Statistical Analysis Plan 13Y-MC-JPEF Version 1

postMONARCH: A Randomized, Double Blind, Placebo-Controlled, Phase 3 Study to Compare the Efficacy of Abemaciclib Plus Fulvestrant to Placebo Plus Fulvestrant in Participants With HR+, HER2-, Advanced or Metastatic Breast Cancer Following Progression on a CDK4 & 6 Inhibitor and Endocrine Therapy

NCT05169567

Approval Date: 22-Sep-2022

Title Page

Protocol Title: postMONARCH: A Randomized, Double-Blind, Placebo-Controlled, Phase 3 Study to Compare the Efficacy of Abemaciclib plus Fulvestrant to Placebo plus Fulvestrant in Participants with HR+, HER2-, Advanced or Metastatic Breast Cancer Following Progression on a CDK4 & 6 Inhibitor and Endocrine Therapy

Protocol Number: I3Y-MC-JPEF

Compound Number: Abemaciclib (LY2835219)

Short Title: Abemaciclib plus Fulvestrant compared to Placebo plus Fulvestrant in HR+, HER2-, Advanced or Metastatic Breast Cancer previously treated with a CDK4/6 Inhibitor and Endocrine Therapy

Acronym: postMONARCH

Sponsor Name: Eli Lilly and Company

Legal Registered Address: Eli Lilly and Company, Indianapolis, Indiana, USA 46285

Regulatory Agency Identifier Number(s)

IND: 106100

Document ID: VV-CLIN-062920

Approved on 22 Sep 2022 GMT

Table of Contents

Version history

This statistical analysis plan for Study I3Y-MC-JPEF is based on the protocol amendment (c) dated 17 February 2022. Version 1 was approved prior to planned unblinding.

Table 1 SAP Version History Summary

SAP Version	Approval Date	Change	Rationale
1	See date on Page 1	Not Applicable	Original version

Abbreviations and Definitions

Term	Definition
AE	adverse event
BID	twice daily
BIRC	blinded independent review committee
CBR	clinical benefit rate
CDK	cyclin-dependent kinase
CFS	chemotherapy-free survival
CI	confidence interval
CR	complete response
CSR	clinical study report
CTCAE	Common Terminology Criteria for Adverse Events
CTR	Clinical Trial Registry
DCR	disease control rate
DMC	Data Monitoring Committee
DoR	duration of response
ECOG	Eastern Cooperative Oncology Group
eCRF	electronic case report form
EDR	early discrepancy rate
EORTC IL-19	European Organization for Research and Treatment of Cancer Item Library
CCI [redacted]	[redacted]
CCI [redacted]	[redacted]
HER2-	human epidermal growth factor receptor 2 negative
HR+	hormone receptor positive
HRQoL	health-related quality of life
IES	intercurrent-event strategies
IM	intramuscularly

Approved on 22 Sep 2022 GMT

Term	Definition
IRC	internal review committee
ITT	intent-to-treat
LDR	late discrepancy rate
mBPI-SF	modified Brief Pain Inventory-short form
MedDRA	Medical Dictionary for Regulatory Activities
NCI-CTCAE	National Cancer Institute - Common Terminology Criteria for Adverse Events
NRS	Numeric Rating Scale
ORR	objective response rate
OS	overall survival
PF	performance status
PFS	progression-free survival
PK	pharmacokinetics
PR	partial response
PRO	patient-reported outcomes
PRO-CTCAE	Patient-Reported Outcomes version of the Common Terminology Criteria for Adverse Events
PT	Preferred Term
RECIST	Response Evaluation Criteria in Solid Tumors
SAP	statistical analysis plan
SD	stable disease
SMD	senior management designee
SOC	System Organ Class
SRE	skeletal-related event
TEAE	treatment-emergent adverse event
TTC	time to chemotherapy
WHO	World Health Organization

Approved on 22 Sep 2022 GMT

1. Introduction

The SAP is an extension of the protocol that contains additional details about the analysis plan for efficacy, safety, PK, PROs, and exploratory endpoints.

There are no changes to the analyses described in the protocol.

1.1. Objectives, Endpoints, and Estimands

Objectives	Endpoints
Primary	
To compare the efficacy of fulvestrant with or without abemaciclib	PFS as determined by investigator assessment using RECIST 1.1
Secondary	
To further compare the efficacy of fulvestrant with or without abemaciclib	• OS • PFS by BICR • ORR • CBR • DCR • DoR
To further characterize the safety profile of abemaciclib in combination with fulvestrant	Safety – including but not limited to TEAEs, SAEs, deaths, and clinical laboratory abnormalities
To compare PRO measures of fulvestrant with or without abemaciclib	• Time to worsening in worst pain via the mBPI-SF worst pain item • Time to deterioration in physical function via the EORTC IL-19
To characterize the PK of abemaciclib in combination with fulvestrant	Concentrations of abemaciclib
Exploratory	
To assess exploratory clinical parameters of fulvestrant with and without abemaciclib	CCI

Objectives	Endpoints
To explore other PRO and HRQoL parameters of fulvestrant with and without abemaciclib	CCI

Abbreviations: AE = adverse event; BICR = blinded independent central review; CBR = clinical benefit rate; CCI CTCAE = Common Terminology Criteria for Adverse Events; DCR = disease control rate; DoR = duration of response; CCI EORTC IL-19 = European Organization for Research and Treatment of Cancer Item Library 19; CCI ; mBPI-SF = modified Brief Pain Inventory-short form; ORR = objective response rate; OS = overall survival; PRO = patient-reported outcome; PFS = progression-free survival; CCI CCI ; RECIST = Response Evaluation Criteria in Solid Tumors; SAE = serious adverse event; CCI

Note: This information was copied from the protocol. Additional details may be added; this especially applies to those protocols with minimal information regarding estimands. At a minimum, the content reuse objectives table should be used as is. The following instructions describe information regarding how estimands should be used.

Primary estimand

The primary research question is: What is the difference in PFS time between Arm A (abemaciclib plus fulvestrant) versus Arm B (placebo plus fulvestrant) following progression/relapse on prior cyclin-dependent kinases 4 and 6 (CDK4 & 6) inhibitor-based therapy in participants with advanced/metastatic HR+, HER2- breast cancer.

The estimand for the primary objective is described by the following attributes:

- Population: adult participants with advanced/metastatic HR+, HER2- breast cancer after progression/relapse on prior treatment with CDK4 & 6 inhibitor-based therapy, randomized to study intervention (primary analysis population).
- Endpoint: investigator-assessed PFS in the primary analysis endpoint, which is defined as the time from randomization until
    - first occurrence of documented disease progression per RECIST 1.1, or
    - death from any cause in the absence of documented progressive disease
- Treatment condition: the randomized study intervention (Arms A and B) will be administered until disease progression, unacceptable toxicity, or another protocol-defined reason for study intervention discontinuation.
- CCI
    - CCI

- CCI
  - CCI

CCI

1.2. Study Design

postMONARCH is a Phase 3, global, multicenter, randomized, double-blind, placebo-controlled study in participants with HR+, HER2- advanced or metastatic breast cancer. This study will enroll adults who experienced relapse or disease progression on/after a CDK4 & 6 inhibitor and an aromatase inhibitor in the first-line setting or on/after CDK4 & 6 inhibitor with endocrine therapy in the adjuvant setting.

Approximately 350 participants will be equally randomized between 2 treatment arms and will be treated until disease progression or other discontinuation criteria are met.

- **Arm A:** abemaciclib plus fulvestrant
  - abemaciclib 150 mg orally BID on Days 1 through 28 of each cycle
  - fulvestrant 500 mg IM on Cycle 1 Day 1 and Cycle 1 Day 15, then on Day 1 of Cycle 2 and beyond
- **Arm B:** placebo plus fulvestrant
  - placebo BID on Days 1 through 28 of each cycle
  - fulvestrant 500 mg IM on Cycle 1 Day 1 and Cycle 1 Day 15, then on Day 1 of Cycle 2 and beyond
- Participants will be randomized using the following stratification factors:
  - geography CCI

- presence of visceral metastases (Yes versus No)
- duration on prior adjuvant/metastatic CDK4 & 6 inhibitor-based regimen

CCI

2. Statistical Hypotheses

Primary Hypothesis (Arm A versus Arm B): **Multiplicity Adjustment**

A gated hypothesis testing procedure will be used to ensure control of the familywise error rate at 0.025 (one-sided) across the 2 endpoints (PFS and OS). The primary endpoint of PFS will first be tested; the secondary OS will be tested only if statistical significance is achieved for PFS. Other endpoints will not be error controlled.

3. Analysis Sets

The populations for analysis are defined as follows.

Population	Description
Entered	All participants who sign the ICF
ITT	All participants randomly assigned to study treatment, regardless of whether they take any doses of study treatment, or if they take the correct treatment. Participants will be analyzed according to the treatment group to which they were assigned
Safety	All participants randomly assigned to study treatment and who take at least 1 dose of study treatment. Participants will be analyzed according to the study treatment they actually received
PK	All participants who have received at least 1 dose of abemaciclib, have at least 1 evaluable PK sample, and have sufficient dosing information

Abbreviations: ICF = informed consent form; ITT = intention to treat; PK = pharmacokinetics.

4. Statistical Analyses

4.1. General Considerations

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

Continuous variables will be summarized using descriptive statistics (ie, number of patients, mean, median, standard deviation, minimum, and maximum). Categorical variables will be summarized by frequency and its corresponding percentage.

All tests of treatment effects will be conducted at a 1-sided alpha level of 0.025, unless otherwise stated, and all CIs will be given at a 2-sided 95% level.

Any change to the data analysis methods described in the protocol will require a protocol amendment only if it changes a principal feature of the protocol. Any other change to the data analysis methods described in the protocol, and the justification for making the change, will be described in the SAP and the CSR. Additional exploratory analyses of the data will be conducted as deemed appropriate.

4.1.1. Definitions and Conventions

Study drug refers to abemaciclib or placebo.

Study treatment refers to abemaciclib plus fulvestrant or placebo plus fulvestrant.

The **date of randomization** is the date the patient was randomly assigned to study treatment using the interactive web response system.

The **date of first dose** is the date of the first dose of study drug or fulvestrant.

The **baseline value of a safety assessment** is the last value observed prior to the first dose of study drug or fulvestrant.

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 08 March 2016 and the date of first dose was 06 March 2016, 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 05 March 2016 and the date of first dose was 06 March 2016, 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.

Approved on 22 Sep 2022 GMT

- 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.

4.1.2. Handling of Dropouts or Missing Data

With the exception of dates, missing data will not be imputed. The method of imputation for any dates that are imputed is described in the relevant section.

4.2. Participant Dispositions

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, rescreened after screen failure, randomized in the study, and 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 predetermined categories.

4.3. Primary Endpoint Analysis

4.3.1. Definition of Endpoint

The primary endpoint is investigator-assessed PFS in the ITT population. Progression-free survival is defined as the time from randomization until the first occurrence of documented disease progression per RECIST version 1.1 criteria (Eisenhauer et al. 2009), or death from any cause in the absence of progressive disease. Participants known to be alive and without disease progression will be censored according to the censoring scheme detailed in Section 4.3.2.

The final analysis of PFS will be conducted after approximately 251 investigator-assessed events have been observed in the ITT population. One interim analysis is planned for the primary endpoint of PFS, after approximately 176 events have been observed. See further details in Section 4.9.

4.3.2. Main analytical approach

The Kaplan-Meier method (Kaplan and Meier 1958) will be used to estimate the PFS curves. Median PFS and PFS rates at various time points with 95% CIs will be estimated for each arm. The comparison of PFS curves between treatment arms will be conducted by a stratified log-rank test as the primary analysis, stratified by the randomization strata. The treatment effect will be estimated by hazard ratio with its corresponding 95% CIs using the stratified Cox proportional hazard model (Cox 1972) with treatment as the only covariate, stratified by the randomization strata. A detailed PFS event/censoring scheme is provided in Table 2 below.

Table 2 PFS Censoring Scheme

Situation	Event/Censor	Date of Event or Censor
Tumor progression or death	Event	Earliest date of PD or death
No tumor progression and no death	Censored	Date of last adequate tumor assessment, per RECIST 1.1 criteria, or date of randomization (whichever is later)
	Unless	
No baseline radiologic tumor assessment available	Censored	Date of randomization
No adequate postbaseline tumor assessment available and death reported after 2-scan intervals following randomization	Censored	Date of randomization
New systemic anticancer therapy prior to tumor progression or death	Censored	Date of last adequate tumor assessment, per RECIST 1.1 criteria, prior to start of new therapy or date of randomization (whichever is later)
Tumor progression or death documented immediately after 2 or more missing scan intervals following last adequate tumor assessment or randomization (whichever is later)	Censored	Date of last adequate tumor assessment prior to 2 or more missing scans, per RECIST 1.1 criteria, or date of randomization (whichever is later)

Abbreviations: CR = complete response; PD = progressive disease; RECIST 1.1 = Response Evaluation Criteria in Solid Tumors, version 1.1; SD = stable disease.

Note: Symptomatic deterioration (ie, symptomatic progression that is not radiologically confirmed per RECIST 1.1 criteria) will not be considered as tumor progression.

Note: Adequate tumor assessment per RECIST 1.1 criteria refers to an assessment with 1 of the following responses: CR, PR, SD, or PD.

Note: The 2-scan interval is counted from the date of last adequate tumor assessment to the date of next 2 scheduled tumor assessments plus 8 days (adjusted by tumor assessment window).

Note: If there are multiple dates associated with 1 assessment, the assessment date will be set to the first date when the overall response is PD and the last date otherwise.

4.3.3. Sensitivity Analyses

Sensitivity analyses will be undertaken for calculation of the primary endpoint in order to evaluate the robustness of the analysis. The following sensitivity analyses will be performed for PFS:

- Using different rules for censoring (details provided in Table 3 below)
- Using an unstratified log-rank test and unstratified Cox model
- Using stratification factors based on the eCRF data if available
- Using a multivariate Cox regression model constructed by selecting variables among all the potential variables such as the variables used in the subgroup analyses, using stepwise selection method, with an entry p-value of 0.05 and an exit p-value of 0.1. The treatment factor will be kept out of the model throughout the covariate selection process and only added to the final model

Table 3 PFS Censoring Scheme for Sensitivity Analyses

Definition	Situation	Event/ Censor	Date of Event or Censor
SA1: Ignoring absence of adequate postbaseline tumor assessment	No adequate postbaseline tumor assessment available and death reported after 2-scan intervals following randomization	Event	Death
SA2: Ignoring missing tumor assessments	PD or death documented after 2 or more missing scan intervals following last adequate tumor assessment or randomization (whichever is later)	Event	Earliest date of PD or death
SA3: Symptomatic progression as a PFS event	Symptomatic progression that is not radiologically confirmed per RECIST 1.1 criteria	Event	Date of symptomatic progression

Abbreviations: PD = progressive disease; PFS = progression-free survival; RECIST 1.1 = Response Evaluation Criteria in Solid Tumors, version 1.1.; SA = sensitivity analysis.

Other sensitivity analyses for PFS may be conducted if deemed appropriate.

4.4. Secondary Endpoints Analysis

4.4.1. Key/Confirmatory Secondary Endpoint: Overall Survival

4.4.1.1. Definition of endpoint

Overall survival is defined as the time from randomization until death from any cause. If the participant is alive or lost to follow-up at the time of analysis, OS data will be censored on the last date the participant is known to be alive.

4.4.1.2. CCI


CCI

4.4.1.3. Sensitivity Analyses

One sensitivity analysis of OS is planned. Overall survival time for this analysis is defined as the time from the date of study enrollment to the date of death due to study disease. Survival time will be censored on the date the patient was last known to be alive for patients who have no reported event. For patients that have died due to reasons not related to the disease, survival time will be censored at the date of death.

4.4.2. Supportive Secondary Endpoints

Objective response rate is defined as the proportion of participants who achieve a best overall response of CR or PR. Confirmation of CR and PR is not required. The ORR will be reported for both the ITT population and the subset of patients with measurable disease. The ORR with 95% CIs will be summarized for each treatment arm and compared between treatment arms using the Cochran-Mantel-Haenszel test adjusting for the randomization strata.

Clinical benefit rate is defined as the number of participants who achieve a best overall response of CR, PR, or SD ≥6 months. Confirmation of CR and PR is not required. The CBR will be reported for both the ITT population and the subset of patients with measurable disease. The CBR with 95% CIs will be summarized for each treatment arm and compared between treatment arms using the Cochran-Mantel-Haenszel test adjusting for the randomization strata.

Disease control rate is defined as the number of participants who achieve a best overall response of CR, PR, or SD. Confirmation of CR and PR is not required. The DCR will be reported for both the ITT population and the subset of patients with measurable disease. The DCR with 95% CIs will be summarized for each treatment arm and compared between treatment arms using the Cochran-Mantel-Haenszel test adjusting for the randomization strata.

Duration of response is defined as the time from the date measurement criteria for CR or PR (whichever is first recorded) until the first date that disease is recurrent or objective progression is observed, per RECIST 1.1 criteria, or the date of death from any cause in the absence of objectively determined disease progression or recurrence. The DoR will be censored according to the same scheme as the main scheme for PFS. Median DoR with 95% CIs and curves for each treatment arm will be estimated using the Kaplan-Meier method. The analysis of DoR will be based on the participants who achieve an objective response (CR or PR) in the ITT population.

Progression-free survival by BIRC is defined the same way as the primary endpoint of PFS. For BIRC analysis, scans will be collected and reviewed for all randomized participants based on RECIST version 1.1. BIRC-assessed PFS intends to evaluate the reliability of the treatment effect based on the investigator-assessed PFS. BIRC-assessed PFS will be analyzed using the same methods as the investigator-assessed PFS. BIRC-assessed PFS is not intended to provide an alternative means of definitive analysis, but it may be useful to evaluate bias in local assessments. Discordance rates (ie, differences in assessment of progression between investigator and blinded independent central review) will be summarized for each arm (Amit et al. 2011). The agreement between BICR and investigator within a treatment arm is represented in a tabular form (Table 5). Specifically, differential discordance will be described using early discrepancy rate and late discrepancy rate differences.

Table 5 BIRC-assessed Versus Investigator-assessed Disease Progression

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

Abbreviations: a1 = number of agreements on timing and occurrence of PD; a2 = number of times investigators declare PD later than BIRC; a3: number of times investigators declare PD earlier than BIRC; BIRC = blinded independent review committee; PD = progressive disease.

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

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

The LDR quantifies the frequency with which the investigator assessment declares progression later than BIRC within each arm and is defined as:

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

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

The ORR, CBR, DCR, and DoR by BIRC will also be summarized for each arm.

Pharmacokinetic and Exposure-Response Analyses

Pharmacokinetic analyses will be conducted on all patients who have received

- at least 1 dose of abemaciclib
- at least 1 evaluable PK sample, and
- sufficient dosing information.

Observed concentration data for each analyte will be graphically assessed and may also be summarized by time and/or dose.

Population PK modeling approaches may also be used to compute mean PK parameters (for example, clearance, exposure, volume of distribution) and interindividual PK variability.

A separate Population PK and Exposure-Response Analysis Plan will describe the planned PK and exposure-response analyses.

4.5. Exploratory Endpoints Analysis

4.6. Patient-Reported Outcomes Analyses

The PROs will be used to compare changes in cancer-related symptoms, physical function, CCI outcomes between treatment arms and CCI

CCI The analyses for PROs will be conducted in the ITT population. The following PRO instruments are utilized:

- "Worst pain single item" NRS
- CCI
- EORTC IL-19: Physical Function
- CCI
- mBPI-SF
- CCI

Time to sustained worsening of worst pain (based on the Worst Pain NRS) is defined as the time from randomization to the first increase (≥2 points) in the weekly average of the worst pain score with confirmation in the next consecutive week. Participants not known to have sustained worsening will be censored at the last documented assessment. Time to sustained worsening of worst pain will be summarized for each arm by the Kaplan-Meier method and will be compared between the arms using the stratified log-rank test. Additional cutoff values of increase (eg, ≥3 points, ≥4 points) may be explored as sensitivity analyses if deemed appropriate. Sensitivity analysis will be done by including death as part of the event definition. Additional sensitivity analyses for time to worsening without confirmation in the next consecutive week may be performed if deemed appropriate. Analgesic use will be summarized descriptively by cycle. The WHO-Analgesic Ladder (Jadad and Brownman 1995) may be used to provide supportive interpretation of the pain endpoint in a sensitivity analysis.

Time to worsening of physical function (based on either CCI or EORTC IL-19) is defined as the time from randomization to the first ≥10-point decrease from baseline with confirmation at the next cycle. Participants not known to have worsening will be censored at the last documented assessment. Sensitivity analysis will be done by including death as part of the event definition. Time to worsening of physical function will be summarized for each arm by the Kaplan-Meier method, and the stratified log-rank test will be used to compare between the 2 arms.

CCI

CCI

Compliance will be assessed for each instrument. The compliance rate by instrument will be calculated at baseline, per week (if appropriate) and per cycle and overall, with reasons for noncompliance to be reported. The number of participants with expected assessments at each post-baseline visit is the number of participants who have received the study drugs in a previous visit and will be used as a denominator, the number of participants who completed assessments will be used as a numerator for calculating a compliance rate. For the Worst Pain NRS, participants will be considered compliant if they have completed ≥50% of the daily assessments during each period (ie, week) and will be reported for each cycle. A 7-day average score regardless of dosing will be calculated for each consecutive week starting from Cycle 1 Day 2, if

the participant is compliant for the period (completed 4 days out of the 7-day period). For other instruments, the reason and number of missing and incomplete questionnaires and/or assessments by visit will be summarized for each instrument and treatment arm.

The Cycle 1 Day 1 (pre-dose) visit will be considered as baseline for all PRO analyses. The analysis for change from baseline will be based on the participants who have baseline and at least 1 post-baseline data. Other PRO analyses will be described in a separate PRO SAP.

4.7. Safety Analyses

The safety analyses will be conducted in the safety population.

4.7.1. Extent of Exposure

Drug exposure, dose intensity, and drug adjustment (dose omissions, increases, reductions, interruptions, and delays) for abemaciclib/placebo and fulvestrant will be summarized for all treated patients per treatment arm. Drug exposure will include summaries of cycles received per patient, duration on therapy, and cumulative dose. 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 prescribed times 100% (expressed as a percentage).

For abemaciclib/placebo, extent of exposure will be measured by pill counts and summarized cumulatively. The summary will include total dosage taken, dose intensity, and relative dose intensity. The assigned cumulative dose for each patient during each cycle is 150 mg per dose × 2 doses per day × 28 days = 8400 mg. The assigned cumulative dose while on study is 150 mg per dose × 2 doses per day × number of days on treatment.

For fulvestrant, extent of exposure will be measured using the fulvestrant administration eCRF and summarized by cycle and cumulatively. The summary will include total dosage administered, dose intensity, and relative dose intensity. The assigned cumulative dose for each patient during each cycle is 1000 mg for Cycle 1 and 500 mg for Cycle 2 and beyond. The assigned cumulative dose while on study is 500 mg + 500 mg × number of cycles started.

Dose adjustments and omissions, along with the reason for adjustment or omission, will be summarized for abemaciclib/placebo and fulvestrant.

4.7.2. Adverse Events

The MedDRA PT derived from the verbatim term will be used when reporting AEs by MedDRA terms. The MedDRA Lower Level Term will be used in the treatment-emergent computation. Toxicity grades will be assigned by the investigator using NCI-CTCAE version 5.0.

Pre-existing conditions are defined as AEs that either are ongoing at informed consent or end on or after informed consent. Pre-existing conditions will be included in the listing of AE so that the history of AEs can be traced.

Treatment-emergent adverse events are events that first occur or worsen in CTCAE grade after the first dose of study treatment, and up to 30-day short-term follow-up visit. Treatment-emergent adverse events will be summarized by SOC and by decreasing frequency of PT within SOC.

Adverse event analyses will include summaries of the following:

- overview of AEs
- TEAEs, including toxicity grade (any grade and grade ≥3) and possible relationship to study drug
- serious AEs, including possible relationship to study drug
- AEs leading to dose adjustments/omissions
- discontinuations from study treatment due to AEs or death, and
- time to onset for selected TEAEs.

4.7.3. Additional Safety Assessments

Deaths

- Deaths and their primary cause
- Adverse events leading to death

Laboratory Abnormalities

The severity of laboratory results will be classified according to NCI-CTCAE. The laboratory toxicity by worst NCI-CTCAE grade and shifts in toxicity grading from baseline to the worst post-baseline grade will be summarized. Abnormal laboratory parameters will be listed.

Shift to low/high tables will include the number and percentage of patients within each baseline category (baseline value is low, normal, high, or missing) versus each postbaseline category (worst value is low, normal, high, or missing) by treatment arm.

The analyses of AEs, deaths, and laboratory abnormalities will be conducted in the safety population.

Electrocardiograms

Electrocardiogram will be summarized by visit and by treatment arm. A summary of change from baseline (by visit) and the corresponding AEs will also be provided.

Vital Signs

All vital signs (eg, temperature, blood pressure, pulse rate, height, weight, heart rate) will be summarized by visit and by treatment arm. Treatment emergent abnormal changes in vital signs will also be summarized by treatment arm.

4.8. Other Analyses

4.8.1. Demographics

Patient demographics will be summarized. Patient demographics will include the following:

- race
- ethnicity
- geography
- age
- height

- weight
- body mass index
- baseline ECOG performance status

4.8.2. Baseline Disease Characteristics

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

- initial pathological diagnosis
- study entry pathological diagnosis
- disease stage (Stage IIA, Stage IIB, etc.) at initial diagnosis and study entry
- histopathological diagnosis grade (G1, G2, etc.) at initial diagnosis and study entry
- site of disease (liver, lung, etc.)
- measurable disease at baseline (yes versus no)
- presence of visceral metastases (yes versus no)
- number of organs involved (1, 2, or 3+)
- estrogen receptor status, and
- progesterone receptor status

Number of organs involved will be derived from the "Target Tumor: RECIST 1.1" and "Non Target Tumor: RECIST 1.1" case report forms at baseline. All patients with at least one lesion on the target lesion form will be counted as having measurable disease. The number of organs involved will be derived from the location codes of the target and nontarget lesions.

4.8.3. Medical History

Historical conditions and preexisting conditions (using MedDRA PTs) will be summarized by treatment arm.

4.8.4. Prior Therapy

Prior radiotherapy, surgery, and systemic therapy will be summarized by treatment arm. Prior radiotherapy and surgery will be categorized by reason for regimen. Prior systemic therapies will be categorized by type of regimen (endocrine therapy, chemotherapy, etc.) and reason for regimen (neoadjuvant, adjuvant, locally advanced, or metastatic). Frequency of each specific therapy will be tabulated within each type of therapy and per reason for regimen.

Most recent CDK4 & 6 inhibitor-based regimen and the duration of that therapy will be summarized within each of the following subgroups:

- Patients whose most recent CDK4 & 6 inhibitor-based regimen was an adjuvant therapy
- Patients whose most recent CDK4 & 6 inhibitor-based regimen was for locally advanced or metastatic disease.

This summary will include median duration of treatment (date of end of therapy – date of start of therapy + 1), median time to progression (date of progression – date of first dose + 1), and frequency of each specific therapy. If only the month and year of a treatment date or progression date is available, the day will be imputed to the 15th.

4.8.5. Concomitant Therapy

All medications will be coded to the generic preferred name according to the current World Health Organization drug dictionary. All concomitant medications will be summarized for the ITT population using the preferred name.

4.8.6. Poststudy Treatment Discontinuation Therapy

The numbers and percentages of participants receiving poststudy treatment discontinuation anticancer therapies will be provided by type of therapy (surgery, radiotherapy, or systemic therapy), and by drug class and/or name, overall, and by line of therapy.

4.8.7. Treatment Compliance

Treatment compliance of abemaciclib/placebo will be measured by pill counts and summarized. Compliance will be calculated as the ratio of total dose taken to the total assigned dose (minus any dose adjustments and doses omitted/withheld). The total assigned dose for a patient with no adjustments or omissions is 150 mg per dose × 2 doses per day × 28 days = 8400 mg. Fulvestrant is administered in the clinic. For analysis of fulvestrant exposure, see Section 4.7.1.

4.8.8. Follow-up Time

Follow-up time is defined as the time from the date of randomization until death from any cause or last date the patient is known to be alive and under follow-up. Median follow-up time will be estimated using Kaplan-Meier estimation of potential follow-up ("reverse Kaplan-Meier") (Schemper and Smith 1996). The inverse of the censoring rules for the OS will be used (ie, considering all censoring times for OS as event times [times when the patient is known to be still alive and under follow-up] and censoring patients who had OS events at the date of death).

4.8.9. Medical Resource Utilization

Frequency counts of hospitalizations, emergency room visits, radiation, surgery, transfusion, and analgesic use will be summarized descriptively by cycle and overall for each arm.

Duration of hospital stays and average number of emergency room visits will be reported by treatment arm.

4.8.10. Subgroup analyses

Subgroup analyses of PFS and OS will be performed for potential prognostic subgroup variables, including

- all baseline stratification factors
- measurable disease at baseline (yes versus no)
- age (<65 years versus ≥65 years)
- race (Caucasian, Asian, and Other)
- progesterone receptor status (positive versus negative)
- baseline ECOG PS (0 versus 1)
- number of sites involved (1 versus 2 versus 3+)
- prior CDK4/6 inhibitor (abemaciclib versus palbociclib versus ribociclib), and

- prior CDK4/6 inhibitor setting (adjuvant versus locally advanced/metastatic)

If a level of a factor consists of fewer than 5% of randomized patients, analysis within that level will be omitted. Analyses will be done within subgroup and, separately, across subgroups with a test of interactions of subgroups with treatment performed. Estimated hazard ratios 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. Other subgroup analyses may be performed as deemed appropriate. If any safety analyses identify important imbalances between arms, subgroup analyses of these endpoints may be performed.

4.8.11. Biomarker Analyses

The distributions of biomarkers with continuous measures, such as gene or protein expression, will be described. Summary statistics will include means, medians, corresponding standard errors, quartiles, and ranges. Biomarkers with discrete measures, such as genotype locus, will be summarized in frequency tables. Correlative analyses may be performed to investigate associations between biomarkers and clinical endpoints as deemed appropriate.

4.8.12. Important Protocol Deviations

Important protocol deviations that potentially compromise the data integrity and participants' safety will be summarized. These deviations will include deviations that can be identified programmatically and those which can only be identified by the clinical research associates during monitoring. Important protocol deviations are described in another document within the study Trial Master File.

4.9. Interim Analyses

4.9.1. Data Monitoring Committee

Interim analyses for safety and efficacy will be conducted under the guidance of an independent DMC. The DMC will consist of at least 3 members, including 2 clinicians and 1 statistician. The DMC will communicate any recommendations based on interim analysis to the Sponsor SMD. If necessary, the SMD may form an IRC to review and act upon the recommendations of the DMC. Details will be provided in a separate DMC charter.

4.9.2. Safety Interim Analyses

The DMC will monitor the overall safety of the study. The DMC members will review unblinded safety data at each 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 SMD.

In the event that blinded safety monitoring by the study team 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 the Sponsor representatives external to the study team may be required for evaluation of selected serious adverse events for determination of regulatory reporting.

There will be no prespecified rules for stopping the trial due to safety concerns.

The DMC will meet and review the overall data approximately every 6 months thereafter while patients remain in the on-study intervention periods. At the recommendation of the DMC, the frequency of safety interim analyses may be modified.

4.9.3. Efficacy/Futility Interim Analyses

The interim PFS analysis for efficacy will be conducted once approximately 176 PFS events in total (approximately 70% information fraction) have been observed with Type I error controlled via the sequential monitoring approach of DeMets and Lan (1994) with the O'Brien-Fleming type spending function. The critical p-value boundaries, cumulative Type I error rate, and cumulative power is provided in Table 6. At the interim analysis, the monitoring for futility will be conducted. Futility should be declared if the observed hazard ratio is ≥1.03. Note the boundaries will be updated based on the number of events actually observed at each analysis.

If the futility boundary is met at the interim analysis, the DMC should inform the SMD and recommend that the study be stopped for futility. The SMD may convene an IRC to review the DMC's recommendation prior to making a decision based on the DMC recommendation. If the study is deemed to be futile, patients receiving active treatment may stop the blinded study drug depending on the data. Long-term follow-up of patients, including data collection on outcomes and sample collection, will proceed per protocol. If enrollment has not been completed at the time the study is deemed futile by the IRC, enrollment will be stopped. Planned study enrollment will continue during the conduct of the interim analysis.

The sponsor has no intent to stop the study based on interim analysis of efficacy, and all patients will continue follow-up for PFS until the primary PFS analysis and for OS until study close. If statistical significance is achieved for PFS at interim or final PFS analysis, an interim efficacy analysis of OS will also be conducted; the DMC will be instructed to recommend to the SMD that the results be released to the Sponsor. The SMD may convene an IRC to review the DMC's recommendation prior to Sponsor unblinding.

Table 6 Efficacy Information and Decision Boundaries

Analysis	Number of Events	Information Fraction	HR Boundary for Futility	Critical one-sided P-value Boundary for Efficacy	Critical Hazard Ratio Boundary	Cumulative Type I Error Rate	Cumulative Power under Assumed HR of 0.7
Interim PFS	176	70%	1.03	0.007	0.69	0.007	0.472
Final PFS	251	100%	NA	0.023	0.78	0.025	0.801

Abbreviations: HR = hazard ratio; PFS = progression-free survival; NA = not applicable.

5. Sample Size Determination

The study will randomize approximately 350 participants in 1:1 randomization ratio, with approximately 175 participants per treatment arm.

A 2-look group-sequential design of the primary endpoint of investigator-assessed PFS will be used to accommodate an event-driven plan for an interim PFS and final PFS analyses. The interim efficacy analysis will be performed after approximately 176 PFS events in total have occurred (ie, approximately 70% of the planned PFS events).

The final PFS analysis will be performed after approximately 251 PFS events in total have occurred (that is, an approximately 30% censoring rate). Assuming a hazard ratio of 0.70, this sample size yields approximately 80% statistical power to detect superiority for the abemaciclib plus fulvestrant arm versus the placebo plus fulvestrant arm.

This will be conducted with a one-sided log-rank test and a cumulative Type I error of 0.025.

Approved on 22 Sep 2022 GMT

6. Supporting Documentation

6.1. Appendix 1: Clinical Trial Registry Analyses

Additional analyses will be performed for the purpose of fulfilling the CTR requirements.

Analyses provided for the CTR requirements include the following:

Summary of adverse events, provided as a dataset which will be converted to an XML file. Both Serious Adverse Events and 'Other' Adverse Events are summarized by treatment group and, by MedDRA PT.

- An adverse event is considered 'Serious' whether or not it is a TEAE.
- An adverse event is considered in the 'Other' category if it is both a TEAE and is not serious. For each Serious AE 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 participants/subjects in every treatment group may not be included if a 5% threshold is chosen (5% is the minimum threshold).
- AE reporting is consistent with other document disclosures for example, the CSR, manuscripts, and so forth.

In addition, the following rules apply in order to meet the requirement for participant flow and accurately represent study completion.

Study discontinuation reason	Completed	Not Completed
Participants who had an event (progressive disease or death)	X	
Participants who were off the treatment and were alive at study conclusion	X	
Lost to follow-up[a]		X
Withdrew consent to study participant (participant or physician)*		X
On study treatment at study conclusion		X

[a] Include participants only if not meeting the definition for "Completed."

7. References

Amit O, Mannino F, Stone AM, Bushnell W, Denne J, Helterbrand J, Burger HU. 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

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

DeMets DL, Lan KK. Interim analysis: the alpha spending function approach. *Stat Med*. 1994;13(13-14):1341-1352. https://doi.org/10.1002/sim.4780131308

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(2):228-247. https://doi.org/10.1016/j.ejca.2008.10.026

CCI [redacted]

Jadad AR, Browman GP. The WHO analgesic ladder for cancer pain management. Stepping up the quality of its evaluation. JAMA. 1995;274(23):1870-1873.

Kaplan EL, Meier P. Nonparametric estimation from incomplete observations. *J Am Stat Assoc*. 1958;53(282):457-481. https://doi.org/10.2307/2281868

Schemper M, Smith TL. A note on quantifying follow-up in studies of failure time. *Control Clin Trials*. 1996;17(4):343-346. https://doi.org/10.1016/0197-2456(96)00075-x

Signature Page for VV-CLIN-062920 v1.0

Approval	PPD 22-Sep-2022 08:12:45 GMT+0000

Signature Page for VV-CLIN-062920 v1.0