Document Type:	Statistical Analysis Plan
Official Title:	A randomized, double–blind, placebo–controlled Phase III study of darolutamide (ODM–201) versus placebo in addition to standard androgen deprivation therapy and docetaxel in patients with metastatic hormone–sensitive prostate cancer
NCT Number:	NCT02799602
Document Date:	11-Nov-2021

Title page

A randomized, double-blind, placebo-controlled Phase III study of darolutamide (ODM-201) versus placebo in addition to standard androgen deprivation therapy and docetaxel in patients with metastatic hormone-sensitive prostate cancer

Darolutamide in addition to standard androgen deprivation therapy and docetaxel in metastatic hormone-sensitive prostate cancer

Acronym: ARASENS

Bayer study drug	BAY 1841788 / Darolutamide		
Study purpose:	Efficacy and Safety		
Clinical study phase:	III	**Date:**	11 NOV 2021
Study No.:	17777	**Version:**	4.1

Author: PPD
PPD
PPD Bayer Pharmaceuticals
Whippany, NJ 07981

This Statistical Analysis Plan is produced on a word-processing system and bears no signatures.

The approval of the Statistical Analysis Plan is documented in a separate Signature Document.

Table of Contents

Table of Tables

Table of Figures

List of Abbreviations

ADT	Androgen deprivation therapy
AE	Adverse event
AEE	Actual extent of exposure
ALP	Alkaline phosphatase
ATC	Anatomical-Therapeutic-Chemical
bid	Twice daily
BHA	Bone health agent
BPI-SF	Brief Pain Inventory – Short Form
CI	Confidence interval
CT	Computed tomography
CV	Coefficient of variation
DMC	Data Monitoring Committee
EBRT	External beam radiation therapy
ECG	Electrocardiogram
ECOG PS	Eastern Cooperative Oncology Group Performance Status
EOT	End of treatment
FAS	Full analysis set
FPSI-DRS-P	FACT/NCCN FPSI-17 disease-related symptoms - physical
GCP	Good Clinical Practice
HR	Hazard ratio
ICF	Informed consent form
KM	Kaplan-Meier
LHRH	Luteinizing hormone releasing hormone
LKAD	Last known alive date
LLOQ	Lower limit of quantification
M&S	Modeling & Simulation
MedDRA	Medical Dictionary for Regulatory Activities
mg	Milligram
mHSPC	Metastatic hormone-sensitive prostate cancer
mL	Milliliter
MRI	Magnetic resonance imaging
NCCN	National Comprehensive Cancer Network
NCCN-FACT FPSI-17	Functional assessment of cancer therapy / National Comprehensive Cancer Network prostate cancer symptom index 17 item questionnaire
NCI-CTCAE v4.03	National Cancer Institute-Common Terminology Criteria for Adverse Events; version 4.03
ng	Nanogram
OEE	Overall extent of exposure
OS	Overall survival
PCWG3	Prostate Cancer Clinical Trials Working Group
PK	Pharmacokinetics
PKS	Pharmacokinetic Analysis Set
PRO	Patient-reported outcomes
PSA	Prostate-specific antigen
PT	Preferred Term
QoL	Quality of life
RECIST	Response Evaluation Criteria In Solid Tumors
SAE	Serious adverse event
SAF	Safety analysis set
SAP	Statistical Analysis Plan
SD	Standard deviation
SOC	System Organ Class
SSE	Symptomatic skeletal event
SSE-FS	Symptomatic skeletal event free survival
TEAE	Treatment emergent adverse events

TESAE	Treatment emergent serious adverse events
t_{last}	Time to the last quantifiable concentration
t_{max}	Time to the maximum plasma concentration
ULN	Upper limit of normal
WHO	World Health Organization
WHO-DD	WHO Drug Dictionary
WPS	Worst pain subscale

1. Introduction

Study 17777 (ARASENS) is an international, randomized, double-blind, placebo-controlled, Phase III efficacy and safety study of darolutamide in addition to standard androgen deprivation therapy (ADT) and docetaxel in patients with metastatic hormone-sensitive prostate cancer (mHSPC).

This statistical analysis plan (SAP) describes the pre-specified analyses and data presentations for the primary completion analyses, which will be performed when the targeted number of death events for the primary efficacy endpoint of overall survival (OS) is reached.

The SAP contains definitions of analysis sets, derived variables, and statistical methods for the analysis of efficacy and safety.

For SAP version history and changes in the planned statistical analysis, please refer to Section 7.

2. Study Objectives

The primary objective of this study is:

- To demonstrate the superiority in overall survival (OS) of darolutamide in addition to standard androgen deprivation therapy (ADT) and docetaxel over placebo in addition to standard ADT and docetaxel

The secondary objectives of this study are to evaluate:

- Time to castration-resistant prostate cancer
- Time to pain progression
- Symptomatic skeletal event free survival (SSE-FS)
- Time to first symptomatic skeletal event (SSE)
- Time to initiation of subsequent systemic antineoplastic therapy
- Time to worsening of disease-related physical symptoms based on functional assessment of cancer therapy / National Comprehensive Cancer Network prostate cancer symptom index 17 item questionnaire (NCCN-FACT FPSI-17)
- Time to initiation of opioid use for ≥7 consecutive days
- Safety

The secondary objective of the efficacy endpoints will be tested in the hierarchical order as given above, except for safety.

The exploratory objectives of this study include the evaluation of:

- Quality of life (QoL), measured using the NCCN-FACT FPSI-17 questionnaire
- Medical resource use
- Prostate-specific antigen (PSA) assessments
- PK and exposure-response analysis
- Evaluate biomarkers to investigate the drug (i.e. mode-of-action-related effect and / or safety) and / or the pathomechanism of the disease

3. Study Design

Design overview

This is an international, randomized, double-blind, placebo-controlled, multicenter Phase III study of darolutamide in patients with mHSPC. Approximately 1,300 patients will be randomized (1:1 ratio) to receive one of the following study drugs:

- Darolutamide, 600 mg (2 x 300 mg tablets) bid orally with food
- Placebo matching darolutamide tablets in appearance, bid orally with food

All patients must receive ADT of Investigator's choice (LHRH agonist/antagonists or orchiectomy) as standard therapy, started ≤12 weeks before randomization. Six cycles of docetaxel will be administered as background treatment after randomization. The first cycle of docetaxel should be administered within 6 weeks after start of study drug.

For at least the first 20 randomized patients who have received at least 1 cycle of docetaxel, and for whom a detailed mandatory PK analysis will be performed, docetaxel should be administered at least 14 days after randomization.

Patients will be stratified at randomization as follows:

- Extent of disease
    - Non-regional lymph nodes metastases only
    - Bone metastases with or without lymph node metastases
    - Visceral metastases with or without lymph node metastases or with or without bone metastases
- Alkaline Phosphatase (ALP)
    - ALP < upper limit of normal (ULN)
    - ALP ≥ ULN

Note: Blood samples to measure ALP levels for stratification will be analyzed in a central laboratory.

The study will comprise 4 consecutive periods: Screening, Treatment, Active Follow-up, and Long-term (survival) Follow-up.

Screening period:

All trial-related procedures and evaluations will only be performed after the patient has agreed to participate and has signed the informed consent form (ICF).

The Screening period will consist of multiple evaluations that will take place within 28 days prior to randomization to ensure that all eligibility criteria are met.

Once eligibility is confirmed and documented, eligible patients will be randomized in a ratio of 1:1 to treatment with darolutamide or placebo.

Treatment period:

The start of the treatment period is defined by the first administration of study drug (darolutamide or placebo). Study treatment will be provided for all patients to be taken twice

daily until disease progression (symptomatic progressive disease, change of systemic antineoplastic therapy), unacceptable toxicity, consent withdrawal or withdrawal from the study at the discretion of the Investigator or his/her designated associate(s), death, or non-compliance.

Patients will be evaluated every 12 weeks for castration-resistant prostate cancer, initiation of subsequent systemic antineoplastic therapy, SSEs, opioid use for ≥7 consecutive days, pain progression, worsening of physical symptoms of disease based on NCCN-FACT FPSI-17, adverse events (AEs) and serious adverse events (SAEs), QoL and PSA assessment.

Active Follow-up period:

The Active Follow–up period is the interval from the end of darolutamide or placebo intake to the end of all protocol–specified post–treatment interventions.

It includes the End of Treatment Visit (EOT). An EOT Visit will be conducted 30 (+ window of 7) days after the last dose of darolutamide or placebo. The following assessments will be performed at the EOT Visit: QoL, pain assessment, analgesic consumption, subsequent systemic antineoplastic treatments for prostate cancer, SSEs, and all AEs and SAEs regardless of causality.

During the Active Follow-up period, the following assessments will be performed during standard of care clinic visits approximately every 12 weeks for up to 1 year: QoL, pain assessment, analgesic consumption, survival status, subsequent systemic antineoplastic treatments for prostate cancer (which should be provided with start and stop dates and reason for change, SSEs, and study drug-related SAEs. After approximately 1 year of Active Follow-up, patients will be transitioned to Long-term (survival) Follow-up. The Active Follow-up period extends from the discontinuation of darolutamide or placebo treatment period for up to 1 year or until the patient can no longer travel to the clinic, dies, is lost to follow-up, or withdraws informed consent and actively objects to collection of further data.

Long-term (survival) Follow-up period:

After Active Follow-up, patients will continue to be contacted approximately every 12 weeks (by phone) to capture all systemic antineoplastic treatments for prostate cancer with start and stop date and reasons for change (PSA progression, clinical progression, radiological progression, toxicity, other), study drug-related SAEs and survival status. The end of Long-term Follow-up period is death, lost to follow-up, consent withdrawal, or end-of-study.

Survival sweep:

For every analysis of OS, survival data will be collected through additional survival sweeps. All patients considered alive at the database cut-off date and prior to any subsequent additional analysis will be contacted for survival status.

Primary completion

The first analysis of OS will be performed as an interim futility analysis when approximately 153 deaths have occurred. In case the study continues after the futility analysis, a final efficacy analysis of OS will be performed when approximately 509 deaths have occurred and will be considered the primary completion. The database cut–off date for this analysis will be the date of primary completion.

4. General Statistical Considerations

4.1 General Principles

Statistical analyses will be performed by using the software package SAS® Version 9.2 or higher (SAS Institute Inc., Cary, NC, USA).

All variables will be analyzed by descriptive statistical methods by treatment arm and overall, unless otherwise specified. The number of data available and missing data, mean, standard deviation (SD), minimum, quartiles, median, and maximum will be calculated for continuous variables. Frequency tables will be generated for categorical variables.

4.2 Handling of Dropouts

A patient who discontinues study participation prematurely for any reason is defined as a "dropout" if the patient has already been randomized. Patients who discontinue treatment or study will not be replaced.

4.3 Handling of Missing Data

All missing or partial data will be presented in the patient data listing as they are recorded on the Case Report Form. Except as noted, missing data will not be estimated or carried forward in any statistical analysis.

When appropriate, the following rules will be implemented so as not to exclude patients from statistical analyses due to missing or incomplete data:

Overall survival and other time-to-event endpoints:

Every effort should be made to resolve incomplete or missing dates during the course of the study (i.e. edit checks, data cleaning/monitoring etc.). However, in rare circumstances, missing parts of either the censoring date or the event date may occur where an imputation algorithm has to be defined. In general the following rule should be followed: Missing month or year is not acceptable. To impute a value for day, day 15 of the month will be used for the calculation. This applies to all OS and time-to-event endpoints dates if the dates are partially missing in the selected data panels.

AE and Concomitant Medication (CM) partial or missing dates

Refer to SAP appendix 9.1 imputation rule for incomplete AE/CM dates.

Laboratory Assessments

If laboratory assessment results contain an inequality sign "<" or ">", then the lower or higher detection limit will be used as the imputed laboratory assessment value for both safety and efficacy analyses.

ePRO Questionnaire

Two scores will be derived for Brief Pain Inventory-Short Form (BPI-SF): (1) for the pain severity score if one answer is missing then scoring will be set to missing, (2) for the pain interference score if four or more answers are missing out of the seven questions then the score will be set to missing (see scoring of BPI-SF in Appendix 9.3).

For NCCN-FACT FPSI-17, total scores and subscales scores (disease-related physical symptoms, disease–related emotional symptoms, treatment side-effects, function/well-being) will be assessed. Where there are missing items, subscale scores can be prorated if >50% of items on subscale are completed. If ≤50% of the items are answered for any domain, then the score of that domain is set to missing. The total score is set to missing if the related overall item response rate is less than or equal to 80% (see scoring of NCCN-FACT FPSI-17 in Appendix 9.4). In addition, a total score should only be calculated if all of the component subscales have valid scores.

4.4 Determination of Sample Size

The sample size of the study is based on the primary endpoint of OS. The study is designed to have 90% power to detect a 25% decrease in risk of death with darolutamide compared to placebo, corresponding to a HR of 0.75 with a one-sided test with a type I error of 0.025 (equivalent a two-sided test with a type I error 0.05). The OS data will be considered mature when approximately 509 deaths are observed.

With the additional assumptions that patients will enroll at a rate of 50 patients per month, exponential distributions of the OS event times, median time of OS in the control group of 60 months, 5% dropout rate of patients, and a 6-months enrollment ramp-up period, it follows that approximately 1,300 patients need to be randomized to observe 509 deaths after approximately 70 months.

4.5 Interim Analyses and Data Monitoring

4.5.1 Interim Analyses

For the futility interim and final analyses together, a one-sided overall beta of 0.1 will be used. Stopping boundaries will be calculated with an O'Brien-Fleming beta-spending function using the actual number of events observed up to the cut-off date. The interim futility analysis is planned when approximately 153 deaths were observed (information fraction=0.3). The critical boundary for the futility analysis will be calculated separately from the efficacy boundary, in order to not interfere with the type I error of the efficacy analysis.

The following scenario in Table 4–1 serves as an example.

Table 4–1 Example of stopping boundaries for interim analyses

		Stopping Boundaries Hazard Ratio Scale (darolutamide+docetaxel arm vs. placebo+docetaxel arm) (% Improvement) / Z Score			
Analysis Time	**# Events**	**Crossing Lower Bound (efficacy)**	**Crossing Upper Bound (futility - lack of efficacy)**	**Cumulative Alpha Spent**	**Cumulative Beta Spent**
1st Interim[a]	153	-	1.166 (-14.2 %) / 0.952	-	0.003
Final	509	0.841 (19.0%) / 1.96	0.841 (19.0%) / 1.96	0.025	0.100

Software ADDPLAN neo V10.0.4 was used for this design

a: the 1st interim analysis was completed and its stopping boundaries were calculated before the decision was made to remove the 2nd interim analysis, which was originally planned but removed with version 3.0 of the SAP. Please refer to Section 7 for additional details.

If the estimated hazard ratio (darolutamide over placebo) is greater than or equal to 1.166 at the interim analysis, stopping the trial for lack of efficacy (i.e. futility) would be considered. Otherwise, i.e. if the estimated hazard ratio is less than 1.166, the trial would continue to the final analysis, providing there were no other safety considerations to warrant discontinuation of patient treatment.

4.5.2 Independent Data Monitoring Committee

A DMC will be instituted to ensure ongoing safety of study patients with respect to a risk/benefit assessment during periodic data review meetings, review results from planned interim analyses and provide a formal recommendation for continuation/termination of the study and monitor study conduct to ensure the overall integrity of study is maintained. The DMC will operate independently of the Sponsor and Investigators.

4.6 Data Rules

Baseline

Baseline is defined as the last measurements performed in Screening or Visit 1 (regardless the scheduled or the unscheduled visits), unless otherwise specified prior to the first study drug administration.

Repeated measures

If there are repeated measurements per time point on the same day (e.g. laboratory values, vital signs, etc.), the following rules will be used (unless otherwise specified):

- Before the start of the study drug administration (i.e., for screening and baseline value), the latest measurement at scheduled visits will be used. Unscheduled visits will be used if there are no measurements at scheduled visits. If the latter is the case, the last unscheduled visit will be used.
- In case of repeated measurements at any post baseline time point, the first measurement at scheduled visits will be used. Unscheduled visits will be used if there are no measurements at scheduled visits. If the latter is the case, the first unscheduled visit will be used.

Stratification variables

Since differences may occur between the values of the stratification factors entered by the investigator at the time of randomization into IxRS and those collected on the eCRF, the stratified analysis of OS will be performed using both assignments to the strata. The primary stratified analysis for OS will be based on the information collected in the IxRS whereas stratification information entered in the eCRF (considered the ‘true’ information) will be used for the sensitivity analysis. In rare circumstances, if eCRF data cannot be documented, then IxRS data will be used.

5. Analysis Sets

5.1 Assignment of Analysis Sets

Patients will be excluded from all below-mentioned analysis set if they are related to or associated with any critical GCP violations that result in fraudulent patient data (please refer to Section 7 for rationale). The FAS will be used for the analysis of all efficacy endpoints and all other endpoints. The SAF will be used for the analyses of all safety endpoints. The pharmacokinetic data will be analyzed in the PKS. Patients who signed the informed consent but for any reasons did not proceed to randomization will be considered screening failures. They will be listed separately.

Full analysis set (FAS)

All patients who were randomized are included in the Full Analysis Set, except for cases with critical GCP violations. Following the intent-to-treat (ITT) principle, the patients in this set will be grouped according to the planned treatment they were allocated to receive at randomization, irrespective of actual treatment.

Safety analysis set (SAF)

All randomized patients who have received at least 1 dose of darolutamide or placebo represent the SAF, except for cases with critical GCP violations. This safety population will be used in the analyses of all safety endpoints. Patients will be included in the analyses according to the treatment they actually received. Patients will be included in the darolutamide+docetaxel arm if they have received any dose of darolutamide, otherwise they will be included in the placebo+docetaxel arm if they only received placebo.

Pharmacokinetic Analysis Set (PKS)

At least the first 20 randomized patients who have received at least 1 cycle of docetaxel, and for whom mandatory dense PK sampling was done, will be included in the PKS, as long as they have received at least 3 days of uninterrupted study drug treatment as well as one cycle of docetaxel and have at least one post dose PK measurement, except for cases with critical GCP violations.

In addition, a subset of Chinese patients who were randomized and have received at least 1 dose of darolutamide, and for whom mandatory dense PK sampling was done prior to the first dose of docetaxel, will also be included in the PK analysis set.

6. Statistical Methodology

6.1 Population characteristics

6.1.1 Disposition

The number of patients enrolled and included in each population will be tabulated by region, country, and center. The number of patients starting, completing, and discontinuing together with the primary reason (including COVID-19 pandemic related reason) for discontinuing of the Screening, Treatment, Active Follow-up, and Long-term Follow-up periods will be presented overall and by treatment arm (except for the Screening period). The number of patients currently on treatment at primary completion analysis cut-off will also be summarized by treatment arm. Disposition of Screening will be summarized for all enrolled patients, and

disposition of other study periods for the FAS. Patients who enter Active Follow-up are defined as patients who discontinued study treatment not due to death or lost to follow-up and had Active Follow-up visits after the end of treatment visit. The number of patients affected by COVID-19 pandemic related study disruptions (protocol deviations, premature discontinuations, and AEs) will be summarized in a separate table.

In addition, the number of patients with important protocol deviations will be presented by treatment arm and overall. The frequencies of each important protocol deviation will be presented by treatment arm and overall. All COVID-19 pandemic related protocol deviations are considered important protocol deviations and will be presented by treatment arm and overall. Listings of important protocol deviations associated with COVID-19 pandemic will be presented as well. The number of patients with missed visits or assessments in the key assessments will be presented for each treatment arm and overall.

6.1.2 Demographic and Other Baseline Characteristics

Descriptive summaries of demographics and baseline characteristics are presented by treatment arm and overall, for the FAS. Comparability of the treatment arms with respect to demographics and baseline characteristics is assessed using descriptive summaries.

The following demographic data will be summarized:

- Age at screening (years)
- Age category (<65, 65-74, 75-84, ≥85 years)
- Geographical region (North America, Asia Pacific, Rest of the World)[1]
- Race (White, Asian, Black or African American, Other, Not Reported) and ethnicity (Hispanic or Latino, Not Hispanic or Latino, Not Reported)
- Weight at baseline
- Body Mass Index (BMI) (kg/cm^2) (<20, 20 to <25, 25 to <30, ≥30)
- Renal function - eGFR at baseline:
  - Normal: eGFR ≥ 90 mL/min
  - Mild impairment: 60 ≤ eGFR < 90 mL/min
  - Moderate impairment: 30 ≤ eGFR < 60 mL/min
  - Severe impairment: 15 ≤ eGFR < 30 mL/min

  eGFR will be calculated according to the Modification of Diet in Renal Disease (MDRD) Formula:

  $eGFR\ (mL/min/1.73m^2) = 186 \times SCR^{-1.154} \times age^{-0.203} \times (0.742, \text{if female}) \times (1.212 \text{ if blacks or African American}) \times (0.881 \text{ if Japanese}) \times (1.227 \text{ for Chinese (mainland China, Taiwan and Hongkong)}$, where SCR= serum creatinine measured in mg/dL
- Hepatic function at baseline:

[1] Geographic region: North America (Canada and United States), Asia Pacific (China, Japan, South Korea, and Taiwan) and the rest of the world (Australia, Belgium, Bulgaria, Brazil, Czech Republic, Germany, Spain, Finland, France, United Kingdom, Israel, Italy, Mexico, Netherlands, Poland, Russian Federation, and Sweden).

- Normal: Total bilirubin and AST ≤ upper limit of normal (ULN)
- Mild impairment: Total bilirubin > ULN to 1.5 x ULN or (Total bilirubin ≤ ULN and AST > ULN)
- Moderate impairment: Total bilirubin > 1.5 to 3 x ULN, any AST
- Severe impairment: Total bilirubin > 3 x ULN, any AST.

The following baseline characteristics will be summarized:

- Extent of metastatic disease at study entry (Non-regional lymph nodes metastases only, Bone metastases with or without lymph node metastases, Visceral metastases with or without lymph node metastases or with or without bone metastases), from both IxRS and eCRF
- ALP (U/L) at baseline (<ULN, ≥ULN) - central laboratory, from both IxRS and eCRF
- PSA (ug/L) at baseline (<median of overall population, ≥median of overall population) - central laboratory
- Eastern Cooperative Oncology Group Performance Status (ECOG PS) at baseline
- Stage of prostate cancer (TNM Classification) at initial diagnosis (Stage I, IIA, IIB, III, IV (M1, M0))
  The study used AJCC 7 and therefore the stage 4 M0 and M1 were not captured in the eCRF. As the CRFs captured the date of initial diagnosis and the date of initial diagnosis of metastases, the following assumptions will be used to estimate the M0 or M1 status at initial diagnosis:
  - Patients who have a time interval > 3 months between initial diagnosis and initial metastases diagnosis will be assumed to be M0.
  - Patients who have a time interval of ≤ 3 months between initial diagnosis and initial metastases diagnosis will be assumed to be M1.
- Gleason score of prostate cancer (<8, ≥8) at initial diagnosis
- Time from initial diagnosis to the date of the first actual dose of darolutamide or placebo (months)
- Time from initial diagnosis of metastases to the date of the first actual dose of darolutamide or placebo (months)
- Time from initial diagnosis to initial diagnosis of metastases (months)
- Stage of prostate cancer (TNM Classification) at study entry
- Testosterone (<0.5, ≥0.5 ng/mL) at baseline - central laboratory
- Chromogranin A (ug/L) at baseline – local laboratory.
- The disease-related physical symptoms subscale of the NCCN-FACT FPSI-17 at baseline
- Worst pain subscale, average pain subscale of the BPI-SF (0: no pain, 1-3: mild pain, 4-7: moderate pain, 8 to 10: severe pain)

- Pain severity score of the BPI-SF

Demographic and baseline characteristics tables will be presented by geographical regions (North America, Asia Pacific, Rest of the World).

6.1.3 Medical History

For data coding, the Medical Dictionary for Regulatory Activities (MedDRA) will be used for medical history. Medical history findings (i.e. previous diagnoses, diseases or surgeries) not pertaining to the study indication, start before signing of the ICF and considered relevant to the study will be presented for each MedDRA Primary System Organ Class (SOC), and Preferred Term (PT) by treatment arm and overall for the FAS.

6.1.4 Prior and Concomitant Medication and Procedures

Prior therapy is defined as all medications taken before the start of darolutamide or placebo treatment. Concomitant therapy is defined as all medications taken between the start of darolutamide or placebo treatment and the last dose date of darolutamide or placebo.

- Prior and concomitant medications excluding ADT, systemic antineoplastic therapies captured in the anti-cancer therapy eCRF page: frequency of patients for each drug category
- Concomitant statin medications: frequency of patients
- Prior systemic antineoplastic therapy, including ADT by preferred drug name: frequency of patients for each drug category. The prior systemic antineoplatic therapy, including ADT, are selected in CM dataset using ATC codes: L02BX, L02BB, H01CC, V98, G03HB, L02AE, H01CA.
- A summary of subsequent systemic antineoplastic therapy will be presented by preferred drug name based on WHO-DD drug record number.
- A summary of life-prolonging subsequent systemic antineoplastic therapy will be presented by preferred drug name based on WHO-DD drug record number. The following medications after the study treatment period are life-prolonging subsequent therapies for prostate cancer: Abiraterone, Apalutamide, Enzalutamide, Docetaxel, Cabazitaxel, Radium 223, Sipuleucel-T, Lutetium-177.
- Hormonal theray and orchiectomy at study entry (by LHRH agonist/antagonist only, orchiectomy only, LHRH agonist/antagonist and orchiectomy).
- Prior local treatment for prostate cancer at study entry (by prostatectomy, surgery, TURP, other, radiation, no surgical treatment and no radiation).
- Concurrent prostate cancer related-invasive procedures (e.g. procedures for urinary symptoms, surgical procedures for prostate cancer): frequency of patients by grouped term and procedure. Prostate cancer-related invasive procedure is defined as any procedure needed for alleviation of symptoms, signs or findings caused by progression of prostate cancer (e.g. catheterization of the bladder, percutaneous drainage of hydronephrosis, palliative electro resection of the prostate, etc.).

The dictionary used for coding medications was the ATC Class/Subclass and the World Health Organization Drug Dictionary (WHO-DD). For selection of statins medications, please refer to Appendix 9.5.

6.1.5 Extent of Exposure

Study drug is defined as darolutamide or placebo. Six cycles of docetaxel will be administered as background treatment after randomization and is not considered as study drug. As a general rule, trailing "0 mg" records, which are not followed by any positive amount of drug, will not be included in the calculation of any drug duration or amount. Similarly, trailing "drug interruptions" will not be used in statistical tables. Interruption becoming permanent study treatment discontinuation is not accounted as an interruption.

Descriptive statistical summaries will be provided for darolutamide or placebo and docetaxel separately by treatment arm in the SAF population for the following variables:

- Overall time under treatment (or treatment duration)
  Including time off drug and dose interruptions. It will be calculated in days and presented in months as (date of the last dose of any study treatment – date of the first dose of any study treatment + 1) / 30.44.
- Actual dose per day
  Actual dose per day = total amount of dose / number of days with intake > 0
- Total amount of dose
  Total amount of dose = sum of dose received over total time under treatment
- Percent of planned dose received.
  Percent of planned dose received = total amount of dose [mg] / planned dose [mg] * 100%; planned dose is sum of the intended initial dose according to protocol over total time under treatment.
- From 1 day after the last dose of docetaxel to the last dose of darolutamide or placebo.
  Including time off drug and dose interruptions. It will be calculated in days and presented in months as (date of the last dose of darolutamide or placebo – date of the last dose of docetaxel + 1) / 30.44. If the last dose of docetaxel date is after the last dose of darolutamide or placebo, then 0 day of duration will be included in the calculation.

For patients with dose reduction, interruption, or delay, the number of dose reductions, re-escalations, interruptions, or delays per patient and their reasons will be summarized. Only dose modifications after start of treatment and before end of treatment are included in the dose modification summary table. For example, dose interruption or delay documented without further dose resumption afterwards.

Extent of exposure will be described by age groups (<65, 65 - 74, 75 - 84, ≥85 years) and geographical regions (North America, Asia Pacific, Rest of the World).

6.2 Efficacy

For time-to-event analyses the censoring mechanism is assumed to be non-informative. Patients without events will be handled as right-censored in time-to-event analyses. Time to

event will be calculated in days as date of event or censoring minus date of randomization and presented in months (=days/30.44). If event or censoring date is on date of the randomization, then time to event =1 (or day 1).

6.2.1 Primary Efficacy Endpoint

Definition of OS

The primary efficacy variable is OS, defined as the time from date of randomization until death from any cause.

The following censoring rules in Table 6–1 will be applied:

Table 6–1 Censoring rules for OS

Situation	End Date	Censored
Documented death during study before or at data cut-off date	Death date	No
No documented death with no contacts after randomization and before or at data cut-off date	Date of randomization (Day 1)	Yes
No documented death before or at data cut-off date	Last known alive date (LKAD) or at the data cut-off date, whichever comes earlier	Yes

Last Known Alive Date

The last known alive date (LKAD) is derived from the main data sources. The last available date across all selected data panels listed below will be used as the LKAD by patient. Information from selected data, i.e. visit dates, exposure information, laboratory measurements, tumor assessment dates, SSE dates, demographics, survival status date, vital signs and disposition events or follow up assessments will be used to determine survival status.

Analysis of OS

All randomized patients (FAS) will be included in the primary analysis of OS.

The primary analysis of OS will be a stratified log-rank test with the same IxRS stratification factors as used for randomization. The null hypothesis that there is no difference in OS between treatment arms, which is equivalent to a hazard ratio (HR) of 1, will be tested against the alternative hypothesis that the HR of darolutamide over placebo is below 1. If the p-value from the one-sided log-rank test is less than 0.025 (corresponding to a two-sided log-rank test less than 0.05) with the HR (darolutamide+docetaxel arm vs. placebo+docetaxel arm) less than 1, the null hypothesis will be rejected in favor of the alternative hypothesis.

The SAS code will be similar to the following:

```
PROC LIFETEST DATA=<DATASET>;
TIME EFFVAL * CENSORNY;
STRATA {strata variables} / GROUP=TREATMGR test=(logrank);
RUN;
```

The HR (darolutamide or placebo) for OS and its 95% confidence interval (CI) will be calculated using the Cox model, stratified by the same factors as were used for randomization. The SAS code similar to the following will be used:

```
PROC PHREG DATA = <DATASET>;
MODEL EFFVAL * CENSORNY = TREATMGR;
STRATA {strata variables};
RUN;
```

Kaplan-Meier (KM) estimates for the median time of OS (including 95% CI) and 25% and 75% percentiles will be presented for each treatment arm. The KM estimates at time points such as 12 months, 24 months, etc., together with corresponding 95% CIs and the differences of these estimates between the darolutamide+docetaxel arm and the placebo+docetaxel arm will be presented. The SAS code will be similar to the following:

```
PROC LIFETEST DATA = <DATASET>;
TIME EFFVAL * CENSORNY;
STRATA TREATMGR;
RUN;
```

Descriptive statistics of survival follow-up time will be calculated by treatment arm and total. Follow-up time is the time to LKAD or death as described for the OS endpoint for censored and uncensored patients.

Sensitivity analyses

OS will be evaluated with the unstratified log-rank test and Cox model for the FAS population (OS sensitivity analysis 1 – unstratified analysis).

OS will also be evaluated with the stratified log-rank test and Cox model for the FAS population using stratification factors collected from the eCRF, in case there are more than 5% of patients with different values in any stratification variable between IxRS and eCRF (OS sensitivity analysis 2 – eCRF-variables stratified analysis).

In addition, OS will also be evaluated with the stratified log-rank test and Cox model for the FAS population using extent of disease stratification factors collected from the central imaging review (OS sensitivity analysis 3 – central imaging review extent of disease as stratification factor).

6.2.2 Secondary Efficacy Endpoints

If the primary endpoint OS is significant, secondary efficacy endpoints will be tested with the hierarchical gatekeeping procedure. The order of secondary endpoints is as follows:

1) Time to castration-resistant prostate cancer
2) Time to pain progression
3) Symptomatic skeletal event free survival (SSE-FS)

4) Time to first symptomatic skeletal event (SSE)
5) Time to initiation of subsequent systemic antineoplastic therapy
6) Time to worsening of disease-related physical symptoms based on functional assessment of cancer therapy / National Comprehensive Cancer Network prostate cancer symptom index 17 item questionnaire (NCCN-FACT FPSI-17)
7) Time to initiation of opioid use for ≥ 7 consecutive days

If the primary endpoint OS is statistically significant at a 0.025 level (one-sided), the secondary endpoints will be tested using a hierarchical test procedure in the order as described above at the same nominal significance level. If the OS or a secondary endpoint is not statistically significant, the hierarchical procedure is stopped and all subsequent analyses of the secondary endpoints will be considered exploratory.

6.2.2.1 Time to castration-resistant prostate cancer

Time to castration-resistant prostate cancer is defined as the time from randomization to the time of the occurrence of the following events, whatever comes first (date of first occurrence of castration-resistant prostate cancer minus date of randomization):

- PSA progression, according to PCWG3 criteria, with serum testosterone being at castrate level <0.50 ng/mL, is defined as the date that a 25% or greater increase and an absolute increase of 2 ng/mL or more from the nadir (lowest at or after baseline) is documented, which both are confirmed by a second value obtained at least 3 weeks later, including all potential PSA values ≥2 ng/mL above nadir and ≥25% increase above nadir between initial assessment date and confirmation assessment date. This definition requires serum testosterone at castrate levels <0.50 ng/mL and a first assessment date at least 12 weeks from randomization.
- Radiological progression by soft tissue and visceral lesions, is defined according to Response Evaluation Criteria In Solid Tumors (RECIST) version 1.1 (1), based on MRI / CT scans of the chest, abdomen, and pelvis performed by the Investigator, as recommended by Prostate Cancer Clinical Trials Working Group (PCWG3) (2).
- Radiological progression by bone lesions, is defined according to PCWG3 criteria based on whole body ^{99m}Tc methylene diphosphonate bone scans performed by the Investigator. Bone lesions will be recorded separately from soft tissue and visceral lesions.

The following censoring rules in Table 6–2 will be applied:

Table 6–2 Censoring rules for time to castration-resistant prostate cancer

Situation	End Date	Censored
Castration-resistant prostate cancer due to any component event during study	The earliest event date among the three components	No
No baseline or post-baseline event assessment for all three components #	Date of randomization (Day 1)	Yes
PSA progression event immediately after two or more consecutive missing assessments, without any prior radiological	Date of the last PSA assessment before the consecutive missing PSA assessments or date of last radiological	Yes

progression event and before or at data cut-off	assessment, whichever comes later	
No castration-resistant prostate cancer (or no event among the three components) before or at data cut-off date	Date of the latest date among the three components' last assessment before discontinuation or randomization date (censored at Day 1 if there is no follow-up available), whichever is later	Yes
Subsequent systemic antineoplastic therapy without any prior components event, without post PSA progression event and before or at data cut-off	Date of last radiological assessment before or on subsequent systemic antineoplastic therapy start date or the last PSA assessment date or randomization date (censored at Day 1 if there is no follow-up available), whichever comes later	Yes
# This includes subjects with superscan at baseline for radiological bone lesion assessment		

A summary of first castration-resistant prostate cancer events will be provided. Patients with multiple events are only counted for the category in which the first event occurred. If multiple CRPC events (component events) occur on the same date for one subject, the subject is only counted into one category in the order of: radiological soft tissue/ visceral lesion progression > radiological bone progression > PSA progression.

The primary analysis of time to castration-resistant prostate cancer will be based on central PSA assessments only. The sensitivity analysis of time to castration-resistant prostate cancer will be based on both central and local PSA laboratory data.

6.2.2.2 Time to pain progression

Time to pain progression is defined as the time from randomization to the first date a patient experiences pain progression (first date of pain progression minus date of randomization). Pain will be assessed with the BPI-SF questionnaire. The main analysis is based on the ePRO device questionnaire data only. Pain progression is defined as follows:

- For asymptomatic patients (worst pain subscale [WPS] 0 at baseline), pain progression is defined as
    - an increase of 2 or more points in the "worst pain in 24 hours" score (i.e. 2 or more point increase in WPS score) from nadir (i.e. zero) observed at 2 consecutive evaluations ≥4 weeks apart, or
    - initiation of short- or long-acting opioid use for pain.
- For symptomatic patients (WPS >0 at baseline), pain progression is defined as
    - an increase of 2 or more points in the "worst pain in 24 hours" (i.e. 2 or more point increase in WPS score) from nadir observed at 2 consecutive evaluations ≥4 weeks apart and a WPS of ≥4, or
    - initiation of short- or long-acting opioid use for pain.

The following censoring rules in Table 6–3 will be applied:

Table 6–3 Censoring rules for time to pain progression

Situation	End Date	Censored
Pain progression during study	Start date of pain progression	No
No baseline or post-baseline event assessment	Date of randomization (censored at Day 1 if there is no follow-up available)	Yes
No pain progression before or at data cut-off date	The last BPI-SF assessment date or randomization date (censored at Day 1 if there is no follow-up available), whichever comes later	Yes
Opioids taken for any reason within 4 weeks before or on randomization	Date of randomization (Day 1)	Yes

Sensitivity analysis for time to pain progression

- Pain progression will be evaluated with an increase of 2 or more points in the "worst pain in 24 hours" score (i.e. 2 or more point increase in WPS score) from baseline.
- Pain progression will be evaluated with an increase of 2 or more points in the "worst pain in 24 hours" score (i.e. 2 or more point increase in WPS score) from nadir after completion of docetaxel.
- Pain progression will be evaluated with an increase of 2 or more points in the "worst pain in 24 hours" score (i.e. 2 or more point increase in WPS score) from baseline after completion of docetaxel.
- The main analysis is based on the ePRO device questionnaire data only. Pain progression will be evaluated in this sensitivity analysis with source data from both the ePRO device and the paper questionnaire, based on the main analysis definition (change from nadir).

6.2.2.3 Symptomatic skeletal event free survival

Symptomatic skeletal event free survival (SSE-FS) is defined as the time from randomization to the first occurrence of an SSE or death from any cause, whichever comes first (date of first occurrence of an SSE or death minus date of randomization).

An SSE is defined as the occurrence of one of the followings, whichever comes first:

- external beam radiation therapy (EBRT) to relieve skeletal symptoms, or
- new symptomatic pathologic bone fracture, or
- occurrence of spinal cord compression, or
- tumor-related orthopaedic surgical intervention.

The following censoring rules in Table 6–4 will be applied:

Table 6–4 Censoring rules for SSE-FS

Situation	End Date	Censored
SSE during study	Date of the first assessment reporting event	No
Death during study without SSE event	Date of death	No
No SSE before or at data cut-off date	Last SSE assessment before or at data cut-off	Yes
Patient lost to follow-up before or at data cut-off date	Date of last SSE assessment or randomization date (censored at Day 1 if there is no follow-up available), whichever comes later	Yes

A summary of SSE-FS events will be provided. Patients with multiple events are only counted for the category in which the first event occurred. If multiple SSEs (component events) occur on the same date for one patient, the patient is only counted into one category in the order of: spinal cord compression > bone fracture > orthopedic surgery > EBRT.

6.2.2.4 Time to first SSE

Time to first SSE is defined as the time from randomization to the first occurrence of an SSE. Death is not considered as event in this endpoint (date of first SSE minius date of randomization); otherwise the same rules as for SSE-FS apply.

The following censoring rules in Table 6–5 will be applied:

Table 6–5 Censoring rules for first SSE

Situation	End Date	Censored
SSE during study	Date of the first assessment reporting event	No
No SSE before or at data cut-off date	Last SSE assessment before or at data cut-off	Yes
Patient lost to follow-up before or at data cut-off	Date of last SSE assessment or randomization date (censored at Day 1 if there is no follow-up available), whichever comes later	Yes

A summary of first SSE will be provided. Patients with multiple events are only counted for the category in which the first event occurred. If multiple SSEs (component events) occur on the same date for one patient, the patient is only counted into one category in the order of: spinal cord compression > bone fracture > orthopedic surgery > EBRT.

6.2.2.5 Time to initiation of subsequent systemic antineoplastic therapy

Time to initiation of subsequent systemic antineoplastic therapy[2] is defined as the time from randomization to initiation of the first subsequent systemic antineoplastic therapy (first subsequent systemic antineoplastic therapy date minus date of randomization).

Systemic antineoplastic therapy treatment was selected in CM dataset, which includes "SYSTEMIC ANTI-CANCER THERAPY (MEDICATION) DURING FOLLOW-UP" eCRF pages and CONCOMITANT MEDICATION" and "PRIOR AND CONCOMITANT MEDICATION" eCRF pages. Docetaxel administrated concomitantly with study drug that continued after study drug stopped will not be counted as subsequent systemic antineoplastic therapy.

From above 3 eCRF pages, subsequent systemic antineoplastic therapy is identified with criteria available in Appendix 9.2. The sponsor's assessment of systemic antineoplastic therapy is presented in a supplementary document. The final list of systemic antineoplastic therapy will be based on the sponsor's assessment and medical review before database lock.

The following censoring rules in Table 6–6 will be applied:

Table 6–6 Censoring rules for time to initiation of subsequent systemic antineoplastic therapy

Situation	End Date	Censored
Initiation of first subsequent systemic antineoplastic therapy during study	Start date of first subsequent systemic antineoplastic therapy	No
No subsequent systemic antineoplastic therapy before or at data cut-off date	The last know alive date or death date or randomization date (censored at Day 1 if there is no follow-up available), whichever comes later	Yes

6.2.2.6 Time to worsening of disease-related physical symptoms

Time to worsening of disease-related physical symptoms is defined as the time from randomization to the first date a patient experiences an increase in disease-related physical symptoms based on the NCCN-FACT-FPSI-17 questionnaire (date of first worsening of disease-related physical symptoms minus date of randomization). An increase in disease-related physical symptoms is defined as 3-point decrease of FPSI-DRS-P subscale (a lower score = higher symptom burden, a score of "0" is a severely symptomatic patient and the highest possible score is an asymptomatic patient) from baseline in the disease-related physical symptoms subscale (FPSI-DRS-P subscale of the NCCN-FACT-FPSI-17 questionnaire) observed at 2 consecutive evaluations ≥4 weeks apart.

[2] Antineoplastic therapy is the same as "anti-cancer therapy" in this study.

The following censoring rules in Table 6–7 will be applied:

Table 6–7 Censoring rules for time to worsening of disease-related physical symptoms

Situation	End Date	Censored
Worsening of disease-related physical symptoms during study	Start date of worsening of disease-related physical symptoms	No
No baseline or post-baseline event assessment	Date of randomization (censored at Day 1 if there is no follow-up available)	Yes
No worsening of disease-related physical symptoms before or at data cut-off date	Date of last assessment, or randomization date (censored Day 1 if there is no follow-up available), whichever comes later	Yes

The main analysis is based on the ePRO device questionnaire data only. In addition, time to worsening of disease-related physical symptoms will be evaluated as a sensitivity analysis with source data from both the ePRO device and the paper questionnaire.

6.2.2.7 Time to initiation of opioid use for ≥7 consecutive days

Time to initiation of opioid use for ≥7 consecutive days is defined as the time from randomization to the date of first opioid use for ≥7 consecutive days (date of first opioid use for ≥7 consecutive days minus date of randomization). Data of opioid use related to prostate cancer pain will be included. Opioid use for non-malignant cause will be excluded. Time to opioid use will be determined by opioid use captured via electronic case report form.

The following censoring rules in Table 6–8 will be applied:

Table 6–8 Censoring rules for time to initiation of opioid use for ≥7 consecutive days

Situation	End Date	Censored
Opioid use for ≥7 consecutive days after randomization during study	Start date of initiation of opioid use for ≥7 consecutive days for prostate cancer	No
No opioid use for ≥7 consecutive days before or at data cut-off date	Date of last visit at which analgesic consumption question was collected or randomization date (censored at Day 1 if there is no follow-up available), whichever comes later.	Yes
Opioid use for ≥7 consecutive days at/before randomization date	Date of randomization (Day 1)	Yes

A summary of initiation of opioid use for ≥7 consecutive days will be provided.

6.2.2.8 Analyses of Secondary efficacy endpoints

The secondary efficacy variables will be analyzed for the FAS population unless otherwise specified. Time-to-event endpoints will be analyzed using stratified log-rank test with

randomization stratification factors using IxRS data. Hazard ratio and 95% CI will be provided using the Cox model stratified by the same factors as were used for randomization.

Median time, the 25th and 75th percentiles and associated 95% CI of KM estimates will be presented by treatment arm as well as the number and percentage of censored observations. A KM curve will be generated for each treatment arm.

6.2.3 Exploratory Efficacy Endpoints

6.2.3.1 Rate of absolute PSA response at 6 and 12 months

Absolute PSA response is defined as baseline PSA value above the detection limit and a post baseline PSA level below 0.2 ng/mL, confirmed by a second subsequent PSA value below 0.2 ng/mL 3 or more weeks later, with all potential PSA values between initial date and confirmation date below 0.2 ng/mL.

Rate of absolute PSA response is defined as the number of patients with absolute PSA response, divided by the total number of patients randomized. Rate of absolute PSA response will be evaluated on patient data up to 6 months and 12 months after randomization.

6.2.3.2 Rate of relative PSA response at 3, 6 and 12 months

Relative 30% PSA response is defined as baseline PSA value above the detection limit and a post baseline ≥30% reduction of the PSA level compared to the baseline value, confirmed by a second subsequent PSA value with a ≥30% reduction from baseline 3 or more weeks later, with all potential PSA values between initial date and confirmation date showing a ≥30% reduction from baseline. Relative 50% and 90% PSA response is defined in the same way.

Rate of relative PSA response is defined as the number of patients with relative PSA response, divided by the total number of patients randomized. Rate of relative PSA response will be evaluated on patient data up to 3 months, 6 months, and 12 months after randomization.

The comparison of absolute PSA response at 6 and 12 months between two treatment arms will be done using the Cochran-Mantel-Haenszel test stratified by the IxRS stratification factors. Same comparison will be performed for relative PSA response at 3, 6 and 12 months.

In addition, descriptive statistics and frequency distribution (no decline, < 30%, 30% − < 50%, 50% − < 90%, ≥ 90%) will be provided for PSA maximum percent decline from baseline at any time on study.

6.2.3.3 Time to PSA progression

The time to PSA progression is defined as the time from the date of randomization to the date of first PSA progression with testosterone at castrate level < 0.5 ng/mL. The same definition of PSA progression with testosterone at castrate level < 0.5 ng/mL and censoring rules applied to time to castration-resistant prostate cancer in Section 6.2.2 will be applied here. Baseline PSA value is the last non-missing observation on or before the first day of study drug intake for treated patients. For not treated patients the baseline value will be the last non-missing observation on or before the randomization date . Time to PSA progression will be descriptively evaluated using KM estimates for quantiles, including 95% two-sided CIs. The following censoring rules in Table 6–9 will be applied:

Table 6–9 Censoring rules for PSA progression

Situation	End Date	Censored
PSA progression event during study	Event assessment date	No
No baseline or post-baseline event assessment	Date of randomization (censored at Day 1 if there is no follow-up available)	Yes
PSA progression event immediately after two or more consecutive missing assessments	Date of the last PSA assessment before the consecutive missing ones	Yes
No PSA progression before or at data cut-off date	Date of last PSA assessment before discontinuation or randomization date (censored at Day 1 if there is no follow-up available), whichever comes later	Yes

The primary analysis of PSA response and progression will be based on central PSA assessments only. A sensitivity analysis of PSA response and progression will be based on both central and local laboratory data.

6.2.3.4 ECOG PS

ECOG PS will be summarized with descriptive statistics and frequency tables. Changes from baseline to worst post-baseline ECOG PS score during treatment will be summarized in a shift tables by treatment arm. In addition, changes from worst post-baseline ECOG PS score during treatment to EOT score will be summarized in a shift table by treatment arm.

Baseline ECOG value is the last non-missing observation on or before the randomization date.

6.2.3.5 Quality of life

PRO data as measured by the NCCN-FACT FPSI-17 and the BPI-SF will be analyzed. Quality of Life and PRO data analyses specified in this section are based on the ePRO device questionnaire data only.

The PRO analyses will be performed for the patients in the FAS. Statistical tests will be performed with a one-sided test with type I error 0.025 (equivalent a two-sided test with type I error 0.05). The p-values are for descriptive purposes only.

Descriptive statistics on observed data will be presented for the NCCN-FACT FPSI-17 questionnaire (total score and each domain subscale score), and for BPI-SF questionnaire (pain severity and pain interference scores) at each assessment time and for change from baseline by treatment arm. Questionnaires under unscheduled visits and not planned visits per protocol will not be displayed in the descriptive tables. Analyses will be done for patients with baseline assessments.

The frequency for missing health related quality of life (HRQoL) assessment by treatment arm will be summarized.

An analysis of covariance (ANCOVA) model (a mixed linear model, with a random coefficient) for BPI-SF pain severity or interference scores in the time-adjusted Area under Curve (AUC) will be used to estimate the mean difference and treatment effect between the two treatment arms.

Least-square mean estimates, standard errors and 95% CIs will be estimated for each treatment arm and for the treatment arm difference.

Calculation of Time (AUC):

AUC will not be calculated if baseline data is missing.

The trapezoidal rule will be used to derive the AUC for a patient for BPI-SF pain severity and interference scores. The time-adjusted AUC will be calculated by dividing the AUC by the duration (in days) over the treatment period. When calculating the AUC, the exact date of completion of the questionnaires should be plotted and the duration will be calculated in days. The time adjusted AUC for BPI-SF pain severity and interference scores for an individual patient over a period of time [Ta,Tb] will be calculated as follows:

$$AUC_{a\text{-}b} = \frac{1}{2(T_b - T_a)} \sum_{i=a}^{b-1} (PRO_i + PRO_{i+1})(T_{i+1} - T_i)$$

Where PRO_i is the PRO measurement at time T_i (i=a… b).

6.2.4 Subgroup Analyses for Efficacy Endpoints

Subgroup analyses will be conducted for the primary efficacy endpoint OS, based on the FAS population. Descriptive statistics and HR estimates with 95% CI will be provided at least for the subgroups listed below, provided that at least 10 total events are observed within the subgroup across the treatment arms. HRs will be presented in forest plots. All subgroups analyses will be performed using an unstratified Cox model. Number of patients, number of events and KM estimates for the median per arm for all planned subgroups will be presented in the forest plot.

- Stratification Factor based on eCRF: Extent of disease (non-regional lymph nodes metastases only, bone metastases with or without lymph node metastases, or visceral metastases with or without lymph node metastases or with or without bone metastases)
- Stratification Factor based on eCRF: ALP at baseline (<ULN, ≥ULN)
- Age category (<65, 65-74, 75-84, ≥85 years)
- Race (White, Asian, Black or African American, Other)
- Geographical region (North America, Asia Pacific, Rest of the World)
- PSA values (<median of overall population, ≥median of overall population) at baseline
- ECOG PS at baseline (0, 1)
- Gleason score (<8, ≥8) at initial diagnosis
- Metastases at initial diagnosis (Yes: Stage IV-M1, No: Stage I, IIA, IIB, III, IV-M0)

Further important baseline cancer characteristics may also be considered.
Subgroup analyses will not be performed for the sensitivity analyses.

6.3 Safety

The safety endpoints include:

- AEs until the End of Treatment visit
- SAEs until the End of Treatment visit
- Study drug-related SAEs in Active Follow-up for up to one year
- Study drug-related SAEs until the end of Long-term Follow-up
- Laboratory safety assessments (hematology, clinical chemistry, urinalysis)
- Vital signs: blood pressure and heart rate
- 12-lead electrocardiogram (ECG)

6.3.1 Adverse Events

AEs will be coded using MedDRA 24.0 or most recent version. The severity (or intensity) of AEs will be documented using the National Cancer Institute-Common Terminology Criteria for Adverse Event version 4.03 (NCI-CTCAE v 4.03). Descriptive summary tables will be presented on all safety parameters by treatment arm.

The AEs will be presented with their worst NCI-CTCAE v 4.03 grade CTC 1 = mild, CTC 2 = moderate, CTC 3 = Severe, CTC 4 = life threatening, CTC 5 = fatal. The causal relationship is based on whether there was a reasonable causal relationship to darolutamide or placebo and docetaxel. AEs will be classified by the investigator as related or not related to darolutamide or placebo and docetaxel separately.

Treatment-emergent AE (TEAE) is defined as any event arising or worsening after the first dose of darolutamide or placebo until 30 days after the last dose of darolutamide or placebo.

An overview of TEAEs will be summarized by treatment arm in the overall treatment period. The following summaries will be created separately for study drug-related (darolutamide or placebo-related) and docetaxel-related TEAEs, by treatment arm: TEAEs leading to permanent discontinuation of study treatment, TEAEs leading to dose reduction and/or dose interruption, and drug-related TEAEs. Summary statistics (frequency and percentage of patients, not of events) will be presented by treatment arm, MedDRA SOC and PT, and NCI-CTCAE worst grade for the following:

- Incidence of all TEAEs
- Incidence of TEAEs of NCI-CTCAE grade 3, 4 or 5
- Incidence of TEAEs leading to permanent discontinuation of darolutamide or placebo and docetaxel separately
- Incidence of TEAEs leading to dose reduction
- Incidence of TEAEs leading to dose interruption
- Incidence of TEAEs with incidence of at least 5%
- Incidence of all drug-related TEAEs
- Incidence of drug-related TEAEs of NCI-CTCAE grade 3, 4 or 5
- Incidence of drug-related TEAEs leading to permanent discontinuation of darolutamide or placebo and docetaxel separately

- Incidence of drug-related TEAEs leading to dose reduction
- Incidence of drug-related TEAEs leading to dose interruption
- Incidence of drug-related TEAEs with incidence of at least 5%
- Incidence of pre-treatment AEs and post-treatment AEs (after 30 days from the last dose of study drug).
- Incidence and prevalence of most common TEAEs ≥ 10% incidence in either treatment arm, analyzed within pre-specified time intervals. The same summary will be presented by geographical regions (North America, Asia Pacific, Rest of the World) as well.
- Incidence of COVID-19 related TEAEs
- The following subgroup analyses will be performed for overview of TEAEs and for TEAEs showing at least 5% point differences in incidence proportions between any of the subgroup categories .
    - Age category (<65, 65–74, 75–84, ≥85 years)
    - Geographical region (North America, Asia Pacific / Rest of the World
    - Renal function – eGFR at baseline (normal, mild impairment, moderate impairment, severe impairment)
    - Hepatic function at baseline (normal, mild, moderate, severe impairment)
    - Concomitant statin use (no, yes) to be determined by concomitant medication with any medication in the Standardized Drug Grouping (SDG) 'Statins'.
- Incidence of the TEAEs in the following will be presented by groupings and worst CTCAE grade.
    - Patients with present medical history of cardiac disorder (SOC).
    - Patients with present medical history of renal impairment (SMQ Acute renal failure; SMQ Chronic kidney disease).
    - Patients with present TEAE interstitial lung disease (not meidcal history, SMQ interstitial lung disease (narrow search)).
    - Patients with present TEAE neutropenia (SMQ of Haematopoietic leukopenia (narrow search)), by concomitant G-CSF use, where G-CSF is defined by ATC code: L03AA. The same summary will be presented by geographical regions (North America, Asia Pacific, Rest of the World) as well.

TEAE Groupings Considered as Special Topics

The following TEAE groupings are considered as special topics and shown in Appendix 9.6, Table 9–2:

- Bone fractures excluding pathological fractures
- Diabetes mellitus and hyperglycemia
- Fall
- Fatigue/ asthenic conditions

- Weight decreased
- Rash
- Seizure
- Hypertension
- Vasodilatation and flushing
- Mental impairment disorders
- Depressed mood disorders
- Breast disorders/gynecomastia
- Cardiac disorders
- Cerebral ischaemia
- Cerebral and intracranial hemorrhage

The following tables will be created for TEAEs of special topics:

- Incidence of TEAEs
- Incidence of TEAEs leading to permanent discontinuation of (darolutamide or placebo and docetaxel separately)
- Incidence of TEAEs leading to dose reduction
- Incidence of TEAEs leading to dose interruption
- Incidence of treatment-emergent SAEs (TESAEs)
- Incidence and prevalence of most common TEAEs of special topics ≥ 5% incidence in either treatment arm, analyzed within pre-specified time intervals. The same summary will be presented by geographical regions (North America, Asia Pacific, Rest of the World) as well.

Fracture events will be described by a cumulative incidence plot of fracture. A summary of treatment-emergent fracture by bone health agent use (for selection of bone health agent see Appendix 9.7) at study entry will be provided. Association with weight decrease will be presented by histogram of fracture events by patient weight change (the weight collected closest to the start of fracture will be considered and compared to baseline weight). A listing of patients with fracture events breakdown by relevant events occurring within 7 days window: fall, accident, syncope and/or loss of consciousness, dizziness will be presented.

A listing will be generated for patients with fall treatment-emergent events with syncope and/or loss of consciousness (using the MLG Syncope). Timing of occurrence of fracture events based on first dose of study drug will be presented. A summary of time to fall and fractures events will be presented in descriptive and inferential statistics, respectively. In addition, Kaplan-Meier plots will be created to display the time to fall and time to fracture.

An overview of patients at risk for developing an AE seizure will be tabulated by displaying medical history.

To adjust for unequal lengths of study treatment (darolutamide or placebo and docetaxel) between the treatment arm, additional summaries based on event rate per 100 patient year will be performed for all TEAEs, special topics TEAEs, and all TESAEs occurring after the first

dose of darolutamide or placebo. The exposure adjusted incidence rate is calculated as the total number of patients with a given TEAE divided by the total darolutamide or placebo treatment duration of all patients in years. The rate is expressed in 100 patient years.

The treatment duration in years will be calculated as treatment duration in days divided by 365.25. The risk difference 'darolutamide+docetaxel arm – placebo+docetaxel arm' and risk ratio 'darolutamide+docetaxel arm vs. placebo+docetaxel arm' will be calculated together with 95% confidence intervals (CIs). No zero-cell correction will be applied for calculation of the risk ratio, i.e. in case of no events in the placebo+docetaxel arm the risk ratio will not be calculated. The incidence rate ratio for darolutamide+docetaxel arm vs. placebo+docetaxel arm will also be calculated.

6.3.2 Deaths and Serious Adverse events

Serious adverse events (SAEs) will be classified using the NCI-CTCAE version 4.03 and MedDRA version 24.0 or most recent version.

- Treatment-emergent SAEs
- Treatment-emergent SAEs leading to permanent discontinuation of study treatment
- Treatment-emergent SAEs leading to dose reduction
- Treatment-emergent SAEs leading to drug interruption
- Treatment-emergent drug-related SAEs
- COVID-19 related treatment-emergent SAEs
- Listing of treatment-emergent SAEs
- Listing of non-treatment-emergent SAEs

The incidence of deaths in the study: within 30 days after first dose of study drug / docetaxel, during study treatment from first to last dose of study drug / docetaxel, within 30 days after last dose of study drug/ docetaxel, later than 30 days after last dose of study drug/ docetaxel, will be summarized by each treatment arm and cause of death. All deaths before treatment start, during treatment and within 30 days after last dose of study drug, and later than 30 days after last dose of study drug, will be listed by patient with start and stop date of study medication, date of death, and cause of death.

6.3.3 Additional Primary Malignancies

Summary statistics (frequency and percentage of patients, not of events) will be presented by treatment arm, MedDRA SOC and PT for additional primary malignancies. A listing will be generated for patients with additional primary malignancies.

The additional primary malignancies are defined by: SMQ Malignant tumors, excluding: the following terms:

- PT: Prostate cancer
- PT: Prostate cancer metastatic
- PT: Prostate cancer recurrent

- PT: Prostate cancer stage 0
- PT: Prostate cancer stage I
- PT: Prostate cancer stage II
- PT: Prostate cancer stage III
- PT: Prostate cancer stage IV
- PTs from HLGT: Metastases
- PT: Cancer in remission
- LLT: Progression of pre-existing cancer

6.3.4 Laboratory Safety Assessments

Descriptive statistics (number of observations, mean, standard deviation, minimum, median and maximum values) will be presented for clinical laboratory tests (hematology, clinical biochemistry and urinalysis), their changes from baseline (including baseline value), and their percent changes from baseline by treatment arm at applicable visits.

Hematological and biochemical laboratory values will be graded based on NCI CTCAE version 4.03. CTCAE severity grading for laboratory abnormalities are based on applicable laboratory threshold values outlined in NCI CTCAE v4.03. It should be noted that in the present analysis of those laboratory parameters for which additional clinical information potentially can also influence the toxicity grade, this clinical information is in general not available and only the laboratory measurements are used for grading.

Any additional specific handling of the NCI CTCAE v4.03 toxicity grading assignments will be noted in the footnotes of the corresponding tables as applicable per the data collection in the study.

- In the event of overlapping CTCAE criteria ranges for specific lab tests, the algorithm assigns the worst grade
- If calcium type is not recorded (i.e. only “calcium” is recorded), then grading is done as if the calcium is total calcium. “Calcium corrected” is computed from total calcium and serum albumin (if ≤ 4.0 g/dl) from the same time point based on CTCAE v3.0 guidance. If serum albumin (if ≤ 4.0 g/dl) from the same time point is not available or if “calcium, unspecified” was collected then grading is done as if the calcium is ”corrected calcium.”
- Results with special characters (such as “>” and “<”) will be imputed with the lower or higher detection limit (refer to Section 4.3).

The frequency of laboratory abnormalities regarding hematology, international normalized ratio (INR),clinical chemistry, and urinalysis will be tabulated by visit and treatment arm. Worst grades for hematological and biochemical toxicities will be calculated according to CTCAE, v4.03 based on laboratory measurements, and will be summarized by treatment arm and NCI CTCAE v4.03 category and worst grade.

Clinical laboratory toxicities during treatment including a period of 30 days after last dose of treatment will be considered as “treatment-emergent.

The last non-missing value before or on the first day of study drug will be retained as "baseline" data. If several assessments are performed on the same day (without timing information) the average of the values will be considered.

Incidence tables (frequency and percentage of patients) as well as tables with change in NCI CTCAE v4.03 worst grade from baseline will be presented as following:

- Hematological and biochemical toxicity during screening (the last evaluation available before assignment to treatment is taken into account).
- Treatment-emergent hematological and biochemical toxicity.
- Treatment-emergent hematological and biochemical toxicities with incidence rate above 5% in either treatment arm.
- Change in worst grade for hematological and biochemical toxicity from baseline.

The laboratory values will be also categorized into low, normal and high according to their reference ranges.

Descriptive statistics will be calculated by treatment arm and visit.

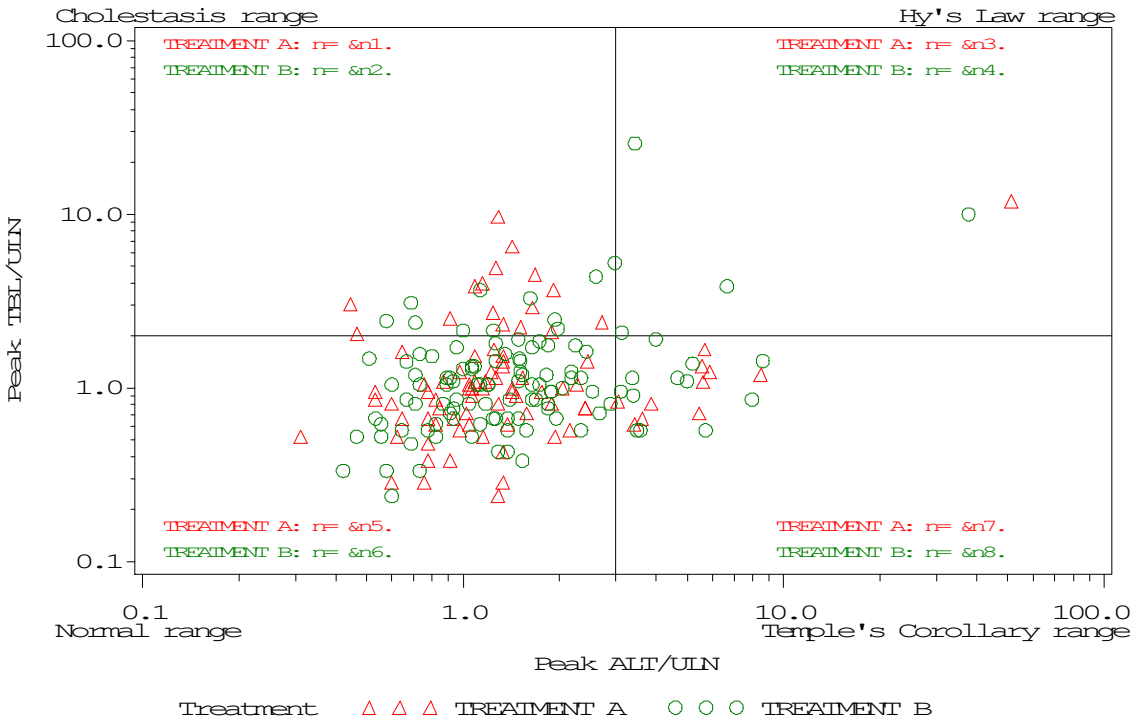
Only for ALT, AST and creatinine, descriptive statistics will be displayed for subgroups of statin users and non-statin users. Clinical laboratory tables will not be described by subgroups.

The data will also be displayed graphically using box plots and laboratory shift plots. These will show the baseline value and post baselines values up-to end of treatment.

In addition, the mean of lab values (from both scheduled and unscheduled assessments) for ALT, AST, Bilirubin, Hemoglobin, Neutrophils and Platelets will be displayed in line plots by treatment arm and time interval (1 week or less) relative to the first concomitant docetaxel treatment date, from baseline up-to Visit 6 (EOT visit will be included if the last visit number before EOT visit is < 7) for safety patients who had concomitant docetaxel with darolutamide or placebo. In the line plots, if one patient has the repeated measures within the same interval, the mean of the repeated measures for the same patient will be used in the mean calculation.

A listing will be provided for all patients possibly fulfilling Hy's Law criteria, i.e. patients with elevated AST and / or ALT > 3xULN, alkaline phosphatase < 2xULN and bilirubin ≥2xULN (ref. 3). For possible Hy's Law cases relevant laboratory data will also be displayed graphically within actual patient profiles (presenting total bilirubin, ALT, AST and ALP values in terms of ULN over time) by treatment arm. Below example of Hy's law plot will be provided for peak total bilirubin vs ALT. If a patient has any total bilirubin ≥ 2xULN then peak bilirubin will be plotted versus the maximum ALT amongst the total bilirubin ≥ 2xULN. Otherwise, peak bilirubin will be plotted versus peak ALT.

Figure 6–1 Example for Hy's law plot



Unscheduled laboratory data will be listed (in Section 14 listings and Section 16 of the CSR),but will not be displayed in the descriptive tables.

6.3.5 Other Safety Measures

12-lead ECG

Analyses of ECG and QTc will be performed on the SAF population; no subgroup tables will be created.

Corrected QT (QTc) will be calculated using Bazett's (QTcB) and Fridericia's (QTcF) formula as below:

- Fridericia's correction: $QTc = QT/RR^{0.33}$
- Bazett's correction: $QTc = QT/RR^{0.5}$

where RR interval = 60 / Ventricular Rate

Standard 12-lead ECGs were performed at screening and at each visit in a supine position after at least 10 minutes rest.

Descriptive statistics including arithmetic mean, SD, median, minimum, and maximum will be presented for the following ECG parameters: Ventricular rate, PR, QRS, RR (derived), QT, QTcB (derived) and QTcF (derived). Parameters will be summarized for actual results and the change from baseline for the safety analysis set by treatment arm at each scheduled visit. Baseline for 12-lead ECG will be the mean of separate recordings. Similarly, post baseline values of 12-lead ECG for patients will be the mean of separate recordings.

Mean time courses +/- one standard deviation will be displayed graphically.

The last non-missing value before the first study drug intake will be considered as "baseline" data.

If several assessments are performed on the same day (without timing information) the average of the values will be considered.

In case the timing of ECG is missing at day 1 then the screening assessment will be considered as baseline data.

The number and percent of patients with absolute QTc interval prolongation or QTc increase from baseline will be defined and summarized by treatment arms. The criteria are:

QTc prolongation: QTc interval ≤ 450 msec, QTc interval 451-480 msec, QTc interval 481-500 msec and QTc interval > 500 msec.

Increase from baseline in QTc interval: QTc interval increases from baseline > 30 msec and QTc interval increases from baseline > 60 msec.

The number of patients with QTC interval (max value) ≤ 450 msec, 451-480 msec, 481-500 msec and > 500 msec data will also be displayed graphically using a bar chart.

The number and percent of patients with new clinically significant abnormalities on ECGs per the investigator's assessment at post-baseline time points will be summarized by treatment arms.

ECG data will be presented in a listing. Values with QTc interval ≤ 450 msec, 451-480 msec, 481-500 msec and > 500 msec, and change from baseline in QTc interval > 30 msec and > 60 msec will be flagged. In addition, abnormalities or clinically significant abnormalities found in the 12-lead ECGs at each scheduled time point will be provided in a listing.

Unscheduled ECG data will be listed (in Section 14 listings and Section 16 of the CSR), but will not be displayed in the summary tables.

Vital signs

For each treatment arm, vital signs (i.e. blood pressure, heart rate, weight and BMI) will be tabulated and summarized by visit for observed values and changes from baseline using descriptive statistics, as appropriate. If more than one baseline assessment is collected, the most recent one will be used. If several assessments are performed on the same day (without timing information) the average of the values will be considered.

In addition to descriptive analysis box plots will be created.

For weight parameter a graph representing mean changes in weight from baseline by visit will be created.

Outlier analyses will be conducted using the following limits:

- low systolic blood pressure: ≤ 90 mmHg and a decrease of ≥ 20 mmHg
- high systolic blood pressure: >190 mmHg and an increase of ≥ 20 mmHg
- low diastolic blood pressure: ≤ 50 mmHg and a decrease of ≥20 mmHg
- high diastolic blood pressure: > 105 mmHg and an increase of ≥ 20 mmHg
- low heart rate: < 50 bpm and a decrease of ≥ 15 bpm

- high heart rate: > 120 bpm and an increase of ≥ 15 bpm

The number and percentage of patients with outlying values will be tabulated by treatment arm and time interval.

Unscheduled vital signs data will be listed (in Section 14 listings and Section 16 of the CSR), but will not be displayed in the summary tables.

Concomitant statins (BCRP substrates) of interest

The number of patients reported with the below concomitant statins which are BCRP substrates will be described with frequency and percentage for the SAF population.

Concomitant statins (BCRP substrates) of interest are defined as:

Atorvastatin (ATC code C10AA05)

Fluvastatin (ATC code C10AA04)

Rosuvastatin (ATC code C10AA07)

Simvastatin (ATC code C10AA01).

More details are available in Appendix 9.8.

A listing with intervals will be created for these subjects when above statins were taken concomitantly with the study-treatment.

Interval will be created as follow: number of days when statins and study-treatment were concomitantly taken plus 14 days.

For partial dates of concomitant statins, the imputation rules in Appendix 9.1will be applied. If the end date is completely missing it will be assumed that the concomitant medication is ongoing at study treatment end. If the start date is completely missing it is assumed that concomitant medication was taken before the start of study treatment.

The frequency of pre-defined adverse events (see Appendix 9.9) will be detailed if occurring during the interval defined above, for the subset of patients taking statins that are BCRP substrates.

6.4 Other Exploratory Variables

- **Medical resource use**

Summary statistics (descriptive statistics and frequency) will be presented by treatment arm for days in hospital and equivalent facility, and days in ICU (intensive care unit) facility, separately. Location/place of healthcare visit will be presented in frequency. Listings of patients with healthcare resource use will be provided.

- **Biomarker evaluations**

Biomarker analyses will be described in a separate biomarker analysis plan. Results from exploratory biomarker analyses will be reported in a separate biomarker report.

6.5 Pharmacokinetics

Pharmacokinetics in two sub-groups: 1) at least the first 20 randomized patients who have received at least 1 cycle of docetaxel and 2) Chinese patients in the PK substudy

The concentration-time courses of all analytes will be tabulated separated by treatment and subgroup. The following statistics will be calculated for each of the sampling points: arithmetic mean, SD, and coefficient of variation (CV); geometric mean, geometric SD (re-transformed SD of the logarithms), and CV; minimum, median, maximum value, and the number of measurements.

Means at any time will only be calculated if at least 2/3 of the individual data were measured and were above the lower limit of quantification (LLOQ). For the calculation of the mean value, a data point below LLOQ will be substituted by one-half of this limit.

In tables showing mean values, where values below LLOQ are included in the calculation of mean values, these means will be marked.

Individual and geometric mean concentration versus time curves of all analytes (using the actual sampling times for individual plots and the planned sampling times for mean plots) will be plotted by treatment using both a linear and semilogarithmic scale. For darolutamide and its diastereomers and metabolite all analytes will be plotted together.

Pharmacokinetic parameters (t_{max}, t_{last} and parameters listed under 'other parameters' excluded) will be summarized by the statistics mentioned above. The t_{max}, t_{last} will be described utilizing minimum, maximum, and median as well as frequency counts.

The diastereomer/darolutamide ratios will be calculated using the following equation:

Diastereomer/darolutamide ratios =

AUC(0-t_{last}) (*S,R*)-darolutamide / AUC(0-t_{last}) darolutamide and

AUC(0-t_{last}) (S,*S*)-darolutamide / AUC(0-t_{last}) darolutamide

Metabolite keto-darolutamide to parent darolutamide AUC(0-t_{last}) ratios will be calculated using the following equation:

Metabolite Ratio =

(AUC(0-t_{last}) metabolite / MW metabolite) / (AUC(0-t_{last}) parent / MW parent)

Molecular weight (MW) keto-darolutamide = 396.8 g/mol

Molecular weight (MW) darolutamide =398.8 g/mol

Individual data points of docetaxel PK parameters Cmax and AUC(0-t_{last}) for treatments of treatments of docetaxel in combination with darolutamide and docetaxel with placebo will be compared with a scatter plot.

Only dense PK sampling data from the first Visit (as well as Day 15 for the Chinese subgroup) will be used for statistical analysis. Sparse PK samples from later visits will only be listed.

Pharmacokinetics in all patients

Individual concentration-time data of darolutamide (BAY 1841788), (S,*S*)-darolutamide (BAY 1896952), (*S,R*)-darolutamide (BAY 1896951), keto-darolutamide (BAY 1896953), and docetaxel will be provided in a clinical study report appendix. The PK data will be used for population PK data anlaysis that will be described in detail in a separate M&S report.

7. Document History and Changes in the Planned Statistical Analysis

- Approval of SAP v1.0, dated 07 APR 2017, was based on the integrated clinical study protocol v2.0 (amendment 1), dated 04 OCT 2016.
- Approval of interim futility SAP v1.0, 11 MAR 2019 was based on SAP v1.0, dated 07 APR 2017 and the integrated clinical study protocol v2.0 (amendment 1), dated 04 OCT 2016.
- Approval of SAP v2.0, dated 03 JUL2019, was mainly an update of the efficacy censoring rules from SAP v1.0 and was based on the integrated clinical study protocol v3.0 (amendment 5), dated 12 FEB 2018
- Approval of interim futility SAP v2.0, 03 JUL 2019, was mainly an update of the efficacy censoring rules from SAP v1.0 and was based on SAP v2.0, dated 03 JUL2019 and the integrated clinical study protocol v3.0 (amendment 5), dated 12 FEB 2018.
- Approval of SAP v3.0, dated 26 MAY 2020, describes the changes of removing interim analysis 2 for futility and efficacy from the testing plan due to data collection limitations during the COVID-19 pandemic and was based on the integrated clinical study protocol v5.0 (amendment 7), dated 26 MAY 2020.
- Approval of SAP v4.0, dated 22 SEP 2021, based on the SAP v3.0, dated 26 MAY 2020, describes the changes as a result of the following:
    - Efficacy changes included applying a hierarchical gatekeeping procedure to the secondary efficacy endpoints, inclusion of a summary of ECOG PS, additional relative 50% and 90% PSA response rate and PSA maximum percent decline from baseline and
    - Safety changes included the analysis of special topics adverse events, additional primary malignancies, liver function laboratory test, 12-lead ECG, etc.
- Approval of SAP v4.1, dated 11 NOV 2021, based on the SAP v4.0, dated 22 SEP 2021, describes the changes as a result of the following:
    - Analysis set changes with an additional statement: Patients will be excluded from all FAS, SAF and PKS if they are related to or associated with any critical GCP violations that result in fraudulent patient data.

      Rationale: After detection of issues with investigator fraud at one site, it was decided to exclude one affected patient from all analysis sets as these data cannot be trusted.
    - Efficacy change included a sensitivity analysis of OS with central imaging review extent of disease as stratification factor.
    - Safety changes included summary of incidence and prevalence of most common TEAEs and special topics TEAEs by geographical regions, and incidence of TEAE neutropenia by concomitant G-CSF use and geographical regions.

8. References

1. Eisenhauer EA, Therasse P, Bogaerts J, Schwartz LH, Sargent D, Ford R, et al. New response evaluation criteria in solid tumours: Revised RECIST guideline (version 1.1). Eur J Cancer. 2009;45:228-47.
2. Scher HI, Morris MJ, Stadler WM, Higano C, Basch E, Fizazi K, 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.
3. U.S. Department of Health and Human Services, Food and Drug Administration, Guidance for Industry Drug-Induced Liver Injury: Premarketing Clinical Evaluation, July 2009.

9. Appendix

9.1 Imputation rule for partial missing Adverse Event/ Concomitant medication dates

Table 9–1 Imputation rule for partial missing dates

Partial Dates Imputation Rule	Impute partial AE/CM Start Date	Impute partial AE/CM Stop Date
The day missing only	IF AESTDT year and months is same as TRTSDT year and months, then impute AESTDT= TRTSDT	IF AEENDT year and months is same as last known alive date (LKAD) year and months, then impute AEENDT= LKAD
	ELSE IF AESTDT year and month is before TRTSDT year and months, then AESTDT= last date of the months	ELSE impute AEENDT= last date of the months
	ELSE IF AESTDT year and month is after TRTSDT year and months, then AESTDT= first date of the months	
Both day and months missing	IF AESTDT year is same as TRTSDT year, then impute AESTDT=TRTSDT	IF AEENDT year is same as last known alive date (LKAD) year, then impute AEENDT= LKAD
	ELSE IF AESTDT year is before TRTSDT year, then impute AESTDT=31DECYYYY	ELSE impute AEENDT=31DECYYYY
	ELSE IF AESTDT year is after TRTSDT year , then impute AESTDT=01JANYYYY	
Completely missing	No need to impute, try to query the sites by DM	No need to impute, try to query the sites by DM
Additional criteria to meet	1. AE/CM start date <= AE/CM stop date 2. The imputed dates <= last known alive date (LKAD) 3. If TRTSDT is missing, use RANDDT as reference date	

9.2 Systemic antineoplastic therapy: time to initiation of subsequent systemic antineoplastic therapy

1. Systemic antineoplastic therapy treatment was selected from eCRF page in 2 and 3 below in CM dataset
2. From “SYSTEMIC ANTI-CANCER THERAPY (MEDICATION) DURING FOLLOW-UP” eCRF pages
    - ATC code class L (antineoplastic and immunomodulating agents): L01 Antineoplastic agents, L02 endocrine therapy, L03 immunostimulants and L04 immunosuppressants.
    - ATC code class V10.
    - ATC code: G03CA, G03CB, V03AX, V, V98.
3. From “SYSTEMIC ANTI-CANCER THERAPY (MEDICATION) DURING FOLLOW-UP” eCRF pages or CONCOMITANT MEDICATION” or “PRIOR AND CONCOMITANT MEDICATION” eCRF pages
    - If the WHO-DD drug record number is 062951(ABIRATERONE), 900315 (ANTINEOPLASTIC AGENTS), 012711 (BICALUTAMIDE), 082361 (RADIUM RA 223 DICHLORIDE.
    - If ATC code or code class is same as any above, but the WHO-DD drug record number is 093203 (MARSDENIA TENACISSIMA STEM), 000826 (MEGESTROL ACETATE), 007269 (LEUPRORELIN ACETATE), 007321 (GOSERELIN ACETATE), 009759 (TRIPTORELIN EMBONATE), 017648 (DEGARELIX ACETATE), 010277 (UBENIMEX), 017660 (THYMALFASIN), 901166 (OTHER THERAPEUTIC PRODUCTS), then do not consider as antineoplastic therapy.

The sponsor’s assessment of antineoplastic therapy is presented in a supplementary document.

9.3 Brief Pain Inventory – Short Form (BPI-SF) questionnaire and scoring information

STUDY ID# _____ HOSPITAL # _____

DO NOT WRITE ABOVE THIS LINE

Brief Pain Inventory (Short Form)

Date: _____/_____/_____ Time:____

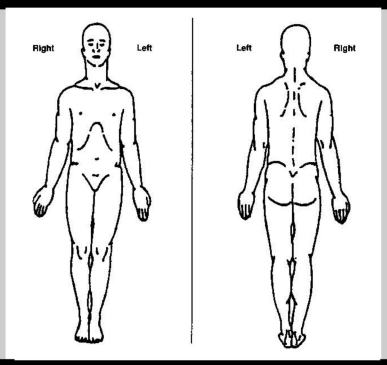
Name: _____ _____ _____

Last First Middle Initial

1. Throughout our lives, most of us have had pain from time to time (such as minor headaches, sprains, and toothaches). Have you had pain other than these every-day kinds of pain today?

   1. Yes 2. No

2. On the diagram, shade in the areas where you feel pain. Put an X on the area that hurts the most.



3. Please rate your pain by circling the one number that best describes your pain at its worst in the last 24 hours.

0	1	2	3	4	5	6	7	8	9	10
No Pain										Pain as bad as you can imagine

4. Please rate your pain by circling the one nuimber that best describes your pain at its least in the last 24 hours.

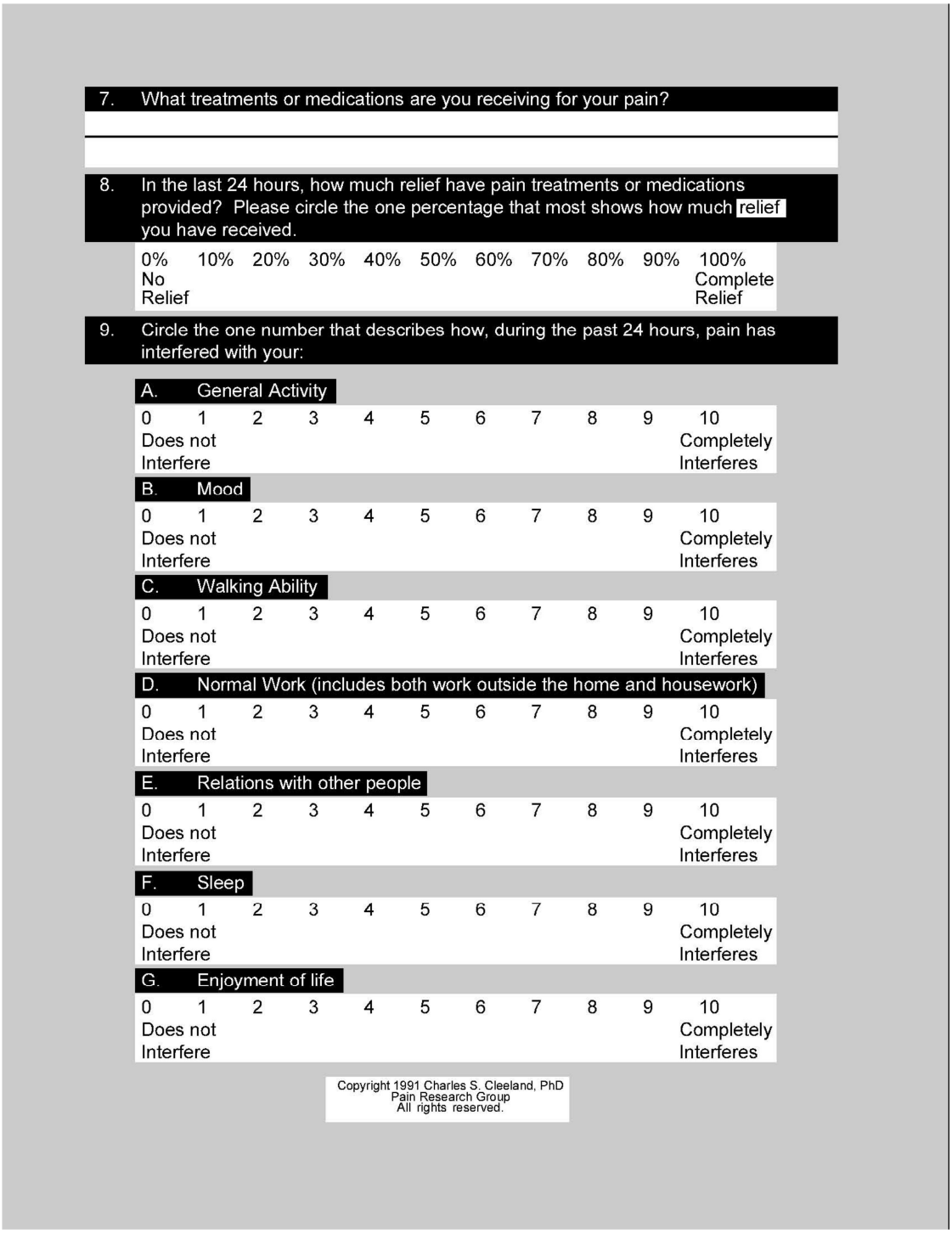
0	1	2	3	4	5	6	7	8	9	10
No Pain										Pain as bad as you can imagine

5. Please rate your pain by circling the one number that best describes your pain on the average.

0	1	2	3	4	5	6	7	8	9	10
No Pain										Pain as bad as you can imagine

6. Please rate your pain by circling the one number that tells how much pain you have right now.

0	1	2	3	4	5	6	7	8	9	10
No Pain										Pain as bad as you can imagine



7. What treatments or medications are you receiving for your pain?

8. In the last 24 hours, how much relief have pain treatments or medications provided? Please circle the one percentage that most shows how much relief you have received.

0% No Relief	10%	20%	30%	40%	50%	60%	70%	80%	90%	100% Complete Relief

9. Circle the one number that describes how, during the past 24 hours, pain has interfered with your:

A. General Activity

0 Does not Interfere	1	2	3	4	5	6	7	8	9	10 Completely Interferes

B. Mood

0 Does not Interfere	1	2	3	4	5	6	7	8	9	10 Completely Interferes

C. Walking Ability

0 Does not Interfere	1	2	3	4	5	6	7	8	9	10 Completely Interferes

D. Normal Work (includes both work outside the home and housework)

0 Does not Interfere	1	2	3	4	5	6	7	8	9	10 Completely Interferes

E. Relations with other people

0 Does not Interfere	1	2	3	4	5	6	7	8	9	10 Completely Interferes

F. Sleep

0 Does not Interfere	1	2	3	4	5	6	7	8	9	10 Completely Interferes

G. Enjoyment of life

0 Does not Interfere	1	2	3	4	5	6	7	8	9	10 Completely Interferes

Scoring:

Pain Severity Score = mean of items 3-6 (pain at its worst, pain at its least, pain on the average, pain for right now)

Pain Interference Score = mean of items 9A-9G (interference of pain with: general activity, mood, walking ability, normal work, relations with other people, sleep, and enjoyment of life).

9.4 NCCN-FACT FPSI-17 questionnaire and scoring information

Below is a list of statements that other people with your illness have said are important.

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

			Not at all	A little bit	Some what	Quite a bit	Very much
DRS-P	GP1	I have a lack of energy.....	0	1	2	3	4
	GP4	I have pain	0	1	2	3	4
	P7	I have difficulty urinating	0	1	2	3	4
	C2	I am losing weight	0	1	2	3	4
	BP1	I have bone pain.....	0	1	2	3	4
	HI7	I feel fatigued.....	0	1	2	3	4
	NCCN3	I have weakness in my legs	0	1	2	3	4
	P3	My pain keeps me from doing things I want to do	0	1	2	3	4
	C6	I have a good appetite.....	0	1	2	3	4
	GF5	I am sleeping well.....	0	1	2	3	4
DRS-E	GE6	I worry that my condition will get worse .	0	1	2	3	4
	GP2	I have nausea	0	1	2	3	4
	P6	I have trouble moving my bowels	0	1	2	3	4
TSE	GS7	I am satisfied with my sex life.....	0	1	2	3	4
	GP5	I am bothered by side effects of treatment	0	1	2	3	4
FWB	GF3	I am able to enjoy life.....	0	1	2	3	4
	GF7	I am content with the quality of my life right now.....	0	1	2	3	4

DRS-P=Disease-Related Symptoms Subscale – Physical
DRS-E=Disease-Related Symptoms Subscale – Emotional
TSE=Treatment Side Effects Subscale
FWB=Function and Well-Being Subscale
English (Universal), Copyright 2001 03 March 2010

NCCN/FACT Prostate Symptom Index-17 (FPSI-17) Scoring Guidelines
(Version 2)

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 symptom index score.
4. As with all FACIT questionnaires, a high score is good. Therefore, a score of “0” is a severely symptomatic patient and the highest possible score is an asymptomatic patient.

Scale	Item Code	Reverse item?		Item response	Item Score
FPSI-17 Total	GP1	4	-	_____	=_____
	GP4	4	-	_____	=_____
	P7	4	-	_____	=_____
Score range: 0-68	C2	4	-	_____	=_____
	BP1	4	-	_____	=_____
	HI7	4	-	_____	=_____
	NCCN3	4	-	_____	=_____
	P3	4	-	_____	=_____
	C6	0	+	_____	=_____
	GF5	0	+	_____	=_____
	GE6	4	-	_____	=_____
	GP2	4	-	_____	=_____
	P6	4	-	_____	=_____
	GS7	0	+	_____	=_____
	GP5	4	-	_____	=_____
	GF3	0	+	_____	=_____
	GF7	0	+	_____	=_____

Sum individual item scores: _____
Multiply by 17: _____
Divide by number of items answered: _____=**FPSI-17 score**

Subscale	Item Code	Reverse item?		Item response	Item Score
FPSI-DRS-P (Disease Related Symptoms-Physical)	GP1	4	-	_____	=_____
	GP4	4	-	_____	=_____
	P7	4	-	_____	=_____
Score range: 0-40	C2	4	-	_____	=_____
	BP1	4	-	_____	=_____
	HI7	4	-	_____	=_____
	NCCN3	4	-	_____	=_____
	P3	4	-	_____	=_____
	C6	0	+	_____	=_____
	GF5	0	+	_____	=_____

Sum individual item scores: _____
Multiply by 10: _____
Divide by number of items answered: _____=**FPSI-DRS-P score**

Subscale	Item Code	Reverse item?		Item response	Item Score
FPSI-DRS-E (Disease Related Symptoms-Emotional)	GE6	4	-	_____	=_____=**FPSI-DRS-E score**
Score range: 0-4					
FPSI-TSE (Treatment Side Effects)	GP2	4	-	_____	=_____
	P6	4	-	_____	=_____
	GS7	0	+	_____	=_____
	GP5	4	-	_____	=_____
Score range: 0-16					
				Sum individual item scores: _____	
				Multiply by 4: _____	
				Divide by number of items answered: _____	=**FPSI-TSE score**
FPSI-F/WB (Function/ Well-Being)	GF3	0	+	_____	=_____
	GF7	0	+	_____	=_____
Score range: 0-8					
				Sum individual item scores: _____	
				Multiply by 2: _____	
				Divide by number of items answered: _____	=**FPSI-F/WB score**

Missing data:

For NCCN-FACT FPSI-17, total scores and subscales scores (disease-related physical symptoms, disease–related emotional symptom, treatment side-effects, function/well-being) will be assessed. Where there are missing items, subscale scores can be prorated if >50% of items on subscale are completed. If ≤50% of the items are answered for any domain, then the score of that domain is set to missing. The total score is set to missing if the related overall item response rate is ≤80%. (e.g., at least 14 of 17 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 >50% of items are answered. In addition, a total score should only be calculated if ALL of the component subscales have valid scores.

9.5 Statins medications

Statins will be identified using the Standardized Drug Grouping (SDG) classification ‘Statins’.

9.6 Special topic AE definition

Table 9–2 Special topic AE definitions

Grouped term	MedDRA search criteria for special topics
Bone fractures excluding pathological fractures	HLT: Fractures and dislocations NEC (without PTs: Joint dislocation, Joint dislocation pathological) HLT Limb fractures and dislocations (without PT: Radial head dislocation) HLT Pelvic fractures and dislocations HLT Skull fractures, facial bone fractures and dislocations HLT Spinal fractures and dislocations (without PT: Dislocation of vertebra) HLT Thoracic cage fractures and dislocations (without PT: Dislocation of sternum)) Exclude PT pathological fractures
Diabetes mellitus and hyperglycemia	MLG Hyperglycemia PT: Diabetes mellitus; Diabetes mellitus inadequate control; Diabetic metabolic decompensation; Type 2 diabetes mellitus; Diabetic ketoacidosis
Fall*	PT Fall PT Accident
Fatigue/ asthenic conditions	MLG Decreased general strength and energy PT: Lethargy; Chronic fatigue syndrome; Malaise
Weight decreased	MLG: Weight decreased
Rash	MLG: Rash MLG: Skin erythema PT: Dermatitis
Seizure	MLG Seizures
Hypertension	MLG: Hypertension
Vasodilatation and flushing	MLG: Vasodilatation and flushing
Mental impairment disorders	HLGT: Mental impairment disorders
Depressed mood disorders	HLGT: Depressed mood disorders and disturbance
Breast disorders/gynecomastia	HLGT: Breast disorders
Cardiac disorders	HLGT Cardiac arrhythmias HLGT Coronary artery disorders HLGT Heart failures
Cerebral ischaemia	MLG: Cerebral infarction and stroke not specified as hemorrhagic or ischemic MLG: Cerebral ischemic infarction and stroke PT: Cerebral ischaemia, Transient ischaemic attack
Cerebral and intracranial hemorrhage	MLG: Cerebral and intracranial hemorrhage

MLG: MedDRA Labeling Grouping; PT: Preferred Term; HLT: High Level Term; HLGT: High Level Group Term

* Based on review of verbatims the search criteria for “Fall” was extended by the PT “Accident”.

9.7 Bone health agent (BHA)

Bone health agent were selected in CM dataset, using ATC codes:

M05BA, M05BB: Bisphosphonates
M05BX: Denosumab

9.8 Concomitant statins (BCRP substrates) of interest

Concomitant statins (BCRP substrates) of interest

The co-medication statins (BCRP substrates) were selected in CM dataset using ATC codes:

Atorvastatin - ATC code C10AA05
Fluvastatin -ATC code C10AA04
Rosuvastatin -ATC code C10AA07
Simvastatin -ATC code C10AA01

As in CM dataset, ATC code (variable CM5CL01) contains only 5 digits, the first 5 digits of the above ATCs with 7 digits will be used.

In addition, the drug name below will be selected from "standardized medication name" variable (variable CMDECOD). This list was validated per medical team based on the output from the prior selection of ATC.

CMDECOD to select:

ATORVASTATIN
ATORVASTATIN CALCIUM
FLUVASTATIN
FLUVASTATIN SODIUM
ROSUVASTATIN
ROSUVASTATIN CALCIUM
SIMVASTATIN

9.9 Pre-defined adverse events for analysis of statin (BCRP substrates)

Pre-defined adverse events of special interest were selected from AE dataset using below MedDRA PT code:

Alanine aminotransferase increased (PT 10001551)
Aspartate aminotransferase increased (PT 10003481)
Blood alkaline phosphatase increased (PT 10059570)
Blood bilirubin increased (PT 10005364)
Blood creatinine increased (PT 10005483)
Blood lactate dehydrogenase increased (PT 10005630)
Hyperbilirubinaemia (PT 10020578)
Muscular weakness (PT 10028372)
Renal failure (PT 10038435)
Renal impairment (PT 10062237)
Transaminases increased (PT 10054889).

Document Type:	Additional Statistical Analysis Plan
Official Title:	A randomized, double–blind, placebo–controlled Phase III study of darolutamide (ODM–201) versus placebo in addition to standard androgen deprivation therapy and docetaxel in patients with metastatic hormone–sensitive prostate cancer
NCT Number:	NCT02799602
Document Date:	04-Feb-2022

Title Page

A randomized, double-blind, placebo-controlled Phase III study of darolutamide (ODM-201) versus placebo in addition to standard androgen deprivation therapy and docetaxel in patients with metastatic hormone-sensitive prostate cancer

Protocol Number: 17777

SAP Supplement 1 (incl. version and date): V1.0, 04 FEB 2022

Compound Number: BAY 1841788

Sponsor Name and Legal Registered Address:

Non-US: Bayer AG, 51368 Leverkusen, Germany

US territory: Bayer HealthCare Pharmaceuticals Inc., 100 Bayer Boulevard, P.O. Box 915, Whippany NJ 07981-0915, USA

Regulatory Agency Identifier Number(s):

Registry	**ID**
IND	114769

Date: 4 FEB 2022
Version: 1.0

This Statistical Analysis Plan is produced on a word-processing system and bears no signatures.
The approval of the Statistical Analysis Plan is documented in a separate signature document.

Table of Contents

1. Introduction

This Supplemental Statistical Analysis Plan (SAP) describes analyses that were not included in the main SAP but may be used for Clinical Study Report (CSR) and marketing authorization / NDA.

This Supplemental SAP version 1.0 is based on main SAP version 4.1 dated 11 NOV 2021.

2. Study Objectives

Refer to main SAP V4.1 dated 11 NOV 2021.

3. Study Design

Refer to main SAP V4.1 dated 11 NOV 2021.

4. General Statistical Considerations

Refer to main SAP V4.1 dated 11 NOV 2021.

5. Analysis Sets

Refer to main SAP V4.1 dated 11 NOV 2021. All tables, figures and listings as described in this supplemental SAP will be based on the safety analysis set (SAF).

6. Statistical Methodology

6.1 Population characteristics

6.1.1 Disposition

Patients with treatment unblinded during double-blind study period will be listed.

6.1.2 Demographic and Other Baseline Characteristics

Not applicable.

6.1.3 Medical History

Not applicable.

6.1.4 Prior and Concomitant Medication and Procedures

A summary of duration of start of prior ADT within 12 weeks from randomization to randomization by geographical region based on WHO-DD drug record number. Patients who started ADT prior to 12 weeks before randomization will be listed. ADT is defined by WHO-DD drug record number: 1. Gonadotropin-releasing hormone agonist: Buserelin or Buserelin

acetate (007716), Gonadorelin (004865), Leuprorelin or Leuprorelin acetate (007269), Goserelin or Goserelin acetate (007321), Histrelin acetate (012465), Triptorelin or Triptorelin acetate or Triptorelin embonate (009759); 2. Gonadotropin-releasing hormone antagonist: Degarelix or Degarelix acetate (017648).

Patients with prior ADT, based on WHO-DD drug record number, before 12 weeks prior to randomization will be listed.

6.1.5 Extent of Exposure

Patients with actual dose of darolutamide >1800 mg per day will be listed.

Patients with darolutamide/placebo dose interruption > 28 consecutive days during study treatment will be listed.

6.2 Efficacy

6.2.1 Primary Efficacy Endpoint

Sensitivity analyses

The following OS sensitivity analyses by docetaxel cycles and including patients with violation of GCP are planned, respectively:

- OS analyses by 6 cycles versus 5 and less cycles (including 26 patients who did not receive any docetaxel, based on full analysis set).
- OS analyses by 6 and 5 cycles versus 4 and less cycles (including 26 patients who did not receive any docetaxel, based on full analysis set).
- OS analysis including one patient with violation of GCP, based on all the randomized patients.

Refer to main SAP V4.1 dated 11 NOV 2021 for analysis of OS.

Number of patients, number of events and KM estimates for the median per arm for OS sensitivity analyses by docetaxel cycles will be presented in the forest plot.

6.2.2 Secondary Efficacy Endpoint

6.2.2.2 Time to pain progression

Time to pain progression is defined as the time from randomization to the first date a patient experiences pain progression (first date of pain progression minus date of randomization). Pain will be assessed with the BPI-SF questionnaire. The main analysis is based on the ePRO device questionnaire data only. Additional sensitivity analyses for time to pain progression are defined as following:

Pain progression including unconfirmed worst pain score is defined as follows:

- For asymptomatic patients (worst pain subscale [WPS] 0 at baseline), pain progression is defined as
    - an increase of 2 or more points in the “worst pain in 24 hours” score (i.e. 2 or more point increase in WPS score) from nadir (i.e. zero), or
    - initiation of short- or long-acting opioid use for pain.

- For symptomatic patients (WPS >0 at baseline), pain progression is defined as
    - an increase of 2 or more points in the "worst pain in 24 hours" (i.e. 2 or more point increase in WPS score) from nadir and a WPS of ≥4, or
    - initiation of short- or long-acting opioid use for pain.

Pain progression based on 4 or more-point increase is defined as follows:

- For asymptomatic patients (worst pain subscale [WPS] 0 at baseline), pain progression is defined as
    - an increase of 4 or more points in the "worst pain in 24 hours" score (i.e. 4 or more point increase in WPS score) from nadir (i.e. zero) observed at 2 consecutive evaluations ≥4 weeks apart, or
    - initiation of short- or long-acting opioid use for pain.
- For symptomatic patients (WPS >0 at baseline), pain progression is defined as
    - an increase of 4 or more points in the "worst pain in 24 hours" (i.e. 4 or more point increase in WPS score) from nadir observed at 2 consecutive evaluations ≥4 weeks apart and a WPS of ≥4, or
    - initiation of short- or long-acting opioid use for pain.

The censoring rule in Table 6-3 of main SAP V4.1 will be applied.

6.2.3 Exploratory Efficacy Endpoint

Not applicable.

6.3 Safety

6.3.1 Adverse Events

The following additional tables for treatment-emergent adverse events (TEAEs) will be created. In all tables, results will be shown by treatment group.

Adverse events will be analyzed by using of predefined customized MedDRA queries denoted as MedDRA Labeling Groupings (MLG). These MLG are created and centrally maintained by Bayer internal coding experts.

Incidence proportions will be calculated as number of subjects experiencing an event per number of subjects exposed. Incidence proportions will be presented by MLG, PT and worst CTCAE grade (grade 1 to grade 5, missing grade, any grade). A separate table will be created for PTs not covered by MLG or special topic adverse events grouping. A definition table for the groupings will be provided, showing the AE grouping, the type of grouping entity used for definition (e.g. MLG name), and the corresponding preferred terms. Preferred terms which occurred in the data are marked with preceding asterisks (***).

To adjust for potential differences in study drug treatment duration (darolutamide or placebo) between treatment arms, additional summaries based on event rate per 100 patient years (PY) will be performed for the MLGs occurring after the first dose of darolutamide or placebo. This additional table will show columns for AE grouping, incidence proportion darolutamide, incidence proportion placebo, risk difference 'Darolutamide – placebo' with 95% confidence interval (CI), risk ratio 'Darolutamide/placebo' with 95% CI, exposure-adjusted incidence

rate darolutamide, exposure-adjusted incidence rate placebo, incidence rate ratio. No zero-correction will be applied for calculation of the risk ratio, i.e. in case of no events in the placebo arm the risk ratio will not be calculated. The exposure-adjusted event rate is calculated as the total number of patients with a given TEAE divided by the total darolutamide or placebo treatment duration of all patients in years. The treatment duration in years will be calculated as treatment duration in days divided by 365.25.

Incidence and prevalence of most common MLGs will be created for the MLGs with an incidence proportion of at least 10% in any treatment arm. The worst grade per MLG will be considered within each time interval. The incidence proportion of the MLGs by time interval is defined as number of patients with any TEAE within the MLG arising (new onset) or worsening in a specific interval divided by number of patients in that time interval of the treatment-emergent period. The prevalence of the MLGs by time interval is defined as number of patients with any TEAE within the MLG starting or ongoing in the specific time interval divided by the number of patients in that time interval of the treatment-emergent period.
The above described tables will also be created for the data driven grouping of the MLG hypertension plus the PTs 'Hypertensive crisis' and 'Hypertensive emergency'.

A listing will be generated of all adverse events for a patient (patient ID PPD) who was assigned to placebo+docetaxel arm but took both darolutamide and placebo treatment during the study.

Incidence of the TEAEs by MedDRA SOC, PT and worst CTCAE grade for subgroup of patients with medical history of hepatic impairment present (PT: Alcoholic liver disease, Cirrhosis alcoholic, Drug-induced liver injury, Endoscopic retrograde cholangiopancreatography, Hemangioma of liver, Hepatic cirrhosis, Hepatic function abnormal, Hepatic steatosis, Hepatitis, Hepatitis alcoholic, Hepatobiliary disease, Hepatomegaly, Liver abscess, Liver disorder, Liver transplant, Steatohepatitis) will be summarized by treatment arm.

Incidence of the TEAEs by MedDRA MLG hypertension, PT and worst CTCAE grade for subgroup of patients with medical history of MLG hypertension present will be summarized by treatment arm.

Per EMA guidance an overview of treatment-emergent adverse events (TEAE) according to specific categories by age classes (<65, 65-74, 75-84 and > 85) created in the main set of tables per SAP v4.1 dated 11 NOV 2021 will be updated in order to add the two following categories:

- Anticholinergic syndrome
- Sum of postural hypotension, falls, black outs, syncope, dizziness, ataxia, fractures

The following selections will be used:

- Anticholinergic syndrome: SMQ 'Anticholinergic syndrome',
- Postural hypotension: PT Orthostatic hypotension,
- Falls: PT Fall and PT Accident,
- Black outs: PT Loss of consciousness,
- Syncope: MLG Syncope,
- Dizziness: MLG Dizziness and PT Vertigo,
- Ataxia: MLG Gait disturbance and gait inability,

- Fractures: Bone fracture (selected using HLT: Fractures and dislocations NEC (without PTs: Joint dislocation, Joint dislocation pathological), Limb fractures and dislocations (without PT: Radial head dislocation), Pelvic fractures and dislocations, Skull fractures, facial bone fractures and dislocations, Spinal fractures and dislocations (without PT: Dislocation of vertebra),Thoracic cage fractures and dislocations (without PT: Dislocation of sternum)).

6.3.2 Deaths and Serious Adverse events

Not applicable.

6.3.3 Additional Primary Malignancies

A listing will be generated for patients with additional primary malignancies who had subsequent systemic antineoplastic therapies.

6.3.4 Laboratory Safety Assessments

The mean of lab values (from both scheduled and unscheduled assessments) for ALT, AST, Bilirubin, Hemoglobin, Neutrophils and Platelets will be displayed in line plots by treatment arm and new grouped time interval relative to the first concomitant docetaxel treatment date, from baseline up-to Visit 6 (EOT visit will be included if the last visit number before EOT visit is < 7) for safety patients who had concomitant docetaxel with darolutamide or placebo.

- Week 0 – 15, mean values are presented by each week.
- Week 16 (grouped with week 16-23),
- Week 24 (grouped with week 24-35),
- Week 36 (grouped with week 36-47),
- Week 48 (grouped with week 48-59),
- Week 60(grouped with week 60-83).

If a patient had multiple measurements of a lab parameter in a specific interval defined above, mean value of a lab parameter will be calculated for a patient in that interval. In the line plots, the mean value of available patients' multiple (or single) measures mean, will be presented for each interval defined above, by treatment arm and lab parameters.

To provide clearer presentation of the data, the selected lab parameters: ALT, AST, Bilirubin, Hemoglobin, Neutrophils, Platelets, PSA and Testosterone will also be displayed graphically using laboratory shift plots within the ranges of the majority of the data, including both scheduled and unscheduled assessments. These plots will show the baseline value and mean post baselines values up-to end of treatment. Clinical laboratory data for patients not displayed in the post-hoc shift plots will be provided in a patient listing.

6.3.5 Other Safety Measures

Incidence of ECG QTc > 500 msec occurring between the first dose of darolutamide or placebo and last dose of darolutamide or placebo plus 30 days will be summarized. To adjust for potential differences in study drug treatment duration (darolutamide or placebo) between the treatment arms, additional summaries based on event rate per 100 patient years (PYs) were performed for ECG QTc > 500 msec occurring between the first dose of darolutamide or placebo and the last dose of darolutamide or placebo plus 30 days. The exposure-adjusted incidence rate was calculated as the total number of patients with an event of ECG QTc > 500

msec divided by the total darolutamide or placebo treatment duration of all patients in years. The rate is expressed in number of patients with events per 100 PYs.

6.4 Other Exploratory Variables

Not applicable.

6.5 Pharmacokinetics/pharmacodynamics

Not applicable.

7. Document history and changes in the planned statistical analysis

- Supplemental SAP version 1.0 dated 04 FEB 2022 without attachments

8. References

Refer to main SAP v4.1 dated 11 NOV 2021.

Document Type:	Additional Statistical Analysis Plan for CSR addendum
Official Title:	A randomized, double-blind, placebo-controlled Phase III study of darolutamide (ODM-201) versus placebo in addition to standard androgen deprivation therapy and docetaxel in patients with metastatic hormone-sensitive prostate cancer
NCT Number:	NCT02799602
Document Date:	25 APR 2023

Title Page

A randomized, double-blind, placebo-controlled Phase III study of darolutamide (ODM-201) versus placebo in addition to standard androgen deprivation therapy and docetaxel in patients with metastatic hormone-sensitive prostate cancer

Protocol Number: 17777

SAP Supplement 2 (incl. version and date): V1.0, 25 APRIL 2023

Compound Number: BAY 1841788

Sponsor Name and Legal Registered Address:

Non-US: Bayer AG, 51368 Leverkusen, Germany

US territory: Bayer HealthCare Pharmaceuticals Inc., 100 Bayer Boulevard, P.O. Box 915, Whippany NJ 07981-0915, USA

Regulatory Agency Identifier Number(s):

Registry	ID
IND	114769
EudraCT	2015–002590–38

Date: 25 APRIL 2023
Version: 1.0

This Statistical Analysis Plan is produced on a word-processing system and bears no signatures.
The approval of the Statistical Analysis Plan is documented in a separate signature document.

Table of Contents

List of Abbreviations

ADT	Androgen deprivation therapy
AE	Adverse event
ALP	Alkaline phosphatase
ATC	Anatomical-Therapeutic-Chemical
BHA	Bone health agent
CI	Confidence interval
FAS	Full analysis set
MedDRA	Medical Dictionary for Regulatory Activities
mg	Milligram
mHSPC	Metastatic hormone-sensitive prostate cancer
NCI-CTCAE v4.03	National Cancer Institute-Common Terminology Criteria for Adverse Events; version 4.03
PSA	Prostate-specific antigen
PT	Preferred Term
SAE	Serious adverse event
SAF	Safety analysis set
SAP	Statistical Analysis Plan
SOC	System Organ Class
TEAE	Treatment emergent adverse events
TESAE	Treatment emergent serious adverse events
ULN	Upper limit of normal
WHO	World Health Organization
WHO-DD	WHO Drug Dictionary

1. Introduction

Study 17777 (ARASENS) is an international, randomized, double-blind, placebo-controlled, Phase III efficacy and safety study of darolutamide in addition to standard androgen deprivation therapy (ADT) and docetaxel in patients with metastatic hormone-sensitive prostate cancer (mHSPC).

This supplement to the statistical analysis plan (SAP) describes the analyses and data presentations for final clinical study report addendum when all patients have discontinued the trial.

This SAP Supplemental 2 version 1.0 is based on integrated protocol version 5.0 (amendment 7) dated 26 MAY 2020, on SAP version 4.1 dated 11 NOV 2021 and Supplemental 1 SAP version 1.0 dated 04 FEB 2022.

2. Study Objectives

Refer to main SAP V4.1 dated 11 NOV 2021.

3. Study Design

Refer to main SAP V4.1 dated 11 NOV 2021.

4. General Statistical Considerations

Refer to main SAP V4.1 dated 11 NOV 2021.

5. Analysis Sets

Refer to main SAP V4.1 dated 11 NOV 2021. All tables, figures and listings as described in this supplemental SAP will be based on either the full analysis set (FAS) or the safety analysis set (SAF).

6. Statistical Methodology

The statistical analyses will be descriptive. Summaries will be provided for both treatment arms, darolutamide+docetaxel arm and placebo+docetaxel arm by period of “At Primary completion cut-off - from FPFV to 25OCT2021” and “Whole Study - from FPFV to LPLV”.

The number of patients will be summarized in disposition tables, concomitant medications, study-drug exposure tables and adverse events tables.

6.1 Population characteristics

6.1.1 Disposition

The number of patients starting, completing, and discontinuing together with the primary reason for discontinuing of the Treatment, Active Follow-up, and Long-term Follow-up periods will be presented by treatment arm.

In addition, the number of patients with important protocol deviations will be presented by period and treatment arm. The frequencies of each important protocol deviation will be presented by period and treatment arm. All COVID-19 pandemic related protocol deviations are considered important protocol deviations and will be presented by period and treatment arm. Median follow-up time during study will be summarized.

6.1.2 Demographic and Other Baseline Characteristics

Refer to main SAP V4.1 dated 11 NOV 2021. Only changed data will be summarized and listed overall and by treatment arm for whole study period. No subgroup analysis will be presented for demographic and baseline characteristics tables.

6.1.3 Medical History

Refer to main SAP V4.1 dated 11 NOV 2021.

6.1.4 Prior and Concomitant Medication

The following tables and listing will be created:

- Concomitant medications excluding ADT, systemic antineoplastic therapies captured in the anti-cancer therapy eCRF page: frequency of patients for each drug category.
- A summary of subsequent systemic antineoplastic therapy will be presented by preferred drug name based on WHO-DD drug record number.
- A summary of life-prolonging subsequent systemic antineoplastic therapy will be presented by preferred drug name based on WHO-DD drug record number or by number of regimens . The following medications after the study treatment period are life-prolonging subsequent therapies for prostate cancer: Abiraterone, Apalutamide, Enzalutamide, Docetaxel, Cabazitaxel, Radium 223, Sipuleucel-T, Lutetium-177.
- Concurrent prostate cancer related-invasive procedures (e.g. procedures for urinary symptoms, surgical procedures for prostate cancer): frequency of patients by grouped term and procedure. Prostate cancer-related invasive procedure is defined as any procedure needed for alleviation of symptoms, signs or findings caused by progression of prostate cancer (e.g. catheterization of the bladder, percutaneous drainage of hydronephrosis, palliative electro resection of the prostate, etc.).

The dictionary used for coding medications was the ATC Class/Subclass and the World Health Organization Drug Dictionary (WHO-DD). All tables will be presented with the most recent version (version 2022SEP) of the dictionary.

6.1.5 Extent of Exposure

After primary completion (25 OCT 2021 cut-off), all patients on darolutamide treatment have transitioned into a separate program to continue receiving darolutamide or have discontinued

the study for any other reason (e.g. death, lost to follow-up, consent withdrawal with no further data collection), and all patients on placebo have discontinued treatment. Extent of exposure will be summarized for the SAF by treatment arm and by period of "At Primary completion cut-off - from FPFV to 25OCT2021" and "Whole Study - from FPFV to LPLV", using descriptive statistics. for the following variables:

- Overall time under treatment (or treatment duration)
  Including time off drug and dose interruptions. It will be calculated in days and presented in months as (date of the last dose of any study treatment – date of the first dose of any study treatment + 1) / 30.44.
- Actual dose per day(or average daily dose received)
  Actual dose per day (or average daily dose received) = total amount of dose / number of days with intake > 0
- Total amount of dose
  Total amount of dose = sum of dose received over total time under treatment
- Percent of planned dose received.
  Percent of planned dose received = total amount of dose [mg] / planned dose [mg] * 100%; planned dose is sum of the intended initial dose according to protocol over total time under treatment.
- From 1 day after the last dose of docetaxel to the last dose of darolutamide or placebo.
  Including time off drug and dose interruptions. It will be calculated in days and presented in months as (date of the last dose of darolutamide or placebo – date of the last dose of docetaxel + 1) / 30.44. If the last dose of docetaxel date is after the last dose of darolutamide or placebo, then 0 day of duration will be included in the calculation.

Patients with actual dose of darolutamide >1800 mg per day will be listed. Patients with darolutamide/placebo dose interruption > 28 consecutive days during study treatment will be listed. No subgroup analysis will be presented for extent of exposure.

6.2 Efficacy

6.2.1 Primary Efficacy Endpoint

Not applicable.

6.2.2 Secondary Efficacy Endpoint

Not applicable.

6.2.3 Exploratory Efficacy Endpoint

Not applicable.

6.3 Safety

6.3.1 Adverse Events

6.3.1.1 Treatment-emergent Adverse Events

AEs will be coded using MedDRA 25.1 or most recent version. The severity (or intensity) of AEs will be documented using the National Cancer Institute-Common Terminology Criteria for Adverse Event version 4.03 (NCI-CTCAE v 4.03). Descriptive summary tables will be presented on all safety parameters by treatment arm.

The AEs will be presented with their worst NCI-CTCAE v 4.03 grade CTC 1 = mild, CTC 2 = moderate, CTC 3 = Severe, CTC 4 = life threatening, CTC 5 = fatal. The causal relationship is based on whether there was a reasonable causal relationship to darolutamide or placebo and docetaxel. AEs will be classified by the investigator as related or not related to darolutamide or placebo and docetaxel separately.

Treatment-emergent AE (TEAE) is defined as any event arising or worsening after the first dose of darolutamide or placebo until 30 days after the last dose of darolutamide or placebo.

An overview of TEAEs will be summarized by treatment arm in the overall treatment period. The following summaries will be created separately for study drug-related (darolutamide or placebo-related) and docetaxel-related TEAEs, by treatment arm: TEAEs leading to permanent discontinuation of study treatment, TEAEs leading to dose reduction and/or dose interruption, and drug-related TEAEs. Summary statistics (frequency and percentage of patients, not of events) will be presented by treatment arm, MedDRA SOC and PT, and NCI-CTCAE worst grade for the following:

- Incidence of all TEAEs
- Incidence of TEAEs of NCI-CTCAE grade 3, 4 or 5
- Incidence of TEAEs leading to permanent discontinuation of darolutamide or placebo and docetaxel separately
- Incidence of TEAEs leading to dose reduction
- Incidence of TEAEs leading to dose interruption
- Incidence of TEAEs with incidence of at least 5%
- Incidence of all drug-related TEAEs
- Incidence of drug-related TEAEs of NCI-CTCAE grade 3, 4 or 5
- Incidence of drug-related TEAEs leading to permanent discontinuation of darolutamide or placebo and docetaxel separately
- Incidence of drug-related TEAEs leading to dose reduction
- Incidence of drug-related TEAEs leading to dose interruption
- Incidence of drug-related TEAEs with incidence of at least 5%
- Incidence of post-treatment AEs (after 30 days from the last dose of study drug).
- Incidence and prevalence of most common TEAEs ≥10% incidence in either treatment arm, analyzed within pre-specified time intervals. The same summary will be presented by geographical regions (North America, Asia Pacific, Rest of the World) as well.

- Incidence of COVID-19-related TEAEs
- Incidence of TEAEs by MedDRA SOC, HLGT, HLT, PT and worst CTCAE grade

Subgroup analyses of TEAEs

The following subgroup analyses will be performed for overview of TEAEs and for TEAEs showing at least 5%-point differences in incidence proportions between any of the subgroup categories:

- Age category (<65, 65–74, 75–84, ≥85 years)
- Age category (<65, 65–74, 75–84, ≥85 years) per EMA guidance for following categories: Psychiatric disorders (SOC), Nervous system disorders (SOC), Accidents and injuries (SMQ), Cardiac disorders (SOC), Vascular disorders (SOC), Central nervous system vascular disorders (SMQ), Anticholinergic syndrome (SMQ), Infections and infestations (SOC), Quality of life decreased (PT), Sum of postural hypotension, falls, black outs, syncope, dizziness, ataxia, fractures. The selection criteria of the grouped terms are defined in Appendix 9.1.
- Geographical region (North America, Asia Pacific / Rest of the World

TEAEs by medical conditions

The incidence of TEAEs will be presented by groupings and worst CTCAE grade for the following:

- Patients with present medical history of cardiac disorder (SOC).
- Patients with present medical history of renal impairment (SMQ Acute renal failure; SMQ Chronic kidney disease).
- Patients with present TEAE interstitial lung disease (not medical history, SMQ interstitial lung disease (narrow search)).
- Patients with present medical history of hepatic impairment (PT: Alcoholic liver disease, Cirrhosis alcoholic, Drug-induced liver injury, Endoscopic retrograde cholangiopancreatography, Hemangioma of liver, Hepatic cirrhosis, Hepatic function abnormal, Hepatic steatosis, Hepatitis, Hepatitis alcoholic, Hepatobiliary disease, Hepatomegaly, Liver abscess, Liver disorder, Liver transplant, Steatohepatitis) will be summarized by treatment arm.
- Patients with present medical history of MLG hypertension will be summarized by treatment arm.

Exposure-adjusted incidence rates of TEAEs

To adjust for unequal lengths of study treatment (darolutamide or placebo and docetaxel) between the treatment arms, additional summaries based on event rate per 100 patient years (PYs) will be performed for all TEAEs, special topics TEAEs (Section 6.3.1.2), and all TESAEs occurring after the first dose of darolutamide or placebo. The exposure-adjusted incidence rate is calculated as the total number of patients with a given TEAE divided by the total darolutamide or placebo treatment duration of all patients in years. The rate is expressed in 100 PYs).

The treatment duration in years will be calculated as treatment duration in days divided by 365.25. The risk difference 'darolutamide+docetaxel arm – placebo+docetaxel arm' and risk ratio 'darolutamide+docetaxel arm vs. placebo+docetaxel arm' will be calculated together with 95% confidence intervals (CIs). No zero-cell correction will be applied for calculation of the risk ratio, i.e. in case of no events in the placebo+docetaxel arm the risk ratio will not be calculated. The incidence rate ratio for darolutamide+docetaxel arm vs. placebo+docetaxel arm will also be calculated.

6.3.1.2 Treatment-emergent Adverse Events Groupings Considered as Special Topics

The following TEAE groupings are considered as special topics (see definitions in Appendix 9.2):

- Bone fractures excluding pathological fractures
- Diabetes mellitus and hyperglycemia
- Fall
- Fatigue/ asthenic conditions
- Weight decreased
- Rash
- Seizure
- Hypertension
- Vasodilatation and flushing
- Mental impairment disorders
- Depressed mood disorders
- Breast disorders/gynecomastia
- Cardiac disorders
- Cerebral ischaemia
- Cerebral and intracranial hemorrhage

The following tables will be created for TEAEs of special topics:

- Incidence of TEAEs
- Incidence of TEAEs leading to permanent discontinuation of (darolutamide or placebo and docetaxel separately)
- Incidence of TEAEs leading to dose reduction
- Incidence of TEAEs leading to dose interruption

- Incidence of treatment-emergent SAEs (TESAEs)
- Incidence and prevalence of most common TEAEs of special topics ≥ 5% incidence in either treatment arm, analyzed within pre-specified time intervals. The same summary will be presented by geographical regions (North America, Asia Pacific, Rest of the World) as well.

Fracture events will be described by a cumulative incidence plot of fracture. A summary of treatment-emergent fracture by bone health agent use (for selection of bone health agent see Appendix 9.3) at study entry will be provided. Association with weight decrease will be presented by histogram of fracture events by patient weight change (the weight collected closest to the start of fracture will be considered and compared to baseline weight). A listing of patients with fracture events breakdown by relevant events occurring within 7 days window: fall, accident, syncope and/or loss of consciousness, dizziness will be presented.

A listing will be generated for patients with fall treatment-emergent events with syncope and/or loss of consciousness (using the MLG Syncope). Timing of occurrence of fracture events based on first dose of study drug will be presented. A summary of time to fall and fractures events will be presented in descriptive and inferential statistics, respectively. In addition, Kaplan-Meier plots will be created to display the time to fall and time to fracture.

An overview of patients at risk for developing an AE seizure will be tabulated by displaying medical history.

6.3.1.3 MedDRA Labeling Groupings

The following additional tables for treatment-emergent adverse events (TEAEs) will be created. In all tables, results will be shown by treatment group.

Adverse events will be analyzed by using predefined customized MedDRA queries denoted as MedDRA Labeling Groupings (MLG). These MLG are created and centrally maintained by Bayer internal coding experts.

Incidence proportions will be calculated as number of subjects experiencing an event per number of subjects exposed. Incidence proportions will be presented by MLG, PT and worst CTCAE grade (grade 1 to grade 5, missing grade, any grade). A separate table will be created for PTs not covered by MLG or special topic adverse events grouping. A definition table for the groupings will be provided, showing the AE grouping, the type of grouping entity used for definition (e.g. MLG name), and the corresponding preferred terms. Preferred terms which occurred in the data are marked with preceding asterisks (***).

To adjust for potential differences in study drug treatment duration (darolutamide or placebo) between treatment arms, additional summaries based on event rate per 100 PYs will be performed for the MLGs occurring after the first dose of darolutamide or placebo. This additional table will show columns for AE grouping, incidence proportion darolutamide, incidence proportion placebo, risk difference ‘Darolutamide – placebo’ with 95% confidence interval (CI), risk ratio ‘Darolutamide/placebo’ with 95% CI, exposure-adjusted incidence rate darolutamide, exposure-adjusted incidence rate placebo, incidence rate ratio. No zero-correction will be applied for calculation of the risk ratio, i.e. in case of no events in the placebo arm the risk ratio will not be calculated.

Incidence and prevalence of most common MLGs will be created for the MLGs with an incidence proportion of at least 10% in any treatment arm. The worst grade per MLG will be considered within each time interval. The incidence proportion of the MLGs by time interval is defined as number of patients with any TEAE within the MLG arising (new onset) or worsening in a specific interval divided by number of patients in that time interval of the treatment-emergent period. The prevalence of the MLGs by time interval is defined as number of patients with any TEAE within the MLG starting or ongoing in the specific time interval divided by the number of patients in that time interval of the treatment-emergent period.

The above-described tables: incidence and prevalence by time intervals, exposure-adjusted incidence rate and risk ratio will also be created for the data-driven grouping of Hypertension (MLG hypertension plus the PTs 'Hypertensive crisis' and 'Hypertensive emergency').

6.3.1.4 Additional safety AE analyses

The additional safety AE analyses will be performed for AE grouped term of Interstitial lung disease, data-driven grouping Hypertension, and Hepatic disorders by treatment arm. The selection criteria of the grouped terms are defined in Appendix 9.3.

- Exposure-adjusted treatment-emergent Interstitial lung disease (SMQ [narrow]), data-driven grouping Hypertension, Hepatic disorders by Grouping and PT per 100 subject years
- Treatment-emergent Interstitial lung disease (SMQ [narrow]), data-driven grouping Hypertension, Hepatic disorders by AE grouping, PT and worst CTCAE grade.

6.3.2 Deaths and Serious Adverse events

Serious adverse events (SAEs) will be classified using the NCI-CTCAE version 4.03 and MedDRA version 25.1 or most recent version.

- Treatment-emergent SAEs
- Treatment-emergent SAEs leading to permanent discontinuation of study treatment
- Treatment-emergent SAEs leading to dose reduction
- Treatment-emergent SAEs leading to drug interruption
- Treatment-emergent drug-related SAEs
- COVID-19 related treatment-emergent SAEs
- Listing of treatment-emergent SAEs
- Listing of non-treatment-emergent SAEs
- Non-treatment-emergent drug-related SAEs

The incidence of deaths in the study: within 30 days after first dose of study drug / docetaxel, during study treatment from first to last dose of study drug / docetaxel, within 30 days after last dose of study drug/ docetaxel, later than 30 days after last dose of study drug/ docetaxel, will be summarized by each treatment arm and cause of death. All deaths before treatment start, during treatment and within 30 days after last dose of study drug, and later than 30 days after last dose of study drug, will be listed by patient with start and stop date of study medication, date of death, and cause of death.

6.3.3 Additional Primary Malignancies

Summary statistics will be presented by treatment arm, MedDRA SOC, HLGT, HLT and PT for additional primary malignancies.

- Exposure-adjusted additional primary malignancies during treatment and follow-up per 100 PYs by MedDRA HLGT, HLT, and PT.
- Additional primary malignancies during treatment and follow-up by MedDRA HLGT, HLT, PT and worst CTCAE grade.
- Incidence proportion and exposure adjusted incidence rate for data-driven grouping additional primary malignancies excluding superficial skin cancers during treatment or follow-up.

A listing will be generated for patients with additional primary malignancies and subsequent systemic antineoplastic medications used for additional primary malignancy.

The additional primary malignancies are defined Appendix 9.3.

6.3.4 Laboratory Safety Assessments

Descriptive statistics (number of observations, mean, standard deviation, minimum, median, and maximum values) will be presented for clinical laboratory tests (hematology, clinical biochemistry, and urinalysis), their changes from baseline (including baseline value), and their percent changes from baseline by treatment arm at applicable visits.

Hematological and biochemical laboratory values will be graded based on NCI CTCAE version 4.03. CTCAE severity grading for laboratory abnormalities are based on applicable laboratory threshold values outlined in NCI CTCAE v4.03.The lab parameters will be displayed graphically using laboratory shift plots within the ranges of the majority of the data, including both scheduled and unscheduled assessments. These plots will show the baseline value and mean post baselines values up-to end of treatment. Clinical laboratory data for patients not displayed in the shift plots will be provided in a patient listing.

A listing will be provided for all patients possibly fulfilling Hy's Law criteria, i.e. patients with elevated AST and / or ALT >3xULN, alkaline phosphatase <2xULN and bilirubin ≥2xULN (ref. 3). For possible Hy's Law cases relevant laboratory data will also be displayed graphically within actual patient profiles (presenting total bilirubin, ALT, AST, and ALP values in terms of ULN over time) by treatment arm. Below example of Hy's law plot will be provided for peak total bilirubin vs ALT. If a patient has any total bilirubin ≥2xULN then peak bilirubin will be plotted versus the maximum ALT amongst the total bilirubin ≥2xULN. Otherwise, peak bilirubin will be plotted versus peak ALT. Liver function values for patients with a medical history of Gilbert' syndrome will be listed.

6.3.5 Other Safety Measures

Not applicable.

6.4 Other Exploratory Variables

Not applicable.

6.5 Pharmacokinetics/pharmacodynamics

Not applicable.

7. Document history and changes in the planned statistical analysis

- Supplemental SAP version 1.0 dated 04 FEB 2022
- Main SAP version 4.1 dated 11 NOV 2021

8. References

Refer to main SAP v4.1 dated 11 NOV 2021.

9. Appendix

9.1 The selection criteria for EMA guidance on TEAEs are:

- Anticholinergic syndrome: SMQ ‘Anticholinergic syndrome’,
- Postural hypotension: PT Orthostatic hypotension,
- Falls: PT Fall and PT Accident,
- Black outs: PT Loss of consciousness,
- Syncope: MLG Syncope,
- Dizziness: MLG Dizziness and PT Vertigo,
- Ataxia: MLG Gait disturbance and gait inability,
- Fractures: Bone fracture - selected using HLT: Fractures and dislocations NEC (without PTs: Joint dislocation, Joint dislocation pathological), Limb fractures and dislocations (without PT: Radial head dislocation), Pelvic fractures and dislocations, Skull fractures, facial bone fractures and dislocations, Spinal fractures and dislocations (without PT: Dislocation of vertebra), Thoracic cage fractures and dislocations (without PT: Dislocation of sternum).

9.2 Special topic AE definition

Table 9–1 Special topic AE definitions

Grouped term	MedDRA search criteria for special topics
Bone fractures excluding pathological fractures	HLT: Fractures and dislocations NEC (without PTs: Joint dislocation, Joint dislocation pathological) HLT Limb fractures and dislocations (without PT: Radial head dislocation) HLT Pelvic fractures and dislocations HLT Skull fractures, facial bone fractures and dislocations HLT Spinal fractures and dislocations (without PT: Dislocation of vertebra) HLT Thoracic cage fractures and dislocations (without PT: Dislocation of sternum)) Exclude PT pathological fractures
Diabetes mellitus and hyperglycemia	MLG: Hyperglycemia PT: Diabetes mellitus; Diabetes mellitus inadequate control; Diabetic metabolic decompensation; Type 2 diabetes mellitus; Diabetic ketoacidosis
Fall*	PT: Fall PT: Accident
Fatigue/ asthenic conditions	MLG: Decreased general strength and energy PT: Lethargy; Chronic fatigue syndrome; Malaise
Weight decreased	MLG: Weight decreased
Rash	MLG: Rash MLG: Skin erythema PT: Dermatitis
Seizure	MLG: Seizures
Hypertension	MLG: Hypertension
Vasodilatation and flushing	MLG: Vasodilatation and flushing
Mental impairment disorders	HLGT: Mental impairment disorders
Depressed mood disorders	HLGT: Depressed mood disorders and disturbance
Breast disorders/gynecomastia	HLGT: Breast disorders
Cardiac disorders	HLGT: Cardiac arrhythmias HLGT: Coronary artery disorders HLGT: Heart failures
Cerebral ischaemia	MLG: Cerebral infarction and stroke not specified as hemorrhagic or ischemic MLG: Cerebral ischemic infarction and stroke PT: Cerebral ischaemia, Transient ischaemic attack
Cerebral and intracranial hemorrhaege	MLG: Cerebral and intracranial hemorrhage

MLG: MedDRA Labeling Grouping; PT: Preferred Term; HLT: High Level Term; HLGT: High Level Group Term

* Based on review of verbatims the search criteria for “Fall” was extended by the PT “Accident”.

9.3 The selection criteria for additional safety AE analysis

- Interstitial lung disease: SMQ Interstitial lung disease (SMQ narrow)
- Data-driven grouping Hypertension: data-driven grouping of the MLG hypertension plus the PTs 'Hypertensive crisis' and 'Hypertensive emergency'.
- Hepatic disorders:
  Five MLGs:
    - MLG Increase in bilirubin,
    - MLG Increase in transaminases,
    - MLG Hepatic impairment,
    - MLG Hepatic failure acute,
    - MLG Hepatic failure unspecified,

  plus selected PTs per MedDRA v25.1 from the SMQ Drug related hepatic disorders - comprehensive search (SMQ) as follows
    - AST/ALT ratio abnormal
    - Alanine aminotransferase abnormal
    - Aspartate aminotransferase abnormal
    - Cholestasis
    - Cholestatic liver injury
    - Coma hepatic
    - Drug-induced liver injury
    - Hepatitis
    - Hepatitis acute
    - Hepatitis cholestatic
    - Hepatic encephalopathy
    - Hepatitis fulminant
    - Hepatitis toxic
    - Hepatorenal failure
    - Hyperammonaemia
    - International normalised ratio increased
    - Jaundice
    - Jaundice cholestatic
    - Jaundice hepatocellular
    - Liver disorder
    - Non-alcoholic steatohepatitis
    - Non-cirrhotic portal hypertension
    - Prothrombin time prolonged
    - Prothrombin time ratio increased
- Additional primary malignancy (per PSUR):

- SMQ Malignant tumors (SMQ) excluding: PT: Hormone-dependent prostate cancer; PT: Hormone-refractory prostate cancer, PT: Prostate cancer, PT: Prostate cancer metastatic, PT: Prostate cancer recurrent, PT: Prostate cancer stage 0, PT: Prostate cancer stage I, PT: Prostate cancer stage II, PT: Prostate cancer stage III, PT: Prostate cancer stage IV, PTs from HLGT: Metastases, PT: Cancer in remission, LLT: Progression of pre-existing cancer

9.4 Bone health agent (BHA)

Bone health agent were selected in CM dataset, using ATC codes:
M05BA, M05BB: Bisphosphonates
M05BX: Denosumab

Document Type:	Additional Statistical Analysis Plan for CSR addendum
Official Title:	A randomized, double-blind, placebo-controlled Phase III study of darolutamide (ODM-201) versus placebo in addition to standard androgen deprivation therapy and docetaxel in patients with metastatic hormone-sensitive prostate cancer
NCT Number:	NCT02799602
Document Date:	20 JUL 2023

Title Page

A randomized, double-blind, placebo-controlled Phase III study of darolutamide (ODM-201) versus placebo in addition to standard androgen deprivation therapy and docetaxel in patients with metastatic hormone-sensitive prostate cancer

Protocol Number: 17777

SAP Supplement 3 (incl. version and date): V1.0, 20 JULY 2023

Compound Number: BAY 1841788

Sponsor Name and Legal Registered Address:

Non-US: Bayer AG, 51368 Leverkusen, Germany

US territory: Bayer HealthCare Pharmaceuticals Inc., 100 Bayer Boulevard, P.O. Box 915, Whippany NJ 07981-0915, USA

Regulatory Agency Identifier Number(s):

Registry	ID
IND	114769
EudraCT	2015–002590–38

Date: 20 JULY 2023
Version: 1.0

This Statistical Analysis Plan is produced on a word-processing system and bears no signatures.
The approval of the Statistical Analysis Plan is documented in a separate signature document.

Table of Contents

List of Abbreviations

AE	Adverse event
ATC	Anatomical-Therapeutic-Chemical
CI	Confidence interval
MedDRA	Medical Dictionary for Regulatory Activities
NCI-CTCAE v4.03	National Cancer Institute-Common Terminology Criteria for Adverse Events; version 4.03
PT	Preferred Term
SAP	Statistical Analysis Plan
TEAE	Treatment emergent adverse events
WHO	World Health Organization
WHO-DD	WHO Drug Dictionary

1. Introduction

This Supplemental 3 Statistical Analysis Plan (SAP) describes post-hoc analyses that were not included in the pre-defined Supplemental 2 SAP but may be used for Clinical Study Report (CSR) addendum and marketing authorization / NDA.

This Supplemental 3 SAP version 1.0 is based on integrated protocol version 5.0 (amendment 7) dated 26 MAY 2020, Supplemental 2 SAP version 1.0 dated 25 APR 2023, Supplemental 1 SAP version 1.0 dated 04 FEB 2022 and main SAP version 4.1 dated 11 NOV 2021.

2. Study Objectives

Refer to Supplemental 2 SAP V1.0 dated 25 APR 2023.

3. Study Design

Refer to Supplemental 2 SAP V1.0 dated 25 APR 2023.

4. General Statistical Considerations

Refer to Supplemental 2 SAP V1.0 dated 25 APR 2023.

5. Analysis Sets

Refer to Supplemental 2 SAP V1.0 dated 25 APR 2023.

6. Statistical Methodology

Refer to Supplemental 2 SAP V1.0 dated 25 APR 2023.

6.1 Population characteristics

6.1.1 Disposition

Refer to Supplemental 2 SAP V1.0 dated 25 APR 2023.

6.1.2 Demographic and Other Baseline Characteristics

Refer to Supplemental 2 SAP V1.0 dated 25 APR 2023.

6.1.3 Medical History

Refer to Supplemental 2 SAP V1.0 dated 25 APR 2023.

6.1.4 Prior and Concomitant Medication

A summary of prior systemic antineoplastic therapy will be presented by preferred drug name based on WHO-DD drug record number.

The dictionary used for coding medications was the ATC Class/Subclass and the World Health Organization Drug Dictionary (WHO-DD). All tables will be presented with version 2022SEP of the dictionary.

6.1.5 Extent of Exposure

Refer to Supplemental 2 SAP V1.0 dated 25 APR 2023.

6.2 Efficacy

6.2.1 Primary Efficacy Endpoint

Not applicable.

6.2.2 Secondary Efficacy Endpoint

Not applicable.

6.2.3 Exploratory Efficacy Endpoint

Not applicable.

6.3 Safety

6.3.1 Adverse Events

6.3.1.1 Treatment-emergent Adverse Events

The following additional tables for treatment-emergent adverse events (TEAEs) will be created. In all tables, results will be shown by treatment group.

Adverse events will be analyzed by using of predefined customized MedDRA queries denoted as MedDRA Labeling Groupings (MLG). These MLG are created and centrally maintained by Bayer internal coding experts. AEs will be coded using MedDRA 25.1 or most recent version. The severity (or intensity) of AEs will be documented using the National Cancer Institute-Common Terminology Criteria for Adverse Event version 4.03 (NCI-CTCAE v 4.03). Descriptive summary tables will be presented on all safety parameters by treatment arm.

6.3.1.2 MedDRA Labeling Groupings

Adverse events will be analyzed by using predefined customized MedDRA queries denoted as MedDRA Labeling Groupings (MLG). These MLG are created and centrally maintained by Bayer internal coding experts.

Incidence proportions will be calculated as number of subjects experiencing an event per number of subjects exposed. Incidence proportions, overall risk ratio, risk difference 'Darolutamide – placebo' with 95% confidence interval (CI), risk ratio 'Darolutamide/placebo' with 95% CI, exposure-adjusted incidence rate darolutamide, exposure-adjusted incidence rate placebo, incidence rate ratio will be presented by MedDRA Labeling Groupings (MLG), PT for darolutamide/placebo-related and docetaxel-related treatment-emergent adverse events, separately. No zero-correction will be applied for calculation of the risk ratio, i.e., in case of no events in the placebo arm the risk ratio will not be calculated.

6.3.2 Laboratory Safety Assessments

Refer to Supplemental 2 SAP V1.0 dated 25 APR 2023.

6.3.3 Other Safety Measures

Vital signs data, ECG measurements and ECG findings data based on whole study period will be listed in Section 16.

6.4 Other Exploratory Variables

Not applicable.

6.5 Pharmacokinetics/pharmacodynamics

Not applicable.

7. Document history and changes in the planned statistical analysis

- Supplemental 2 SAP version 1.0 dated 25 APR 2023.
- Supplemental 1 SAP version 1.0 dated 04 FEB 2022
- Main SAP version 4.1 dated 11 NOV 2021

8. References

Refer to Supplemental 2 SAP V1.0 dated 25 APR 2023.

9. Appendix

Refer to Supplemental 2 SAP V1.0 dated 25 APR 2023.