Statistical Analysis Plan I3Y-MC-JPCM Version 4

A Phase 2/3, Randomized, Double-Blind, Placebo-Controlled Study of Abiraterone Acetate Plus Prednisone With or Without Abemaciclib in Patients With Metastatic Castration-Resistant Prostate Cancer

NCT03706365

Approval Date: 25-Oct-2023

1. Statistical Analysis Plan: I3Y-MC-JPCM CYCLONE 2: A Phase 2/3, Randomized, Double-Blind, Placebo Controlled Study of Abiraterone Acetate plus Prednisone with or without Abemaciclib in Patients with Metastatic Castration-Resistant Prostate Cancer

Abemaciclib (LY2835219)

Study I3Y-MC-JPCM is a Phase 2/3, multicenter, multinational, randomized, double-blind, placebo-controlled study designed to evaluate the clinical benefit of abiraterone acetate plus prednisone with or without abemaciclib in patients with metastatic castration-resistant prostate cancer.

Eli Lilly and Company
Indianapolis, Indiana USA 46285

Protocol I3Y-MC-JPCM
Phase 2/3

Statistical Analysis Plan electronically signed and approved by Lilly: 30 November 2018

Statistical Analysis Plan Version 2 electronically signed and approved by Lilly: 15 May 2020

Statistical Analysis Plan Version 3 electronically signed and approved by Lilly: 19 March 2021

Statistical Analysis Plan Version 4 electronically signed and approved by Lilly on date provided below.

Document ID: VV-CLIN-081628

2. Table of Contents

Section **Page**

Table of Contents

Table **Page**

Table of Contents

Figure **Page**

3. Revision History

SAP Version 1 was approved prior to unblinding.

SAP Version 2 was approved prior to any analysis of efficacy data and prior to the planned futility analysis. The SAP was updated to Version 2 following the approval of amendment (b) to the protocol, and maintained alignment with the protocol by:

- the establishment of an independent data monitoring committee (DMC) responsible for all interim efficacy analyses and reference to the corresponding DMC Charter
- the requirement for regular blinded safety reviews to be conducted by the study team
- the incorporation of central collection of radiographic images to enable blinded independent review of imaging-based endpoints

SAP Version 3 was approved after the futility analysis planned at 30 radiographic progression-free survival (rPFS) events and prior to any other analysis of efficacy data, in conjunction with the approval of amendment (c) to the protocol. These enable an adaptive Phase 2/3 study design by changing the timing and number of interim analyses and adding Part 3, which will be triggered based on prespecified expansion rules at adaptive interim analysis 1 and/or 2. If triggered, Part 3 will randomize an additional 170 patients. Time to prostate-specific antigen (PSA) progression is not considered an appropriate validated surrogate efficacy endpoint, hence, is being removed as a co-primary endpoint and changed to a secondary endpoint. With this change, rPFS becomes the sole primary efficacy endpoint. Study-wise type I error rate will now be controlled at 0.025, by:

- Splitting alpha at each interim analysis and the primary analysis of rPFS
- Testing overall survival in a gated, hierarchical fashion

Testing for time to PSA progression will no longer be included in the study-wise type I error control. The overall survival (OS) testing scheme has been updated to account for these changes. Additionally, rPFS by blinded independent central review (BICR) (if Part 3 is opened for enrollment) and time to worst pain progression are being added as secondary endpoints. CCI [redacted] Response-based endpoints such as objective response rate (ORR) and duration of response (DoR) were updated to require confirmation of a response. Appendices 1 through 3 were added to provide more detail on the expansion criteria for the adaptive interim analyses, to demonstrate that this design adequately controls type I error, and to describe simulation results that illustrate the design's operating characteristics.

SAP Version 4 was approved prior to the primary analysis of rPFS, before Lilly was unblinded to any treatment assignments, but after the last patient had entered treatment. This update included

- The addition of Appendix 4 to document Notable Patient Criteria, which were previously maintained in the trial's Patient Narrative Planning Tool

- Clarifications to the rPFS censoring rules, as well as refinements to the planned sensitivity analyses and details on the assessment of concordance between rPFS by investigator assessment and rPFS by BICR
- A change to the OS analysis plan (a gated secondary endpoint), requiring more events before the final OS analysis is performed and adding a second interim analysis. This change was motivated by recent clinical trial results coming from similar patient populations. Waiting for more mature OS data will enable a better characterization of the true OS effect.
- Refinements to the planned subgroup analyses and baseline disease characteristics
- Additional prespecified exploratory HRQoL analyses
- Other minor clarifications throughout

4. Study Objectives

Table JPCM.1. shows the objectives and endpoints of the study.

Table JPCM.1. Objectives and Endpoints

Objectives Part 1	Endpoints
Primary	
To determine the RP2D of abemaciclib that may be safely administered to patients with mCRPC in combination with abiraterone acetate and prednisone.	Safety (including, but not limited to): TEAEs, SAEs, deaths, and clinical laboratory abnormalities.
Objectives Parts 1, 2, and 3: Combined Populations	**Endpoints**
Primary	
To compare the rPFS of patients receiving abiraterone acetate plus prednisone with or without abemaciclib.	rPFS
Secondary	
To further characterize the safety profile of the combination of abemaciclib and abiraterone acetate plus prednisone.	The safety endpoints evaluated will include, but are not limited to, the following: AEs, TEAEs, SAEs, clinical laboratory tests, ECGs, vital signs, and physical examinations.
To compare the efficacy in patients receiving abiraterone acetate plus prednisone with or without abemaciclib.	• ORR and DoR • OS • Time to PSA progression • (If Part 3 is opened) rPFS by blinded independent central review (BICR)
Time to symptomatic progression.	Time from randomization to any of the following (whichever occurs earlier): • Symptomatic skeletal event: symptomatic fracture, surgery or radiation to bone, or spinal cord compression • Pain progression or worsening of disease-related symptoms requiring initiation of a new systemic anticancer therapy • Development of clinically significant symptoms due to loco-regional tumor progression requiring surgical intervention or radiation therapy.
To characterize the PK of abemaciclib and abiraterone acetate when administered in combination.	Abemaciclib and abiraterone acetate steady-state plasma concentrations.
To assess patient-reported pain.	Time to worst pain progression, using the Worst Pain NRS score and the WHO-AL.
CCI	

Abbreviations: AEs = adverse events; DoR = duration of response; ECG = electrocardiogram; CCI mCRPC = metastatic castration-resistant prostate cancer; NRS = numerical rating scale; ORR = objective response rate; OS = overall survival; PK = pharmacokinetics; PSA = prostate-specific antigen; ; RP2D = recommended Phase 2 dose; rPFS = radiographic progression free survival; SAE = serious adverse event; TEAE = treatment-emergent adverse event; CCI ; WHO-AL = World Health Organization – Analgesic Ladder.

5. Study Design

5.1. Summary of Study Design

Study I3Y-MC-JPCM (JPCM) is a Phase 2/3, multicenter, multinational, randomized, double-blind, placebo-controlled study designed to evaluate the clinical benefit of abiraterone acetate plus prednisone with or without abemaciclib in patients with metastatic castration-resistant prostate cancer (mCRPC). Planned analyses with detailed descriptions are listed below in Table JPCM.2..

Table JPCM.2. Planned Analyses

Analysis	Timing	Details
RP2D Selection	CCI	AC review of safety and PK data only to select RP2D (Section 6.14.1).
Futility	Approximately CCI rPFS events	Performed by independent DMC to assess if stopping rule for futility has been met (Section 6.14.2).
Adaptive Interim 1	Approximately CCI rPFS events	Performed by independent DMC to assess if criteria for opening enrollment to Part 3 have been met (Section 6.14.2).
Adaptive Interim 2	Approximately CCI rPFS events	(If applicable) Performed by independent DMC to assess if criteria for opening enrollment to Part 3 have been met (Section 6.14.2).
Primary Analysis of rPFS	Approximately CCI rPFS events if Part 3 is not opened. Approximately CCI rPFS events if Part 3 is opened for enrollment	Final analysis of primary endpoint (Section 6.7.1), and if applicable, interim analysis of OS (Section 6.7.2)
Interim Analysis of OS	See Section 6.7.2	See Section 6.7.2 for details.
Final Analysis of OS	See Section 6.7.2	See Section 6.7.2 for details.

Abbreviations: AC = assessment committee; DMC = data monitoring committee; OS = overall survival; PK = pharmacokinetic; RP2D = recommended Phase 2 dose; rPFS = radiographic progression-free survival.

This study will include a lead-in portion (Part 1) to evaluate safety and tolerability, and to determine the recommended Phase 2 dose (RP2D) of abemaciclib in combination with abiraterone acetate plus prednisone.

Part 1 is a double-blind placebo-controlled safety lead-in period that will randomize approximately 30 patients to explore 2 doses of abemaciclib, 150 mg twice daily (Arm A1; 3 tablets/capsules twice daily) and 200 mg twice daily (Arm A2; 4 tablets/capsules twice daily), in combination with abiraterone acetate 1000 mg once daily plus prednisone 5 mg twice daily. Both Arms A1 and A2 will be placebo-controlled (Arm B1 and B2; patients will receive 3 or 4 placebo tablets/capsules twice daily and abiraterone acetate 1000 mg once daily plus prednisone

5 mg twice daily). Thus, there will be 4 arms in Part 1 and randomization will be based on a 2:2:1:1 ratio.

CCI [redacted] patient has entered treatment, an internal assessment committee (AC) will review safety and available pharmacokinetic (PK) data and determine the RP2D. The AC will, at a minimum, be composed of Lilly members not involved in the day-to-day study conduct (the medical director, a Global Patient Safety physician, a PK/pharmacodynamics scientist, and a statistician). In the case of unacceptable and/or unmanageable toxicity of the combination, the AC may decide to discontinue the study upon completion of Part 1. During the evaluation period, patients may still enroll in Part 1. Once RP2D has been selected, patients enrolled in Part 1 will switch to the selected RP2D (or matching placebo) to continue the study. In the situation where blinded study treatment has been dose reduced below the selected RP2D, dose escalation to RP2D will not be implemented and dose-reduced patients will continue study at the same lower dose level. Subsequent patients will then start to enroll in Part 2 and will be randomized to receive abemaciclib at RP2D or matching placebo.

Part 2 will randomize approximately 150 patients in a 1:1 ratio between 2 arms:

- Arm A (experimental arm): Abiraterone acetate 1000 mg once daily plus prednisone 5 mg twice daily and abemaciclib at RP2D twice daily
- Arm B (control arm): Abiraterone acetate 1000 mg once daily, prednisone 5 mg twice daily, and placebo (matching number of tablets/capsules twice daily).

After approximately CCI [redacted] rPFS events, an independent data monitoring committee (DMC) will perform a futility analysis. As described in the DMC Charter, the independent DMC will consist of at least 3 members external to Lilly, none of whom are involved as study investigators, including at least 1 clinician and 1 statistician. CCI [redacted], the DMC will perform an adaptive interim analysis. If the prespecified adaptive expansion criteria are met, Part 3 will be opened for enrollment. If the expansion criteria are not met, the independent DMC will perform another adaptive interim analysis CCI [redacted]. If the prespecified adaptive expansion criteria are met, Part 3 will be opened for enrollment. If opened for enrollment, Part 3 will randomize approximately 170 additional patients in a 1:1 ratio. If the expansion criteria are not met in either interim adaptive analysis, Part 3 will not be opened for enrollment, and scientific evaluation will continue until study completion (Section 7.8 of the protocol).

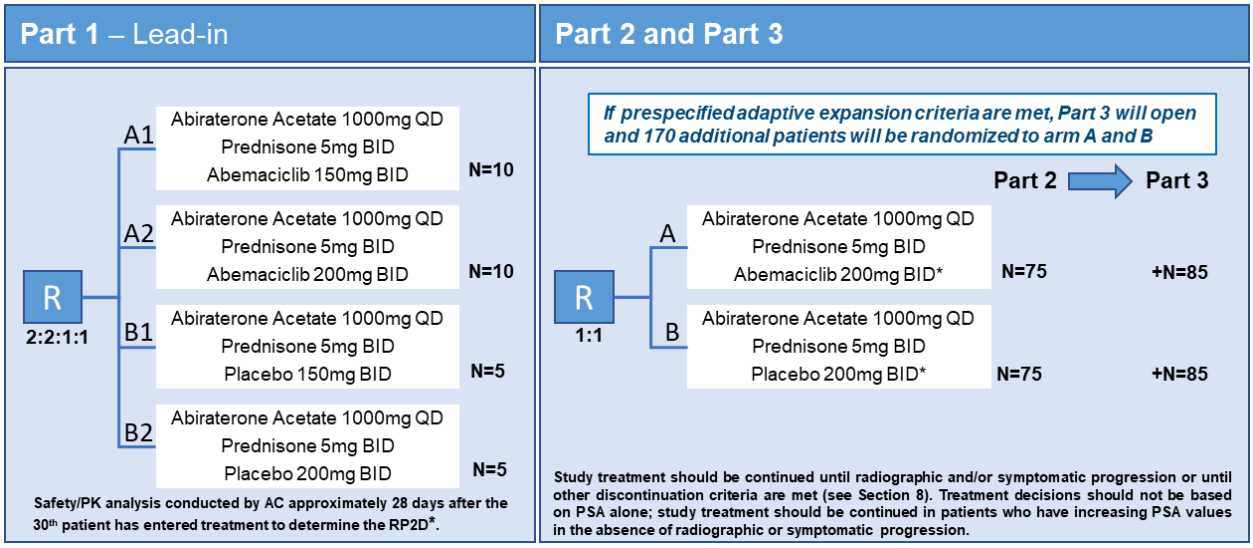
Abemaciclib, matching placebo, abiraterone acetate, and prednisone will be administered orally. Where prednisone is not commercially available, prednisolone may be substituted. Throughout the document, prednisone refers to prednisone or prednisolone.

Patients (enrolled in Part 1, Part 2, or Part 3) will be centrally randomized to the abemaciclib group(s) and the placebo group(s) according to the following stratification factors:

- Radiographic progression* at study entry (yes/no)
- Measurable disease** (yes/no)
- Prior docetaxel for metastatic hormone-sensitive prostate cancer (mHSPC) (yes/no)

* Per Response Evaluation Criteria in Solid Tumors, Version 1.1 (RECIST 1.1) for soft tissue and/or per Prostate Cancer Working Group 3 (PCWG3) for bone.

** Per RECIST 1.1.



Abbreviations: AC = assessment committee; BID = twice daily; N = number of patients; Part 1: N=30, CCI [redacted]; PK = pharmacokinetic; CCI [redacted]; QD = once daily; R = randomization; RP2D = recommended Phase 2 dose.

CCI [redacted]

Figure JPCM.1. Illustration of study design.

5.2. Determination of Sample Size

Patients enrolled in Part 1, Part 2, and Part 3 (if applicable) will be combined to form the experimental group (approximately 95 patients from Parts 1 and 2 in abemaciclib plus abiraterone acetate and prednisone, with an additional 85 patients from Part 3, if opened) and control group (approximately 85 patients from Parts 1 and 2 in the placebo plus abiraterone acetate and prednisone group, with an additional 85 patients from Part 3, if opened).

Radiographic progression-free survival is the primary endpoint for the study. The overall type I error is controlled at a 1-sided alpha level of 0.025 for the study. Group-sequential design and an extension of Chen's 2-in-1 adaptive design (Chen 2018) will be used to accommodate the event-driven plan for the futility, interims, and final rPFS analyses (see Section 6.7.1.2 for details). If Part 3 is opened, the final rPFS analysis will be performed after approximately CCI rPFS events have occurred (ie, an approximate CCI censoring rate). If Part 3 is not opened, the final rPFS analysis will be performed after approximately CCI rPFS events have occurred (ie, an approximate 44% censoring rate). CCI [redacted]

CCI [redacted].

Further details on the marginal and conditional operating characteristics of this design can be found in Appendix 2 of the SAP, which describes simulated operating characteristics for varying true effects. Furthermore, Appendix 3 contains a proof that the expansion rules and design do not inflate overall type I error.

CCI [redacted]

6. A Priori Statistical Methods

6.1. General Considerations

6.1.1. Populations

The following populations will be defined for this study:

Intention-to-treat (ITT) population: will include all randomized patients. The ITT analysis of efficacy data will consider allocation of patients to treatment groups as randomized and not by actual treatment received. Here, treatment groups refer to patients randomized to abemaciclib (Arms A1, A2, or A) or patients randomized to placebo (Arms B1, B2, or B). This population will be used for all baseline, efficacy (except for soft tissue tumor response related endpoints such as best overall response), and health economics analyses.

Measurable disease population: will include all randomized patients who have measurable disease in soft tissue at baseline according to modified RECIST v1.1. This population will be used for efficacy analyses based on soft tissue tumor response related endpoints such as best overall response.

Safety population: will include all randomized patients who received at least one dose of any study drug, regardless of their eligibility for the study. The safety evaluation will be performed by treatment group based on the first dose of study treatment a patient actually received, regardless of the treatment group to which he was randomized. The safety population will be used for all dosing/exposure, safety, and resource utilization analyses.

Pharmacokinetic population: will include all randomized patients who received at least 1 dose of any study drug and have evaluable PK samples and sufficient dosing information.

Biomarker population: will include the subset of patients from the ITT population from whom a valid assay result has been obtained.

6.1.2. Definitions and Conventions

Study drug refers to abemaciclib or placebo.

Study treatment refers to abemaciclib plus abiraterone acetate and prednisone, or placebo plus abiraterone acetate and prednisone.

The **date of randomization** is the date the patient was randomly assigned to

- Arm A1, A2, B1, or B2 if enrolled in Part 1;
- Arm A or Arm B if enrolled in Part 2 or Part 3

using the IWRS.

The **date of first dose** is the date of the first dose of study drug, abiraterone acetate, or prednisone.

The **baseline value of a safety assessment** is the last value observed prior to the first dose. This may occur on the day of first dose.

The **baseline value of an efficacy assessment** is the last value observed prior to the date of randomization. If a patient's first assessment occurs after randomization but prior to the first dose, this assessment will be used as the baseline.

The **study day of a safety event or assessment** will be calculated as:

- the difference between the date of the event or assessment and the date of first dose plus 1 for all events or assessments occurring on or after the day of first dose. For example, if an event occurs on 08JUN2018 and the date of first dose was 06JUN2018, the study day of the event is 3.
- the difference between the date of the event or assessment and the date of first dose for all events or assessments occurring before the day of first dose. For example, if an event occurs on 05JUN2018 and the date of first dose was 06JUN2018, the study day of the event is -1.

The **study day of an efficacy event or assessment** will be calculated as:

- the difference between the date of the event or assessment and the date of randomization plus 1 for all events or assessments occurring on or after the date of randomization.
- the difference between the date of the event or assessment and the date of randomization for all events or assessments occurring before the date of randomization.

One **month** is defined as 365/12 days.

Unless otherwise noted, **summaries of continuous variables** will include a mean, median, standard deviation, minimum, and maximum.

Unless otherwise noted, **summaries of categorical variables** will include the frequency and percentage (relative to the population being analyzed) of each category.

Statistical analysis of this study will be the responsibility of Lilly or its designee.

Unless otherwise stated, all tests of treatment effects will be conducted at a 1-sided alpha level of .025, and all confidence intervals (CIs) will be given at a 2-sided 95% level.

6.2. Handling of Dropouts or Missing Data

With the exception of dates, missing data will not be imputed. Partial dates may be recorded, particularly for events prior to study entry, such as prior therapies, prior procedures, or medical history events. Generally, for start dates, if only the day is unknown, it will be imputed as 1. If the month is also unknown, it will be imputed as January. However, for end dates, if only the day is unknown, the last day of the month will be imputed, and if the month is unknown, it will be imputed as December. Any other method of imputation for any dates that are imputed is described in the relevant section.

6.3. Patient Disposition

A detailed description of patient disposition will be provided. It will include a summary of the number and percentage of patients entered into the study, enrolled in the study, treated in the study, reasons for discontinuation from study treatment (safety population only), and reasons for discontinuation from study (ITT population only). Reason for discontinuation from both study treatment and the study will be summarized by pre-determined categories. If the reason for discontinuation is adverse event (AE), the associated AE term will be reported.

6.4. Patient Characteristics

6.4.1. Demographics

Patient demographics will be summarized for all randomized patients. Patient demographics will include sex, race, ethnicity, country, age, height, weight, and body mass index.

6.4.2. Baseline Disease Characteristics

Disease characteristics will be summarized. Disease characteristics will include the following:

- Initial histopathological diagnosis
- Tumor, node, metastasis (TNM) stage at initial diagnosis
- Gleason score at initial diagnosis
- Time from initial diagnosis to
  - Castration-resistant prostate cancer (CRPC)
  - Study treatment start date
- Type of progression at study entry (PSA progression only versus radiographic progression)
- Extent of disease: nodal disease only, bone disease (with or without nodal disease), visceral disease (with or without nodal disease and/or bone disease), other
- Measurable vs nonmeasurable disease in soft tissue
- Number of organs involved (1, 2, or 3+)
- Site(s) of disease (Local, Nodal-Regional, Nodal-Distant, Liver, Lung, Soft Tissue-Other, Bone)
- Baseline Eastern Cooperative Oncology Group (ECOG) performance status (PS)
- Prior docetaxel use (Yes/No)
- Baseline pain score (0 to 1, 2 to 3, or >3)
- Baseline laboratory values (PSA, testosterone, lactate dehydrogenase [LDH], alkaline phosphatasemax [ALP], hemoglobin [Hb], albumin)

Extent of disease and number of organs involved will be derived from the baseline electronic clinical (case) report forms (eCRFs): 'Target Tumor: RECIST 1.1 (Tumor Assessment)' and 'Non-Target Tumor: RECIST 1.1 (Tumor Assessment)' for soft tissue, and 'Bone Lesion – Baseline (Tumor Assessment).' The number of organs involved in soft tissue will be derived from the location codes of the target and non-target lesions.

Exploratory analyses using PSA at initial diagnosis and/or PSA doubling time (PSADT) prior to study entry may be performed as deemed appropriate.

6.4.3. Historical Illnesses

Historical illnesses are clinically relevant events in the past that ended before the screening visit. Historical illnesses (using Preferred Term(s) [PTs] from the most current version of the Medical Dictionary for Regulatory Activities [MedDRA™]) will be summarized.

6.4.4. Prior Therapies

Prior radiotherapy, surgery, and systemic therapy will be summarized. Prior radiotherapy and surgery will be categorized by reason for regimen. Prior systemic therapies will be categorized by type of regimen (hormone therapy, chemotherapy, targeted therapy, immunotherapy, radiopharmaceuticals, or other) and reason for regimen ([neo]adjuvant therapy or therapy for locally advanced, nonmetastatic castration resistant, metastatic hormone sensitive, or metastatic castration resistant disease). Frequency of each specific therapy will be tabulated within each type of regimen and reason for regimen.

Additionally, type of androgen deprivation therapy at study entry will be summarized by type (surgical or chemical), with the frequency of each specific therapy tabulated within each type.

6.4.5. Post-Study Treatment Discontinuation Therapies

Therapies received following study treatment discontinuation will be summarized by treatment group. Therapies will be summarized overall and by category: hormone therapy, chemotherapy, targeted therapy, immunotherapy, radiopharmaceuticals, or other.

6.5. Treatment Compliance

Treatment compliance of abemaciclib/placebo will be measured by tablet/capsule counts and summarized cumulatively. Compliance will be calculated as the ratio of total dose taken to the total assigned dose (plus or minus any dose adjustments and doses omitted/withheld). The total assigned dose for a patient with no adjustments or omissions is as follows:

- 150 mg per dose × 2 doses per day × number of days on study drug, if enrolled in Part 1 and randomized to Arm A1 or B1;
- 200 mg per dose × 2 doses per day × number of days on study drug, if enrolled in Part 1 and randomized to Arm A2 or B2;
- RP2D per dose × 2 doses per day × number of days on study drug, if enrolled in Part 2 or Part 3.

Treatment compliance of abiraterone acetate and prednisone will be calculated using the ‘Exposure Compliance: Study Treatment Abiraterone Acetate’ and ‘Exposure Compliance: Study Treatment Prednisone’ forms, respectively. Since abiraterone acetate and prednisone will not be centrally supplied by Lilly, compliance will only be collected and calculated on the scale of number of doses, without considering detailed dose level (mg per dose). The estimate of

percent compliance will be calculated as ratio of total number of doses taken to the total number of assigned dose (minus number of doses omitted/withheld).

6.6. Concomitant Therapy

All medications will be coded to the generic preferred name according to the current World Health Organization (WHO) drug dictionary. All concomitant medications will be summarized by number and percentage of patients for the ITT population using the base name (without esters or salts).

6.7. Efficacy Analyses

Unless otherwise noted, all efficacy analyses will be performed on the ITT population.

The stratification factors for the analysis of primary and secondary analyses are

- Radiographic progression at study entry (yes/no)
- Measurable disease (yes/no)
- Prior docetaxel for mHSPC (yes/no)

The stratification factors are captured in the IWRS and are also derived from information collected on eCRFs. Unless otherwise specified, all stratified analyses will be based on the stratification factor per eCRFs. A cross tabulation of the frequency of each level of the stratification factor per IWRS and eCRF will be produced.

Unless otherwise stated, all tests of treatment effects will be conducted at a 1-sided alpha level of .025, and all confidence intervals (CIs) will be given at a 2-sided 95% level.

Radiographic progression-free survival (rPFS) is the primary efficacy endpoint for the study. The overall type I error is controlled at a 1-sided alpha level of 0.025 for the study. The study will be considered positive only if rPFS is positive at the primary analysis described below. If positive, OS will be hierarchically tested as described in Section 6.7.2.

6.7.1. Primary Endpoint: Radiographic Progression-Free Survival

6.7.1.1. Definition

The primary endpoint of this study is rPFS. The rPFS time is measured from the date of randomization to the earliest date of investigator determined radiographic disease progression (by objective radiographic disease assessment per RECIST 1.1 for soft tissue AND/OR radionuclide bone scan using PCWG3 criteria for bone) or death from any cause, whichever occurs first. Patients who have neither progressed nor died will be censored at the day of their last radiographic tumor assessment (if available) or date of randomization if no post-initiation (ie, post-baseline) radiographic assessment is available. The detailed censoring rules are described in Table JPCM.3..

Table JPCM.3. Radiographic Progression-Free Survival Event/Censoring Scheme

Situation[a]	Event/Censor	Date of Event or Censor
Radiographic disease progression or death	Event	Earliest date of radiographic disease

Situation[a]	Event/Censor	Date of Event or Censor
		progression or death[b]
No radiographic disease progression and no death	Censored	Date of last adequate radiographic assessment showing no evidence of disease progression or date of randomization (whichever is later)[c]
Unless		
No baseline radiographic tumor assessment available	Censored	Date of randomization
No adequate postbaseline radiographic tumor assessment available and death reported after 2 scan intervals following randomization[c,d]	Censored	Date of randomization
New systemic anticancer treatment started and no radiographic disease progression or death within 14 days[e]	Censored	Date of last adequate radiographic assessment showing no evidence of disease progression prior to (start of new therapy +14 days) or date of randomization (whichever is later)[c]
Radiographic disease progression or death documented after 2 or more scan intervals following last adequate radiographic tumor assessment or randomization (whichever is later)[c,d]	Censored	Date of last adequate radiographic assessment showing no evidence of disease progression or date of randomization (whichever is later)[c]

Abbreviations: CR = complete response; LHRH = ; luteinizing hormone-releasing hormone; PD = progressive disease; rPFS = radiographic progression-free survival; PR = partial response; SD = stable disease.

[a] Clinical/symptomatic deterioration that is not radiologically confirmed will not be considered as radiographic disease progression.

[b] To be specific, once the radiographic disease progression is observed, especially for confirmed bone progression, the date of disease progression is the date at which initial bone progression (or soft tissue progression, whichever is earlier) is observed, not the date that the bone progression is confirmed by a subsequent assessment. Censoring will be performed according to the date of progression. For example, if a patient has observed initial bone disease progression, initiated a new systemic anticancer treatment, and only later had the bone progression confirmed, this will still be considered as an rPFS event.

[c] Adequate radiographic tumor assessment refers to an assessment of one of the following:

- Soft tissue disease with one of the following responses: CR, PR, SD; and adequate bone scan showing no or initial evidence of disease progression.
- Soft tissue disease with a response of PD.
- Bone scan showing confirmed disease progression.

To be specific, “last adequate radiographic assessment showing no evidence of disease progression” does not include the assessment showing initial/unconfirmed evidence of bone progression. If tumor scans for one assessment were performed on multiple days, use the earliest scan date for that visit as the date of assessment.

[d] Two scan intervals refer to the time period of next two protocol scheduled radiographic tumor assessments. Time is measured from the last adequate radiographic tumor assessment date. The window of the tumor assessment is also considered. For example, if the last adequate assessment occurs during Cycle 3, the next two protocol mandated assessments are Cycle 5 and Cycle 7 and the window around each assessment is 7 days, then two scan intervals will be considered as 2*(28*2+7) =126 days.

[e] New systemic anticancer treatment refers to treatment with intended oncological benefits. For example, the following would not be considered as new systemic anticancer treatments: LHRH agonist/antagonist, first generation antiandrogen, bone-modifying agents, and supportive care. Additionally, surgery and radiotherapy would not be considered as new systemic anticancer treatments.

6.7.1.2. Primary Analysis

The primary analysis of rPFS to test the superiority of experimental group to control group in improving rPFS time will be performed on the ITT population at an experiment-wise 1-sided alpha level of 0.025 and will use the log-rank test stratified by the randomization factors of:

- Radiographic progression at study entry (yes/no)
- Measurable disease (yes/no)
- Prior docetaxel for mHSPC (yes/no)

There is 1 planned futility analysis, up to 2 planned adaptive interim analyses, and 1 final analysis for rPFS in this study. CCI

CCI

The study will only be positive if the final analysis of the primary endpoint, rPFS, is positive. If positive, testing of secondary endpoints will proceed as described in Section 6.7.2.

The interim rPFS analyses will be performed by the DMC. The requirements for unblinding the study team at the interim analysis are found in Protocol Section 10.3.5.

6.7.1.3. Additional Analyses of rPFS

6.7.1.3.1. rPFS Curves and Hazard Ratio (HR)

The Kaplan-Meier (KM) method (Kaplan and Meier 1958) will be used to estimate the rPFS curves as well as rPFS rates at 3, 6, 9, 12, 18, and 24 months for each treatment group. These rates will be compared based on a normal approximation for the difference between the rates.

The corresponding HR between treatment groups will be estimated using a stratified Cox regression model (Cox 1972), stratified by the randomization strata. An additional unstratified Cox regression model will be employed to explore the effects of the stratification variables on treatment response.

6.7.1.3.2. Secondary Endpoint: rPFS by Blinded Independent Central Review

If Part 3 is opened for enrollment, the above analyses of rPFS will be repeated for rPFS by BICR, defined as the time from the date of randomization to the earliest date of radiographic disease progression determined by BICR (by objective radiographic disease assessment per RECIST 1.1 for soft tissue AND/OR radionuclide bone scan using PCWG3 criteria for bone) or death from any cause, whichever occurs first. The same censoring rules, stratification factors, and planned analyses will be applied.

BICR-assessed rPFS is not intended to provide an alternative means of definitive analysis, but it may be useful to evaluate bias in local assessments. Discordance rates (that is, differences in assessment of progression between investigator and BICR) will be summarized for each group (Amit et al. 2011). The agreement between BICR and investigator within a treatment group is represented in a tabular form (Table JPCM.4.) and will be summarized for each group and overall. Additionally, differential discordance will be described using early discrepancy rate (EDR) and late discrepancy rate (LDR) differences.

Table JPCM.4. BICR-assessed Versus Investigator-assessed Disease Progression

	BICR	
Investigator	PD	No PD
PD	a=a1+a2+a3	b
No PD	c	d

Abbreviations: a = number of agreements on occurrence of PD, regardless of timing; a1 = number of agreements on timing and occurrence of PD; a2 = number of times investigators declare PD later than BICR; a3 = number of times investigators declare PD earlier than BICR; b = number of times investigators declare PD and BICR never declares PD; BICR = blinded independent central review; c = number of times BICR declares PD and investigators never declare PD; d = number of agreements on patients with no occurrence of PD; PD = progressive disease.

The EDR quantifies the frequency with which the investigator assessment declares progression earlier relative to BICR within each group and is defined as

$$EDR = (b+a3)/(a+b)$$

The LDR quantifies the frequency with which the investigator assessment declares progression later than BICR within each group and is defined as

$$LDR = (c+a2)/(b+c+a2+a3)$$

The EDR and LDR will be summarized for each treatment group and the differential discordance around each measure can be defined as the rate on the experimental group minus the rate on the control group. A negative differential discordance for the EDR and/or positive differential discordance for the LDR are suggestive of a bias in the investigator-assessed rPFS favoring the experimental group.

6.7.2. Gated Secondary Endpoint: Overall Survival

Overall survival (OS) is an important secondary endpoint for this study. The OS time is measured from the date of randomization to the date of death from any cause. If the patient is alive or lost to follow-up at the time of data analysis, OS data will be censored on the last date the patient is known to be alive.

There are 2 planned interim analyses and a final analysis for OS that will occur at the following time points:

- Interim OS Analysis 1: At the time of the final rPFS analysis (CCI rPFS events if Part 3 is opened for enrollment, or CCI rPFS events if Part 3 is not opened for enrollment)
- Interim OS Analysis 2: Approximately CCI months after Interim OS Analysis 1 or when events have been observed in approximately CCI of patients in the final ITT population, whichever comes first
- Final OS analysis: When events have been observed in approximately CCI of patients in the final ITT population (CCI OS events if exactly 350 patients were randomized or CCI OS events if exactly 180 patients were randomized)

To maintain the experiment-wise type I error rate, OS will be hierarchically tested in the following way: only if rPFS is significant will OS also be tested inferentially for significance (Glimm et al. 2010), specifically

- The first potential time point for OS analysis will be at the time of the primary rPFS analysis. If rPFS is significant at this stage, the first interim analysis of OS will also be performed. If OS is not significant at this stage, a second interim analysis of OS will be performed. If OS is still not significant at this stage, a final analysis will be performed as specified above.
- If rPFS is not significant at primary analysis, OS will not be statistically evaluated.

The cumulative 1-sided type I error rate of 0.025 for OS analysis will be maintained using the Lan-Demets method with the following O'Brien-Fleming like alpha-spending function:

$$\alpha^*(t_k) = 2\left(1 - \Phi\left(\Phi^{-1}(1-\alpha/2)\Big/\sqrt{t_k}\right)\right)$$

Here, t_k is the information fraction at time k, Φ is the standard normal cumulative distribution function, and Φ^{-1} is the standard normal quantile function. The boundary p-value at each analysis will be calculated based on the actual number of events observed at the time of analysis using software that implements this alpha-spending function (for example, ADDPLAN 6.0 or SAS 9.2, or East 6.0).

The OS analysis to test the superiority of the experimental group to the control group in improving OS time will use the log-rank test stratified by the randomization factors.

The following additional analyses will be conducted for OS:

- Kaplan-Meier curves (Kaplan and Meier 1958) will be generated; medians, quartiles, and appropriate point probabilities will be calculated. Interval estimates will be calculated. The OS rates at 1, 2, and 3 years for each treatment group will be estimated and compared using a normal approximation for the difference between the rates.
- The Cox regression stratified by the randomization factors will be used to estimate the HR between the 2 treatment groups, along with CI.

In addition, follow up time for OS will be defined from the date of randomization and will use the inverse of the censoring rules for OS. The median follow-up time will be calculated using the KM method.

6.7.3. Other Secondary Endpoints

6.7.3.1. Time to Prostate-Specific Antigen Progression (PSA)

6.7.3.1.1. Definition

Prostate-specific antigen progression is defined as a ≥25% increase and an absolute increase of ≥2 ng/mL above the nadir (or baseline value if baseline is the smallest on study), which is confirmed by a second value obtained 3 or more weeks later. Any post-baseline PSA measurements within 12 weeks since baseline will be ignored in determining PSA progression. Time to PSA progression is measured from the date of randomization to the date of first observation of PSA progression. Patients who have not had PSA progression will be censored at the day of their last PSA assessment (if available) or date of randomization if no post initiation (ie, postbaseline) PSA assessment is available. The detailed censoring rules are described in Table JPCM.5..

Table JPCM.5. Time to PSA Progression Event/Censoring Scheme

Situation	Event/Censor	Date of Event or Censor
Confirmed PSA progression	Event	Date of first observation of PSA progression[a]
Not known to have had confirmed PSA progression confirmed	Censored	Date of last PSA assessment showing no evidence of PSA progression or date of randomization (whichever is later)[b]
Unless		
No baseline PSA assessment, or no postbaseline PSA assessments available	Censored	Date of randomization

Situation	Event/Censor	Date of Event or Censor
beyond 12 weeks since baseline		

Abbreviations: PSA = prostate-specific antigen.

[a] To be specific, for confirmed PSA progression, the date of first observation of PSA progression is the date at which initial PSA progression is observed, not the date that the PSA progression is later confirmed by a subsequent assessment.

[b] To be specific, "last PSA assessment showing no evidence of PSA progression" does not include the assessment showing initial/unconfirmed evidence of PSA progression.

The time to PSA progression analysis to test the superiority of the experimental group to the control group in improving time to PSA progression will be performed on the ITT population and will use the log-rank test stratified by the randomization factors.

6.7.3.2. Analyses of Time to PSA Progression

The Kaplan-Meier method (Kaplan and Meier 1958) will be used to estimate the time to PSA progression curves as well as survival rates at 3, 6, 9, 12, and 18 months for each treatment group. These rates will be compared based on a normal approximation for the difference between the rates.

The corresponding HR between treatment groups will be estimated using a stratified Cox regression model (Cox 1972), stratified by the randomization strata. An additional unstratified Cox regression model will be employed to explore the effects of the stratification variables on treatment response.

6.7.3.3. Objective Response Rate

Objective response rate (ORR) is a summary measure of best overall response (BOR) as defined by RECIST v1.1 for soft tissue per investigator assessment. BOR is derived from time point responses. All time point responses observed while on study treatment and during the short-term follow-up period (but before the initiation of postdiscontinuation systemic anticancer therapy) will be included in the derivation.

Each patient's BOR will be categorized as complete response (CR), partial response (PR), stable disease (SD), progressive disease (PD), or not evaluable (NE). A BOR of CR or PR will require confirmation, but sensitivity analyses of response-based endpoints may be performed where confirmation of a BOR of CR or PR is not required. Objective response rate is the proportion of patients with a BOR of CR or PR; point estimates and CIs (using the normal approximation to the binomial) will be calculated by treatment group, based on the measurable disease population. Stratified tests comparing these rates between treatment groups will be conducted using a Cochran-Mantel-Haenszel test.

6.7.3.4. Duration of Response

The DoR time is defined only for responders (patients with a soft tissue BOR of CR or PR) in the measurable disease population. It is measured from the date of first evidence of soft tissue CR or PR to the earliest date of investigator determined radiographic disease progression (by objective radiographic disease assessment per RECIST 1.1 for soft tissue AND/OR radionuclide bone scan

using PCWG3 criteria for bone) or death from any cause, whichever is earlier. It is calculated as date of radiographic progression or death – date of first soft tissue response evaluation of CR or PR + 1. The DoR will be censored according to the same rules as rPFS.

A KM analysis of DoR will be performed to estimate the DoR curve for each treatment group. Point estimates and CIs for DoR quartiles and DoR rates will be calculated every 6 months for the first 18 months.

6.7.3.5. Time to Symptomatic Progression

Time to symptomatic progression is defined as the time from randomization to any of the following (whichever occurs earlier):

- Symptomatic Skeletal Event (SSE), defined as symptomatic fracture, surgery or radiation to bone, or spinal cord compression
- 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.

For patients not known to have had symptomatic progression at the time of data analysis, data will be censored on the last date at which no symptomatic progression is indicated.

A KM analysis will be performed to estimate the time to symptomatic progression survival curve, as well as quartiles and rates at 3-month intervals for each treatment group. In addition, the Cox regression stratified by the randomization factors will be used to estimate the HR between the 2 treatment groups, along with CI.

6.7.4. Sensitivity Analyses

Sensitivity analyses will be undertaken for calculation of the primary or key secondary endpoints in order to evaluate the robustness of the analysis. The following sensitivity analyses will be performed:

Radiographic Progression-Free Survival Sensitivity Analysis 1 (symptomatic progression as an rPFS event): if a patient is discontinued from study treatment due to investigator-determined symptomatic progression, then the patient's rPFS time will be calculated using the date of nonradiographic progression as the progression date. This endpoint is sometimes referred to as clinical progression-free survival.

Radiographic Progression-Free Survival Sensitivity Analysis 2 (forward-dating progressions at unscheduled assessments): if a patient has objective progression at an unscheduled disease assessment, then the PFS time for that patient will be forward-dated to the next scheduled disease assessment.

Radiographic Progression-Free Survival Sensitivity Analysis 3: rPFS will also be analyzed after adjusting for selected prognostic factors. Potential prognostic factors include the stratification factors and other factors as outlined in Section 6.12. The HR for treatment effect will be estimated using a multivariate Cox proportional hazard model constructed by selecting

variables among all the potential variables, using stepwise selection method with entry p-value .05 and exit p-value .1. The treatment factor will be kept out of the model throughout the covariate selection process and only added into the final model.

Radiographic Progression-Free Survival Sensitivity Analysis 4 (removing new anticancer treatment started as censoring condition): rPFS time will be omitting the "new systemic anticancer treatment started and no radiographic disease progression within 14 days" censoring condition.

Radiographic Progression-Free Survival Sensitivity Analysis 5 (symptomatic progression or PSA progression as an rPFS event): if a patient is discontinued from study treatment due to investigator-determined symptomatic progression, or if a patient experiences a confirmed PSA progression, then the patient's rPFS time will be calculated using the date of the earliest progression (radiographic or nonradiographic) as the progression date.

Radiographic Progression-Free Survival Sensitivity Analysis 6 (treatment discontinuation before radiographic progression as an rPFS event): if a patient is discontinued from study treatment before any radiographic progression is observed, the patient will be considered as an rPFS event with progression date of the date of treatment discontinuation.

Blinded Independent Central Review: as appropriate, in addition to the secondary endpoint of rPFS by BICR, imaging-based endpoints (including ORR and DoR) will also be assessed by a blinded independent review of centrally stored imaging. Radiographic progression-free survival, ORR, and DoR by investigator-assessment may then be compared to those by BICR.

Proportional Hazards Assumption: a formal evaluation of the proportional hazard assumption for rPFS and OS will be conducted. This will be done visually through inspection of the graph of log(-log[S(t)]) versus log(t) for the 2 treatment groups. In the event that the proportional hazards assumption is violated, additional analyses will be performed.

Analysis by Study Part: as appropriate, efficacy analyses and analyses of patient characteristics will also be performed by study part, and results compared across study part. Analyses may be done for each study part, or for patients enrolled into Part 1 or Part 2 vs Part 3.

Analysis by Initial Dose: as appropriate, analyses may be performed on only those patients who were randomized to abemaciclib at the RP2D, or to any dose of the placebo.

6.7.5. Exploratory Analyses

CCI [redacted]

6.8. Health Outcomes/Quality-of-Life Analyses

Patient-reported outcomes are measured through the following:

- single-item Worst Pain Numerical Rating Scale (NRS)
- CCI [redacted]

- CCI [redacted]

For each instrument, the compliance rate will be calculated by visit for each treatment group. Reasons for noncompliance will also be summarized.

Exploratory analyses may be performed to investigate associations between patient-reported data and clinical efficacy endpoints as appropriate.

If deemed necessary, further analysis of health outcomes, quality of life, and health utilization may be described in a separate SAP.

6.8.1. Patient-Reported Pain

Patient-reported pain intensity will be measured by the single-item Worst Pain NRS: mild pain is defined as scores 1 to 4, moderate pain as scores 5 to 6, and severe pain as a score ≥7 (Serlin et al. 1995). Analgesic use will be categorized according to the WHO analgesic ladder (WHO-AL). Descriptive statistics will be summarized by treatment group and cycle. A mixed-effects, repeated measures model may also be applied to explore differences between treatment groups by cycle with respect to change from baseline.

Time to worst pain progression is a secondary endpoint, and defined as the time from randomization to any of the following (whichever occurs earlier):

- For patients without opioid use at baseline (WHO-AL ≤2):
  - Worst pain progression (an increase of 2 points from baseline on the Worst Pain NRS item on 2 consecutive evaluations)
  - Initiation of weak or strong opioids (WHO-AL ≥3)
- For patients with weak or strong opioid use at baseline (WHO-AL ≥3):
  - Worst pain progression (an increase of 2 points from baseline on the Worst Pain NRS item on 2 consecutive evaluations) without concurrent decreased opioid use (a decrease in WHO-AL of 1 or more)
  - Increased opioid use (an increase in WHO-AL of 1 or more)

For patients not known to have had worst pain progression at the time of data analysis, data will be censored on the last date at which no worst pain progression is indicated.

A KM analysis will be performed to estimate the time to worst pain progression survival curve, as well as quartiles and rates at 3-month intervals for each treatment group. In addition, the Cox regression stratified by the randomization factors will be used to estimate the HR between the 2 treatment groups, along with CI.

Additionally, the components of time to worst pain progression—time to worst pain progression without consideration of analgesic use, time to initiation of opioids, and time to increased opioid use—will be analyzed using similar time-to-event analysis methods. Time to strong opioid use may also be analyzed using the same time-to-event analysis methods.

CCI [redacted]

CCI

6.8.2. Patient-Reported Quality-of-Life

CCI

Time to deterioration in HRQoL is a set of exploratory endpoints defined as the time from

CCI


CCI
CCI

CCI [redacted]

[redacted]

CCI [redacted]

6.8.3. Health Economics

Utilization data will be summarized by category across treatment groups, including:

- The use of all concomitant medications, particularly analgesics, bisphosphonates, and RANK-L targeted agents
- Blood product transfusions
- Radiation therapy
- Surgery
- Hospitalization days

6.9. Bioanalytical and Pharmacokinetic/Pharmacodynamic Methods

Pharmacokinetic and pharmacodynamic analyses will be performed according to a separate PK analysis plan.

6.10. Biomarker Analyses

Protocol Sections 9.8.1 and 9.8.2 describe the biomarker samples that will be collected during the course of the study and will be used to carry out biomarker specific exploratory analyses. In addition to characterizing the biomarker data distribution by treatment group and overall, these analyses may include the assessment of prognostic or predictive relationships between the biomarker(s) being investigated and the clinical efficacy endpoints of the study.

If deemed necessary, further analysis of biomarkers may be described in a separate biomarker SAP.

6.11. Safety Analyses

6.11.1. Extent of Exposure

Drug exposure, drug adjustments and omissions for abemaciclib/placebo, abiraterone acetate and prednisone will be summarized for all treated patients per treatment group. Drug exposure will include summaries of cycles received per patient, duration on therapy, and cumulative dose. The summary of dose adjustments and omissions will include the reason for adjustment or omission. Dose intensity will be summarized for all treated patients per treatment group for abemaciclib/placebo. Dose intensity will be calculated as the actual cumulative amount of drug taken divided by the duration of treatment. Relative dose intensity will be calculated as the actual amount of drug taken divided by the amount of drug assigned times 100% (that is, expressed as a percentage).

For abemaciclib/placebo, extent of exposure will be measured by tablet/capsule counts. Dose intensity will be expressed in mg/day. The assigned cumulative dose for abemaciclib/placebo while on study is

- 150 mg per dose × 2 doses per day × number of days on study drug, if enrolled in Part 1 and randomized to Arm A1 or B1;
- 200 mg per dose × 2 doses per day × number of days on study drug, if enrolled in Part 1 and randomized to Arm A2 or B2;
- RP2D per dose × 2 doses per day × number of days on study drug, if enrolled in Part 2 or Part 3.

For abiraterone acetate and prednisone, extent of exposure will be collected using the 'Exposure Compliance: Study Treatment Abiraterone Acetate' and 'Exposure Compliance: Study Treatment Prednisone' forms, respectively. Since abiraterone acetate and prednisone will not be centrally supplied by Lilly, the amount of dose will only be collected and calculated as number of doses, without considering detailed dose level (mg per dose).

6.11.2. Adverse Events

Adverse event (AE) verbatim text will be mapped by the sponsor or designee to corresponding terminology within Medical Dictionary for Regulatory Activities (MedDRA). Severity grades will be assigned by the investigator using Common Terminology Criteria for Adverse Events (CTCAE) Version 5. MedDRA preferred terms (PTs) identified by Medical as clinically identical or synonymous will be grouped together under a single consolidated PT. For example, 'Neutropenia' and 'Neutrophil count decreased' will be reported as 'Neutropenia'. Adverse events will be reported according to PT resulting from this process.

Pre-existing conditions are defined as AEs that begin prior to the first dose of study drug.

A TEAE is defined as any AE that begins between the day of first dose and 30 days after treatment discontinuation (or up to any time if serious and related to study treatment), or any pre-existing condition that increases in CTCAE grade between the day of first dose and 30 days after treatment discontinuation (or up to any time if serious and related to study treatment).

The MedDRA lower level term (LLT) will be used for comparisons of pre-existing conditions to on-treatment events in the treatment-emergent computation.

A serious adverse event (SAE) is any AE during this study that results in one of the following outcomes:

- death
- initial or prolonged inpatient hospitalization
- a life-threatening experience (that is, immediate risk of dying)
- persistent or significant disability/incapacity
- congenital anomaly/birth defect
- considered significant by the investigator for any other reason.

Important medical events that may not result in death, be life-threatening, or require hospitalization may be considered SAEs when, based on appropriate medical judgment, they may jeopardize the patient and may require medical or surgical intervention to prevent one of the outcomes listed in this definition.

The following AE listings and summaries will be produced:

- Overview of TEAEs
- Summary of TEAEs by PT (all grade and grade ≥3)
- Summary of TEAEs by SOC and PT (all grade and grade ≥3)
- Summary of TEAEs by PT and maximum grade (1-5)
- List of SAEs
- Summary of SAEs by SOC and PT (all grade and grade ≥3)
- Summary of AEs as reason for study treatment discontinuation by PT

The TEAE and SAE summaries will be produced for all TEAEs/SAEs and repeated for TEAEs/SAEs related to any study treatment, where relationship of the AE to the study treatment will be assessed by the investigator (yes or no).

Additionally, a listing of all COVID-19 AEs or deaths will be produced.

6.11.3. Deaths

All deaths on study not attributed to study disease by the investigator will be listed along with the reason for death, if known. For those deaths attributed to an AE, the listing will include the PT of the AE. A summary of deaths including reasons for death will be produced.

6.11.4. Clinical Laboratory Evaluation

All relevant hematology and chemistry laboratory values will be graded according to CTCAE Version 5. These calculated grades will be summarized by cycle and maximum post-baseline grade over the entire study for each treatment group. Treatment-emergent changes will be summarized by the maximum post-baseline grade, and a shift table of baseline grade by maximum post-baseline grade will be produced.

6.11.5. Vital Signs and Other Physical Findings

Temperature, blood pressure, pulse rate, respiration rate, weight, and ECOG PS will be summarized by cycle.

6.11.6. Electrocardiograms

Local electrocardiograms (ECGs) will be summarized by cycle and overall. The summary by cycle will classify patients as having normal or abnormal ECG and summarize AEs identified by ECG within each cycle. The overall summary will classify patients as having an abnormal ECG at any point and summarize all AEs identified by ECG.

6.12. Subgroup Analyses

Subgroup analyses of rPFS and OS will be performed for each of the following potential prognostic subgroup variables:

- Radiographic progression at study entry (yes/no)
- Measurable disease at baseline (yes/no)
- Prior docetaxel for mHSPC (yes/no)
- Gleason score (<8 vs ≥8)
- Metastasis stage at initial diagnosis (M0, M1)
- Extent of disease at baseline (node only vs bone [with or without node] vs visceral [with or without node and with or without bone])
- Age (<70, ≥70 years)
- Race (White, Asian, Black, Other)
- Baseline ECOG PS (0 versus 1)
- Baseline PSA (above or below median)

If a level of a factor consists of fewer than 10% of randomized patients, analysis within that level may be omitted.

Analyses will be done within subgroup and separately, across subgroups with a test of interactions of subgroups with treatment performed. Estimated HRs and CIs for the within subgroup analyses will be presented as a Forest plot along with p-values for tests of interactions between subgroup variables and treatment.

If data warrants, other exploratory subgroup analyses may be performed, including subgroups by ethnicity, region of enrollment, PSA at initial diagnosis, baseline worst pain NRS score (≤3 vs >3), baseline bone disease burden (no bone involvement, 1 to 4 bone lesions, 5 to 9 bone lesions, >9 lesions or superscan), and baseline superscan for bone disease (yes vs no). If any safety analyses identify important imbalances between treatment groups, subgroup analyses of these endpoints may be performed.

6.13. Protocol Deviations

Significant protocol deviations that potentially compromise the data integrity and patients' safety will be summarized by treatment group for all randomized patients. These violations will include

deviations that can be identified programmatically and those that can only be identified by the clinical research associate during monitoring. Significant protocol deviations are described in another document within the study Trial Master File.

As deemed appropriate, sensitivity analyses may be performed to assess the impact of significant protocol deviations on key endpoints. For example, rPFS may be analyzed excluding those patients who did not meet enrollment criteria and were inadvertently enrolled.

Additionally, listings of all protocol deviations and all important protocol deviations related to COVID-19 will be produced.

6.14. Interim Analyses and Data Monitoring

There are planned interim analyses for safety, one planned interim analysis for futility, and up to 2 planned adaptive interim analyses. The first safety interim analysis described below will be reviewed by an AC made up of Lilly members not involved in the day-to-day study conduct to identify the RP2D using only safety and PK data. All other interim analyses will be reviewed by an independent DMC made up of external members, none of whom are involved as study investigators. Unblinding details are specified in a separate unblinding plan document. The DMC Charter contains further details on the activities and responsibilities of the DMC.

6.14.1. Safety Interim Analyses

The first safety interim analysis is planned at CCI [redacted] The AC members will review safety and available PK data at the interim analysis to determine the dose of abemaciclib (RP2D) in combination with abiraterone acetate plus prednisone for Part 2 and Part 3. Specifically, the safety evaluation will be based, at least, on the following data reports:

- summary of treatment discontinuations and reasons for discontinuation
- listing of treatment discontinuations due to AE or death
- summary of SAEs
- summary of TEAEs
- summary of CTCAE-graded central laboratory results
- summary of drug adjustments and omissions, including reasons for adjustment/omission
- Lilly Safety System reports for all patients with SAEs (as requested).

In the case of unacceptable and/or unmanageable toxicity of the combination, the AC may recommend termination of the study upon completion of Part 1. This analysis will only use PK and safety data and will be the only unblinded analysis the AC will review. After dose selection, only the DMC will be authorized to review unblinded data.

Additionally, the DMC will review safety as described in the DMC charter, no less than approximately every 6 months. At each applicable interim analysis, the DMC may recommend the trial continue without modifications, continue with specific modifications, or be stopped for safety concerns. There will be no prespecified rules for stopping the trial due to safety concerns. The DMC members will review unblinded safety data at each applicable interim analysis. If a

significant safety signal is identified, the DMC may recommend a protocol amendment, termination of enrollment, and/or termination of study treatment. The recommendations of the DMC will be communicated to the Lilly Senior Management Designee and, if necessary, an Internal Review Committee.

In the event that safety monitoring uncovers an issue that needs to be addressed by unblinding at the treatment group level, members of the DMC can conduct additional analyses of the safety data. Additionally, unblinding of a limited number of Lilly representatives external to the study team may be required for evaluation of selected SAEs for determination of regulatory reporting.

6.14.2. Futility/Adaptive Interim Analyses

The interim analysis for futility is planned after approximately CCI rPFS events have been observed. Futility for the interim analysis will be determined in terms of rPFS. The DMC should recommend stopping the trial for futility if the observed hazard ratio is CCI. The DMC will consider all evidence, including safety and other efficacy parameters, in making this decision.

The planned adaptive interim analyses are based on rPFS, and will occur at approximately CCI rPFS, and if necessary, at CCI rPFS events. The adaptive interim analyses will be conducted to determine whether Part 3 of the study should be opened for enrollment. The expansion decision criteria will be based on the conditional probability that the final rPFS analysis, given Part 3 will be opened for enrollment, will be positive. See Appendix 1 for statistical details on the definition and derivation of conditional probability of success. Part 3 will only be opened for enrollment if that conditional probability is at least 70% at either of the adaptive interim analyses. As an illustration, if the adaptive interim analyses were to occur at exactly CCI rPFS, then the criteria to expand would be if the observed rPFS hazard ratios were no more than 0.666 and 0.686, respectively. Appendix 2 reports on simulations that demonstrate the operating characteristics of this design, including control of type I error. Furthermore, Appendix 3 contains a proof that shows this design does not inflate type I error. Should Part 3 be opened for enrollment after the first adaptive interim analysis at CCI rPFS events, a second adaptive interim analysis will not be performed.

The sponsor has no intent to stop the study based on the adaptive interim analyses and all patients will continue follow-up for all study objectives until study close. In addition, patients will remain blinded for the duration of the study unless the criteria in Section 7.3 of the protocol are met. Only the DMC is authorized to evaluate unblinded adaptive interim and futility analyses. Study sites will receive information about interim results only if they need to know for the safety of their patients.

6.15. Annual Report Analyses

Annual report analyses, including Developmental Safety Update Report and Investigator's Brochure analyses, are described in the LY2835219 Program SAP.

6.16. Clinical Trial Registry Analyses

Additional analyses will be performed for the purpose of fulfilling the Clinical Trial Registry (CTR) requirements.

Analyses provided for the CTR requirements include the following:

Summary of AEs provided as a dataset which will be converted to an XML file. Both SAEs and 'Other' AEs will be summarized by treatment group and MedDRA PT.

- An AE is considered 'Serious' whether or not it is a TEAE.
- An AE is considered in the 'Other' category if it is both a TEAE and is not serious. For each SAE and 'Other' AE, for each term and treatment group, the following are provided:
  - the number of participants at risk of an event
  - the number of participants who experienced each event term
  - the number of events experienced.
- Consistent with www.ClinicalTrials.gov requirements, 'Other' AEs that occur in fewer than 5% of patients/subjects in every treatment group may not be included if a 5% threshold is chosen (5% is the minimum threshold).
- Adverse event reporting is consistent with other document disclosures for example, the clinical study report (CSR), manuscripts, and so forth.

In addition, a participant flow will be created that will describe how many enrolled patients completed the study, and for those who did not, the frequency of each reason for not completing. This analysis will be based on study discontinuation, not treatment discontinuation. A patient will be identified as having completed the study if the patient dies while on the study or the patient had discontinued study treatment and is in follow up at the time of the final OS analysis. Patients who withdraw consent before the final OS analysis or who are still on treatment at the time of the final OS analysis will be identified as not completing the study.

7. Unblinding Plan

Randomization will occur using an IWRS system. Assignment to treatment groups will be determined by a computer-generated random sequence. Security measures will be taken so that treatment group code that can link patients to study arm will be blinded in the database.

In order to maintain the scientific integrity of this double-blind trial, access to study data will be strictly controlled prior to the interim and final analyses of OS. Dummy treatment assignment will be used in the reporting database until the database lock for the final analysis of overall survival. During this time, analyses using unblinded treatment codes will be performed only at the interim analysis points specified in the protocol/SAP. For those safety and efficacy analyses assigned to the AC or DMC, only the designated Statistical Analysis Center (SAC) will perform analyses on unblinded data. For any PK analysis to occur prior to the final rPFS analysis, the list of individuals that will have access to unblinded data will be provided with the PK/pharmacodynamic analysis plan, and documentation concerning their access to the data will be retained.

Data sets will be created for the purpose of aggregate data review by the sponsor in which treatment assignment is scrambled so that personnel involved in the day-to-day conduct of the trial and development and validation of analysis programs will be blinded to patient treatment assignment.

While every effort will be made to blind both the patient and the investigator to the identity of the treatment, the inadvertent unblinding of a patient may occur. This unblinding will not be sufficient cause (in and of itself) for that patient to be excluded from any safety or efficacy analyses.

8. References

Amit O, Mannino F, Stone AM, et al. Blinded independent central review of progression in cancer clinical trials: results from a meta-analysis. *Eur J Cancer*. 2011;47(12) :1772-1778. https://doi.org/10.1016/j.ejca.2011.02.013

CCI

Chen C, Anderson K, Mehrotra DV, et al A 2-in-1 adaptive phase 2/3 design for expedited oncology drug development. *Contemp Clin Trials*.2018;64:238-242. https://doi.org/10.1016/j.cct.2017.09.006

CCI

CCI

Open. 2016;6(12):e013096. https://doi.org/10.1136/bmjopen-2016-013096

Cox DR. Regression models and life-tables. *J Royal Stat Soc Ser B*. 1972;34(2):187-220.

CCI

Glimm E, Maurer W, Bretz F. Hierarchical testing of multiple endpoints in group-sequential trials. *Stat Med*. 2010;29(2):219-228. https://doi.org/10.1002/sim.3748

Kaplan EL, Meier P. Nonparametric estimation of incomplete observations. *J Amer Stat Assoc*. 1958;53:457-481. http://www.jstor.org/stable/2281868 .

Serlin RC, Mendoza TR, Nakamura Y et al. When is cancer pain mild, moderate or severe? Grading pain severity by its interference with function. *Pain*. 1995;61(2):277-284. https://doi.org/10.1016/0304-3959(94)00178-H

CCI

9. Appendices

Appendix 1. Definition of Conditional Probability of Success

As described in Section 6.14.2, the expansion decision criteria at each adaptive interim analysis will be based on the conditional probability that the final radiographic progression-free survival analysis, given Part 3 will be opened for enrollment, will be positive. CCI

Appendix 2. Simulation Details and Results

In this appendix, we simulate trial outcomes under varying underlying treatment effects to assess the results operating characteristics, with a particular focus on the impact of the adaptive interim analyses on the overall type I error rate. Assume that survival times follow exponential distributions (constant hazards). Without loss of generality, consider unstratified analyses, and assume adaptive interim analysis 1 occurs exactly at CCI radiographic progression-free survival (rPFS) events, adaptive interim analysis 2 (if applicable) occurs exactly at CCI rPFS events, and the primary analysis occurs after exactly CCI events, given that Part 3 is not opened for enrollment or is opened for enrollment, respectively. Furthermore, assume patients are randomized to treatment arms in the exact ratios specified in the protocol, meaning of the 30 patients enrolled to Part 1, 20 patients are randomized to the abemaciclib arms and 10 to the placebo arms, and so forth. Finally, we assume the median survival time in the placebo arms is 16 months, and we evaluate trial operating characteristics for median survival times in the abemaciclib arms ranging from 16 months to 32 to months. For N=10,000 simulated trials, the following steps are performed:

CCI


CCI


CCI


CCI

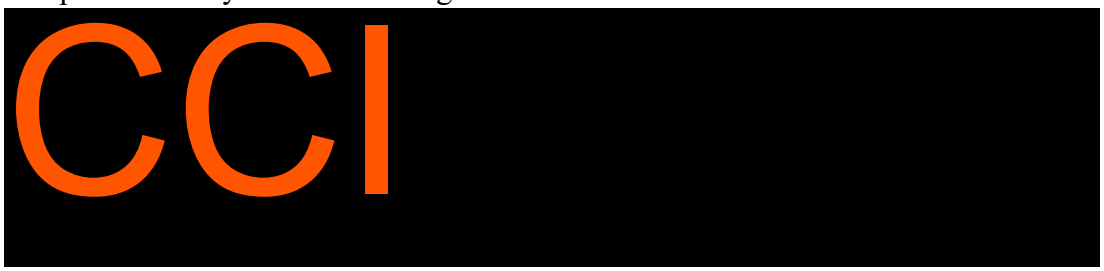
Appendix 3. Proof of Type I Error Control

Extending the work of Chen (2018), we show that under general conditions, type I error is not inflated given 2 opportunities to expand with 1-sided decision rules, assuming the same endpoint is used at each analysis. CCI [redacted]

Appendix 4. Notable Patient Criteria

Prior to the approval of Version 4 of this SAP, the notable patient criteria were documented in the Patient Narrative Planning Tool. The following notable criteria supersede the ones in the Patient Narrative Planning Tool:

- Discontinued trial treatment (at least 1 of the study drugs) due to an AE, serious or nonserious
- Died while on trial treatment or within 30 days after the date of trial treatment discontinuation, regardless of whether the death was due to trial disease
- Experienced a SAE not described in Table 7.1 of the investigator's brochure (Serious Adverse Reactions for Abemaciclib Considered Expected for SUSAR Reporting Purposes, as a Single Agent)
- Experienced any of the following:



Narratives will be provided for patients in the safety population with at least 1 notable event.

Signature Page for VV-CLIN-081628 v1.0

Approval	PPD Statistician 25-Oct-2023 09:18:39 GMT+0000

Signature Page for VV-CLIN-081628 v1.0