolled, the Sponsor will terminate the study. Any patient still under treatment in Cohort 1 can be switched to commercially available i.v. docetaxel. Patients still treated in Cohort 2 and who experience clinical benefit, will be offered the opportunity for compassionate use ModraDoc006/r.

For the purpose of data summarization, data analyses will be performed after the last patient has completed their End of Trial (EOT) visit or the last enrolled patient has completed six (6) months of treatment, whichever occurs first.

Figure 1: Schematic of study design. Metastatic Castration Resistant Prostate Cancer (mCRPC) patients for whom treatment with i.v. docetaxel is indicated according to the standard of care will be randomized 1:1 to one of the two cohorts depicted below.



Tumor assessment will be performed (following PCWG3) by PSA every cycle and radiological assessment (i.e., Computed Tomography (CT)/ Magnetic Resonance Imaging (MRI) and bone scan) after every 8 weeks for first 24 weeks (i.e. during week 9, 17 and 25), thereafter every 12 weeks and at End Of Treatment (EOT), according to the RECIST v1.1 criteria. Both PSA response and radiological response (PR/CR) must be confirmed after 3-4 weeks for PSA and after 4-6 weeks for RECIST v1.1 in patients with measurable disease. Confirmation of changes in bone lesions by PCWG3 criteria will occur after a minimum of 16 weeks (i.e. scheduled assessments during week 17 and beyond).

The study flowchart can be found in Appendix 1 (Section 11.1).

2.3 Sample Size and Power

In total, 100 patients with evaluable disease according to RECIST v1.1 and PCWG3 criteria will be treated in this trial. Patients may be replaced (see Section 2.2), therefore, the total number of treated patients may be higher than the anticipated number of evaluable patients. This phase Iib study is being conducted to gather preliminary efficacy and safety information on the new ModraDoc006/r oral treatment. The sample size is not based on a formal hypothesis testing. However, sample size sensitivity assessments were performed based on precision (width of 95% confidence interval) for the difference in the proportions of radiographic progression free survival (rPFS) between treatment at time of the 6 months follow-up. The average median rPFS in patients with mCRPC with docetaxel treatment in previous studies (Table 1) was 8.3 months, corresponding to the proportion of patients with rPFS at 6 months of around 61% (assuming rPFS follows an exponential distribution). It is expected that ModraDoc006/r will be at least as effective as i.v. docetaxel. The following table provides 95% confidence intervals for the treatment differences in proportion of rPFS at 6 month follow-up.

Docetaxel Response		ModraDoc006/r Response		Treatment Difference (ModraDoc006/r – Docetaxel)		
Median PFS	% rPFS at 6 Month	Median PFS	% rPFS at 6 Month	Difference	Lower 95% CI	Upper 95%CI
8.3	61%	8.3	61%	0%	-19.1%	19.1%
8.3	61%	9.1	63%	2%	-17.0%	21.0%
7	55%	7	55%	0%	-19.5%	19.5%
7	55%	7.7	58%	3%	-16.4%	22.4%
6	50%	6	50%	0%	-19.6%	19.6%
6	50%	6.6	53%	3%	-16.6%	22.6%

3. Efficacy and Safety Variables

3.1 Primary Efficacy Endpoint(s)

The primary efficacy endpoint for this study is rPFS according to PCWG3 criteria.

rPFS is defined as the time from the date of randomization to the date of the first objective evidence of radiologic progression or death due to any cause after treatment discontinuation, whichever occurs first will be assessed, per PCWG3 criteria.

Radiographic disease progression is defined by local assessment:

- Progressive disease by RECIST 1.1 for soft tissue disease (soft tissue disease is defined as at least 1 target lesion per RECIST 1.1 at study entry),
- or the appearance of two or more new bone lesions on bone scan (PCWG3), that is defined as question "Is there a progression of bone metastases according to PCWG3?" answered "Yes" in the CRF page of Bone Scintigraphy.

The documentation required for the determination of radiographic disease progression is listed below (Table 16 of study protocol, section 6.9.2).

Date Progression Detected (Visit)[a]	Criteria for Progression	Criteria for Confirmation of Progression (requirement and timing)	Criteria for Documentation of Disease Progression on Confirmatory Scan
Week 9	Bone lesions: 2 or more new lesions compared to **baseline** bone scan by PCWG3.	Timing: at least 6 weeks after progression identified or at Week 17 visit.[b]	Two or more new bone lesions on bone scan (compared to Week 9 scan)
	Soft tissue lesions: Progressive disease* on CT or MRI by RECIST 1.1	Confirmation required for soft tissue disease (scan of same modality as demonstrated progression)[b]	Confirmation of progressive soft tissue disease by RECIST 1.1
Week 17	Bone lesions: Two or more new lesions on bone scan compared to Week 9 bone scan.	Timing: at least 6 weeks after progression identified or at Week 25 visit. Required for bone lesions observed on bone scan[b]	Persistent[c] or increase in number of bone lesions on bone scan compared to Week 17 scan
	Soft tissue lesions: Progressive disease* on CT or MRI by RECIST 1.1	No confirmatory scan required for soft tissue disease progression.	n/a
Week 25 and Later	Bone lesions: Two or more new lesions compared to Week 9 bone scan.	Timing: at least 6 weeks after progression identified. Required for bone lesions observed on bone scan[b]	Persistent[c] or increase in number of lesions on bone scan compared to prior scan
	Soft tissue lesions: Progressive disease* on CT or MRI by RECIST 1.1	No confirmatory scan required for soft tissue disease progression	n/a

[a] Progression detected by bone scan at an unscheduled visit either prior to Week 9 or between scheduled visits will require a confirmatory scan at least 6 weeks later and should follow confirmation criteria outlined in the table for the next scheduled scan. Progression detected by CT/MRI at an unscheduled visit prior to Week 13 will require a confirmatory scan at least 6 weeks later whereas progression on or after Week 13 does not require confirmation.
[b] Confirmation must occur at the next available scan.
[c] For confirmation, at least two of the lesions first identified as new must be present at that next available scan (confirmation scan).
n/a, not applicable.
* Progressive disease is defined as overall response ticked as "Progressive Disease" on CRF page of response according to RECIST 1.1.

For computing the rPFS, the earliest documented radiologic progression will be selected, that fulfills above definition and confirmation requirement. For derivation purpose:

- "prior to Week 13" corresponds to an analysis day < D85
- "on or after Week 13" corresponds to an analysis day >=D85
- "confirmatory scan at least 6 weeks" corresponds to a consecutive scan >=42 days after the considered assessment.

Subjects without radiologic progression or death at time of analysis will be censored at the date of their last evaluable radiologic assessment as defined thereafter.
An evaluable assessment for radiological progression as per PCWG3 is defined as two components: CT/MRI scan for soft tissue disease and a bone scan for bone lesions. The date of censoring will be the latest of date of last RECIST 1.1 assessment showing no progression and date of last bone scan showing no progression, i.e. the last evaluable radiologic assessment. Evaluable according to RECIST 1.1 means with response in {CR, PR, SD, PD}. Evaluable bone scan means "Is there a progression of bone metastases according to PCWG3?" answered "Yes" or "No". If there is no such evaluable radiologic assessment, date of randomization should be used instead (e.g. for computing the delay of 168 days).
Patients with several events reasons (if that may happen) will be analyzed whichever occurs first. Patients with several censoring reasons (if that may happen) will be censored whichever occurs first.

Definition of date of rPFS or censoring refers to Food and Drug Administration (FDA) guidelines for rPFS [3] (definition including documented rPFS only), as per the table thereafter.

	Situation	Event/Censoring Date	Outcome	Description
A] No post-baseline tumor assessments (without new anticancer treatment given) i.e. No RECIST1.1 and no bone scan at all, i.e. no assessments performed for both parameters post baseline.				
A1	Death <= 168 days after rando	Date of death	Event	Death
B] C] D] With post-baseline tumor assessments (without new anticancer treatment given prior to event/censor)				
B1	Progression [i.e documented radiologic progression] after {0 or 1} missing radiologic assessment or within 168 days after previous evaluable radiologic assessment	Date of radiologic progression = Min (Date of earliest assessment showing progression based on PCWG3-modified RECIST v1.1 fulfilling confirmation criteria required by SAP/protocol ; Date of earliest bone scintigraphy showing progression based on PCWG3 fulfilling confirmation criteria required by SAP/protocol)	Event	Progression
B2	Else, Death within 168 days after last evaluable radiologic assessment (last (bone scan , RECIST))	Date of death	Event	Death
C1	Else, Progression after ≥2 consecutive missing radiologic assessments	Date of last evaluable radiologic assessment before missing assessments (either RECIST if considered progression was identified based on RECIST, either bone scan if considered progression was identified based on bone scan)	Censored	Progression after 2 or more missed assessments
C2	Else, Death > 168 days after last evaluable radiologic assessment (last (bone scan ,RECIST))	Date of last evaluable radiologic assessment = max(last recist, last bone scan)	Censored	Death after 2 or more missed assessments
D	Else [i.e. No progression, no death]	Date of last evaluable radiologic assessment = max(last recist, last bone scan)	Censored	No progression nor Death

E] F] New anticancer treatment started prior to observed event/censor				
E	And No post-baseline tumor assessment	Date of randomization	Censored	No post-baseline tumor assessments
F	Else:	Date of last evaluable radiologic assessment (i.e. max(last recist, last bone scan)) prior to initiation of new anticancer treatment	Censored	Use of new anticancer treatment

Notes:

* New anticancer treatment is defined from medications collected in CRF form "Prior and Concomitant Medications" that started after first study drug administration while not allowed by protocol, as follows: Anatomical Therapeutic Chemical (ATC) code L02 (other than L02AE, L02BB and L02BX) and L01 (antineoplastic agents).

** This 168-day interval represents 2 consecutive tumor assessments, 12 weeks apart (larger interval during study).

3.2 Secondary Efficacy Endpoints

Secondary efficacy endpoints are:

- **ORR as assessed by RECIST v1.1 (local assessment)**

The best overall soft tissue response as assessed by investigators using PCWG-3-modified RECIST 1.1 will be summarized. Only patients with measurable soft tissue disease at screening (i.e., at least 1 target lesion per RECIST 1.1) will be included in this analysis.

To compute the ORR, Best Overall Response (BOR) of each subject will be computed, using investigator's assessment of overall response at each visit, according to RECIST v1.1 criteria ([2]):

- Use of overall response at each visit when imaging is done per protocol schedule:

At each efficacy assessment (imaging), investigator is requested to assess the overall response according to RECIST v1.1 criteria. Overall response will be assessed from the evaluation of target lesions, non-target lesions, and appearance of new lesions at the given visit:

Target lesions	Non-target lesions	New lesions	Overall response
CR	CR	No	CR
CR	Non-CR/non-PD	No	PR
CR	Not evaluated	No	PR
PR	Non-PD or not all evaluated	No	PR
SD	Non-PD or not all evaluated	No	SD
Not all evaluated	Non-PD	No	NE
PD	Any	Yes or No	PD
Any	PD	Yes or No	PD
Any	Any	Yes	PD

CR = complete response, PR = partial response, SD = stable disease, PD = progressive disease, and NE = inevaluable.

Assessment of "Overall response" will be either Complete Response (CR), Partial Response (PR), Stable disease (SD), Progressive Disease (PD) or Non Evaluable (NE), as entered by the investigator in the eCRF.

- Computation of BOR

The BOR is the best response recorded during study period, including follow-up visit. Per RECIST 1.1 guidelines:

- "Confirmation of response is required for trials with response primary endpoint but is no longer required in randomized studies since the control arm serves as appropriate means of interpretation."
- "Best response in these trials is defined as the best response across all time points (for example, a patient who has SD at first assessment, PR at second assessment, and PD on last assessment has a best overall response of PR). When SD is believed to be best response, it must also meet the protocol specified minimum time from baseline. If the minimum time is not met when SD is otherwise the best time point response, the patient's best response depends on the subsequent assessments. For example, a patient who has SD at first assessment, PD at second and does not meet minimum duration for SD, will have a best response of PD. The same patient lost to follow-up after the first SD assessment would be considered inevaluable."

Therefore, for this study, computation of BOR will be derived as:

- BOR="CR" if there is a time point with OR (overall response)="CR"
- Else BOR="PR" if there is a time point with OR ="PR"
- Else BOR="SD" if there is a time point with OR="SD" where the date of assessment is >= 6 weeks after first study drug intake
- Else BOR="PD" if there is a time point with OR="PD"
- Else BOR="NE"
    - That should be all remaining cases where there is no time point having both the [date of assessment >= Y weeks] and [OR in {CR, PR, SD, PD}], i.e.:
        - Patients without any assessment post-baseline
        - Patients with a time point with OR="NE" (and not previously considered for other BOR)
        - Patients with a time point with OR="SD" where the date of assessment is < 6 weeks after first study drug intake (and not previously considered for other BOR)

Minimum duration for SD is set to 6 weeks as it corresponds to two theoretical cycles=42 days).

- Computation of ORR

ORR is defined as the proportion of subjects having CR or PR as BOR during the study treatment period, among the population evaluable for radiological response.

- **Disease control rate (DCR)**

DCR is defined as the proportion of subjects who have a CR, PR, or stable disease (SD) as best overall response during study, based on RECIST v1.1 (see computation of BOR above).

- **Duration of response (DOR) based on RECIST v1.1**

DOR is defined as the time from the documentation of tumor response using RECIST 1.1 (CR or PR, whichever is first recorded) until the first objective evidence of radiologic progression (i.e. according to RECIST 1.1 or bone scintigraphy as defined in section 3.1).

DOR will be calculated in the subpopulation of subjects experiencing a response (CR or PR) among the population evaluable for radiological response.

Subjects without documented radiologic progression at time of analysis will be censored at the date of their last evaluable radiologic assessment as defined in section 3.1.

- **rPFS at 6 months based on RECIST v1.1**

This will be obtained using same variable as in section 3.1.

- **Time to progression (TTP)**

TTP is defined as the time from the date of randomization to the date of the first documented radiologic progression.

The documentation required for the determination of radiographic disease progression is defined in section 3.1, according to PCWG3 criteria.

For computing the TTP, the earliest documented radiologic progression will be selected, that fulfills above definition and confirmation requirement.

Subjects without radiologic progression at time of analysis will be censored at the date of their last evaluable radiologic assessment as defined in section 3.1.

Definition of date of progression or censoring referred for TTP is as follows:

	Situation	**Event/Censoring Date**	**Outcome**	**Description**
A] No post-baseline tumor assessments (without new anticancer treatment given)				
A	All cases	Date of randomization	Censored	No post-baseline tumor assessments
B] C] D] With post-baseline tumor assessments (without new anticancer treatment given prior to event/censor)				
B1	Progression [i.e documented radiologic progression] after {0 or 1} missing radiologic assessment or within 168 days after previous evaluable radiologic assessment	Date of radiologic progression = Min (Date of earliest assessment showing progression based on PCWG-3-modified RECIST v1.1 fulfilling confirmation criteria required by SAP/protocol ; Date of earliest bone scintigraphy showing progression based on PCWG3 fulfilling confirmation criteria required by SAP/protocol)	Event	Progression
C1	Else, Progression after ≥2 consecutive missing radiologic assessments	Date of last evaluable radiologic assessment before missing assessments (either RECIST if considered progression was identified based on RECIST, either bone scan if considered progression was identified based on bone scan)	Censored	Progression after 2 or more missed assessments

D	Else [i.e. No progression]	Date of last evaluable radiologic assessment = max(last recist, last bone scan)	Censored	No progression
E] F] New anticancer treatment started prior to observed event/censor				
E	And No post-baseline tumor assessment	Date of randomization	Censored	No post-baseline tumor assessments
F	Else:	Date of last evaluable radiologic assessment (i.e. max(last recist, last bone scan)) prior to initiation of new anticancer treatment	Censored	Use of new anticancer treatment

Notes:
* New anticancer treatment is defined from medications collected in CRF form "Prior and Concomitant Medications" that started after first study drug administration while not allowed by protocol, as follows: Anatomical Therapeutic Chemical (ATC) code L02 (other than L02AE, L02BB and L02BX) and L01 (antineoplastic agents).
** This 168-day interval represents 2 consecutive tumor assessments, 12 weeks apart (larger interval during study).

- **PSA response rate (PSA RR) according to PCWG3 criteria**

A PSA response is defined as PSA decline of ≥50% from baseline to lowest post-baseline PSA result, with confirmatory read ≥3 weeks later, based on the PCWG3 recommendations.

A subject will be considered a responder for PSA if confirmed PSA response occurred at any time during study.

Baseline PSA value to be used as reference is the last value measured prior to study drug administration.

It will be computed for patients with PSA values at the baseline assessment and with at least 1 post baseline PSA assessment. The analysis of the tumor marker PSA will be performed on the efficacy population (FAS) as per PCWG3 criteria. The variables will be evaluated using appropriate descriptive statistics for each assessment point and for the changes from baseline.

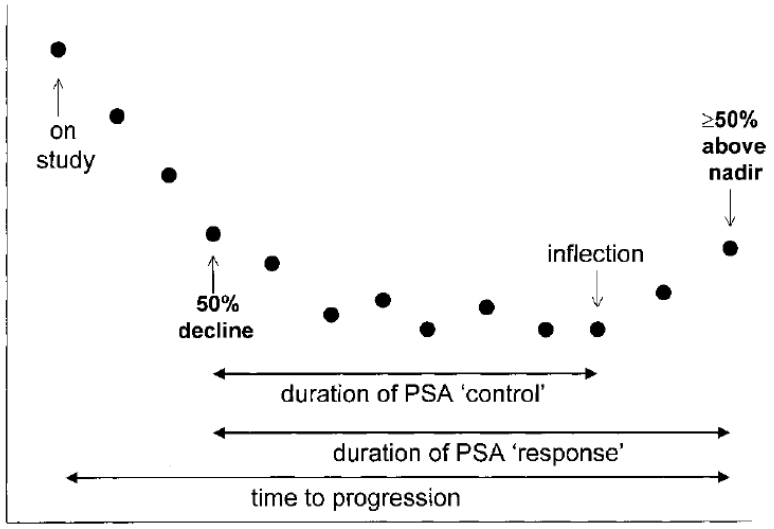
- **PSA-PFS according to PCWG3 criteria**

PSA-PFS is defined as the time from the date of randomization to the date of the first PSA progression or death from any cause during study, whichever occurs first.
PSA progression will be defined as per PCWG3 guidance:

- if subject presented first a decline from baseline, progression is defined as the first PSA increase that is ≥ 25% and ≥ 2 ng/mL above the nadir, and which is confirmed by a consecutive second value ≥ 3 weeks later that fulfils same criteria (ie, a confirmed rising trend);
- if subject did not present a decline from baseline, progression is defined as the first PSA increase that is ≥ 25% and ≥ 2 ng/mL increase from baseline beyond 12 weeks

The nadir is the lowest PSA value presented by the subject up to the point of PSA measurement, as shown by Figure 2 below.

Figure 2: Definitions of PSA control, PSA response, time to PSA progression ([4]).



Subjects without PSA progression or death at time of analysis will be censored at the date of their last PSA assessment.
Same logic as PFS table presented in Section 3.1 should be used for analysis of PSA-PFS. PSA progression then replaces radiographic progression. If date of PSA assessment is missing (but value is available), the date of corresponding visit Will be used instead.

- **Time to PSA progression**

Time to PSA progression is defined as the time from the date of randomization to the PSA progression as defined by PCWG3 (see previous criteria).

Subjects without PSA progression at time of analysis will be censored at the date of their last PSA assessment.
Same logic as TTP table should be used for analysis of Time to PSA progression.
PSA progression then replaces radiographic progression. If date of PSA assessment is missing (but value is available), the date of corresponding visit Will be used instead.

- **Time to first skeletal-related event**

Time to first skeletal-related event is defined as the time from randomization to the occurrence of the first skeletal-related event (i.e., radiation therapy or surgery to bone, pathological bone fracture, spinal cord compression).

If not presenting event, patient will be censored at the [date of last intake + 30 days] or date of death, whichever comes first.

Skeletal-related events will be available from the following eCRF forms:

- Adverse events (AEs) ticked as "Classed as a skeletal event". Bone surgeries are expected to be reported there, along with pathological bone fracture, spinal cord compression. Date of occurrence of skeletal-related event will be computed as the AE start date.
- Concomitant procedures:
    - Bone radiotherapies are expected to be reported there, with indication ticked as "Treatment of Bone Pain".
    - Date of occurrence of skeletal-related event will be computed as the "concomitant procedure/therapy" start date.

3.3 Health Related Quality of Life (HRQoL) Variables

Several questionnaires will be employed to assess Health Related Quality of Life (HRQoL).

- Functional Assessment of Cancer Therapy-General (FACT-G) Version 4. It consists of 27 items which assess patient function in four domains: Physical, Social/Family, Emotional, and Functional well-being [5]. Each item is rated on a 0 to 4 Likert type scale. Negatively stated items are reversed by subtracting the response from "4". After reversing proper items, all subscale items are combined to produce subscale scores, as well as a global QoL score (range 0-108). Scoring instructions are shown in Appendix 2. Higher scores represent better QoL.

- FACT-Prostate (FACT-P) Version 4, a validated and frequently used questionnaire in metastatic prostate cancer trials which has led to HRQoL-based approvals by the European Medicines Agency. It comprises the same 4 domains as FACT-G questionnaire that measures general HRQoL in cancer patients, and a 12-items prostate cancer subscale (PCS) to assess for prostate related symptoms. The PCS is designed specifically to measure prostate cancer-specific quality of life. Each item is rated on a 0 to 4 Likert type scale, and then combined to produce subscale scores for each domain, as well as a global QoL score. Higher scores represent better QoL. Scoring instructions are shown in Appendix 2. From the FACT-P questionnaires, following score will be defined:
  - PCS (12 items, score range 0-48)
  - The FACT–P total score, computed as the sum of the FACT–G and the PCS (range 0-156).
- FACT-taxane Version 4, for assessing the HRQoL related to toxicities. It comprises the same 4 domains as FACT-G questionnaire, further supplemented by 16 specific items of Taxane subscale. The Taxane subscale combines the previously validated 11-item Neurotoxicity subscale and 5 additional questions assessing symptoms related to arthralgia, myalgia, and skin discoloration [6]. Each item is rated on a 0 to 4 Likert type scale, and then combined to produce subscale scores for each domain (all 16 items of the taxane subscale are negatively stated so should be reverted by subtracting the response from "4", as done for FACT-G physical well-being), as well as a global QoL score. Higher scores represent better QoL. From the FACT-taxane questionnaires, following score will be defined:
  - Taxane-specific domain score (16 items, range 0-64)
  - The FACT–taxane total score, computed as the sum of the FACT–G and the Taxane subscale (range 0-172).
- Treatment Satisfaction according to Treatment Satisfaction Questionnaire for Medication (TSQM) Version 1.4. This questionnaire comprises 14 items across four domains focusing on effectiveness (three items), side effects (five items), convenience (three items), and global satisfaction (three items) of the medication over the previous 2–3 weeks, or since the patient's last use [7]. With the exception of item 4 (presence of side effects; yes or no), all items have five or seven responses, scored from one (least satisfied) to five or seven (most satisfied). Item scores are summed to give four domain scores, which are in turn transformed to a scale of 0–100 (no computed score should be lower or higher than these limits). Item 4 is not included for scoring. If an item score is missing, domain scores may be imputed as defined in the scoring algorithm [8]:

- EFFECTIVENESS: ([(Item 1 + Item 2 + Item 3) − 3] divided by 18) × 100. If one item is missing: ([(Sum of Item 1? + Item 2? + Item 3?)) − 2] divided by (12) × 100
- SIDE EFFECTS: ([Sum of Item 5 to Item 8) − 4] divided by 16) × 100. If one item is missing: ([(Sum of Item 5? to Item 8?)) − 3] divided by 12) × 100
- CONVENIENCE: ([Sum of Item 9 to Item 11) − 3] divided by 18) × 100. If one item is missing: ([(Sum of Item9? to Item11?)) − 2] divided by (12) × 100
- OVERALL SATISFACTION First recode Item14_recode = (Item14 − 1) × 5/6 Then: ([Sum of Item 12 to Item 14) − 3] divided by (12) × 100. If any one Item is missing: ([Sum of Item 12? to Item 14?) − 2] divided by (8) × 100

- EQ-5D-5L, a standardized instrument developed by the EuroQol (EQ) Group as a measure of health-related quality of life. It consists of:
  - a descriptive system. It comprises five dimensions: mobility, self-care, usual activities, pain/discomfort and anxiety/depression. Mobility dimension asks about the person's walking ability. Self-care dimension asks about the ability to wash or dress by oneself, and usual activities dimension measures performance in "work, study, housework, family or leisure activities". In pain/discomfort dimension, it asks how much pain or discomfort they have, and in anxiety/depression dimension, it asks how much anxious or depressed they are. Each dimension has 5 levels: no problems, slight problems, moderate problems, severe problems, and extreme problems. It is scored as a 1-digit number expressing the level selected for that dimension (1 to 5).
  - a visual analogue scale (VAS). The VAS records the patient’s self-rated health on a vertical VAS

  The 5-level EQ-5D version (EQ-5D-5L) will be used, with 5 levels of severity for each of the 5 dimensions.

3.4 Safety Variables

The following secondary safety and tolerability endpoints will be assessed:

- AEs and serious adverse events (SAEs), according to the current National Cancer Institute Common Terminology Criteria for Adverse Events (NCI-CTCAE) v5.0
- Assessments of physical examinations, body weight, vital signs, World Health Organization (WHO) performance status, changes in hematology and biochemistry results, and electrocardiogram (ECG) readings.

4. Pharmacokinetic/Pharmacodynamic variables

Not applicable.

5. Analysis populations

5.1 Screened population

All subjects screened will be included in the Screened population.

5.2 Randomized population

All subjects randomized will be included in the Randomized population.

5.3 Safety population

The Safety Population (SAF): all subjects receiving at least one dose of trial medication in either study arm. The SAF will be used for the evaluation of safety. Safety data will be analyzed according to the treatment actually received.

5.4 Full Analysis Set

The Full Analysis Set (FAS): all subjects who received at least one dose of i.v. docetaxel (Cohort 1) or one full cycle of ModraDoc006/r (Cohort 2) and have at least one post-baseline tumor assessment. To be included in FAS, there is no requirement to have PSA measurement. The FAS will be used for the evaluation of primary and secondary efficacy criteria (otherwise stated) and the HRQoL evaluation.

Efficacy data will be analyzed according to the randomized treatment assignment and additionally according to the protocol version in which subject is randomized.

5.5 Per Protocol population

The Per Protocol (PP) population will include all subjects in the FAS population, with the exclusion of subjects who are not compliant to study treatment or have at least one major protocol deviation that would affect the interpretation of efficacy results.

For consistency between study protocol and terminology used by Covance and International Council on Harmonization (ICH E3 Q&A, June 2012), the SAP considers "major protocol deviation" equivalent to "important protocol deviation".

Section 5.5.1 details the deviations.

5.5.1 Important Protocol Deviations Leading to Exclusion from the PP population Analysis

All potential protocol deviations are defined in the study "Protocol Deviation Management Plan", and classified as "important" or "non-important".

Only important protocol deviations considered to have a major effect on efficacy will lead to exclusion of the subject from the PP population.

At time of writing this SAP and for the purposes of this study, the following criteria have been identified as important protocol deviations that can lead to exclusion from the PP population, as it is considered that the occurrences of any of these criteria might have an important influence on the primary efficacy endpoint. Subjects will be assessed purely by comparison of their eCRF data with the criteria below:

- Failure to complete or comply with inclusion/exclusion criteria. As far as possible, the determination should be made on the basis of actual data values rather than tick boxes on the Inclusion and Exclusion Criteria eCRF pages. In particular, following subjects will be excluded from PP:
    - Inclusion Criteria 2 not met: testosterone > 50 ng/dL or > 0.50 ng/mL or > 1.73 nmol/L, or no evidence of progressive metastatic disease)
    - Inclusion Criteria 3 not met: subjects without evaluable disease (as per RECIST v1.1 and/or PCWG3)
- Incorrect study drug dose, frequency, timing or method of drug delivery:
    - Consideration should be given to excluding subjects from the PP population whose compliance percentage does not fall within [70-120].
    - Will be excluded from PP:
        - Cohort 1 subjects not exposed to treatment
        - Cohort 2 subjects without one full cycle of ModraDoc006/r
        - Cases of misrandomization (allocation of an incorrect treatment compared to the randomization central system)
- Excessive number of Assessments/Visits performed outside the allowed time window

- Non-compliance with study procedures. Subjects without any assessment of response as per RECIST v1.1 (for nodal or visceral lesions) nor as per PCWG3 (for bone metastases), or without any post-baseline PSA assessment will be excluded from PP.
- Failure to receive at least 6 weekly administrations of ModraDoc006/r or 2 standard three-weekly cycles of i.v. docetaxel
- Unauthorized medication: Non-compliance with study restrictions (e.g., use of prohibited concomitant medication or other anti-neoplastic treatment while receiving study therapy, refer to protocol Appendix VII)
- Other, e.g., consent not signed

Regular review of deviations by Covance and the Sponsor, as defined in the Protocol Deviation Management Plan, will identify important protocol deviations that may lead to exclusion of PP population.

Should additional important protocol deviations leading to exclusion from the PP population, not anticipated at the time of preparing this SAP, be identified during the study (or amendment of the above categories), they will be documented in a specific document and included in all relevant protocol deviation reviews and approvals prior to database lock.

All important protocol deviations leading to exclusion from the PP population occurring during the study will be reviewed and approved by the Sponsor prior to database lock.

5.6 Special Subpopulations

5.6.1 Population evaluable for radiological response

To be considered evaluable for radiological response:

- Subjects should have received at least 6 weekly administrations of ModraDoc006/r or 2 standard 3-weekly cycles of i.v. docetaxel
- Subjects should have measurable lesions according to RECIST v1.1. This will be based on question "Does the patient have target lesions" in CRF page "Target Lesion Measurement (Baseline)": Yes.
- Response should be evaluated according to RECIST v1.1 and PCWG3.

6. Data Handling

As patients may still be on trial at the time of data summarization (see study design section 2.2), a cut-off management macro will be used at statistical programming level. Based on raw data, this macro will exclude from analysis all events having its date posterior to the cut-off date.

Other data handling is described below.

6.1 Time points and Visit Windows

Day 1 is defined as the day of first dose of treatment.

Relative days after Day 1 are calculated as (assessment date – Day 1 date) + 1. Relative days prior to Day 1 are calculated as (assessment date – Day 1 date).

The day prior to Day 1 is Day -1.

For QOL questionnaires, due to heterogeneity of the post-baseline visits where data was collected in CRF, following visit windows will be applied in order to pool assessments performed around each visit as scheduled in protocol:

Original visit	**Analysis visit used in summary tables**
CYCLE 03 DAY 01 CYCLE 03 DAY 15 CYCLE 04 DAY 01	End of Cycle 3
CYCLE 06 DAY 01 CYCLE 06 DAY 15 CYCLE 07 DAY 01	End of Cycle 6
CYCLE 10 DAY 01 CYCLE 10 DAY 15 CYCLE 11 DAY 01	End of Cycle 10

Visit windows do not apply to Baseline and End of Treatment visits.

If several QOL assessments are available within the same window, the assessment with the worst score (indicating worst health status) will be kept for analysis, i.e.:

Score	**Definition of Worst value**
FACT-G Physical well-being	Lowest value
FACT-G Social/Family well-being	Lowest value
FACT-G Emotional well-being	Lowest value
FACT-G Functional well-being	Lowest value

FACT-G Total score	Lowest value
FACT-P PCS	Lowest value
FACT-P Total score	Lowest value
FACT-taxane Taxane-specific domain score	Lowest value
FACT-taxane Total score	Lowest value
EQ-5D-5L mobility, self-care, usual activities, pain/discomfort, anxiety/depression	Highest value
EQ-5D-5L VAS	Lowest value
Treatment Satisfaction (Effectivness, Side Effects, Convenience, Overall Satisfaction)	Lowest value

Visit windows will be applied on subscale scores after their derivation. It will apply only for summary tables. Listings will show the original visits only, i.e. as collected in CRF.

Except for QOL, all data will be analyzed using nominal study visits as defined in the Study Schedule and eCRF. No visit windows will be applied for summary and analysis.

6.2 Handling of Dropouts, Missing Data, and Outliers

For FACT questionnaires, missing data will be handled as following [5]:

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, using the formula below:

Prorated subscale score = [Sum of item scores] x [N of items in subscale] / [N of items answered]

Of note, prorating by subscale in this way is acceptable as long as more than 50% of the items are answered (e.g., a minimum of 4 of 7 items, 4 of 6 items, etc). If less than 50% of the items are answered for a given subscale, the subscale score remains empty.

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.

Otherwise, no specific handling of missing data is planned.

For computing the time to HRQoL deterioration, if date of QOL assessment is missing, corresponding visit date will be used instead.

Incomplete dates:

In all listings, incomplete dates should be left as they have been recorded. However, for calculation / sorting / assignation based on dates, the following methods will be used:

- As general rule:
  - Incomplete start dates will be imputed as first day of the month/year
  - Incomplete end dates will be imputed as last day of the month/year
- The most conservative approach will be systematically considered (i.e., if the onset date of an AE/concomitant medication is missing / incomplete, it is assumed to have occurred during the study treatment phase (i.e., a TEAE for AEs) except if the partial onset date or other data [stop date, …] indicates differently).
- A missing/incomplete date of medical history or disease diagnosis will be assumed to have occurred before any study treatment.
- If a partial date and the associated information do not allow a statement about the assignment to a group / category, all the possible groups / categories will be considered (i.e.: an AE could be assigned to several possible doses at event onset according to its partial onset date and stop date. Particularly an AE with missing start date will be assigned to each dose received before its end date.).

Where possible, the derivations based on a partial date will be presented as superior inequalities (i.e., for an AE started in NOV2019 after the administration performed on 31OCT2019, the days since last administration will be "≥2". Similarly the duration of ongoing AEs or medication will be"≥xx" according to the start and last visit dates).

No rules for outlier detection are planned.

7. Statistical Methods

7.1 General Principles

All data collected in this trial will be documented using summary tables, figures, and subject data listings.

Listings will present data from all collected visits/timepoints included unscheduled or unplanned visits/timepoints, otherwise stated in the listing shell. Listings will be using Screened population, and randomization arm as treatment group per protocol version.

Table and figures will display randomization arm, both per protocol version and for protocol versions combined, i.e. will be repeated showing treatment arms splitted by protocol version.

Tables and figures will present only scheduled visits and their scheduled timepoints per protocol (schedule of assessment), including scheduled visits performed after Cycle 10 if any.

Differences between the cohorts will be compared using parametric techniques for continuous variables (only in case of severe deviations of normality will the Wilcoxon-Mann Whitney test be used) and Fisher's exact test for categorical variables.

Time to event data (PFS, TTP, DOR, PSA-PFS, time to PSA progression, time to HRQoL deterioration) will be summarized using Kaplan-Meier estimates and differences will be determined via the log-rank test. Univariate Cox model will provide estimates of hazard ratios (HR).

Log-rank test will be performed under the hypothesis of proportional hazards. This hypothesis will be checked through the following means:

- Visually by means of plots of log(-log(survival)) versus log of survival time;
- And/or by the inclusion of time-dependent covariates in the Cox model.

If the performed checks do not suggest the validity of the proportional hazards assumption, the Wilcoxon test will be used instead of the Log-rank test. In this case Cox models will not be performed and HRs will not be provided.

All tests will be two-sided and considered significant at $P<0.05$.

All data processing, summarization and analyses will be performed using Covance's SAS Environment / Version 9.4 (or later) of the SAS® statistical software package.

The SAS codes are listed in Appendix 3.

The following principles will be applied to all TFLs unless otherwise stated:

Principle	Value
Significant tests	Two-sided and use a 5% significance level for main effects and 10% significance level for interaction terms.
Treatment group labels and order presented	Tables and Figures by randomization arm: Docetaxel i.v. ModraDoc006/r Tables and Figures repeated by Protocol Version: Docetaxel i.v. PV2.0 ModraDoc006/r PV2.0 Docetaxel i.v. PV3.0 ModraDoc006/r PV3.0 Listings: Docetaxel i.v. PV2.0 ModraDoc006/r PV2.0 Docetaxel i.v. PV3.0 ModraDoc006/r PV3.0 Screening Failures
Tables	Data in summary tables presented by treatment group (see above), assessment and visit (where applicable).
Listings	All data collected presented by Treatment group (see above), country, site, subject, visit, date/time of assessment and visit (where applicable), unless otherwise specified.
Descriptive summary statistics for continuous variables	Number of subjects/observations (N), mean, standard deviation (SD), first and third quartiles (Q1;Q3), median, minimum, and maximum.
Descriptive summary statistics for categorical variables	Frequency counts and percentages [n (%)]
Denominator for percentages	Number of subjects in the analysis population, unless stated otherwise in table shell(s)
Include "Missing" as category	Demographics and Other Baseline Characteristics only, when the number missing is greater than zero for at least one treatment group.
Display for 0 percentages	0%
Display to one more decimal place than collected value	Mean, SD, Median, Q1, Q3
Limit of precision for displays	3 decimal places
Date Format	DDMMMYYYY

7.2 Subject Disposition and Data Sets Analyzed

Subject disposition will be summarized by treatment group, both per protocol version and for protocol versions combined, and overall and will include the number and percentage of subjects for the screened population:

- screened;
- screened and not randomized;
- randomized;
- randomized and not treated;
- included in each study population (Safety, FAS, PP, evaluable for radiological response).

In addition, the number and percentage of subjects who are still ongoing, who complete the study and who discontinue early, including a breakdown of the primary reasons for discontinuation, will be presented for the randomized population.

A corresponding listing will be provided with appropriate level of detail.

A subject will be considered a completer if:

- he discontinued the study due to radiographic disease progression (according to RECIST v1.1 and/or PCWG3)
- he was still under study treatment when the Sponsor terminated study as planned by protocol.

A summary of subject enrollment by country and site will also be provided by treatment group and overall for the randomized population.

7.3 Protocol Deviations

All protocol deviations will be listed and summarized by treatment group for the FAS population. It will allow to identify events related to Coronavirus Disease of 2019 (COVID-19) and the type of COVID-19 disruption (e.g. missed visits).

Important protocol deviations leading to exclusion from the PP population (see Section 5.5.1) will be listed and summarized by treatment group for the FAS population.

7.4 Demographics and Other Baseline Characteristics

Demographic and baseline characteristics will be listed and summarized by treatment group, both per protocol version and for protocol versions combined, and overall for FAS population.

As a general rule, the baseline value of study assessment (e.g. weight) will be defined as last available value collected prior to the first dose of treatment.

Standard descriptive statistics will be presented for the continuous variables of:

- Age (years) as captured in the eCRF
- Weight (kg)
- Height (cm)
- Body mass index (kg/m^2) as computed in the eCRF
- BSA (m^2)=[height (cm) x weight (kg) / 3600]^0.5
- Time since diagnosis of prostate cancer (PC) in months, computed from Baseline visit
- Time since diagnosis of mCRPC in months, computed from Baseline visit
- PSA level in ng/mL
- Serum testosterone level in ng/dL. If needed, the following conversion factor will be used: 50 ng/dL = 0.50 ng/mL = 1.73 nmol/L

The total counts and percentages of subjects will be presented for the categorical variables of:

- Age group (< vs ≥ 65 years)
- Country
- Race and ethnicity
- Eastern Cooperative Oncology Group (ECOG)
- Disease Stage at diagnosis with PC (TNM staging)
- Type of progression at study entry: PSA only, radiological only, both
- Measurable disease at study entry, based on question "Does the patient have target lesions" in CRF page "Target Lesion Measurement (Baseline)": Yes, No.
- Presence of bone metastases at baseline, assessed with bone scintigraphy (if available), otherwise as assessed at diagnosis with mCRPC.
- Number of bone metastases at baseline assessed with bone scintigraphy, if available (0, 1, 2-4, 5-9, 10-20, and > 20)

- Disease site at baseline:
    - Lymph nodes only, when lymph node(s) reported as target or non-target lesion at baseline, and does not fulfill next condition:
    - Bone/Bone+lymph nodes (no visceral involvement), when presence of bone metastases at baseline is derived as YES (see definition above), and does not fulfill next condition:
    - Any visceral (regardless bone or lymph nodes involvement), when at least one visceral lesion reported as target or non-target lesion at baseline
- Prior therapies for PC, by type of therapy: surgery, radiotherapy, hormonal therapy, enzalutamide, abiraterone, regular androgen deprivation therapy (ADT).

All demographics and baseline characteristics will be presented in listings, by subject.

Other baseline measurements, such as vital signs, hematology, biochemistry, urinalysis, physical examination, ECG, quality of life, will be summarized by treatment group and listed with the post-baseline measurements.

7.4.1 Medical History

Medical history will be coded by Covance using the Medical Dictionary for Regulatory Activities (MedDRA) Version 22.0 (or a later version if updated during the study). All medical history will be listed, and the number and percentage of subjects with any medical history will be summarized for FAS population by system organ class (SOC) and preferred term (PT) for each treatment group and overall.

7.4.2 Previous and Concomitant Medications

Medications received prior to or concomitantly with treatment will be coded by Covance using the WHODrug Dictionary Version March2019-Global B3 (or a later version if updated during the study) and Anatomical Therapeutic Chemical (ATC) Classification codes.

Previous medications and concomitant medications are defined as follows:

- Previous medications are those with a stop date prior to the first dose date of treatment.
- Concomitant medications are those with a start date on or after the first dose date of treatment, or those with a start date before the first dose date of treatment and a stop date on or after the first dose date of treatment or ongoing end of study.

If a medication cannot be classified as "previous" or "concomitant" after applying imputation rules for missing/incomplete dates (see Section 6.2), it will be classified as concomitant.

Previous medications and concomitant medications will be listed together using therapeutic class (ATC-Level 2), chemical subgroup (ATC-Level 4) and generic term, and sorted by medication start date.

Concomitant medications will be summarized separately, for FAS population. The number and percentage of subjects using each medication will be displayed together with the number and percentage of subjects using at least one medication within each therapeutic class (ATC-Level 2) and chemical subgroup (ATC-Level 4).

7.5 Measurements of Treatment Compliance

For each subject, compliance will be computed for overall study period.

For subjects treated with ModraDoc006/R, compliance of ModraDoc006 and ritonavir will be computed separately, based on tablet counts.

Percentage compliance is calculated as: 100 * actual tablets taken/expected tablets taken.

Where:

- Actual taken is defined as number of tablets dispensed minus number of tablets returned.
- Expected taken is the sum of theoretical tablets count over all scheduled administrations between C1D1 and last scheduled administration of last cycle (inclusive), considering weekly administrations at D1, D8, D15 of each 21-days cycle. For subjects who have dose modifications for safety reason as defined in protocol (e.g., dose reduction or dose withholding), the expected number of capsules will be calculated according to the dose modification.

	Dosage at a scheduled visit (mg)	Theoretical tablets count
ModraDoc006 under protocol V2.0 *	50	5
	Reduced: 40	4
	Reduced: 30	3
ModraDoc006 under protocol V3.0*	40	4
	Reduced: 30	3
Ritonavir	300	3

* Per study protocol, patients enrolled under protocol amendment 1 (Protocol V2.0) who are dosed with 30 mg of ModraDoc006 combined with 200 mg ritonavir in the morning and 20 mg of ModraDoc006 combined with 100 mg ritonavir in the evening, will, at the investigator's discretion, be able to continue at this dose. For computation of the compliance, if investigator chooses to reduce the dosage to comply with amended protocol at a given visit/timepoint, theoretical tablets count is calculated according to the dose modification.

For subjects treated with docetaxel i.v., percentage compliance is calculated as:

100 * number of injections performed/number of injections scheduled,

where number of injections scheduled is defined as the number of 21-days theoretical cycles between C1D1 (included) and EOT visit.

Percentage compliance will be summarized descriptively by treatment group for the Safety population.

The number and percentage of compliant subjects will be presented, where compliant is defined as percentage compliance between 70.0% and 120.0%, inclusive. The following percentage compliance categories will also be presented:

- <70.0%
- >120.0%

7.6 Efficacy

7.6.1 Primary Efficacy Analysis

rPFS will be summarized by treatment group, both overall and by protocol version, using Kaplan-Meier estimates (median and rates at 6 and 12 months, with associated 95% CI) and illustrated graphically using a Kaplan-Meier plot. Under the condition of proportional hazards, differences between the cohorts will be determined using a non-stratified log-rank test, and univariate Cox models will provide estimation of HR with 95% CI.

Analysis of rPFS will be performed on FAS and repeated on PP population as sensitivity analysis (see Section 7.6.3).

7.6.2 Secondary Efficacy Analysis

All secondary efficacy endpoints will be analyzed on the FAS, both overall and specified per protocol version, unless otherwise stated.

No adjustments for multiplicity will be made as this is a phase IIb study, conducted to gather preliminary efficacy and safety information on the new oral treatment.

- ORR based on RECIST v1.1

ORR will be summarized by treatment group, with 95% exact Clopper-Pearson CI for binomial proportions.

Population of analysis will be the subjects evaluable for radiological response as defined in Section 5.5.1.

The comparison between the cohorts will be performed using Fisher's exact test.

In addition, the following analyses will be presented based on response assessment with RECIST v1.1, on the subjects evaluable for radiological response:

- Response to therapy according to RECIST v1.1 will be presented as the numbers and proportions of the subjects evaluable for radiological response who had a progressive disease, stable disease, partial radiological response or complete radiological response as BOR.
- Number and proportion of subjects per cohort who have presented response (CR or PR) at fixed time intervals (i.e., by assessment point) and separately at the "End of Trial" visit
- Tumor sizes will be summarized by assessment point and for changes during the study: sum of diameters for target lesions will be described by assessment point, in raw value and in change from baseline.
- DCR

DCR will be summarized by treatment group, with 95% CI. The comparison between the cohorts will be performed using Fisher's exact test.

- DOR based on RECIST v1.1

DOR will be summarized by treatment group using Kaplan-Meier estimates (median and rates at 6 and 12 months, with associated 95% CI) and illustrated graphically using a Kaplan-Meier plot, among the population evaluable for radiological response. Under the condition of proportional hazards, differences between the cohorts will be determined using a non-stratified log-rank test, and univariate Cox models will provide estimation of HR with 95% CI.

- rPFS at 6 months based on RECIST v1.1

See section 7.6.1.

- TTP based on RECIST v1.1, incorporating the bone metastatic variable according to PCWG3

TTP will be summarized by treatment group using Kaplan-Meier estimates (median and rates at 6 and 12 months, with associated 95% CI) and illustrated graphically using a Kaplan-Meier plot. Under the condition of proportional hazards, differences between the cohorts will be determined using a non-stratified log-rank test, and univariate Cox models will provide estimation of HR with 95% CI.

- PSA RR according to PCWG3 criteria

PSA RR will be summarized by treatment group, with 95% CI.

The comparison between the cohorts will be performed using Fisher's exact test.

- PSA-PFS according to PCWG3 criteria

PSA-PFS will be summarized by treatment group using Kaplan-Meier estimates (median and rates at 6 and 12 months, with associated 95% CI) and illustrated graphically using a Kaplan-Meier plot. Under the condition of proportional hazards, differences between the cohorts will be determined using a non-stratified log-rank test, and univariate Cox models will provide estimation of HR with 95% CI.

- Time to PSA progression

Time to PSA progression will be summarized by treatment group using Kaplan-Meier estimates (median and rates at 6 and 12 months, with associated 95% CI) and illustrated graphically using a Kaplan-Meier plot. Under the condition of proportional hazards, differences between the cohorts will be determined using a non-stratified log-rank test, and univariate Cox models will provide estimation of HR with 95% CI.

- Time to first skeletal-related event

Time to first skeletal-related event will be summarized by treatment group using Kaplan-Meier estimates (median and rates at 6 and 12 months, with associated 95% CI) and illustrated graphically using a Kaplan-Meier plot. Under the condition of proportional hazards, differences between the cohorts will be determined using a non-stratified log-rank test, and univariate Cox models will provide estimation of HR with 95% CI.

- PSA values

PSA levels will be summarized by visit and by treatment group, along with percentage change from baseline.

Lowest PSA value during study will be identified as the best response. Absolute change in PSA from baseline to best response will be also summarized by treatment group.

7.6.3 Sensitivity Analysis

Primary efficacy analysis (rPFS) will be repeated on PP population as sensitivity analysis.

As described in 7.6.1, median rPFS will be provided with 95%CI, and cohorts will be compared using log-rank test.

In addition, time to treatment failure (TTF) will be analyzed on Safety Population, using treatment duration as defined in section 7.7.1. Subjects who haven't discontinued the treatment at time of analysis will be censored. Method for time-to-event data will be applied: median TTF will be provided with 95%CI, and cohorts will be compared using log-rank test. Related figure will be provided.

Besides, the analysis of ORR will be repeated while excluding the subjects not belonging to the PP population, still among patients evaluable for radiological response.

7.6.4 Subgroup Analysis

Not applicable.

7.6.5 Exploratory Analysis

For PSA-PFS, PFS and time to PSA progression, multivariate Cox proportional hazards models may be constructed to evaluate the effect of confounding variables on those time-to-event variables. Following prognostic factors will be tested ([9]):

- Age (< 65yrs versus ≥ 65yrs);
- Time since diagnostic of PC (in months)
- ECOG at baseline (0, >0);
- Type of progression at study entry: PSA only, radiological only, both
- Disease site at baseline as defined in Section 7.4
- PSA level at baseline (< vs ≥ median)
- Hemoglobin level at baseline (< vs ≥ median)
- Lactate dehydrogenase (LD) level at baseline (< vs ≥ median)
- Alkaline phosphatase at baseline (< vs ≥ median)
- Albumin level at baseline (< vs ≥ median)
- Any other covariate as discussed and approved by the Sponsor.

These prognostic factors will be first analysed in separate univariate analyses and then in multivariate analysis. More specifically, each potential predictor will be tested in a Cox model where the considered predictor will be the only covariate included (treatment group will not be included in the model).

Potential predictors with a p-value below 0.10 will be selected for the multivariate analysis. The selected predictors will be then all included in a multivariate Cox model (treatment will not be included in the model) and further selected with a backward selection procedure eliminating covariates showing a p-value above 0.10 in presence of the other covariates. Treatment group will be finally introduced as an additional covariate in the reduced model obtained at the end of the backward selection procedure. The covariate by treatment interactions will be tested by adding all interactions corresponding to the finally retained covariates in a separate model.

7.7 Safety

7.7.1 Extent of Exposure

Duration of exposure will be defined in months as:

(last dose date – first dose date + 1) / (365.25/12).

For subjects in Cohort 1, dates of docetaxel injections will be considered.

For subjects in Cohort 2, dates of ModraDoc006 administrations will be considered.

If date of first dose date is missing then the date of first dose dispensed will be used. If last dose date is missing then date of last known administration of study drug will be used.

Duration of exposure will be listed and summarized using descriptive statistics for each treatment group for the Safety population.

The number and percentage of subjects with duration of exposure in the following categories will be summarized for the Safety population:

-]0-3] months
-]3-6] months
-]6-12] months
- >12 months

Also, number of complete cycles will be described:

- For Cohort 1, number of cycles = number of injections
- For Cohort 2, number of cycles will be estimated as the number of days with at least one ModraDoc006/r intake, divided by 3 (as ModraDoc006/r administration is planned at D1, D8 and D15 of each cycle)

Data listings for dose administration and dose modifications for safety reasons (including dose reduction and dose withholding) will be provided.

7.7.2 Adverse Events

All AEs recorded on the eCRF will be coded using the MedDRA dictionary Version 22.0 (or a later version if updated during the study) and classified as either baseline signs and symptoms or treatment – emergent AEs (TEAEs) as follows:

- Baseline signs and symptoms are events that start prior to the date of first dose of treatment.
- TEAEs are events with start date and time on or after the date and time of first dose of treatment and up to 30 days after date of last dose of treatment, or events with start date and time prior to the date and time of first dose of treatment whose severity worsens on or after the date and time of first dose of treatment.

All AE data will be listed by treatment group. Treatment-emergence status will be flagged in the listing. In addition, corresponding listings of SAEs, AEs leading to discontinuation of treatment, and AEs resulting in death will be produced.

Summary tables of TEAEs by treatment group will be produced for the Safety population.

Assessment of AE severity will be based on the NCI-CTCAE, version 5.0. The relationship between an AE and treatment is assessed as definite, probable, possible, or not related. A treatment-related AE is an AE considered by the investigator as definitely, possibly, or probably related to treatment or with unknown/missing relationship to treatment.

An overview table will summarize the number and percentage of subjects with at least one of the following TEAEs, where subjects with more than one TEAE in a particular category are counted only once in that category:

- any AE;
- any AE by severity (NCI-CTCAE grade);
- treatment-related AE;
- treatment-related AE by severity (NCI-CTCAE grade);
- AE leading to treatment discontinuation;
- SAE;
- treatment-related SAE;
- SAE leading to death;
- treatment-related SAE leading to death;
- SAE leading to treatment discontinuation.

Overview table will be repeated for subset of patients being below and above median BSA, using the median BSA of the Safety population at baseline.

A second overview table will also summarize the number of the following TEAEs:

- all AEs;
- AEs by severity (NCI-CTCAE grade);
- treatment-related AEs;
- treatment-related AEs by severity (NCI-CTCAE grade);
- SAEs;
- SAEs by severity (NCI-CTCAE grade);
- treatment-related SAEs.
- SAE leading to death;
- AE leading to treatment discontinuation;

Specifically, the incidence of following TEAEs will be summarized and tested between treatment groups using Fisher's exact test:

- grade 3-4 neutropenia (unique PT).
- febrile neutropenia (unique PT)
- toxicity-related hospital admissions (hospital admissions will be defined as an SAE with hospitalization ticked)
- grade 3-4 diarrhea (unique PT)
- grade 3-4 fatigue. Fatigue will include following PTs: Fatigue, asthenia, lethargy, malaise.
- grade 3-4 neuropathy. Acute polyneuropathy, neuropathy peripheral, paresthesia, peripheral sensory neuropathy, polyneuropathy, peripheral motor neuropathy, dysgeusia
- grade 3-4 treatment-related allergic reactions, that include following PTs: flushing, skin reactions, itching, chest tightness; difficulty in breathing, fever or chills, back pain, low blood pressure

List of terms specified above may be further amended upon recommendation from Sponsor or medical expert.

Otherwise, no statistical comparisons of AEs between treatment groups will be performed.

The number and percentage of subjects reporting each AE will be summarized by SOC and PT for the Safety population. Tables will be sorted alphabetically by SOC. PTs will be sorted by descending overall total. The following summaries will be produced:

- AEs, by SOC and PT;
- AEs by PT;
- AEs related to treatment, by SOC and PT;
- AEs related to treatment, by PT;
- AEs by maximum severity (NCI-CTCAE grade), by SOC and PT;
- AEs related to treatment by maximum severity (NCI-CTCAE grade), by SOC and PT;
- SAEs, by SOC and PT;
- SAEs related to treatment, by SOC and PT;
- SAEs by maximum severity (NCI-CTCAE grade), by SOC and PT;
- SAEs related to treatment by maximum severity (NCI-CTCAE grade), by SOC and PT;

In the above summaries, subjects with more than one AE within a particular SOC are counted only once for that SOC. Similarly, subjects with more than one AE within a particular PT are counted only once for that PT.

For summaries by maximum severity, subjects with multiple AEs within a particular SOC or PT will be counted under the category of their most severe AE within that SOC or PT. AEs with missing intensity/severity will be included (as Grade 4) in the overall count of subjects with AEs, but will not be included in the counts of subjects with AEs within a SOC or PT.

7.7.3 Laboratory Evaluations

Data for the following hematology, blood chemistry, and urinalysis analytes recorded in the eCRF will be listed and summarized by treatment group and visit for the Safety population. If data for any additional analytes are also recorded then these will be listed only.

Hematology	Biochemistry	Urinalysis (dipstick)
Hemoglobin	Sodium (Na)	Protein
Hematocrit	Potassium (K)	Glucose
Red blood cell count	Calcium (Ca)	Blood
White blood cell count	Magnesium (Mg)	Ketones
Neutrophils count	Glucose	pH
Lymphocyte count	Urea	
Monocyte count	Blood Urea Nitrogen (BUN)*	
Basophil count	Creatinine	
Eosinophil count	Bilirubin	
Platelet count	Alkaline phosphatase	
	Gamma-glutamyl transferase (GGT)	
	Aspartate aminotransferase (ASAT)	
	Alanine aminotransferase (ALAT)	
	Lactate dehydrogenase (LD)	
	Total protein	
	Albumin	
	Creatinine clearance using the Cockcroft-Gault formula or Modification of Diet in Renal Disease (MDRD).	

* Data recorded for BUN was allowed to be collected after first patient was included.

All laboratory data will be reported in International System of Units (SI) units. Out-of-reference-range values will be flagged as high (H) or low (L) in the listings.

For analysis purposes, values preceded by a "<" or a ">" sign (i.e., those below or above the limits of quantification) will be considered equal to the lower or upper limit of quantification, respectively.

Laboratory data will be summarized by visit using standard descriptive statistics for the Safety population.

For hematology and biochemistry, changes from baseline will also be summarized.

For the quantitative description by visit (row value and change from baseline), BUN results will be summarized as any other parameter. BUN will also be listed as part of the other collected parameters.

For each laboratory analyte, the baseline value will be defined as last scheduled or unscheduled value collected prior to the first dose of treatment. For post-baseline, only data from scheduled visits will be included in the summary tables.

For hematology and biochemistry, shift tables presenting movement in and out of reference range from baseline to each scheduled post-baseline visit will be provided for each treatment group. For this analysis, BUN and urea will be analyzed as a combined parameter labelled "Urea/BUN": per patient and per analysis, one value will be used (urea if available, otherwise BUN) to define abnormality criteria for the combined parameter (low/normal/high).

7.7.4 Vital Signs

The following vital signs will be listed and summarized by treatment group and visit for the Safety population.

- weight (kg)
- systolic and diastolic blood pressure (mmHg)
- body temperature (°C). Temperatures collected in °F will be converted to °C using following formula: (°F – 32) x (5/9) = °C.
- heart rate (bpm)

Vital signs data and changes from baseline in vital signs will be summarized by visit using standard descriptive statistics for the Safety population.

The baseline value will be defined as last scheduled or unscheduled value collected prior to the first dose of treatment. For post-baseline, only data from scheduled visits will be included in the summary tables.

7.7.5 Electrocardiograms

The following quantitative ECG measurements will be taken during the study:

- heart rate (bpm);
- PR interval (ms)
- QRS interval (ms)
- QT interval (ms)
- Bazett corrected QT (QTcB) interval (ms)
- Fridericia corrected QT (QTcF) interval (ms)
- RR interval (ms)

An overall Investigator ECG interpretation will be provided (categories "normal", "abnormal, not clinically significant" and "abnormal, clinically significant"), with specification if clinically significant.

The ECG measurements and changes from baseline in ECG will be listed and summarized by treatment group and visit using standard descriptive statistics for the Safety population.

The Investigator assessment will be listed and the number and percentage of subjects within each assessment category will be tabulated by treatment group and visit for the Safety population.

The baseline value will be defined as last scheduled or unscheduled value collected prior to the first dose of treatment. For post-baseline, only data from scheduled visits will be included in the summary tables.

7.7.6 Physical Examination

Physical examination results (normal/abnormal) and details of abnormalities will be listed for each subject and visits.

7.7.7 Other Safety Variables

WHO performance status (ECOG) will be summarized by treatment group and visit for the Safety population, presenting the number and percentage of subjects in each category.

Listing of all ECOG assessments data will be provided.

7.8 HRQol

HRQol analysis will be performed on the FAS population.

7.8.1 FACT scales

HRQoL scores and absolute changes from baseline will be summarized by visit and treatment group using standard descriptive statistics for the FAS population:

- FACT-G:
  - Physical well-being (7 items; score range 0–28)
  - Social/Family well-being (7 items; score range 0–28)
  - Emotional well-being (6 items; score range 0–24)
  - Functional well-being (7 items; score range 0–28)
  - Total score (score range 0-108), computed only when at least 22 of 27 FACT-G items completed
- FACT-P:
  - PCS (12 items, score range 0-48)
  - Total score (FACT-G total score + PCS, score range 0-156)
- FACT-taxane:
  - Taxane-specific domain score (16 items, 0 to 64)
  - Total score (0 to 172)

For post-baseline visits, only data from scheduled visits will be included in the summary tables.

Visit windows as defined in section 6.1 will be applied for analysis table of FACT (9 scores defined above). Listing will show original visit only.Also, the following categorical variables will be computed and summarized by treatment group:

- Overall HRQoL improvement, at least once during the study: Response is defined as a 10-point or greater increase in the FACT-G total score at a post-baseline assessment compared with baseline.
- Improvements in individual HRQoL domains:
  - ≥3 for FACT-G physical well-being, at least once during the study
  - ≥3 for FACT-G social or family well-being, at least once during the study
  - ≥3 for FACT-G emotional well-being, at least once during the study
  - ≥3 for FACT-G functional well-being, at least once during the study
  - ≥3 for PCS, at least once during the study
  - ≥3 for FACT-Taxane specific items, at least once during the study

7.8.2 Time to HRQoL deterioration

The time to HRQoL deterioration will be computed as the time from date of randomization to the date of first HRQoL deterioration, defined as:

- a ≥10-point decrease in the FACT-G total score at a post-baseline assessment compared with baseline
- or death from any cause, whichever occurs first, regardless date of death.

Subjects without events as defined above will be censored at the date of their last post-baseline assessment of FACT-G total score. In the absence of post-baseline assessment, they will be censored at the date of randomization.

If date of QOL assessment is missing, visit date will be used instead.
Time to HRQoL deterioration will be summarized by treatment group using Kaplan-Meier estimates (median and rates at 6 and 12 months, with associated 95% CI) and illustrated graphically using a Kaplan-Meier plot, for the FAS population. Differences between the cohorts will be determined using a non-stratified log-rank test. Under the condition of proportional hazards, Univariate Cox models will be ran and estimation of HR will be provided with 95% CI.

7.8.3 Overall Health Related Utility

The overall Health Related Utility will be assessed by the EQ-5D-5L.

A frequency distribution of level response for each dimension (mobility, self-care, usual activities, pain/discomfort, and anxiety/depression) and a summary of the visual analogue scale will be described by visit and treatment group.
Absolute change from baseline will be summarized as well.For post-baseline visits, only data from scheduled visits will be included in the summary tables.

Visit windows as defined in section 6.1 will be applied for analysis table of EQ-5D-5L. Listing will show original visit only.

7.8.4 Treatment Satisfaction

Each of the 4 domain scores computed from the TSQM will be summarized by visit and treatment group using standard descriptive statistics for the FAS population.

Only data from scheduled visits will be included in the summary tables. Visit windows as defined in section 6.1 will be applied for analysis table of EQ-5D-5L. Listing will show original visit only.

Of note, study protocol V3.0 does not require the collection at baseline visit. Therefore, any assessment collected at baseline visit will not be presented in the table by visit.

7.9 Interim Analysis

No formal interim analysis will be performed for this study.

8. Changes in Planned Analysis

The interim analysis described in protocol section 8.5 will not be performed.

9. Data Issues

Not applicable.

10. References

[1] Scher et al. Trial Design and Objectives for Castration-Resistant Prostate Cancer: Updated Recommendations From the Prostate Cancer Clinical Trials Working Group 3. J Clin Oncol. 2016; 34(12):1402-18.

[2] Eisenhauer EA, Therasse P, Bogaerts J, et al. New response evaluation criteria in solid tumours: revised RECIST guideline (version 1.1). European journal of cancer (Oxford, England: 1990). 2009;45(2):228-247.

[3] US Food and Drug Administration, Guidance for Industry: Clinical Trial Endpoints for the Approval of Cancer Drugs and Biologics, December 2018.

[4] Bubley GJ, Carducci M, Dahut W, et al. Eligibility and response guidelines for phase II clinical trials in androgen-independent prostate cancer: recommendations from the Prostate-Specific Antigen Working Group. Journal of clinical oncology: official journal of the American Society of Clinical Oncology. 1999;17(11):3461-3467.

[5] Webster K, Cella D, Yost K. The Functional Assessment of Chronic Illness Therapy (FACIT) measurement system: properties, applications, and interpretation. Health Qual Life Outcomes. 2003; 1:79.

[6] Cella D, Peterman A, Hudgens S, Webster K, Socinski MA. Measuring the side effects of taxane therapy in oncology: the Functional Assesment of Cancer therapy-Taxane (FACT-taxane). Cancer. 2003 Aug 15;98(4):822-31.

[7] Atkinson MJ, Sinha A, Hass SL, et al. Validation of a general measure of treatment satisfaction, the Treatment Satisfaction Questionnaire for Medication (TSQM), using a national panel study of chronic disease. *Health Qual Life Outcomes.* 2004;2:12.

[8] Atkinson MJ, Kumar R, Cappelleri JC, et al. Hierarchical construct validity of the treatment satisfaction questionnaire for medication (TSQM version II) among outpatient pharmacy consumers. Value Health, 2005; 8(S1):S9-S24.

[9] Halabi et al. Updated Prognostic Model for Predicting Overall Survival in First-Line Chemotherapy for Patients With Metastatic Castration-Resistant Prostate Cancer. J Clin Oncol. 2014; 32: 671–677.

11. Appendices

The list and shells of Tables, Figures and Listings will be provided in a separate document.

11.1 Schedule of Assessments

Trial procedures	**Baseline**	**Cohort 1:** Docetaxel 75 mg/m² iv Once every 3 weeks (cycle = 21 days with infusion on day 1) Premedication with oral dexamethasone 8mg, 12 hrs, 3 hrs and 1 hr before infusion, or according to local standard of care. Prednisone 2dd 5 mg continuously **Cohort 2:** ModraDoc006/r 20/20mg + 200/100mg Once every week twice daily (cycle = 21 days with ModraDoc006/r intake on day 1, 8, 15) Prednisone 2dd 5 mg continuously					
	≤ Day 28	**Cycle 1**			**Subsequent Cycles**	**EOT*** **	**FU**** **
		Day 1*	**Day 3 ****	**Day 10 ****	**Day 1***		
Visit outpatient clinic	X	X			X	X	X
Compliance assessments			X	X			
Informed Consent	X						
Randomization [1]	X						
In-/exclusion criteria	X	X					
Demographic Data [1]	X						
Medical History[2]	X						
Physical examination[3]	X	X			X	X	X
Vital Signs[4]	X	X			X	X	X
Hematology[5]	X	X			X	X	X
Clinical chemistry[6]	X	X			X	X	X
Serum testosterone	X						
Urinalysis[7]	X					X	
Creatinine Clearance[8]	X	X					
12-lead ECG	X					X	
CT-scan /MRI[9]	X					X	X[13]
Bone scintigraphy[9]	X					X	X[13]
PSA	X	X			X	X	X
Safety assessments[10]	X	X	X	X	X	X	X
Concomitant medication[11]	X	X			X	X	X
HRQoL questionnaires and EQ-5D[12]	X				X[12]	X	X

* Assessments may be performed within 3 days of day 1
**Day 3 and 10 (+/- 1 day) safety (AEs) and compliance assessments via triage by telephone during cycle 1 only
*** End Of Treatment (EOT) visit as soon as possible after discontinuation
**** Follow Up visit (FU) 30 days after the last administration of ModraDoc006/r or i.v. docetaxel

1. Randomization ≤ 3 days; Demographic data: age, ethnic origin
2. Medical History: (including details of malignancy, stage of cancer, radiotherapy and other therapies, with start and stop dates of therapy, number of cycles, cumulative doses, (cancer related) surgery etc., if applicable)).
3. Physical examination: WHO Performance Status, weight in kg, height in cm (height: baseline only).
4. Vital Signs: blood pressure, pulse rate and temperature.
5. Hematology: hemoglobin, hematocrit, RBC, WBC with differential (differential should include: neutrophil, lymphocyte, monocyte, basophil and eosinophil counts), platelet count.
6. Clinical Chemistry: ASAT, ALAT, GGT, LD, alkaline phosphatase, total and direct bilirubin, Na, K, Ca, Mg, glucose, urea, total protein, albumin, serum creatinine,.
7. Urinalysis: dipstick for protein, glucose, blood, pH, and ketones.
8. Creatinine Clearance: will be measured using Cockcroft-Gault formula or MDRD formula.

   For items 3) – 8): Assessments during screening less than 3 days before dosing need not to be repeated at day 1 of Cycle 1.

9. Tumor assessment (following PCWG3) by PSA prior to every cycle and imaging/radiological assessment (i.e. CT/MRI + bone scan) at baseline (≤ 28 days prior to administration of study medication), every 8 weeks for first 24 weeks (i.e. during week 9, 17 and 25), thereafter every 12 weeks and at End of Treatment (EOT), according to the RECIST v1.1 and PCWG3 criteria. Both PSA response and radiological response must be confirmed either after at least 3 weeks for PSA and after 4-6 weeks for RECIST v1.1. Of note: In this randomized comparative trial, the scheduled assessments should not be affected by delays in therapy, drug holidays or any other events that might lead to imbalance in a treatment arm in the timing of disease assessment (see Table 15 in study protocol section 6.9.1 for further details).
10. Toxicity assessments: baseline signs and symptoms and all related and unrelated adverse events will be assessed using the NCI-CTCAE criteria Version 5.0, including start and stop dates, severity, relationship to study drug, outcome and action taken.
11. Concomitant medication: including start and stop dates, dose, frequency, route of administration and indication.
12. Baseline, at the end of cycle 3, 6 and 10 or at EOT (if this would occur earlier): FACT global, FACT-P and FACT-taxane, Treatment Satisfaction (excluding baseline) and EQ-5D-5L questionnaire.
13. Imaging to be performed in the Follow Up visit, in case of unconfirmed PR or CR.

11.2 FACT-G and FACT-P Scoring instructions

FACT-G Scoring Guidelines (Version 4)

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 FACT-G score. ***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 score**

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 score**

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 score**

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 score**

TOTAL SCORE:

Score range: 0-108

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

*For additional guidelines please refer to the Administration and Scoring Guidelines in the manual or at www.facit.org.

FACT-P Scoring Guidelines (Version 4)

Subscale	Item Code	Reverse item?		Item response	Item Score	
PROSTATE CANCER SUBSCALE (PCS)		C2	4	-	_____	=_____
		C6	0	+	_____	=_____
		P1	4	-	_____	=_____
		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

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

To Derive a FACT-P total score:

Score range: 0-156

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

11.3 Sample SAS® code for analyses

- **Tables that need descriptive statistics – continuous variables:**

PROC UNIVARIATE DATA=dset NOPRINT;
VAR var1 var2 var3 …varn;
BY byvar; (optional)
OUTPUT OUT=outname;
N=n MEAN=mean MIN=min MAX=max MEDIAN=median STD=std;
RUN;

- **Tables that need frequency counts:**

PROC FREQ DATA=dset NOPRINT;
BY byvar; (optional)
*TABLES var1*var2;*
OUTPUT OUT=outname;
RUN;

- **Tables that need exact (Clopper-Pearson)95% CIs between groups for proportions:**

PROC FREQ DATA=dset;
BY byvar; (optional)
*TABLES var1 * var2 / binomial(exact) MEASURES ALPHA=0.05;*
RUN;

Notes: 1 Estimates are computed for 2x2 tables only
2 This code also gives exact 95% CIs within group for binomial proportions

- **Tables that need 95% CIs within group for binomial proportions:**

PROC FREQ DATA=dset;
BY byvar; (optional)
TABLES var1;
EXACT BINOMIAL;
RUN;

- **Tables that require analysis of (co)variance and 95/90% CIs between arms for continuous variables:**

PROC GLM DATA= dset OUTSTAT=outset;
CLASS class variables;
MODEL response = <effect> <treatment> / SOLUTION;
LSMEANS treatment / STDERR PDIFF CL; <ADD: ALPHA=0.1 for 90% CIs>
ESTIMATE 'T1 - T2' treatment 1 -1 0; or 'T2 – T1' treatment -1 1 0 if applic.
ESTIMATE 'T1 - T3' treatment 1 0 -1;
ESTIMATE 'T2 – T3' treatment 0 1 -1;
BY byvar; (optional)
WHERE wherever; (optional)
RUN;
QUIT;

Note: (Treatment order: 1=drug 1,2= drug 2,3= drug 3, etc… 5= placebo)

- **Tables that present Fisher's Exact or CMH:**
  PROC FREQ DATA=dset NOPRINT;
  BY byvar; (optional)
  *TABLES var1*var2/CMH score=MODRIDIT EXACT;*
  OUTPUT OUT=outname CMH EXACT;
  RUN;
- **Tables that need to use WILCOXON:**
  PROC NPAR1WAY DATA=dset WILCOXON;
  CLASS class variables;
  VAR variable;
  RUN;
- **Tables that need number of events/censored and probabilities of failure/survival at cut off times:**
  PROC LIFETEST DATA=dset OUTSURV=LIFE METHOD=LT INTERVALS=12, 24,;
  *TIME duration*censor (0 or 1);*
  ID subject;
  STRATA treatment;
  RUN;
- **Tables that need life table with estimates of survival, with CIs and log rank test:**
  PROC LIFETEST DATA=dset OUTSURV=LIFE METHOD=KM;
  *TIME duration*censor (0 or 1);*
  ID subject;
  STRATA treatment;
  RUN;
- **Kaplan-Meier curves for treatment:**
  PROC LIFETEST data=dataset plots=survival(strata=individual);
  *TIME time*event(0);*
  STRATA treatment;
  RUN;
- **Tables that need Cox Proportional Hazards models:**
  PROC PHREG data=dataset;
  CLASS treatment;
  *MODEL time*event(0) = treatment / ties=exact;*
  CONTRAST 'a vs b' treatment 1 / estimate=exp;
  RUN;

Testing PH assumptions:

```
PROC PHREG data=dataset;
    CLASS treatment;
    MODEL time*event(0) = treatment / ties=exact;
    Treatment_t = treatment*log(time);
    Proportionality_test: test treatment_t ;
RUN;

Model for time dependent covariate:PROC PHREG data=dataset;
    CLASS treatment;
    MODEL time*event(0) = treatment vartimedependant / ties=exact;
    CONTRAST 'a vs b' treatment 1 / estimate=exp;
    Vartimedependant=var*(time>x);
RUN;
```

Vartimedependant=var(time>x) will be defined after review the results of model with interaction and HR curves*

11.4 Document History

Document Version, Status, Date	Summary/Reason for Changes
Version 1, Final, 18 October 2019	Not applicable; the first version
Version 2, Final, 23 November 2020	Major updates: - Use of protocol Final Version 3.0 (Amendment 2) dated 31Mar2020, including change from Objective Response Rate to radiographic Progression Free Survival as primary endpoint - BOR/ORR/DCR/DOR: confirmation of response is not required - Section 7.1: Tables and Listings will be split by protocol version - Section 7.7.2: Additional AE summaries according to baseline BSA level - Clarification of some planned statistical analysis mainly for programming purpose - Use of Covance template ST-AD-008 version 05 instead of version 4 (Removed approvals page as this is now a separate associated document, Document history moved to the Appendix)
Version 3, Final, 07 October 2021	Main updates: - Disposition of patients: Add number of ongoing subjects - Efficacy: o Clarification of event/censoring rules for time-to-event data o Analysis of TTF as sensitivity analysis, with additional Table 14.2.1.5 and Figure 14.2.1.6. o Analysis of ORR repeated on the PP population, among the patients evaluable for radiological response - Use of visit windows for analysis of Quality of Life questionnaires - Safety: Consideration of BUN parameter Of note, template ST-AD-008 version 06 is not used because changes motivating this SAP amendment are not critical and do not justify its use. Same template as SAP Version 2.0 is used i.e. ST-AD-008 Version 05.