JNJ-56021927 (apalutamide)
ARN-509-003 Documentation of Statistical Methods and Interim Analysis Plans

# TABLE OF CONTENTS

# Statistical Analysis Plan (SAP)

Version 7.0

26 June 2017

## SPARTAN

(**S**elective **P**rostate **AR** **T**argeting with **ARN-509**)

## A Multicenter, Randomized, Double-Blind, Placebo-Controlled, Phase III Study of ARN-509 in Men with Non-Metastatic (M0) Castration-Resistant Prostate Cancer

### ARN-509-003; Phase 3

Sponsor:                           Aragon Pharmaceuticals, Inc*

*Aragon Pharmaceuticals, Inc. is a wholly-owned subsidiary of Johnson & Johnson. Janssen Research & Development, LLC is part of the Janssen Pharmaceutical Companies of Johnson & Johnson and provides various services to its affiliated company, Aragon Pharmaceuticals, Inc.

**Status:**       Approved
**Date:**         26 June 2017
**Prepared by:**  Janssen Research & Development, LLC
**Document No.:** EDMS-ERI-80096900; 7.0

## TABLE OF CONTENTS

## 1. INTRODUCTION

This document describes the planned statistical analyses for Protocol ARN-509-003 the **SPARTAN** Study **(**A Multicenter, Randomized, Double-Blind, Placebo-Controlled, Phase III Study of ARN-509 in Men with Non-Metastatic (M0) Castration-Resistant Prostate Cancer). This analysis plan is meant to supplement the study protocol. Any deviations from this analysis plan will be described in the clinical study report.

## 2. STUDY DESIGN

This is a randomized (2:1), multicenter, double-blind, placebo-controlled, Phase III clinical trial evaluating the efficacy and safety of ARN-509 versus placebo in approximately 1200 men with high risk non-metastatic (M0) castration-resistant prostate cancer (NM-CRPC), defined as PSA Doubling Time (PSADT) ≤ 10 months.

ARN-509 will be administered orally on a continuous daily dosing schedule, at a starting dose of 240 mg per day in the treatment group. Matched placebo will be administered orally on a continuous daily dosing schedule, at a starting dose of 240 mg per day in the placebo group.

Patients will be followed for safety and efficacy as per the schedule of assessments and will remain on study treatment until documented radiographic progression (development of distant metastases as assessed by blinded independent central review [BICR]) or the development of unacceptable toxicity.

Patients discontinuing treatment due to documented radiographic progression will enter the survival follow-up period, where they will be followed for the development of symptomatic progression, initiation of subsequent anti-cancer therapies (in particular, cytotoxic chemotherapy) every 4 months until death, loss of follow-up, or withdrawal of consent, whichever comes first.

Patients discontinuing treatment prior to documented radiographic progression will also enter the survival follow-up period where they will continue to have scheduled disease assessments every 4 months until documented radiographic progression, and will be followed for the development of symptomatic progression and initiation of subsequent anti-cancer therapies (in particular, cytotoxic chemotherapy) every 4 months until death, loss of follow-up, or withdrawal of consent, whichever comes first.

### 2.1 STUDY OBJECTIVES

**Primary Objective**

- To demonstrate superiority in the metastasis-free survival (MFS) of men with high-risk NM-CRPC treated with ARN-509 versus placebo

**Secondary Objectives**

- To compare the overall survival (OS) of men with high risk NM-CRPC treated with ARN-509 versus placebo

- To compare the time to symptomatic progression in men with high-risk NM-CRPC treated with ARN-509 versus placebo

- To compare the time to initiation of cytotoxic chemotherapy in men with high-risk NM-CRPC treated with ARN-509 versus placebo

- To compare the progression-free survival (PFS) of men with high-risk NM-CRPC treated with ARN-509 versus placebo

- To compare the time to metastasis (TTM) in men with high-risk NM-CRPC treated with ARN-509 versus placebo

- To evaluate the safety and tolerability of ARN-509

**Other Objectives**

- To compare patient reported outcomes (PROs) of health-related quality of life and prostate cancer-specific symptoms in men with high risk NM-CRPC treated with ARN-509 versus placebo

- To evaluate the population pharmacokinetics of ARN-509

- To evaluate the effect of ARN-509 on ventricular repolarization in a subset of patients from selected clinical sites [Appendix 8 of the protocol]

- To evaluate exploratory biomarkers predictive of response and resistance to ARN-509 treatment

## 2.2   STUDY ENDPOINTS

**Primary Endpoint**

- Metastasis-Free Survival (MFS)

**Secondary Endpoints**

- Overall Survival (OS)

- Time to symptomatic progression

- Time to initiation of cytotoxic chemotherapy

- Progression-Free Survival (PFS)

- Time to Metastasis (TTM)

**Other Evaluations**

- Health-related quality of life and prostate cancer-specific symptoms

- Type, incidence,  toxicity, timing, seriousness, and relatedness of adverse events and laboratory abnormalities

- PSA response

- Time to PSA progression

- Population pharmacokinetics

- Exploratory biomarkers

- Assessment of ventricular repolarization [Appendix 8 of the protocol]
- Second progression-free survival (PFS2)

## 3. RANDOMIZATION

Upon verification of inclusion and exclusion criteria, eligible patients will be centrally randomized in a 2:1 ratio to either ARN-509 or placebo. The randomization will be stratified as follows:

- PSADT ( ≤ 6 months vs. > 6 months)
- Bone-sparing agent use: Yes vs. No
- Loco-regional disease: N0 vs. N1

In order to ensure accurate and consistent determination of PSADT, the Interactive Voice Recognition System (IVRS) will also provide PSADT calculations (using a linear regression model of the natural logarithm of PSA and time) based on PSA values entered by the sites prior to randomization.[1]

## 4. SAMPLE SIZE DETERMINATION

The primary efficacy analysis will be event-driven and will take place when approximately 372 MFS events have occurred. The study provides 90% power to detect a 30% reduction in the risk of developing metastases (HR = 0.70) for patients receiving ARN-509, with a 2-sided α of 0.05.[2] Based on an assumed median MFS of 25 months in the placebo group, this treatment effect represents an increase in the median MFS of approximately 11 months (from 25 months to 36 months). Assuming an accrual period of 24 months (with 75% of the patients accrued in the second year), approximately 1200 patients will need to be randomized.

The study was also designed to provide 85% power to detect a 25% reduction (HR = 0.75) in the risk of death for patients receiving ARN-509, based on an assumed median OS of 49 months in the placebo group. This treatment effect represents an increase in the median OS of approximately 16 months (from 49 to 65 months).

## 5. INTERIM ANALYSIS

There will be no interim analysis for the primary endpoint (MFS) and for the secondary endpoints of TTM and PFS. There will be 1 interim analysis for time to symptomatic progression, and up to 2 interim analyses for OS and time to initiation of cytotoxic chemotherapy; see Section 7.5.3 for details.

## 6. INDEPENDENT DATA MONITORING COMMITTEE

An independent third-party Data Monitoring Committee (IDMC) will be established to ensure the overall integrity and conduct of the study.

The IDMC will review the progress of the study and cumulative unblinded safety data on a periodic basis (e.g., a minimum of two review meetings per year) as well as review the efficacy analysis when necessary. In addition to the review meetings, blinded (or unblinded,

if necessary) listings of serious adverse events will be provided to the IDMC on a monthly basis.

Following each review meeting, the IDMC will recommend to the Sponsor whether to continue the trial unchanged, modify the conduct of the study, or terminate the study early. Rules for early termination, modification or continuation of the study, as well as how these recommendations will be made to the Sponsor and Health Authorities will be outlined in a separate IDMC Charter.

The IDMC will be composed of 3 external members [2 physicians and 1 biostatistician] not associated with the conduct of the study. The Sponsor will also designate an independent biostatistician not affiliated with the project to prepare and provide study data to the IDMC. Complete details regarding the composition and governance of the IDMC will be outlined in the IDMC Charter.

Periodic adverse event data review will also be performed by designated members of the Sponsor's primary study team and will be blinded to treatment assignment with adverse event data from both treatment groups combined. Any safety issues of concern identified by the primary study team that require notification of the IDMC will be communicated as described in the IDMC Charter.

## 7. STATISTICAL METHODS

The primary objective of the study is to evaluate the efficacy of ARN-509 compared to placebo in patients with high risk NM-CRPC as measured by metastasis-free survival (MFS), based on blinded independent central review (BICR) of tumor assessments.

### 7.1 ANALYSIS POPULATIONS

**Full Analysis (Intent-to-Treat) Population [ITT]**: All eligible patients who are randomized into the study, with study drug assignments designated according to initial randomization, regardless of whether patients receive study drug or receive a different drug from that to which they were randomized to will be included in the analyses of all efficacy and clinical benefit endpoints and patient characteristics.

**Safety Analysis Population [SAFETY]**: All patients who receive at least one dose of study drug, with treatment assignments designated according to actual study treatment received will be the primary population for evaluating safety and treatment compliance and administration.

- **Population Pharmacokinetics Population [PK]**: Subset of the safety analysis population that was randomized to the ARN-509 treatment group and that has at least one PK sample collected.

- **Biomarker Population**: Subset of the safety analysis population that has at least 1 biomarker sample collected.

### 7.2 DEFINITIONS

**Study Day**: Study day will be calculated in reference to the date of randomization for randomized untreated subjects and in reference to the date of first dose for treated subjects.

Study Day 1 corresponds to the date the subject was randomized into the study or to the date of first dose the subject was treated.

**Baseline Value**: Unless otherwise specified, the baseline value will be defined as the closest measurement prior to the first dose of study drug. Change from baseline will be defined as (post-baseline value – baseline value).

**Treatment Duration**: Treatment duration will be defined as the duration of time from the date of the first dose of study drug to the date of last dose of study drug + 1 day.

**Time to event**: Time to event calculations will be defined as the time from randomization to date of event + 1 day. Time to event or duration of event endpoints will be based on the actual date of the event, not visit number or visit label.

**Survival Follow-Up Phase**: The survival follow-up phase will start from the safety follow-up visit (28 days following the last dose of study drug) and continue through the end of the study.

## 7.3    ANALYSIS OF STUDY CONDUCT

All patients randomized into the study will be summarized by treatment group as randomized. The randomization stratification factors will be listed and tabulated as recorded in the IVRS by treatment group as randomized.

The number of patients who are in the ITT and SAFETY populations will be summarized by treatment group and overall.

Study treatment administration, duration of follow-up, discontinuation from study treatment and the reasons for discontinuation will be summarized by treatment group for all randomized patients. In addition, major protocol deviations and eligibility violations will also be summarized by treatment group. Protocol violations for this study include, but are not limited to, the following:

- Failure to meet inclusion/exclusion criteria
- Use of a prohibited concomitant medication
- Dose modifications that are not within the protocol specifications
- Any other deviation that presents significant risk or safety concerns to the patient

## 7.4    ANALYSIS OF TREATMENT GROUP COMPARABILITY

The evaluation of treatment group comparability between the 2 treatment groups will include summary of demographics, baseline disease characteristics, medical history, and patient treatment history.

Descriptive statistics (mean, standard deviation, median, range) will be presented by treatment group for continuous variables such as age and time from initial diagnosis. Categorical variables will be summarized using frequencies and percentages.

### 7.4.1    Demographics and Baseline Characteristics

The following parameters will be summarized overall by treatment group as randomized for all patients in the ITT population: Age, race, ethnicity and baseline weight and height.

### 7.4.2    Disease Characteristics and Prior Therapy

The following parameters will be summarized overall by treatment group as randomized for all patients in the ITT population:

- ECOG performance status

- Baseline PSA value and PSA Doubling Time (PSADT)

- Time (months) from initial diagnosis of prostate cancer to randomization

- Total Gleason Score at initial diagnosis

- Tumor stage at initial diagnosis

- Lymph nodes stage at initial diagnosis

- Number of prior hormonal therapies

- Use of bone-sparing agent (Yes/No) at baseline

- History of surgical prostate cancer procedures (Yes/No)

- Type of surgical procedures (prostatectomy, orchiectomy, transurethral resection of the prostate (TURP), and other)

- History of radiotherapy (Yes/No)

- Type of radiotherapy

- History of prior adjuvant/neoadjuvant chemotherapy (Yes/No)

### 7.4.3    Study Drug Administration

The SAFETY population will be used to summarize drug exposure, treatment compliance, and dose modifications by treatment group as treated.

Treatment duration will be defined as the duration of time from the date of the first dose of study drug to the date of last dose of study drug + 1 day.

The total cumulative dose in milligrams (mg) will be calculated as 30 or 60 mg multiplied by the number of capsules or tablets taken. The number of capsules or tablets taken will be calculated based on the number of capsules or tablets dispensed at the study visits minus the number of capsules or tablets indicated as having been returned.

The overall treatment compliance will be defined as the total dose in mg taken during the study divided by the expected total dose in mg.  A subject's expected total dose will be calculated as the assigned dose per day multiplied by treatment duration.  Each patient should be taking 8 capsules or 4 tablets per day maximum while on the study.  For patients with dose reductions, the expected number of capsules or tablets will be reflective of the new dose with a reduced number of total capsules or tablets.

Dose reduction or interruption and the reason for the dose reduction or interruption will be summarized by treatment group as treated.

### 7.4.4    Pre-study and Concomitant Medications

Concomitant medications and medications taken prior to starting study treatment will be summarized for all patients in the SAFETY population by treatment group as treated. Medications are considered concomitant if taken during the treatment-emergent period. Prior medications are medications with the start date and/or end date before study drug date. Medications will be summarized by WHO Drug therapeutic class and generic medication name.

### 7.4.5    Subsequent Anti-Cancer Therapy

Subsequent systemic anti-cancer therapy taken after the patient has discontinued study drug will be summarized by treatment group and overall. Medications will be summarized by WHO Drug therapeutic class and generic medication name.

## 7.5    EFFICACY ANALYSES

The following section outlines the planned analyses of the primary and secondary efficacy outcomes of the study.

Efficacy analyses will be performed on the ITT population, incorporating the randomization stratification factors as documented on the IVRS, unless otherwise specified.

Time-to-event endpoints will be summarized using the Kaplan-Meier method [3] and displayed graphically where appropriate. Median event times and 2-sided 95% confidence interval for each median will be provided. Cox proportional-hazard models, including the stratification factors at baseline, will be used to estimate the hazard ratio (HR) and its 95% confidence interval (CI).

Response endpoints (eg, PSA response rate) will be summarized using descriptive statistics for categorical data by treatment group. The relative risk (treatment:control) will be reported along with the associated 2-tailed 95% CIs. The two treatment groups will be compared using the stratified Mantel-Haenszel test; Fisher's exact test may be used if the expected counts in some cells are small.

### 7.5.1    Blinded Independent Central Review (BICR)

Analyses of efficacy endpoints which are based on radiographic tumor assessments (MFS, TTM, and PFS) will be based on the results of the blinded independent central review (BICR), provided via electronic data transfer by the third-party core imaging laboratory.

All scans will be submitted for independent review of disease progression during the study according to an Independent Review Charter to be prepared by the core imaging laboratory in consultation with the Sponsor.

### 7.5.2    Primary Endpoint (MFS)

The primary efficacy endpoint is metastasis-free survival (MFS), defined as the time from randomization to first evidence of BICR-confirmed radiographically detectable bone or soft tissue distant metastasis (simply referred to as "metastasis" from this point forward) or death due to any cause (whichever occurs earlier) + 1 day.

MFS data for patients without metastasis or death will be censored on the date of the last tumor assessment (or, if no tumor assessment was performed after the baseline visit, at the date of randomization + 1 day).

Additional censoring rules will vary according to whether the analysis is performed for US or ex-US regulatory purposes, as follows; both results will be provided in the clinical study report.

| Scenario | US regulatory guidance | ex-US regulatory guidance[4] |
|---|---|---|
| Data from patients who are lost to follow-up or whose disease progression (development of metastasis) or death occurs after 2 or more consecutively missing or unevaluable tumor assessments | Censored on the date of the last tumor assessment that the patient was known to be metastasis-free | Time of progression will be determined using the first date when there is documented evidence of progression or death (whichever occurs earlier) regardless of missed or unevaluable tumor assessments |
| Patients that receive new systemic anti-cancer therapy prior to documented disease progression (development of metastasis) or death | Censored on the date of the last tumor assessment prior to the start of the new systemic anti-cancer therapy | Time of progression will be determined using the first date when there is documented evidence of progression or death (whichever occurs earlier) regardless of change of therapy |

Disease progression (development of metastasis) will be assessed by an independent core imaging laboratory using the Response Evaluation Criteria in Solid Tumors (RECIST v1.1).[5] The appearance of new (or distant) metastatic lesions denotes disease progression. For new bone lesions detected on bone scans, a second imaging modality (e.g., CT or MRI) will be required to confirm progression.

The primary efficacy analysis will be completed when approximately 372 MFS events have occurred. The primary analysis will compare the MFS distributions in the two treatment groups using a two-sided log-rank test, stratified by PSADT ($\leq$ 6 months vs. > 6 months), the use of a bone-sparing agent (Yes vs. No), and the presence of loco-regional disease (N0 vs. N1) at the 0.05 significance level. The non-stratified log-rank test will be provided as a sensitivity analysis. Additional sensitivity analyses are provided in Section 7.8

### 7.5.3 Secondary Endpoints

Time-to-event-based secondary endpoint analyses (OS, time to symptomatic progression, time to initiation of cytotoxic chemotherapy, progression-free survival, and time to metastasis) will be performed using a two-sided stratified log-rank test. Non-stratified log-rank tests will be provided as sensitivity analyses for OS.

For ease of notation, in the sequel, time to symptomatic progression is abbreviated as
"SymProg" and time to initiation of cytotoxic chemotherapy is abbreviated as "CytoChemo".
A hierarchical adaptive group sequential procedure will be used to test the secondary
endpoints (Figure ). The method allows re-estimation of the number of events necessary for
the next best opportunity to achieve the desired conditional power. The method controls the
familywise type I error rate for the primary and all secondary endpoints.

**Figure 1: Hierarchical Adaptive Group Sequential Procedure**
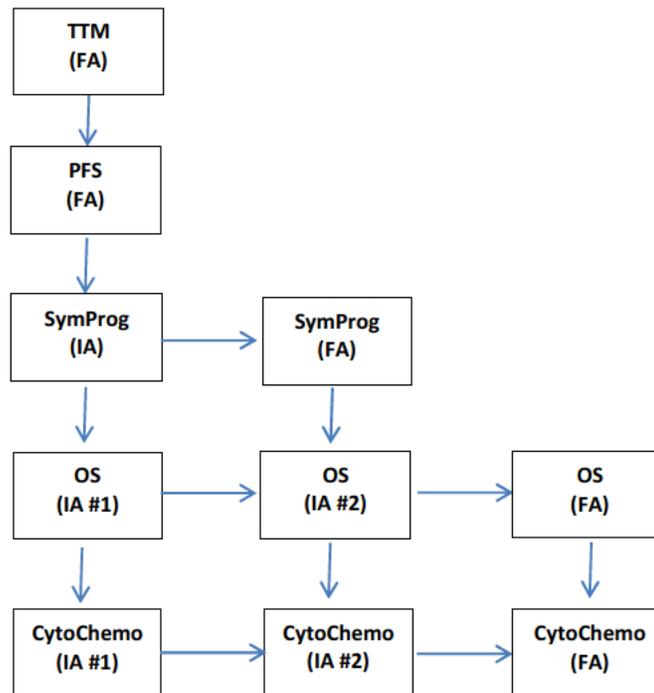


IA=interim analysis; FA=final analysis

Specifically, a hierarchical testing will be performed in the following order: TTM, PFS,
SymProg, OS, and CytoChemo, each at alpha=0.05 (2-sided). As depicted in Figure 1, each
endpoint will have a final analysis (FA) but there will be no interim analysis (IA) for TTM
and PFS, 1 IA for SymProg, and up to 2 IAs for OS and CytoChemo. The testing of
SymProg, OS and CytoChemo endpoints will utilize an adaptive group sequential method,
according to the pre-specified O'Brien-Fleming (OBF)-type alpha spending function[6] with
possible re-estimation of the required number of events necessary for the next analysis to
maintain the desired conditional power. If SymProg is significant at the IA, then there will be
only 1 IA for OS and CytoChemo (ie, with "IA #2" boxes removed from Figure 1); otherwise
there will be 2 IAs for OS and CytoChemo.

13

JNJ-56021927 (apalutamide)
ARN-509-003 Documentation of Statistical Methods and Interim Analysis Plans

The FA of TTM and PFS, IA of SymProg, and the first IAs of OS and CytoChemo will all be performed at the same time as the primary analysis of MFS (approximately 372 events). At the time of the MFS analysis, testing of the SymProg, OS and CytoChemo endpoints will be carried out using the OBF-type alpha spending function and nominal p-values will be provided.

Assuming that the primary test for MFS is significant at alpha=0.05, TTM will then be tested at alpha=0.05. If significant, then PFS will be tested at alpha=0.05. If PFS is significant, then the Sponsor will perform the IA for SymProg. Assuming HR=0.75 for SymProg, then 427 events are required to yield 80% power at alpha=0.05. Suppose that *X1* events will have accumulated at the IA for SymProg, the information fraction would be *t=X1*/427. Efficacy of SymProg will be concluded if the stratified log-rank test statistic exceeds the critical value derived from the OBF-type alpha spending function.

The SymProg events re-estimation will be based on a conditional power of 90% for the next stage, calculated using the observed HR. Note that because of the variability associated with the observed HR, a conditional power of 90% is used in order to maintain an overall power of approximately 80% if the true HR is 0.75. The recommended number of SymProg events for the FA should be in the range of 191 (additional 191-*X1* events set as the minimum required for the next stage) to *X1*+(427-*X1*)*1.1 (not more than 10% increase on the pre-planned number of events for the next stage). Because the true HR for SymProg is unknown, the minimum number of 191 is chosen to yield at least 80% power to detect an HR of 0.65 at an alpha level of 0.05.

The SymProg events re-estimation will be performed by the independent statistician who supports the IDMC activities. The dissemination and review of the specific results of the IA will be limited to the IDMC.

In order to maintain a strong control of the type I error rate for the SymProg analysis, an inverse normal p-value combination method will be used as the final test.[7] The inverse normal p-value combination method allows flexible adaptations at an IA and creates a valid test that controls the type I error rate in a strong sense analytically. In this proposed design the adaptation is the potential adjustment of the required number of events for the next stage.

The final test statistics for the null hypothesis $H_0$: HR for SymProg $\geq 1$ is defined as

$$Z = w_1 F^{-1}(1 - p_1) + w_2 F^{-1}(1 - p_2),$$

where $F^{-1}(x)$ is the inverse of the standard normal cumulative distribution function, $w_1 = \sqrt{t}$, $w_2 = \sqrt{1 - t}$, $p_1$ denotes the first stage p-value and $p_2$ denotes the second stage p-value.

Critical values for success are calculated based on the OBF-type alpha spending function. The study may be stopped early for efficacy if the interim test statistic exceeds the first stage critical value $z_1$, or stops for success at the second stage, if the final test statistic exceeds the second stage critical value $z_2$. Therefore, the null hypothesis $H_0$ will be rejected either at the first analysis if $F^{-1}(1 - p_1) > z_1$, or at the FA if $Z > z_2$.

The adaptive group sequential testing of OS and CytoChemo will be similar to that of SymProg as described above, except that there is a possibility of 1 more IA to be incorporated in the inverse normal p-value combination method. Due to the hierarchical

structure of the testing procedure, an outline of some additional technical details is provided in Appendix 1.

### 7.5.3.1   Overall Survival

OS will be defined as the time from randomization to the date of death due to any cause + 1 day.  Patients who are alive at the time of the analysis will be censored on the last known date that they were alive.  In addition, the following censoring rules will apply:

| Scenario | Date of Censoring |
|---|---|
| Patients with no post-baseline information | Censored on the date of randomization + 1 day |
| Patients who are lost to follow-up or who withdraw consent for further follow-up | Censored on the last known date that they were alive |

Survival rate at 1-, 2-, 3- and 5-year will be estimated using the Kaplan-Meier method. Additional sensitivity analyses are provided in Section 7.8.

### 7.5.3.2   Time to Symptomatic Progression

Time to symptomatic progression will be defined as the time from randomization to documentation in the CRF of any of the following (whichever occurs earlier) + 1 day:

- Development of a skeletal-related event (SRE): pathologic fracture, spinal cord compression, or need for surgical intervention or radiation therapy to the bone.

- Pain progression or worsening of disease-related symptoms requiring initiation of a new systemic anti-cancer therapy.

- Development of clinically significant symptoms due to loco-regional tumor progression requiring surgical intervention or radiation therapy.

Adverse event, concomitant medication, or survival follow-up CRFs may also be the source of these findings.

Time to symptomatic progression for patients who do not experience any of the events described above will be censored on the date on which they were last known to be event-free.

### 7.5.3.3   Time to Initiation of Cytotoxic Chemotherapy

Time to initiation of cytotoxic chemotherapy will be defined as the time from randomization to documentation of a new cytotoxic chemotherapy being administered to the patient (e.g., survival follow-up CRF) + 1 day.

Time to initiation of cytotoxic chemotherapy for patients who do not start a cytotoxic chemotherapy will be censored on the date of last contact.

### 7.5.3.4   Progression-Free Survival

In order to capture loco-regional disease progression, a secondary endpoint of progression-free survival (PFS) will be assessed and defined as the time from randomization to first

JNJ-56021927 (apalutamide)
ARN-509-003 Documentation of Statistical Methods and Interim Analysis Plans

documentation of BICR-confirmed radiographic progressive disease or death due to any cause (whichever occurs first) + 1 day.

Progressive disease (PD) will be based on RECIST v1.1, and further defined as follows:

- For patients with at least one measurable lesion, PD will be defined as at least a 20% increase in the sum of diameters of target lesions taking as reference the smallest sum on study (this includes the baseline sum if that is the smallest on study). In addition to the relative increase of 20%, the sum must also demonstrate an absolute increase of at least 5 mm. Furthermore, the appearance of one or more new lesions is also considered progression.

- For patients with only non-measurable disease observed on CT or MRI scans, unequivocal progression (representative of overall disease status change) or the appearance of one or more new lesions will be considered progression. For new bone lesions detected on bone scans, a second imaging modality (e.g., CT or MRI) will be required to confirm progression.

Progression-free survival data for patients without loco-regional disease will be censored on the date of the last tumor assessment (or, if no tumor assessment was performed after the baseline visit, at the date of randomization + 1 day). Additional censoring rules will vary according to whether the analysis is performed for US or ex-US regulatory purposes, as shown below; both results will be provided in the clinical study report.

| Scenario | US regulatory guidance | ex-US regulatory guidance |
|---|---|---|
| Data from patients who are lost to follow-up or whose disease progression or death occurs after 2 or more consecutively missing or unevaluable tumor assessments | Censored on the date of the last tumor assessment that the patient was known to be progression-free | Time of progression will be determined using the first date when there is documented evidence of progression or death (whichever occurs earlier) regardless of missed or unevaluable tumor assessments |
| Patients that receive new systemic anti-cancer therapy prior to documented disease progression or death | Censored on the date of the last tumor assessment prior to the start of the new systemic anti-cancer therapy | Time of progression will be determined using the first date when there is documented evidence of progression or death (whichever occurs earlier) regardless of change of therapy |

### 7.5.3.5 Time to Metastasis

Time to Metastasis (TTM) will be defined as the time from randomization to first evidence of BICR-confirmed radiographically detectable bone or soft tissue distant metastasis + 1 day.

TTM data for patients without metastasis will be censored on the date of the last tumor assessment (or, if no tumor assessment was performed after the baseline visit, at the date of randomization + 1 day). Additional censoring rules will vary according to whether the analysis is performed for US or ex-US regulatory purposes, as shown below; both results will be provided in the clinical study report.

| Scenario | US regulatory guidance | ex-US regulatory guidance |
|---|---|---|
| Data from patients who are lost to follow-up or whose disease progression (development of metastasis) occurs after 2 or more consecutively missing or unevaluable tumor assessments | Censored on the date of the last tumor assessment that the patient was known to be metastasis-free | Time of progression will be determined using the first date when there is documented evidence of progression regardless of missed or unevaluable tumor assessments |
| Patients that receive new systemic anti-cancer therapy prior to documented disease progression (development of metastasis) | Censored on the date of the last tumor assessment prior to the start of the new systemic anti-cancer therapy | Time of progression will be determined using the first date when there is documented evidence of progression regardless of change of therapy |

## 7.6   SAFETY ANALYSES

### 7.6.1   Adverse Events

Patients will be assessed for adverse events at each monthly clinic visit while on the study. Adverse events (AEs) will be graded according to the NCI Common Terminology Criteria for Adverse Events (CTCAE) Version 4.0 and coded to preferred term and system organ class (SOC) using the most recent version of MedDRA.

All AEs reported during the AE reporting period (inclusive of the 28-day post last dose of study drug period) will be considered as treatment-emergent adverse events and will be summarized by treatment group as treated using all patients in the SAFETY population.

For each treatment group, adverse event incidence rates will be summarized with frequency and percentage by MedDRA SOC and preferred term, with all patients treated in that treatment group as the denominator, unless otherwise specified. In addition, AE incidence rates will also be summarized by severity and relationship to study drug. Treatment-related AEs are those judged by the Investigator to be at least possibly related to the blinded study drug. Patients with multiple occurrences of events will only be counted once at the maximum severity to study drug for each preferred team, SOC, and overall. Deaths that occur within 28 days after the last dose of study drug are defined as on-study deaths.

Summary tables of the following AEs will be provided:

- Overall summary of AEs: the number and percentage of patients who experienced any AE, any serious adverse event (SAE), any treatment-related AE, any treatment-related SAE, any discontinuations due to an AE, and any deaths

- All AEs by SOC and preferred term

- All AEs by SOC, preferred term, and toxicity grade

- All AEs by decreasing frequency of preferred term

- Grades 3 or 4 AEs by SOC, preferred term and toxicity grade

- Drug-related AEs by SOC and preferred term

- Drug-related AEs by SOC, preferred term, and toxicity grade

- Drug-related Grades 3 or 4 AEs by SOC, preferred term and toxicity grade

- AEs that led to study drug discontinuation by SOC and preferred term. Study drug discontinuation will be determined from the End of Treatment CRF (where reason for termination is "Adverse Event") and the specific AE will be determined from the AE CRF page (where action taken is "Withdrawn from Study")

- AEs that led to study drug discontinuation by SOC, preferred term, and toxicity grade.

- All SAEs by SOC and preferred term

- All SAEs by SOC, preferred term, and maximum severity

- Deaths will be summarized by time period (on-study vs. during follow-up) and cause of death.

Patient listings of all Grades 3 or 4 AEs, all SAEs, AEs that led to study drug discontinuation and all deaths will be provided as well.

Narratives will be written for the following patients in the final clinical study report:

- Patients who die within 28 days of the last dose of study drug

- Patients who discontinue study drug due to adverse events

- Patients who have a serious adverse event

- Patients who experience a seizure

- Grade 3 or higher adverse events of special interest

### 7.6.2　Laboratory Abnormalities

Only data collected by the central laboratory will be summarized. Local laboratory data collected for the purposes of planning treatment administration, dose modification, or monitoring adverse events, will not be summarized.

Normal ranges will be used to identify values that are outside the normal ranges and abnormal laboratory results will be graded according to the NCI CTCAE Version 4.0.

Descriptive statistics will be provided for selected test results and for the change from baseline by visit.

A shift summary of baseline grade by maximum post-baseline CTCAE grade will be presented, as appropriate. For each laboratory parameter, the baseline laboratory value will be defined as the last laboratory value collected on or prior to the date of the first dose of study drug. Patients who develop toxicities of Grade $\geq 3$ will be listed.

### 7.6.3   Vital Signs

Each vital sign (temperature, blood pressure (systolic and diastolic), respiration rate, and heart rate) at baseline will be summarized and presented by treatment group. The number and percentage of subjects with marked abnormalities in vital signs as compared to baseline will be summarized and listed.

Data will be summarized and presented according to the following categories:

| Parameter | Criteria for Marked Abnormality |
|---|---|
| Systolic Blood Pressure | Absolute result > 160 mmHg and increase from baseline > 20 mmHg |
| | Absolute result < 90 mmHg and decrease from baseline > 20 mmHg |
| Diastolic Blood Pressure | Absolute result > 100 mmHg and increase from baseline > 10 mmHg |
| | Absolute result < 50 mmHg and decrease from baseline > 10 mmHg |
| Heart Rate | Absolute result > 100 bpm and increase from baseline > 30 bpm |
| | Absolute result < 60 bpm and decrease from baseline > 20 bpm |

## 7.7   OTHER EVALUATIONS

### 7.7.1   Second Progression-Free Survival (PFS2)

PFS2 is defined as the time from randomization to second documentation of investigator-assessed disease progression (PSA, radiographic, symptomatic, or any combination) or death (any cause) on subsequent treatment (whichever occurs first) + 1 day. Subjects without a documented progression will be censored at the last date known to be progression-free or death whichever occur first.

PFS2 will be analyzed using time-to-event analysis methods outlined in section 7.5.

### 7.7.2   PSA

PSA kinetics (e.g., PSA response and time to PSA progression) will be assessed at the time of the primary analysis of MFS according to the Prostate Cancer Clinical Trials Working Group (PCWG2) criteria.[8]

Summary tables and waterfall plots describing change in PSA relative to baseline will be reported at 12 weeks (or earlier for those who discontinue study treatment prior to 12 weeks), and separately, the maximum change at any time on study will also be reported for each patient using summary tables and waterfall plots.

PSA response rate will be summarized by treatment group. The relative risk (treatment:control) will be reported along with the associated 2-tailed 95% CIs. The two treatment groups will be compared using the stratified Mantel-Haenszel test.

The time to PSA progression will be calculated as the time from randomization to the time when the criteria for PSA progression according to PCWG2 are met + 1 day. Kaplan-Meier methods will be used to estimate the median time to PSA progression and 95% confidence intervals for each treatment group

### 7.7.3 Health-Related Quality of Life and Prostate Cancer-Specific Symptoms

The FACT-P and EQ-5D data will be scored and handled as recommended in their respective User's manuals, including handling of missing data both within the subscales and overall.[9,10] All FACT-P and EQ-5D data analyses will be performed in the ITT population. On-study scores and change from baseline scores will be summarized and displayed graphically as appropriate.

A 10-point change in the FACT-P total score is considered clinically meaningful.[11] Therefore, any patient experiencing a 10-point decrease in FACT-P total scores from baseline at any post baseline time point will be considered to have experienced clinically meaningful deterioration in functional status. The proportions of patients with a 10-point decrement in FACT-P total score will be summarized by treatment group, and the two treatment groups will be compared using a Mantel-Haenszel test, stratified by PSADT ($> 6$ months vs. $\leq 6$ months), the use of a bone-sparing agent (Yes vs. No), and the presence of loco-regional disease (N0 vs. N1) at a two-sided 0.05 significance level.

The EQ-5D data will be summarized descriptively by treatment group and study visit.

Details on the PRO analyses will be provided in a separate PRO SAP.

### 7.7.4 Population Pharmacokinetic Analyses

A separate Population PK analysis plan will be prepared prior to database lock. A separate report will be generated.

### 7.7.5 Exploratory Biomarker Analyses

Biomarker analyses are exploratory and data generated will be reported separately.

### 7.7.6 Assessment of Ventricular Repolarization

The assessment of ventricular repolarization will be a sub-study conducted in a subset of patients from selected clinical sites and analyzed by an independent cardiac safety laboratory. Description of the sub-study is provided in Appendix 8 of the study protocol. A separate analysis plan will be prepared.

### 7.7.7 Medical Resource Utilization Analysis

Protocol-mandated procedures, tests, and encounters are excluded. The MRU data may be used to conduct economic analyses.

## 7.8    SENSITIVITY ANALYSES

Sensitivity analyses will be performed on the primary efficacy endpoint of MFS and secondary endpoint of OS to support the results in the primary analysis.

### 7.8.1    Investigator-Derived MFS

A sensitivity analysis will be performed on the primary endpoint of MFS using the investigator-derived progression (development of metastasis or death). Patients without a MFS event will be censored on the known date of progression-free according to the investigator tumor assessment.

### 7.8.2    OS

Sensitivity analyses for the OS may be carried out as appropriate if it is deemed useful to aid in the interpretation of the results.

#### 7.8.2.1    Covariate Effects

A non-stratified multivariate analysis will be performed on the OS endpoint to estimate treatment effect adjusting for important baseline factors. The following baseline covariates will be considered:

- PSADT ( $\leq$ 6 months vs. > 6 months)
- Bone-sparing agent use (Yes vs. No)
- Loco-regional disease (N0 vs. N1)
- ECOG Performance status (0 vs. 1)
- Number of prior hormonal therapies (1 vs. $\geq$ 2)
- Gleason score ($\leq$ 7 vs. $\geq$ 8)
- Age (continuous)
- Logarithm of PSA (continuous)

The adjusted hazard ratio and its 95% confidence interval for treatment and each factor will be estimated using Cox regression.

#### 7.8.2.2    Other Sensitivity Analyses for OS

A large number of subjects are expected to receive life-extending subsequent therapies, the following analyses may be used in estimating the true treatment effect.

1. Inverse Probability of Censoring Weighted (IPCW) log-rank Tests by Robins et al[12,13,14,15] will be used to estimate the treatment effect and its associated confidence interval.

2. Using a time-dependent Cox regression; the HR prior to receiving subsequent anticancer therapy and after receiving the subsequent anticancer therapy will be estimated, the associated 95% confidence interval will also be calculated.

Additional analyses may be performed for OS if appropriate.

### 7.8.3    Subgroup Analyses

In order to assess the consistency of treatment benefit with respect to the primary efficacy endpoint of MFS and secondary endpoint of OS across important subgroups, forest plots will be provided for the following variables:

- ECOG performance status (0 vs. 1) at baseline

- Age category (< 65 vs. ≥ 65 years and < 75 vs. ≥ 75 years)

- Race (white, black, Asian, and others)

- Geographic region (NA, EU, and ROW)

- Number of prior hormonal therapies (1 vs. ≥ 2)

- Baseline PSA value (at or below median vs. above median)

- PSADT (> 6 months vs. ≤ 6 months)

- Bone-sparing agent use (Yes vs. No)

- Loco-regional disease (N0 vs. N1)

The comparison between treatment groups will be evaluated by a single hazard ratio with its 95% confidence interval based on a non-stratified Cox regression model for each subgroup.

## 7.9    ANALYSIS BY FORMULATION

During the conduct of the study, treatment formulation was switched from capsule to tablet, at which time a significant number of subjects have been already enrolled and have received capsule formulation.  The randomization and blinding remained unchanged during the formulation switch.  Newly enrolled subjects began treatment using tablets while subjects already enrolled and receiving capsules made the switch to tablets at the start of a new cycle.

As outlined in previous sections, the primary analyses will be based on the Safety and ITT Populations irrespective of the formulation switch.  However, to show that the two formulations have no clinically relevant impact on key efficacy and safety data, supplemental descriptive analyses will be performed.  For safety analyses, the descriptive analysis by formulation subgroups (capsule only, capsule/tablet, and tablet only) will be performed.

### 7.9.1    Subject Distribution by Formulation

For each formulation subgroup the following summaries by treatment group will be generated:

- Number of subjects randomized, treated, discontinued and ongoing
- Primary reasons for study treatment discontinuation

### 7.9.2    Demographics and Baseline characteristics by Formulation

The parameters outlined in section 7.4.1 will be summarized by treatment group and formulation.  The summary will include subjects as randomized.

### 7.9.3    Study Drug Administration by Formulation

Descriptive statistics will be provided on dose modifications by treatment group and formulation and the subjects will be included as treated.

### 7.9.4    Safety Analyses by Formulation

The by-formulation safety analysis will use subjects as treated and will include the following AEs summaries:

- Overall summary of AEs (See Appendix 2 for table shell)
- All AEs by System Organ Class (SOC), Preferred Term (PT) and toxicity grade
- Grade 3/4 AEs by SOC, PT and toxicity grade

### 7.9.5    Efficacy (MFS and PSA Response) Analyses by Formulation

The primary MFS analysis and statistical inference will be based on the ITT population regardless of the formulation as defined in section 7.5.2.  In order to assess the consistency of treatment benefit (relative to ITT analysis), a supplemental analysis by formulation subgroup will be performed on the BICR derived MFS endpoint and PSA response.

The following formulation subgroups will be used in the analysis:

1. Capsule only, tablet only and capsule + tablet
2. Two groups based on greater duration of exposure on specific formulation: greater duration on capsule versus greater duration on tablet.

The MFS subgroup analysis will include the following:

- Number of events and censored
- Distribution of MFS endpoint using the Kaplan-Meier method
- Non-stratified estimate of the hazard ratio and its associated 95% confidence interval using the Cox model.

## 8. REFERENCES

1. Pound CR, Partin AW, Eisenberger MA, et al.  Natural history of progression after PSA elevation following radical prostatectomy.  JAMA  1999; 281:1591-7.

2. Collett D.  Modeling Survival Data in Medical Research.  Boca Raton:  Chapman and Hall/CRC, 1994, pp 254-5.

3. Kaplan EL and Meier P.  Nonparametric estimation from incomplete observations. J Am Stat Assoc  1958; 53:457-81.

4. Appendix 1 to the guideline on the evaluation of anticancer medicinal products in man: Methodological consideration for using progression-free survival (PFS) or disease-free survival (DFS) in confirmatory trials. European Medicines Agency 2011; Doc. Ref. EMA/CHMP/27994/2008.

5. Eisenhauer EA, Therasse P, Bogaerts J, et al.  New response evaluation criteria in solid tumours: Revised RECIST guideline (version 1.1).  Eur J Cancer  2009; 45:228-47.

6. Reboussin DM, DeMets DL, Kim KM, and Lan KKG. Computations for group sequential boundaries using the Lan-DeMets spending function method. Controlled Clinical Trials  2000; 21:190-207.

7. Wassmer, G.  Planning and analyzing adaptive group sequential survival trials. Biometrical Journal 2006; 48:714-729.

8. Scher HI, Halabi S, Tannock I, et al.  Design and end points of clinical trials for patients with progressive prostate cancer and castrate levels of testosterone: Recommendations of the Prostate Cancer Clinical Trials Working Group.  J Clin Oncol  2008; 26:1148-59.

9. Esper P, Mo F, Chodak G, et al.  Measuring quality of life in men with prostate cancer using the Functional Assessment of Cancer Therapy-Prostate (FACT-P) instrument.  Urology 1997; 50:920-8.

10. The EuroQol Group.  EuroQol – a new facility for the measurement of health-related qualify of life.  Health Policy 1990; 16:199-208.

11. Cella D, Nichol MB, Eton D et al. Estimating clinically meaningful changes for the Functional Assessment of Cancer Therapy-Prostate: Results from a clinical trial of patients with metastatic hormone-refractory prostate cancer. Value in Health 2009; 12:124-129.

12. Herna M, Brumback B, Robins J. Marginal structural models to estimate the causal effect of zidovudine on the survival of HIV-positive men.  Epidemiology 2000; 11: 561-570.

13. Robins J, Finkelstein D. Correcting for noncompliance and dependent censoring in an AIDS clinical trial with inverse probability of censoring wieghted (IPCW) log-rank tests.  Biometrics 2000; 56:779-788.

14. Robins J, Hernan N, Brumback B. Marginal structural models and causal inference in epidemiology. Epidemilogy 2010; 11: 550-560.

15. Westreich D, Cole S et al. Time scale and adjusted survival curves for marginal structural cox models. Am J Epidemiology 2010; 171: 691-700.

## 9. APPENDICES

## 9.1 APPENDIX 1: OUTLINE OF THE ADAPTIVE GROUP SEQUENTIAL TESTING OF OS AND CYTOCHEMO

Both OS and CytoChemo are tested possibly at 3 times each: at the same time as IA for SymProg, at the same time as FA for SymProg, and then at their respective FA.

- ❖ If SymProg is significant at IA
  - ✓ Test OS at the IA for SymProg at significance level alpha(t1) as derived from the OBF-type alpha spending function, where t1 is the information fraction for OS.
    - ➢ If OS is significant at the IA for SymProg, then CytoChemo will be tested at the IA for SymProg.
      - If CytoChemo is significant at the IA for SymProg, then the study stops with superiority established on all endpoints.
      - If CytoChemo is not significant at the IA for SymProg, then the timing of the FA for CytoChemo is determined by the 90% conditional power on CytoChemo and the study continues to this FA. The test for CytoChemo at FA will combine the p-values for CytoChemo for stages 1 and 2 with weights w1=sqrt(t1) and w2=sqrt(1-w1^2), where t1 is the information fraction for CytoChemo.
    - ➢ If OS is not significant at the IA for SymProg, then the timing of the FA for OS is determined by the 90% conditional power on OS. The OBF-type alpha spending will apply to CytoChemo at the IA for SymProg, the weight w1 is determined and the corresponding p-value calculated to be used later by the combination test. However, significance of CytoChemo could not be claimed even if the efficacy boundary is crossed at this time. The study proceeds to the FA for OS. The required adjustment at FA for OS: use p-value combination test with the weights w1 and w2 =sqrt(1-w1^2).
      - If OS is not significant at the FA for OS, then the study stops with superiority established on all endpoints except OS and CytoChemo.
      - If OS is significant at the FA for OS, then CytoChemo will be tested at the time of the FA for OS. The required adjustment for CytoChemo at the FA for OS: use p-value combination test with the weights w1 and w2 =sqrt(1-w1^2).

- o If CytoChemo is significant at the FA for OS, then the study stops with superiority established on all endpoints.
- o If CytoChemo is not significant at the FA for OS, then the study stops with superiority established on all endpoints except CytoChemo.

- ❖ If SymProg is not significant at IA
  - ✓ The OBF-type alpha spending will apply to OS and CytoChemo at the IA for SymProg, the weight w1 is determined for each of them and the corresponding p-values calculated to be used later by the combination test. However, significance of either OS or CytoChemo could not be claimed even if the efficacy boundary is crossed for either of them at this time. The study proceeds to the FA for SymProg determined by the 90% conditional power on SymProg. The required adjustment for SymProg: use p-value combination test with the prefixed w1 and $w2=sqrt(1-w1^2)$.
    - ➢ If SymProg is not significant at FA, the study stops. No testing and next stages for OS/CytoChemo are possible.
    - ➢ If SymProg is significant at FA
      - • Test OS at the FA for SymProg at significance level alpha(t2), where t2=0.8 is the proportion of projected OS events at this time over the pre-planned maximum number of OS events.
        - o If OS is not significant at the FA for SymProg, then the timing of the FA for OS is determined by the 90% conditional power on OS. The OBF-type alpha spending will apply to CytoChemo at the FA for SymProg, the weight $w2=sqrt(t2-t1)$ is determined and the corresponding p-value calculated to be used later by the combination test, where t2=0.66 is the proportion of projected CytoChemo events at this time over the pre-planned maximum number of CytoChemo events. However, significance of CytoChemo could not be claimed even if the efficacy boundary is crossed at this time. The study proceeds to the FA for OS.
          - ▪ If OS is significant at FA, then CytoChemo will be tested at the time of FA for OS, and if this is also significant, the study stops with superiority established on all endpoints. Otherwise, the study stops with superiority established on all endpoints except CytoChemo. The required adjustment at FA for OS and CytoChemo: use

Aragon Pharmaceuticals - Confidential

27

p-value combination test with the weights w1=sqrt(t1) and w2=sqrt(t2-t1), and w3= sqrt(1-w1^2-w2^2).

o If OS is significant at the FA for SymProg, then CytoChemo will be tested at the FA for SymProg and if this is also significant, the study stops with superiority established on all endpoints. Otherwise, the Sponsor has 2 options: (1) The study stops with superiority established on all endpoints except CytoChemo. (2) The study proceeds to the FA for CytoChemo determined by the 90% conditional power on CytoChemo. In Option 2, the required adjustment at FA for CytoChemo: use p-value combination test with the weights w1, w2, and w3=sqrt (1-w1^2-w2^2).

JNJ-56021927 (apalutamide)

ARN-509-003 Documentation of Statistical Methods and Interim Analysis Plans

## 9.2    APPENDIX 2:  TABLE SHELL FOR OVERALL SUMMARY OF ADVERSE EVENTS BY FORMULATION

|  | Placebo | | | Apalutamide | | |
|---|---|---|---|---|---|---|
|  | Capsule | Tablet | Capsule+Tablet | Capsule | Tablet | Capsule+Tablet |
| Analysis Set:  Safety population | ### | ### | ### | ### | ### | ### |
| Number of subjects with AEs | ### (xx x%) | ### (xx x%) | ### (xx x%) | ### (xx x%) | ### (xx x%) | ### (xx x%) |
| Drug-related[a] | ### (xx x%) | ### (xx x%) | ### (xx x%) | ### (xx x%) | ### (xx x%) | ### (xx x%) |
| Number of subjects with grade 3-4 AEs | ### (xx x%) | ### (xx x%) | ### (xx x%) | ### (xx x%) | ### (xx x%) | ### (xx x%) |
| Drug-related[a] | ### (xx x%) | ### (xx x%) | ### (xx x%) | ### (xx x%) | ### (xx x%) | ### (xx x%) |
| Number of subjects with SAEs | ### (xx x%) | ### (xx x%) | ### (xx x%) | ### (xx x%) | ### (xx x%) | ### (xx x%) |
| Drug-related[a] | ### (xx x%) | ### (xx x%) | ### (xx x%) | ### (xx x%) | ### (xx x%) | ### (xx x%) |
| Grade 3-4 | ### (xx x%) | ### (xx x%) | ### (xx x%) | ### (xx x%) | ### (xx x%) | ### (xx x%) |
| Number of subjects with AEs leading to treatment discontinuation | ### (xx x%) | ### (xx x%) | ### (xx x%) | ### (xx x%) | ### (xx x%) | ### (xx x%) |
| Drug-related[a] | ### (xx x%) | ### (xx x%) | ### (xx x%) | ### (xx x%) | ### (xx x%) | ### (xx x%) |
| Number of subjects with AEs leading to death | ### (xx x%) | ### (xx x%) | ### (xx x%) | ### (xx x%) | ### (xx x%) | ### (xx x%) |
| Drug-related | ### (xx x%) | ### (xx x%) | ### (xx x%) | ### (xx x%) | ### (xx x%) | ### (xx x%) |
| All deaths within 28 days of last dose | ### (xx x%) | ### (xx x%) | ### (xx x%) | ### (xx x%) | ### (xx x%) | ### (xx x%) |
| Adverse event | ### (xx x%) | ### (xx x%) | ### (xx x%) | ### (xx x%) | ### (xx x%) | ### (xx x%) |
| Death due to prostate cancer | ### (xx x%) | ### (xx x%) | ### (xx x%) | ### (xx x%) | ### (xx x%) | ### (xx x%) |
| other | | | | | | |

[a] Adverse events reported as related.

Note: Percent is based on the Safety population

Note: Treatment-emergent adverse events are those that occurred between the date of 1st dose of study drug and date of last dose of study drug+28 days.  For each category, subjects are counted only once, even if they experienced multiple events in that category

Aragon Pharmaceuticals - Confidential

## 9.3     APPENDIX 3:  MODIFIED MEDDRA QUERIES AS SEARCH CRITERIA FOR AE OF SPECIAL INTEREST

The search criteria for adverse events of special interest are based on adverse event preferred terms from MedDRA version 19.1 dictionary.

Most categories are based on a MedDRA SMQ, but if one does not exist, a compilation of terms that reflect the event will be proposed for extraction and analysis of the data. Each of these events is defined below.

**Adverse Event of Special Interest Category= Seizure**

**Search Criteria Category= Selected PTs**

| | |
|---|---|
| Acquired epileptic aphasia | Idiopathic generalised epilepsy |
| Acute encephalitis with refractory, repetitive partial seizures | Lafora's myoclonic epilepsy |
| Alcoholic seizure | Lennox-Gastaut syndrome |
| Amygdalohippocampectomy | Migraine-triggered seizure |
| Atonic seizures | Molybdenum cofactor deficiency |
| Atypical benign partial epilepsy | Myoclonic epilepsy |
| Aura | Myoclonic epilepsy and ragged-red fibres |
| Automatism epileptic | Narcolepsy |
| Autonomic seizure | Partial seizures |
| Baltic myoclonic epilepsy | Partial seizures with secondary generalisation |
| Benign rolandic epilepsy | Petit mal epilepsy |
| Biotinidase deficiency | Polymicrogyria |
| Change in seizure presentation | Post stroke epilepsy |
| Clonic convulsion | Post stroke seizure |
| Complex partial seizures | Post-traumatic epilepsy |
| Convulsion prophylaxis | Postictal headache |
| Convulsions local | Postictal paralysis |
| Convulsive threshold lowered | Postictal psychosis |
| Corpus callosotomy | Postictal state |
| Deja vu | Preictal state |
| Double cortex syndrome | Psychomotor seizures |
| Dreamy state | Schizencephaly |
| Drop attacks | Seizure |

Aragon Pharmaceuticals - Confidential

30

| | |
|---|---|
| Drug withdrawal convulsions | Seizure anoxic |
| Epilepsy | Seizure cluster |
| Epileptic aura | Seizure like phenomena |
| Epileptic psychosis | Severe myoclonic epilepsy of infancy |
| Febrile convulsion | Simple partial seizures |
| Foaming at mouth | Status epilepticus |
| Frontal lobe epilepsy | Sudden unexplained death in epilepsy |
| Generalised non-convulsive epilepsy | Temporal lobe epilepsy |
| Generalised tonic-clonic seizure | Tongue biting |
| Glucose transporter type 1 deficiency syndrome | Tonic clonic movements |
| Hemimegalencephaly | Tonic convulsion |
| Hyperglycaemic seizure | Tonic posturing |
| Hypocalcaemic seizure | Topectomy |
| Hypoglycaemic seizure | Uncinate fits |
| Hyponatraemic seizure | |

**Adverse Event of Special Interest Category= Skin rash**

**Search Criteria Category= Selected PTs**

| | |
|---|---|
| Acquired epidermolysis bullosa | Noninfective conjunctivitis |
| Acute generalised exanthematous pustulosis | Oculomucocutaneous syndrome |
| Administration site hypersensitivity | Oral mucosal blistering |
| Administration site rash | Oral mucosal exfoliation |
| Administration site recall reaction | Oral papule |
| Administration site urticaria | Oropharyngeal blistering |
| Application site rash | Papule |
| Blau syndrome | Paraneoplastic rash |
| Blister | Pemphigoid |
| Blister rupture | Pemphigus |
| Bullous impetigo | Penile exfoliation |
| Butterfly rash | Perineal rash |
| Catheter site rash | Pogosta disease |
| Conjunctivitis | Rash |
| Corneal exfoliation | Rash erythematous |
| Cutaneous vasculitis | Rash follicular |
| Dermatitis bullous | Rash generalised |

| | |
|---|---|
| Dermatitis exfoliative | Rash macular |
| Dermatitis exfoliative generalised | Rash maculo-papular |
| Drug eruption | Rash maculovesicular |
| Drug reaction with eosinophilia and systemic symptoms | Rash morbilliform |
| Epidermal necrosis | Rash papular |
| Epidermolysis | Rash papulosquamous |
| Epidermolysis bullosa | Rash pruritic |
| Eruptive pseudoangiomatosis | Rash pustular |
| Erythema multiforme | Rash rubelliform |
| Exfoliative rash | Rash scarlatiniform |
| Eyelid rash | Rash vesicular |
| Familial cold autoinflammatory syndrome | Sea bather's eruption |
| Fixed drug eruption | Skin erosion |
| Genital rash | Skin exfoliation |
| Genital ulceration | Skin necrosis |
| HLA-B*1502 assay positive | Skin reaction |
| HLA-B*5801 assay positive | Skin swelling |
| Hyper IgD syndrome | Staphylococcal scalded skin syndrome |
| Hypopharyngeal synechiae | Stevens-Johnson syndrome |
| Implant site rash | Stoma site rash |
| Incision site rash | Stomatitis |
| Infusion site rash | Symmetrical drug-related intertriginous and flexural exanthema |
| Injection site rash | Systemic lupus erythematosus rash |
| Instillation site rash | Tongue exfoliation |
| Lip exfoliation | Toxic epidermal necrolysis |
| Lupus miliaris disseminatus faciei | Toxic skin eruption |
| Medical device site rash | Urticaria |
| Mouth ulceration | Urticaria papular |
| Mucocutaneous rash | Vaccination site rash |
| Mucocutaneous ulceration | Vaginal exfoliation |
| Mucosa vesicle | Vaginal ulceration |
| Mucosal erosion | Vascular access site rash |
| Mucosal exfoliation | Vessel puncture site rash |
| Mucosal necrosis | Viral rash |

| | |
|---|---|
| Mucosal ulceration | Vulval ulceration |
| Nikolsky's sign | Vulvovaginal rash |
| Nodular rash | Vulvovaginal ulceration |

**Adverse Event of Special Interest Category= Changes in thyroid function**

**Search Criteria Category= Selected PTs**

| | |
|---|---|
| Anti-thyroid antibody | Thyroglobulin present |
| Anti-thyroid antibody positive | Thyroid atrophy |
| Atrophic thyroiditis | Thyroid disorder |
| Autoimmune hypothyroidism | Thyroid dysfunction in pregnancy |
| Autoimmune thyroiditis | Thyroid electron radiation therapy |
| Biopsy thyroid gland abnormal | Thyroid function test abnormal |
| Blood thyroid stimulating hormone abnormal | Thyroid gland scan abnormal |
| Blood thyroid stimulating hormone decreased | Thyroid hemiagenesis |
| Blood thyroid stimulating hormone increased | Thyroid hormone replacement therapy |
| Butanol-extractable iodine decreased | Thyroid hormones increased |
| Butanol-extractable iodine increased | Thyroid operation |
| Congenital hypothyroidism | Thyroid pain |
| Congenital thyroid disorder | Thyroid releasing hormone challenge test abnormal |
| Euthyroid sick syndrome | Thyroid size decreased |
| Free thyroxine index abnormal | Thyroid stimulating hormone deficiency |
| Free thyroxine index decreased | Thyroid therapy |
| Free thyroxine index increased | Thyroidectomy |
| Gamma radiation therapy to thyroid | Thyroiditis |
| Generalised resistance to thyroid hormone | Thyroiditis acute |
| Goitre | Thyroiditis chronic |
| Hashimoto's encephalopathy | Thyroiditis fibrous chronic |
| Hypothyroidic goitre | Thyroiditis subacute |
| Hypothyroidism | Thyroxin binding globulin abnormal |
| Infectious thyroiditis | Thyroxin binding globulin decreased |
| Iodine uptake abnormal | Thyroxin binding globulin increased |
| Iodine uptake decreased | Thyroxine abnormal |
| Iodine uptake increased | Thyroxine decreased |
| Myxoedema | Thyroxine free abnormal |
| Myxoedema coma | Thyroxine free decreased |

| | |
|---|---|
| Photon radiation therapy to thyroid | Thyroxine free increased |
| Polyglandular autoimmune syndrome type II | Thyroxine increased |
| Polyglandular autoimmune syndrome type III | Thyroxine therapy |
| Post procedural hypothyroidism | Transient hypothyroxinaemia of prematurity |
| Primary hypothyroidism | Tri-iodothyronine abnormal |
| Protein bound iodine decreased | Tri-iodothyronine decreased |
| Protein bound iodine increased | Tri-iodothyronine free abnormal |
| Radioactive iodine therapy | Tri-iodothyronine free decreased |
| Radiotherapy to thyroid | Tri-iodothyronine free increased |
| Reverse tri-iodothyronine decreased | Tri-iodothyronine free normal |
| Reverse tri-iodothyronine increased | Tri-iodothyronine increased |
| Secondary hypothyroidism | Tri-iodothyronine uptake abnormal |
| Tertiary hypothyroidism | Tri-iodothyronine uptake decreased |
| Thyreostatic therapy | Tri-iodothyronine uptake increased |
| Thyroglobulin absent | Ultrasound thyroid abnormal |
| Thyroglobulin decreased | X-ray therapy to thyroid |
| Thyroglobulin increased | |

**Adverse Event of Special Interest Category= Fall**

**Search Criteria Category= Selected PTs**

| | |
|---|---|
| Fall | |

**Adverse Event of Special Interest Category= Fracture**

**Search Criteria Category= Selected PTs**

| | |
|---|---|
| Acetabulum fracture | Hip fracture |
| Ankle fracture | Humerus fracture |
| Atypical femur fracture | Ilium fracture |
| Atypical fracture | Impacted fracture |
| Avulsion fracture | Internal fixation of fracture |
| Cervical vertebral fracture | Jaw fracture |
| Chance fracture | Limb fracture |
| Clavicle fracture | Loss of anatomical alignment after fracture reduction |
| Closed fracture manipulation | Lower limb fracture |
| Comminuted fracture | Lumbar vertebral fracture |

| | |
|---|---|
| Complicated fracture | Multiple fractures |
| Compression fracture | Open fracture |
| Costal cartilage fracture | Open reduction of fracture |
| Craniofacial fracture | Open reduction of spinal fracture |
| Elevation skull fracture | Osteochondral fracture |
| Epiphyseal fracture | Osteoporotic fracture |
| External fixation of fracture | Patella fracture |
| Facial bones fracture | Pathological fracture |
| Femoral neck fracture | Pelvic fracture |
| Femur fracture | Periprosthetic fracture |
| Fibula fracture | Pubis fracture |
| Foot fracture | Radius fracture |
| Forearm fracture | Rib fracture |
| Fracture | Sacroiliac fracture |
| Fracture debridement | Scapula fracture |
| Fracture delayed union | Skull fracture |
| Fracture displacement | Skull fractured base |
| Fracture malunion | Spinal compression fracture |
| Fracture nonunion | Spinal fracture |
| Fracture of clavicle due to birth trauma | Spinal fusion fracture |
| Fracture of penis | Sternal fracture |
| Fracture pain | Stress fracture |
| Fracture reduction | Surgical fixation of rib fracture |
| Fracture treatment | Thoracic vertebral fracture |
| Fractured coccyx | Tibia fracture |
| Fractured ischium | Tooth fracture |
| Fractured maxilla elevation | Torus fracture |
| Fractured sacrum | Traumatic fracture |
| Fractured skull depressed | Ulna fracture |
| Fractured zygomatic arch elevation | Upper limb fracture |
| Greenstick fracture | Wrist fracture |
| Hand fracture | |

**Statistical Analysis Plan**

**Ventricular Repolarization Sub-Study**

# SPARTAN
(**S**elective **P**rostate **AR T**argeting with **ARN-509**)

**A Multicenter, Randomized, Double-Blind, Placebo-Controlled, Phase III Study of ARN-509 in Men with Non-Metastatic (M0) Castration-Resistant Prostate Cancer**

**ARN-509-003; Phase 3**

Sponsor:  Aragon Pharmaceuticals, Inc*
*Aragon Pharmaceuticals, Inc. is a wholly-owned subsidiary of Johnson & Johnson. Janssen Research & Development, LLC is part of the Janssen Pharmaceutical Companies of Johnson & Johnson and provides various services to its affiliated company, Aragon Pharmaceuticals, Inc.

**Status:**  Approved
**Date:**  26 June 2017
**Prepared by:**  Janssen Research & Development, LLC
**Document No.:**  EDMS-ERI-145510381

1

# Table of Contents

ARN-509                                        Statistical Analysis Plan ARN-509-003 QT Study

## AMENDMENT HISTORY

N/A

## ABBREVIATIONS

| | |
|---|---|
| ΔQTc | Change from baseline in QTc at each time point |
| ΔΔQTc | Difference in ΔQTc between each treatment and placebo, at each time point |
| AE | Adverse Event |
| ANCOVA | Analysis of Covariance |
| BMI | Body Mass Index |
| CI | Confidence Interval |
| CRF | Case Report Form |
| CSR | Clinical Study Report |
| ECG | Electrocardiogram |
| eCRF | Electronic Case Report Form |
| FDA | Food and Drug Administration |
| HR | Heart rate |
| ICH | International Conference on Harmonization |
| MedDRA | Medical Dictionary for Regulatory Activities |
| msec | Milliseconds |
| NCI-CTCAE | National Cancer Institute - Common Terminology Criteria for Adverse Events |
| PD | Pharmacodynamic |
| PI | Principal Investigator |
| PK | Pharmacokinetic(S) |
| PR | PR interval of ECG |
| PT | Preferred Term |
| QRS | QRS interval of ECG |
| QT | QT interval of ECG |
| QTc | QT interval corrected for heart rate |
| QTcF | QT interval corrected using Fridericia's formula |
| QTcB | QT interval corrected using Bazett's formula |
| QTcP | QT interval corrected using study-specific power correction method |
| RR | RR interval of ECG |
| SAE | Serious Adverse Event |
| SAP | Statistical Analysis Plan |
| SD | Standard Deviation |
| SMQ | Standardized MedDRA Query |
| SOC | System Organ Class |
| TEAE | Treatment-emergent Adverse Event |
| WHO | World Health Organization |

## 1.     INTRODUCTION

An assessment of the potential of a new molecular entity to effect cardiac repolarization is required by regulatory agencies. In accordance with ICH E14 guideline, a thorough QT (TQT) study should be conducted, if possible.  In the case of ARN-509 (JNJ-56021927; hereafter referred as apalutamide), in view of previous FDA advice and the observations to date that nonclinical (hERG and CV safety study) and clinical data (ECG collections in the Phase I/II Study ARN-509-001) suggest no apparent relationship between apalutamide and QT prolongation, an alternative design to the TQT study has been chosen.

The sub-study of ARN-509-003 has been designed to evaluate the potential of apalutamide to prolong the QTc interval in a subset of patients with high risk non-metastatic castration-resistant prostate cancer participating in the main protocol Study ARN-509-003.

This statistical analysis plan (SAP) contains definitions of analysis populations, derived variables and statistical methods for the analysis of pharmacodynamics (PD), pharmacokinetic/pharmacodynamics (PK/PD), and safety for the sub-study (QT study) of ARN-509-003. Any deviations from this analysis plan will be described in the clinical study report.

### 1.1.    Trial Objectives

The primary objective of this sub-study is to evaluate the effect of apalutamide on cardiac repolarization, as detected by changes in electrocardiogram (ECG) QT intervals corrected for heart rate by Fridericia's correction method (QTcF).

The secondary objectives of this sub-study are as follows:

- To investigate the effect of apalutamide on the following ECG parameters: PR, RR, QRS, QT, QTcB (Bazett's correction method), and T-Wave morphology;

- To determine the relationship between the plasma concentrations of apalutamide and its metabolite ARN000308 (M3; hereafter referred as JNJ-56142060) and QT/QTc changes.

- To assess the safety and tolerability of therapeutic exposures of apalutamide.

### 1.2.    Trial Design

The main study of ARN-509-003 is a randomized (2:1), multicenter, double-blind, placebo-controlled, Phase III clinical trial evaluating the efficacy and safety of apalutamide (treatment arm A) versus placebo (treatment arm B) in men with high risk (M0) NM-CRPC, defined as PSA Doubling Time (PSADT) $\leq$ 10 months.

In this sub-study of ARN-509-003, the effect of apalutamide on ventricular repolarization will be centrally analyzed by a third-party cardiac safety laboratory in a subset of 100 patients enrolled at selected sites that will be participating in the main protocol.  Both apalutamide and placebo patients will be enrolled in a blinded manner as per the main protocol randomization criteria.

The subset of 100 patients will undergo the same screening procedures as the main protocol in order to be randomized into the study, following the same Inclusion/Exclusion criteria as per Section 4 of the protocol, with the following additional enrollment criteria:

| Additional Inclusion Criteria | Additional Exclusion Criteria |
|---|---|
| ▪ Enrollment in the main study<br>▪ Obtain separate informed consent for participation in the sub-study | ▪ Heart rate outside of 50 to 100 beats/minute<br>▪ QTcF > 480 msec, determined by central assessment<br>▪ Diagnosed or suspected congenital long QT syndrome, or family history of congenital long QT syndrome or sudden death<br>▪ History of Mobitz II second degree or third degree heart block<br>▪ Implantable pacemaker or automatic implantable cardioverter defibrillator<br>▪ Complete Bundle Branch Block or ventricular conduction delay (QRS > 119 msec)<br>▪ Chronic or persistent atrial arrhythmia, including atrial fibrillation and atrial flutter.<br>▪ Concurrent therapy with medications known to prolong the QT interval and/or associated with TdP (Torsade de Pointes) arrhythmia<br>▪ Smokers and planned nicotine replacement therapy users |

If patients do not qualify for the ventricular repolarization sub-study, they can still participate in the main study provided they meet all other inclusion/exclusion criteria as per Section 4 of the protocol.

A patient may withdraw his consent to participate in the sub-study at any time. If a patient withdraws such consent, the investigator should inform the sponsor (or designee) in writing and document in the Investigator Site File. The patient may continue participating in the main protocol.

## 1.3.    Statistical Hypotheses for Trial Objectives

**Primary Analysis**

For the primary analysis, the quantity of interest is the mean change from baseline ($\Delta$QTcF). The null hypothesis for comparison of the active treatment group apalutamide (T) change from baseline with a threshold value of $\Delta$ is as follows:

*Null hypothesis of an effect of apalutamide (or metabolite) $H_0$: $\mu_{Ti} \geq \Delta$ for at least one i where $\mu$ denotes mean change from baseline in QTcF, i denotes each post-dosing time point and $\Delta = 20$ msec*

5

ARN-509                  Statistical Analysis Plan ARN-509-003 QT Study

*The null hypothesis will be tested against the alternative hypothesis of no effect of apalutamide (or metabolite) $H_1$: $\mu_{Ti} < \Delta$ for all i in the sampling interval*

The upper limit of the one-sided 95% CI for the QTcF change from baseline will be compared to the 20 msec bound (at each time point). Absence of effects of apalutamide (or metabolite) on cardiac repolarization will be concluded if the upper limit of the one-sided 95% confidence intervals (equivalent to the two-sided 90% CI) is lower than 20 msec at all the time points in active treatment group. Because the null hypothesis has to be rejected at every time point, no correction for multiplicity is required.

### Secondary Analysis

For the secondary analysis, the quantity of interest is the difference between the mean changes from baseline in active treatment group apalutamide (T) and placebo (P) groups ($\Delta\Delta$QTcF). The magnitude of the true QTcF prolongation in the secondary analysis (i.e., the expected value of $\Delta\Delta$QTcF) is taken as 5 msec. A two-sided 90% CI for the baseline-adjusted QTcF differences will be presented.

## 1.4. Sample Size Determination

### 1.4.1. Originally Planned Sample Size

To evaluate the effect of apalutamide on ventricular repolarization in a subset of patients, the sub-study was planned to enroll at least 100 patients to ensure that at least 60 patients treated with apalutamide will provide at least 98.7% power to detect a true effect of 10 msec change from baseline considering only the active group. Details for the power calculation and statistical assumptions are provided in protocol.

### 1.4.2. Achieved Sample Size and Change in Planned Analysis

As a consequence of slow patient's enrollment to the QT substudy, approximately 17 subjects were randomized by the clinical cutoff date, which is only 17% of the originally planned sample size. To be consistent with this change, no hypothesis testing will be performed and all planned statistical analysis in the QT substudy will be performed for an estimation purpose.

## 1.5. Randomization and Blinding

The subset of 100 patients at selected sites will undergo the same screening procedures as the main protocol in order to be randomized into the study, following the same Inclusion/Exclusion criteria. Both apalutamide and placebo patients will be enrolled in a blinded manner as per the main protocol randomization criteria.

## 2. GENERAL ANALYSIS DEFINITIONS

## 2.1. Handling of Missing Data and Imputation of Missing Dates

Missing test results or assessments will not be imputed.

For missing dates, in general, imputation of missing dates will be made for AE onset date, AE resolution date, start and end dates of prior and concomitant medications.

- If dates are completely missing, no imputation will be made. For any partial date with missing year, no imputation will be made.

- If only day is missing, then the 15th of the month will be used.

- If only year is present, then June 30th will be used.

- If such imputed date for prior or concomitant medication start date is in the same year and month but on or after date of first dose, then date of first dose - 1 will be used. If the imputed medication start date is after the medication end date, then the medication end date will be used. If the imputed date is for a medication start date and is in the same year and month but after the date of post-study follow-up contact (10±2 days after the last dose of study drug), then the date of post-study follow-up contact will be used.

- If the imputed date for a medication end date is in the same year and month but is before the first dose date, then the first dose date will be used. If the imputed date for a medication end date is after the date of post-study follow-up contact (10±2 days after the last dose of study drug), then the follow-up contact date will be used, or if the imputed medication end date is before the medication start date, then the medication start date will be used.

- If the imputed date for an AE start date is in the same year and month but is before the first dose date, then the first dose date will be used. If the imputed AE start date is after the AE end date, then the AE end date will be used. If the imputed date is for an AE start date and is in the same year and month but after the date of post-study follow-up contact (10±2 days after the last dose of study drug), then the date of post-study follow-up contact will be used.

- If the imputed date for an AE end date is in the same year and month but is before the first dose date, then the first dose date will be used. If the imputed date for an AE end date is after the date of post-study follow-up contact (10±2 days after the last dose of study drug), then the follow-up contact date will be used, or if the imputed AE end date is before the AE start date, then the AE start date will be used.

## 2.2.    Correction Methods for QTc Calculations

The ECG parameters, heart rate (HR), PR, RR, QT, QRS, PR intervals will be measured in triplicates from the 12-lead ECG at the time points specified in protocol. The measured QT data will be corrected for heart rate using Fridericia (QTcF) Bazett (QTcB) and study-specific power (QTcP)correction methods, as per the following formulae/method (with QT, RR and QTc's expressed in msec):

1. Fridericia's Correction

$$QTcF = \frac{QT}{(RR/1000)^{(1/3)}}$$

7

ARN-509           Statistical Analysis Plan ARN-509-003 QT Study

2. Bazett's Correction

$$QTcB = \frac{QT}{(RR/1000)^{0.5}}$$

3. Study-specific Power Correction

A linear regression model will be fitted with logarithm of baseline (predose) QT interval as the dependent variable and logarithm of corresponding RR interval as the predictor variable. The slope of the regression line will be estimated from this model. Using the estimated slope (b) of the regression line, the study-specific power corrected QTcP will be calculated using the formula:

$$QTcP = \frac{QT}{(RR/1000)^{b}}$$

In this study, the primary correction method was the Fridericia's corrected QT interval (QTcF).

## 2.3. Schedule of ECG and PK Data Collections

Digital 12-lead ECG equipment will be provided to each clinical site participating in this sub-study by the central laboratory for the duration of the sub-study.

ECGs should be collected after the patient has rested quietly and is awake in a fully supine (or semi-recumbent, if supine is not tolerated) position for 10 minutes, and prior to any blood collection.  For each patient the same position (eg, supine or semi-recumbent) should be used for all ECGs collected. Starting on Cycle 1 Day 1, time point matched blood samples for PK analyses will be collected immediately following the collection of ECGs and before the collection of blood for all other clinical evaluations.  ECGs will be read by independent cardiologists from the central laboratory in a blinded manner and via single reader paradigm. QT/QTc intervals of each subject will be assessed by using a pre-specified lead (primary Lead V3 or backup Lead II).

The schedule of activities specific to this ventricular repolarization sub-study is as follows. All other assessments at the other time points will follow the Schedule of Activities in the main protocol.

- Cycle 1 Day 1
  - Hour -1 pre-dose:
    - o  Collect a set of triplicate 12-lead ECGs, 2 minutes apart;
  - Hour 0 pre-dose:
    - o  Collect a second set of triplicate 12-lead ECGs, 2 minutes apart;
    - o  Collect one blood sample for PK analysis;

8

- Hour 2 and Hour 4 post-dose:

  o Collect a set of triplicate ECGs, 2 minutes apart;

  o Collect one blood sample for PK analysis;

- Cycle 3 Day 1

  - Hour 0 pre-dose:

    o Collect a set of triplicate 12-lead ECGs, 2 minutes apart;

    o Collect one blood sample for trough PK analysis;

  - Hour 2 and Hour 4 post-dose:

    o Collect a set of triplicate ECGs, 2 minutes apart;

    o Collect one blood sample for PK analysis;

| | Cycle 1 Day 1 | | | | Cycle 3 Day 1 | | |
|---|---|---|---|---|---|---|---|
| Activities and Forms to be Completed Specific to the Sub-Study Only: | Hour -1 pre-dose | Hour 0 pre-dose | 2 Hours post-dose | 4 Hours post-dose | Hour 0 pre-dose | 2 Hours post-dose | 4 Hours post-dose |
| Triplicate 12-lead ECG | X | X | X | X | X | X | X |
| PK samples* | | X | X | X | X | X | X |

\* PK samples are to be collected immediately after each ECG; the main study PK samples on Day 1 of Cycles 1 and 3 between 0.5 and 4 hours post dose do NOT have to be collected for the 100 patients on this sub-study. The 100 patients participating in this sub-study will follow the main study PK sampling schedule for the samples taken on Day 1 of Cycles 2, 6, 12, 18, and 24

## 2.4. Only Use the Baseline and Post-treatment QT/QTc Intervals Assessed by the Same Lead in the Calculation of Change from Baseline QT/QTc Intervals

Per agreement in "Special Protocol Assessment" (SPA) with FDA issued on 11/09/2012, for each subject, only his baseline and post-baseline QT/QTc intervals that were evaluated by the same lead will be acceptable for analysis and submission.

A primary lead (Lead V3) will be used to assess all triplicate baseline QT/QTc intervals, and all triplicate QT/QTc intervals at each post-baseline time-points. For any subject, if his/her QT/QTc intervals based on the primary lead will be unmeasurable or unsuitable at some time-points, a backup lead (Lead II) will be used to assess relevant QT/QTc intervals.

When a subject completed study and had baseline and post-treatment QT/QTc intervals not totally based on the primary lead, a "Lead Harmonization" process will be performed for the subject, in which his triplicate ECG measurements at baseline and at each of post-baseline time-points will be re-analyze by cardiologist using the backup lead (if possible).

In the calculation of change from baseline QT/QTc intervals for each patient, baseline and each post-treatment time-point QT/QTc intervals will be the average of triplicate intervals (2 minutes apart measured at the time-point) that were all assessed by the same lead.

9

## 2.5.    Analysis Sets

PD Analysis Set

Pharmacodynamics (PD) analysis set will include all subjects from the sub-study who were randomized and received at least one dose of the study drug and had both predose and at least one postdose ECG measurement.

PD analysis set will be used for all ECG-related analysis (include Subject Information).

PK Analysis Set

The Pharmacokinetic (PK) analysis set will include all subjects from the sub-study who were randomized and received at least one dose of the study active drug and have estimable PK parameter Cmax. This analysis set combined with Pharmacodynamics (PD) analysis set will be used for PK/PD relation analysis.

Subjects who received placebo will not be part of the PK population.

## 3.    SUBJECT INFORMATION AND SAFETY

All analysis planned in this section will be the same as the planned analysis in the main study except for the analysis set which will be limited to the subjects for this sub-study (defined as in Section 2.5). A list of tables and listings is presented at the DPS.

## 4.    PHARMACOKINETICS (PK)

The apalutamide PK concentration data will be presented in listing and descriptive statistics (mean, SD, median, range). No PK inferential statistical analysis will be performed for this sub-study.

## 5.    PHARMACODYNAMICS (PD)

In this study, the primary variable of interest is the change from baseline in QTcF intervals ($\Delta$QTcF).

For each subject, the baseline ECG parameters will be the average of the 2 sets of predose measurement taken on Cycle 1 Day 1 (Hour -1 and Hour 0).

## 5.1.    QT/QTc Intervals

The average of the triplicate ECG measurements taken at each time point will be used for the analysis. QTc-related analyses will be performed on QTc correction methods QTcF, QTcB and QTcP .

### 5.1.1.    Analysis of QT/QTc

- Mean QT and corrected QTc over time will be listed for each subject by treatment group and will be summarized using descriptive statistics (mean, SD, median, range) at each time point and for each treatment group.

10

- For each treatment group, the incidence count and percent of subjects with any postdose QTc values >450 to ≤480, >480 to ≤500, and >500 msec, will be summarized. They will also be summarized by time point for each treatment group. Subjects with QT/QTc values >500 msec will be listed.

- Mean QT and corrected QTc versus time profiles for both treatment groups will be plotted on the same graph.

### 5.1.2.        Analysis change from baseline of QTc (ΔQTc) in apalutamide

- The change from baseline in QTc (ΔQTc) intervals at each time point will be listed for each subject and treatment. For each treatment and time point of measurement, the change from baseline in QTc intervals (ΔQTc) will be summarized using descriptive statistics (mean, SD, median, range).

- A repeated-measures mixed model will be fitted with change from baseline in QTc (ΔQTc) as the dependent variable, and treatment (Apalutamide vs. Placebo), time point of measurement, and treatment by time point of measurement interaction as fixed effects, baseline value of QTc as covariate, and subject as a random effect.  Using the estimated least square means and intra-subject variance from the model, 2-sided 90% CI for mean ΔQTc of apalutamide active treatment group at each time point will be constructed.  The incidence count and percentage of subjects with QTc increase >30 to ≤60 and >60 msec will be summarized for each treatment group. The incidence count and percentage by time point will also be summarized for each treatment group. Subjects with QTc increase >60 msec will be listed.

- Mean change from baseline in QTc intervals (ΔQTc) versus time profiles for both treatments will be plotted on the same graph.

### 5.1.3.        Analysis of the difference between apalutamide and placebo in ΔQTc (ΔΔQTc)

- The treatment difference in mean (SD) of ΔQTc between apalutamide and placebo will be calculated at each matched time point.

- A repeated-measures mixed model will be fitted with ΔQTc as the dependent variable, and treatment ( apalutamide vs. Placebo), time point of measurement, and treatment by time point of measurement interaction as fixed effects, baseline value of QTc as covariate, and subject as a random effect. Using the estimated least square means and intra-subject variance from the model, 2-sided 90% CI for the difference in mean ΔQTc between apalutamide and placebo at each time point will be constructed.

### 5.1.4.        Analysis of the relationship between QT/QTc and RR at baseline

- The relationship between QT/QTc and RR at baseline will be evaluated graphically by plotting the logarithm of baseline QT/QTc values against the logarithm of corresponding RR intervals.

- A linear regression model will be fitted to the data with logarithm of baseline QT/QTc interval values as the dependent variable and logarithm of RR as the predictor. The slope of the regression line and its standard error, along with 2-sided 95% confidence intervals for the slope, will be estimated from this model.

## 5.2.    Heart Rate (HR)

Mean HR over time will be listed for each subject and treatment. For each treatment and time point of measurement, HR will be summarized using descriptive statistics (mean, SD, median, range).

Mean change from baseline in HR will also be summarized using descriptive statistics (mean, SD, median, range) for each treatment group and time point.

A repeated-measures mixed model will be fitted with $\Delta$HR as the dependent variable, and treatment (Apalutamide vs. Placebo), randomization stratification criteria, time point of measurement, and treatment by time point of measurement interaction as fixed effects, baseline value of HR as covariate, and subject as a random effect.  Using the estimated least square means and intra-subject variance from the model, 2-sided 90% CI for the difference in mean $\Delta$HR between apalutamide and placebo at each time point will be constructed.

The number and percentage of subjects with any postdose heart rate >100 bpm and <50 bpm will be summarized by treatment group and by time point.

To investigate any effect of treatment on heart rate, a mean plot for change from baseline in heart rate ($\Delta$HR) over time by treatment group (on the same page) will be presented.

## 5.3.    QRS and PR intervals

Mean QRS and PR intervals will be listed for each subject and treatment. For each treatment and time point of measurement, QRS and PR intervals will be summarized using descriptive statistics (mean, SD, median, range ).

Mean change from baseline in QRS and PR intervals will also be summarized using descriptive statistics (mean, SD, median, range ) for each treatment group and time point.

A repeated-measures mixed model will be fitted with $\Delta$QRS (or $\Delta$PR) as the dependent variable, and treatment (Apalutamide vs. Placebo), time point of measurement, and treatment by time point of measurement interaction as fixed effects, baseline value of QRS (or PR) as covariate, and subject as a random effect.  Using the estimated least square means and intra-subject variance from the model, 2-sided 90% CI for the difference in mean $\Delta$QRS (or $\Delta$PR) between apalutamide and placebo at each time point will be constructed.

The number and percentage of subjects with QRS interval *>120* msec will be summarized for each treatment group. The incidence count and percentage by time point will also be summarized.

The number and percentage of subjects with PR interval *>200* msec will be tabulated for each treatment group.

NCT01946204

JNJ-56021927 (apalutamide)
ARN-509-003 Documentation of Statistical Methods and Interim Analysis Plans

ARN-509 | Statistical Analysis Plan ARN-509-003 QT Study

## 5.4.    Changes in T Wave Morphology

The incidence of subjects with changes in T-wave morphology from baseline will be tabulated by coded descriptive T-wave morphology and by treatment group.   The codes of descriptive T-wave morphology (based on ECG Chapter of this sub-study) are listed in the following table:

| Code for Descriptive T Wave Morphology | Description |
|---|---|
| Normal T waves | always upright in leads I, II, V5 and V6 and always negative in aVR. |
| Peaked T-waves | one or more T wave appears abnormally tall and sharp |
| Flat T-waves | one or more T wave appears flatter than normal |
| Notched T-waves | one or more T wave is notched (possibly due to T/U fusion) but remains of one polarity |
| Bi-phasic T-waves | one or more T wave has a both a negative and a positive component |
| Inverted T-waves | one or more T wave that should normally be positive is inverted |

In the calculation of the incidence, all abnormal changes in T wave morphology based on any lead (e.g., primary lead, backup lead, and/or other leads) and at any post-baseline time-points should be counted.

## 6.    PHARMACOKINETIC/PHARMACODYNAMIC RELATIONSHIPS

The evaluation of PK/PD relationship will be performed using data from all subjects who received the active drug and have ECG data from at least one time point following the first dose.

- The following plots will be generated to evaluate the relationship between mean and mean change in QTc intervals and plasma concentrations of apalutamide:
  - QTc at each time point of measurement will be plotted against the corresponding plasma concentration of apalutamide.
  - ΔQTc at each time point of measurement will be plotted against the corresponding plasma concentration of apalutamide.

Additional plots will be produced, if deemed necessary.

- For active treatment group apalutamide, the change from baseline in QTc intervals at the individual subject's apalutamide plasma concentrations on Cycle 3 Day 1 (i.e. predose, 2 hr postdose, 4 hr postdose) will be summarized using descriptive statistics mean, SD median, and range.

- Linear mixed effects model will be fitted to the ΔQTc data with apalutamide concentration as a predictor and subject as a random effect (random intercept and random slope model); if the intercept term is not significant, the model will be re-fitted with a zero intercept term. The analyses above will be repeated for principal metabolite (JNJ-56142060).

- Outliers of individual drug concentrations will be identified by pharmacokineticist. If deemed necessary, the above analyses will be repeated with no outliers.

13

48

# REFERENCES

1.  E14 Clinical Evaluation of QT/QTc Interval Prolongation and Proarrhythmic Potential for Non-Antiarrhythmic Drugs. Guidance for Industry. U.S. Department of Health and Human Services, Food and Drug Administration, Center for Drug Evaluation and Research, Center for Biologics Evaluation and Research, ICH. Oct 2005.

2.  Fridericia, L. S. (1920). Die Systolendauer im Elektrokardiogramm bei normalen Menschen und bei Herzkranken. Acta Medica Scandinavica, Stockholm. 57: 469-486

3.  Bazett HC. An analysis of the time-relationship of electrocardiograms. Heart 1919;70:353-370.

# Patient Reported Outcomes (PRO)
# Statistical Analysis Plan (SAP)

## SPARTAN

### (Selective Prostate AR Targeting with ARN-509)

### A Multicenter, Randomized, Double-Blind, Placebo-Controlled, Phase III Study of ARN-509 in Men with Non-Metastatic (M0) Castration-Resistant Prostate Cancer

### ARN-509-003; Phase 3

Sponsor: Aragon Pharmaceuticals, Inc*

*Aragon Pharmaceuticals, Inc. is a wholly-owned subsidiary of Johnson & Johnson. Janssen Research & Development, LLC is part of the Janssen Pharmaceutical Companies of Johnson & Johnson and provides various services to its affiliated company, Aragon Pharmaceuticals, Inc.

**Status:** Approved
**Date:** 26 June 2017
**Prepared by:** Janssen Research & Development, LLC
**Document No.:** EDMS-ERI-144883631; 3.0

**TABLE OF CONTENTS**

## 1. INTRODUCTION

This document describes the planned statistical analyses of the PRO measures collected in ARN-509-003 the **SPARTAN** Study **(**A Multicenter, Randomized, Double-Blind, Placebo-Controlled, Phase III Study of ARN-509 in Men with Non-Metastatic (M0) Castration-Resistant Prostate Cancer). This PRO analysis plan is meant to supplement the study protocol and the study statistical analysis plan (SAP). Any deviations from this analysis plan will be described in the clinical study report.

Study design can be found in the ARN-509-003 protocol and study SAP.

## 2. METHODS

### 2.1 PRO TIME AND EVENTS SCHEDULE

Patients will be required to complete two self-administered quality of life instruments: the Functional Assessment of Cancer Therapy-Prostate (FACT-P) and the Euro-QoL Group EQ-5D[1,Error! Reference source not found.], according to the schedule below:

Treatment Phase:
- Cycle 1 Day 1 (before dose)
- D1 C1-C6, D1 of every 2 cycles starting at C7 to C13, then D1 of every 4 cycles unless otherwise specified

Posttreatment Phase:
- End of treatment visit
- Long-term Follow-up: every 4 months up to 12 months post progression

Patients will complete the FACT-P and EQ-5D at the clinic during the Treatment Phase PRIOR to any other clinical activity. During Long-term Follow-up contact every 4 months via clinic visit or an alternative contact method per institution policy/practice up to 12 months post progression.

### 2.2 FACT-P

The FACT-P will be used to assess health-related quality of life and prostate cancer-specific symptoms. The FACT-P consists of the 27-item FACT-General (FACT-G) and 12 items for the prostate cancer specific concerns. Scoring guidelines for the FACT-P and handling of missing items (as opposed to missing assessments) will be in accordance with methodology described in the Manual of the Functional Assessment of Chronic Illness Therapy (FACIT) Measurement System, version 4.0 (November 1997). See Appendix 1. The scales from the FACT-P are composed of the following items, all with higher scores indicating better functioning:

- Physical Well-Being (PWB): items GP1-GP7 (scale values range from 0 to 28),
- Social/Family Well-Being (SFWB): items GS1-GS7 (scale values range from 0 to 28),

JNJ-56021927 (apalutamide)
ARN-509-003 Documentation of Statistical Methods and Interim Analysis Plans

- Emotional Well-Being (EWB): items GE1-GE6 (scale values range from 0 to 24),

- Functional Well-Being (FWB): items GF1-GF7 (scale values range from 0 to 28),

- FACT-G: items from the PWB, SFWB, EWB, and FWB (scale values range from 0 to 108),

- FACT-P Prostate Cancer Subscale (PCS): the 12 items from the additional concerns section (scale values range from 0 to 48),

- FACT-P Trial Outcome Index (TOI): PWB, FWB, and PCS scales (scale values range from 0 to 104),

- FACT-P Total: scales FACT-G and PCS (scale values range from 0 to 156),

- FACT-P Pain-related Scale: Items P1-P3, and GP4 (scale values range from 0 to 16),

- FAPSI-8: pain (3 items), fatigue, weight loss, urinary difficulties (2 items), and concern about the condition becoming worse (scale values range from 0 to 32).

## 2.3    EQ-5D-3L

The EQ-5D-3L is a validated and reliable self-administered instrument used to assess health status. The EQ-5D-3L can be completed in less than 5 minutes.

The EQ-5D-3L[Error! Reference source not found.] quality of life instrument is a validated tool that consists of a 5-item questionnaire and a visual analogue scale (VAS). The 5-item questionnaire measures mobility, self-care, usual activities, pain/discomfort, and anxiety/depression.

## 2.4    PRO SCALES

The following PRO scales will be included in the analyses.

- FACT-P (See Appendix 1 for the derivation of each scale)
  - FACT-P Total
  - PCS
  - PWB
  - EWB
  - FWB
  - FACT-G
  - TOI
  - FACT-P Pain-related Scale
  - SFWB
  - FAPSI-8

- EQ-5D-3L(See Appendix 2)
  - Health State
  - EQ VAS score

## 2.5    ANALYSIS METHODS

All analysis will be performed by treatment arm. When appropriate, separate analysis may be performed during treatment phase and follow-up phase.

Based on the PRO assessment date, the associated visit will be mapped and used for analysis. For subjects with multiple records at the same visit, the closest one to the visit date will be selected as scheduled assessment, and others will be unscheduled assessments. For those with multiple records on the same assessment date, the first one by time point and sequence number will be selected as scheduled assessment. Only scheduled visits will be presented in the by-visit analysis. All assessments will be included in the time-to-event analysis.

Analysis may be stratified by stratification factors when appropriate. In this study, patients were randomized to the treatment group based on the following stratification factors:

- PSADT: > 6 months vs. ≤ 6 months
- Bone-sparing agent use: Yes vs. No
- Loco-regional disease: N0 vs. N1

In this study, subjects may receive subsequent anti-cancer therapies after discontinuing study drug. For these subjects, only PRO measurements before the initiation of subsequent therapy will be included in analysis unless otherwise stated.

### 2.5.1    Analysis Population

The analysis of patient reported outcomes data described in this SAP will be based on the Intent-to-Treat population as defined below.

**Intent-to-Treat Population [ITT]**: All eligible patients who are randomized into the study, with study drug assignments designated according to initial randomization, regardless of whether patients receive study drug or receive a different drug from that to which they were randomized to will be included in the analyses of all efficacy and clinical benefit endpoints and patient characteristics.

### 2.5.2    Compliance with PRO Assessment

Missing PRO assessments are calculated as the expected number of assessments for a particular visit minus the actual number of assessments received for that visit. Compliance (% received and % missing forms) will be tabulated by treatment group and overall for baseline, each scheduled visit during treatment phase and follow-up phase. Separate tables will be constructed for FACT-P and EQ-5D-3L Forms. Expected number of assessments per visit will be determined by subject-level study completion status. These tables will resemble the example in Table 1 below.

ARN-509-003
PRO Statistical Analysis Plan v2.0 26 June 2017

**Table 1: Sample Table of Compliance with PRO Assessment**

| Timing of Assessment[1] | Placebo (N = xxx) Number of Forms, n (%) | | | Apalutamide (N = xxx) Number of Forms, n (%) | | | Total (N = xxx) Number of Forms, n (%) | | |
|---|---|---|---|---|---|---|---|---|---|
| | Expected | Received | Missing | Expected | Received | Missing | Expected | Received | Missing |
| Baseline | Xxx | xx (xx x) | xx (xx.x) | Xxx | xx (xx x) | xx (xx x) | xxx | xx (xx x) | xx (xx.x) |
| Cycle 2 | Xxx | xx (xx x) | xx (xx.x) | Xxx | xx (xx x) | xx (xx x) | xxx | xx (xx x) | xx (xx.x) |
| Cycle 3 | Xxx | xx (xx x) | xx (xx.x) | Xxx | xx (xx x) | xx (xx x) | xxx | xx (xx x) | xx (xx.x) |
| Cycle 4 | Xxx | xx (xx x) | xx (xx.x) | Xxx | xx (xx x) | xx (xx x) | xxx | xx (xx x) | xx (xx.x) |
| … | Xxx | xx (xx x) | xx (xx.x) | Xxx | xx (xx x) | xx (xx x) | xxx | xx (xx x) | xx (xx.x) |
| Total | Xxx | xx (xx x) | xx (xx.x) | Xxx | xx (xx x) | xx (xx x) | xxx | xx (xx x) | xx (xx.x) |
| Note: Percentages calculated with the number of expected forms in each group as denominator | | | | | | | | | |

### 2.5.3 Descriptive Statistics Over Time

Descriptive statistics (number of observations, mean, standard deviation, minimum, maximum) of scores at baseline and each scheduled visit will be produced by treatment groups for each PRO scale. For these descriptive statistics, and for all analyses specified below, the scores obtained at the end of study treatment visit will be assigned to a cycle number as the next cycle that would have been completed if the subject were to have continued receiving study treatment. Separate summaries will be performed at treatment phase and follow-up phase.

For EQ-5D-3L Scores, frequency count and percentage of each reporting level (1 to 3) over Time will be provided.

### 2.5.4 Time to Degradation for FACT-P

Degradation/deterioration is defined based on a clinically meaningful threshold for each specific PRO scale. Table 2 summarizes the threshold values for determining a FACT-P degradation/deterioration. Time to degradation/deterioration will be defined as the time interval from randomization to the first date a patient experiences a clinically meaningful threshold change. Subjects who did not experience degradation/deterioration will be censored at their last available assessment. Assessments in both treatment and follow-up phases will be included to determine the time of event or censoring. For subjects receiving subsequent anti-cancer therapy, only the assessments on or before the subsequent therapy will be included in analysis. Time to degradation/deterioration will be compared between 2 treatment groups.

Change thresholds for deterioration on the FACT-P PCS, TOI, and FACT-P Total scales are based upon an article,[5] which provided clinically meaningful change estimates in a prostate cancer sample based upon an anchoring methodology. The FACT-G scales (PWB, SFWB, EWB, FWB, and the FACT-G) were not addressed in that article, and so the clinically meaningful change estimates for those scales are derived from an earlier article,[6] which

reports normative values from a large sample from the general population for the FACT-G scales. Taking one half the standard deviation is equivalent to finding a 0.5 effect size. Using this distributional technique, the clinically meaningful change estimates are defined in Table 3 below.

**Table 2:    Time to Deterioration Thresholds for FACT-P Scales**

| Scale/Sub-Scale | Threshold for definition of progression/deterioration |
|---|---|
| Physical Well-Being (PWB) | 3 |
| Social/Family Well-Being (SFWB) | 3 |
| Emotional Well-Being (EWB) | 3 |
| Functional Well-Being (FWB) | 3 |
| FACT-G (General) Scale | 9 |
| Prostate Cancer Subscale (PCS) | 3 |
| Trial Outcome Index (TOI) | 9 |
| FACT-P Total Scale | 10 |
| FACT-P Pain Scale | 2 |
| FAPSI-8 | 3 |

For each PRO scale, the median time to degradation will be estimated using a Kaplan-Meier method. Additionally, hazard ratio (apalutamide / placebo) and associated 95% confidence interval will be estimated by fitting Cox's proportional hazards model stratified by baseline stratification factors (PSADT, Bone-sparing agent use, and Loco-regional disease). Hazard ratio <1 favors apalutamide group. In cases where median values cannot be computed because less than 50% of subjects experienced progression/deterioration, 25th percentiles will be reported and compared instead.

### 2.5.5   Repeated Measures Model of PRO

For each PRO scale, repeated measures mixed effect model[7] will be fitted to estimate the mean PRO scores at each scheduled visit during the treatment phase. In this model the dependent variable will be the PRO score change from baseline. Fixed effects for the models will include, treatment and visit number as discrete parameters, and interaction between time and treatment; subject is included as random effect. Serial correlation is assumed, and a AR(1) correlation matrix will be explored, if appropriate, to account for the correlations between repeated measures within subjects. Lsmeans of change from baseline will be plotted for each PRO scale, by visit and treatment group. T-tests will be performed to test for the differences in mean PRO scores between treatment groups at each visit.

In this repeated measure modeling, a significant amount of missing data occur during final treatment cycles due to the subjects discontinuation of the treatment, truncation will be applied for all subsequent visits at the first visit where 90% or more of the subjects are

missing for any endpoint and from either arm. Once the truncation cycle is determined, this cycle will be applied across both treatment arms. To standardize the measurement points, the End-of-Treatment assessment will be parsed out to the next scheduled cycle visit after the last treatment cycle.
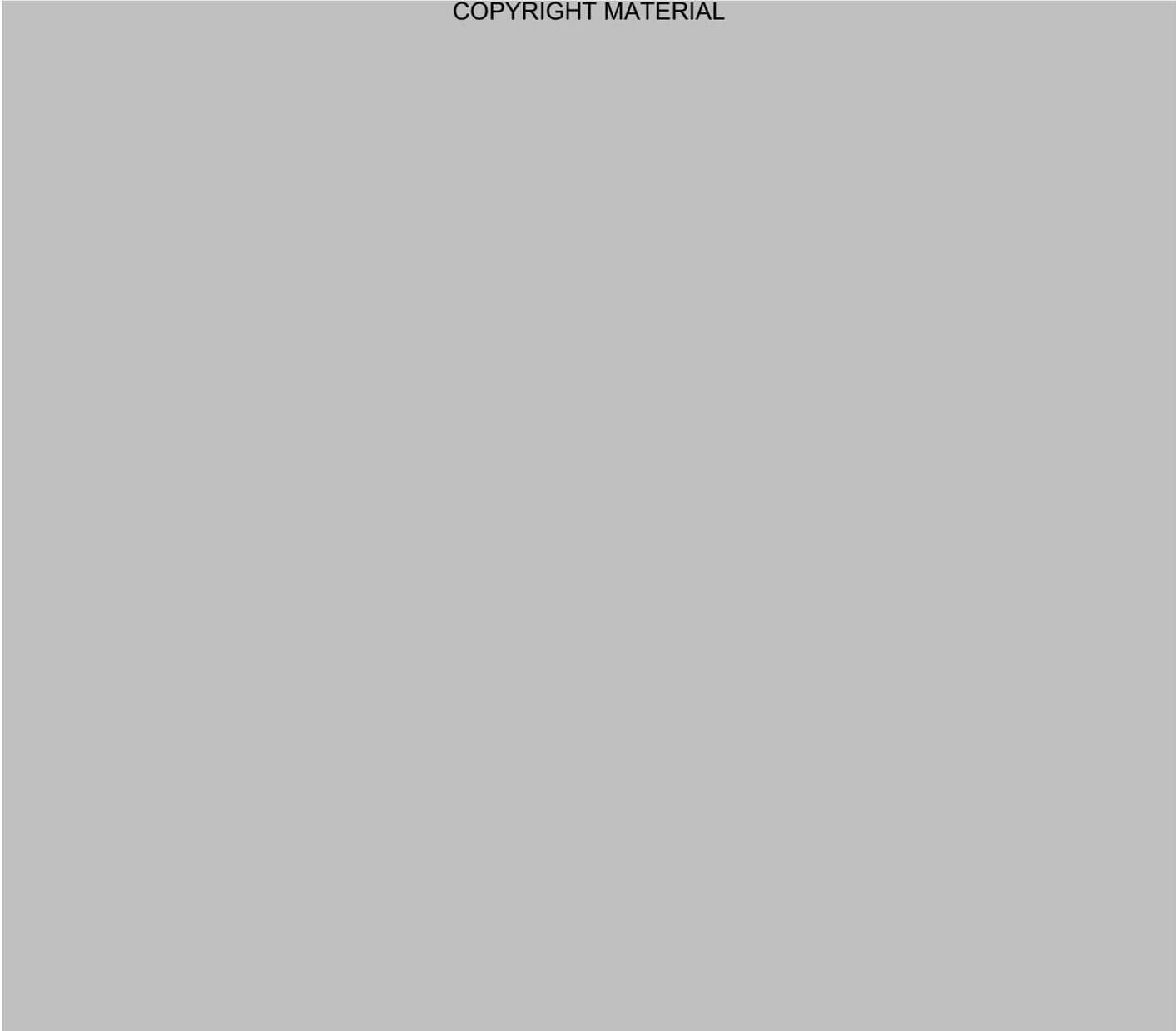
## REFERENCES

1. Esper P, Mo F, Chodak G, et al. Measuring quality of life in men with prostate cancer using the Functional Assessment of Cancer Therapy-Prostate (FACT-P) instrument. Urology 1997;50(6):920-8.

2. The EuroQol Group. EuroQol-a new facility for the measurement of health-related quality of life. Health Policy 1990;16:199-208.

3. EQ-5D-3L User Guide.
http://www.euroqol.org/fileadmin/user_upload/Documenten/PDF/Folders_Flyers/EQ-5D-3L_UserGuide_2015.pdf

4. C. N. Sternberg, A. Molina, S. North, P. Mainwaring, K. Fizazi, Y. Hao, M. Rothman, D. D. Gagnon, T. Kheoh, C. M. Haqq, C. Cleeland, J. S. de Bono, H. I. Scher. Effect of abiraterone acetate on fatigue in patients with metastatic castration-resistant prostate cancer after docetaxel chemotherapy. Annals of Oncology 24(4): 1017-1025, 2013.

5. Cella D, Nichol MB, Eton D, Nelson JB, Mulani P. Estimating Clinically Meaningful Changes for the Functional Assessment of Cancer Therapy—Prostate: Results from a Clinical Trial of Patients with Metastatic Hormone-Refractory Prostate Cancer. Value in Health 12(1): 124-129, 2009.

6. Cella D, Zagari MJ, Vandoros C, Gagnon DD, Hurtz HJ, and Nortier JWR. Epoetin Alfa Treatment Results in Clinically Significant Improvements in Quality of Life in Anemic Cancer Patients When Referenced to the General Population. JCO 21(2): 366-373, 2003.

7. Fitzmaurice G, Laird N, Ware J. Applied Longitudinal Analysis, 2004

## APPENDIX 1: FACT-PROSTATE QUESTIONNAIRES AND SCORING GUIDELINE

Version 4, US English, Copyright 1987, 1997

COPYRIGHT MATERIAL

ARN-509-003
PRO Statistical Analysis Plan v2.0 26 June 2017 Page 11 of 17

COPYRIGHT MATERIAL

COPYRIGHT MATERIAL

ARN-509-003
PRO Statistical Analysis Plan v2.0 26 June 2017 Page 13 of 17

Aragon Pharmaceuticals - Confidential

JNJ-56021927 (apalutamide)
ARN-509-003 Documentation of Statistical Methods and Interim Analysis Plans

COPYRIGHT MATERIAL

ARN-509-003
PRO Statistical Analysis Plan v2.0 26 June 2017                    Page 15 of 17

COPYRIGHT MATERIAL

ARN-509-003
PRO Statistical Analysis Plan v2.0 26 June 2017

## APPENDIX 2: EQ-5D-3L

COPYRIGHT MATERIAL

JNJ-56021927 (apalutamide)
ARN-509-003 Documentation of Statistical Methods and Interim Analysis Plans

COPYRIGHT MATERIAL

**ARN-509-003 (SPARTAN): Summary of Major Changes and Rationale for Each Statistical Analysis Plan Amendment**

| Major Changes | Rationale |
|---|---|
| **v1.0 (dated 05 November 2012) to v2.0 (dated 27 February 2014)** | |
| Overall survival (OS) was changed from being the only key secondary endpoint to being one of the 5 secondary endpoints that included time to symptomatic progression, time to initiation of cytotoxic chemotherapy, progression-free survival (PFS), and time to metastasis (TTM). The statistical testing of secondary endpoints were to be performed by allocating 0.04 to OS and 0.01 for the rest of the endpoints using Bonferroni method to control the overall familywise type I error rate at 0.05. | To incorporate multiple testing method in agreement with the FDA proposal. |
| The number of deaths required for the OS final analysis was revised to reflect the change in the alpha allocation of 0.05 in the original analysis plan to the new alpha allocation of 0.04. One additional IA of OS at approximately 70% of deaths was added. | Modification of the interim analysis of OS, 2 interim analyses and 1 final analysis were planned for OS. |
| The definition and the analysis plan of PFS2 was added. | Added second progression-free survival (PFS2) as an evaluation per agreement with health authorities. |
| | |
| **v2.0 (dated 27 February 2014) to v3.0 (dated 15 April 2015)** | |
| A new section of Analysis by Formulation was added to include additional analysis plans for subject distribution, demographic and baseline characteristics, study drug administration, safety, and the primary endpoint | The protocol was amended to switch softgel capsules to tablets (commercial formulation) for patients receiving the softgel capsules and to administer tablets to newly enrolled patients. |
| | |
| **v3.0 (dated 15 April 2015) to v4.0 (dated 15 March 2017)** | |
| Changed the multiple testing procedure for the secondary endpoints to a hierarchical testing procedure that also included a sample size re-estimation for the required number of events for time to symptomatic progression and overall survival; Details provided in Section 7.5.3 and Appendix 1. | The Bonferroni testing procedure as described in v3.0 of the SAP for secondary endpoints is conservative and lacked power, especially for some key secondary endpoints with low event rates, at the time of the primary analysis. |
| Simplified and removed some unnecessary | Deemed unnecessary in the presence of |

| sensitivity analyses: Removed multivariate analysis of MFS and the analysis of the impact of subsequent therapy on MFS; removed investigator derived PFS and TTM analysis | several other sensitivity analyses of MFS and OS already in the SAP. |
|---|---|
| Removed the by-formulation analysis using the arbitrary cutoff of 6 months (>6 vs <= 6 months of duration of treatment), added analysis by greater tablet duration or greater capsule duration | FDA feedback: Analysis by greater tablet duration or greater capsule duration would be more meaningful |
| Added Appendix 2, table shell for overall summary of AE by formulation | Added for clarity, as to what exactly will be summarized for AE by formulation. |
| | |
| **v4.0 (dated 15 March 2017) to v5.0 (dated 28 March 2017)** ||
| Minor revision in Section 7.5.3 and Appendix 1. Clarified that under the hierarchical testing procedure framework, "not tested" means that no statistical significance can be claimed if the endpoint above in the hierarchy is not found to be positive. | Based on FDA Type A meeting feedback and correspondences. |
| | |
| **v5.0 (dated 28 March 2017) to v6.0 (dated 22 June 2017)** ||
| Removed the definition of the PRO population | ITT population will be used for PRO analysis and hence no need to define a PRO population |
| Revised the criteria for markedly abnormal vital signs in Section 7.6.3 | Clinical proposed new cutoff values based on CTCAE grade and change within grade |
| Added Appendix 3 "Modified MedDRA queries as search criteria for AE of special interest" | For clarity |
| | |
| **v6.0 (dated 22 June 2017) to v7.0 (dated 26 June 2017)** ||
| Minor changes in Appendix 1- clarification on the choice of weights in p-value combination methods. | Fixed weights to be used. |
| | |