I3Y-MC-JPCU Statistical Analysis Plan Version 1

A Multicenter, Open-Label, Randomized-Controlled Study of Abemaciclib, a CDK4 and 6 Inhibitor, in Combination with Fulvestrant Compared to Chemotherapy in Women with HR Positive, HER2 Negative Metastatic Breast Cancer with Visceral Metastases

NCT04031885

Approval Date: 26-Jul-2019

1. Statistical Analysis Plan: I3Y-MC-JPCU: A Multicenter, Open-Label, Randomized, Controlled Phase 4 Study of Abemaciclib, a CDK4 and 6 inhibitor, in combination with Fulvestrant compared to Chemotherapy in Women with HR Positive, HER2 Negative Advanced Breast Cancer with Visceral Metastases

Abemaciclib (LY2835219)

This is a multicenter, open-label, randomized, controlled Phase 4 study of abemaciclib in combination with fulvestrant compared to physician choice of standard chemotherapy in women with HR+, HER2- locally advanced or metastatic breast cancer who have poor prognosis due to visceral metastasis

Eli Lilly and Company
Indianapolis, Indiana USA 46285

Protocol I3Y-MC-JPCU
Phase 4

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

Approval Date: 26-Jul-2019 GMT

2. Table of Contents

Section **Page**

3. Revision History

SAP Version 1 was approved prior to the first visit when a subject receives study drug or any other protocol intervention.

4. Study Objectives

4.1. Primary Objective

To compare the efficacy (ORR) of abemaciclib in combination with fulvestrant to chemotherapy.

4.2. Secondary Objectives

- To further characterize the efficacy of abemaciclib in combination with fulvestrant compared to chemotherapy (PFS, time to response, DoR, PFS2)
- To compare the safety and tolerability of abemaciclib in combination with fulvestrant to chemotherapy

4.3. Exploratory Objectives

- To evaluate PROs and HCRU for abemaciclib in combination with fulvestrant compared to chemotherapy
- To assess the relationship between whole blood and clinical outcome
- To assess the relationship between plasma biomarkers and clinical outcome

5. Study Design

5.1. Summary of Study Design

Study JPCU is a multicenter, open-label, randomized, controlled Phase 4 study of abemaciclib, a CDK 4 and 6 inhibitor, in combination with fulvestrant compared to physician choice of standard chemotherapy in a patient population of post menopausal women with HR+, HER2- locally advanced or MBC who have poor prognosis due to visceral metastasis.

Patients will be randomized using the following stratification factors: the presence of liver metastases (Yes/No), sensitivity to prior ET (sensitive versus primary resistance versus secondary resistance), and postmenopausal status (natural/surgical/therapeutic radiation versus gonadotropin-releasing hormone [GnRH] agonist use).

Sensitivity to prior ET is defined as follows, and will be determined based on the patients' most recent sensitivity status. For example, if a patient met the definition of primary resistance on adjuvant ET and subsequently met the definition of secondary resistance on ET for MBC, the patient would be categorized as having secondary resistance.

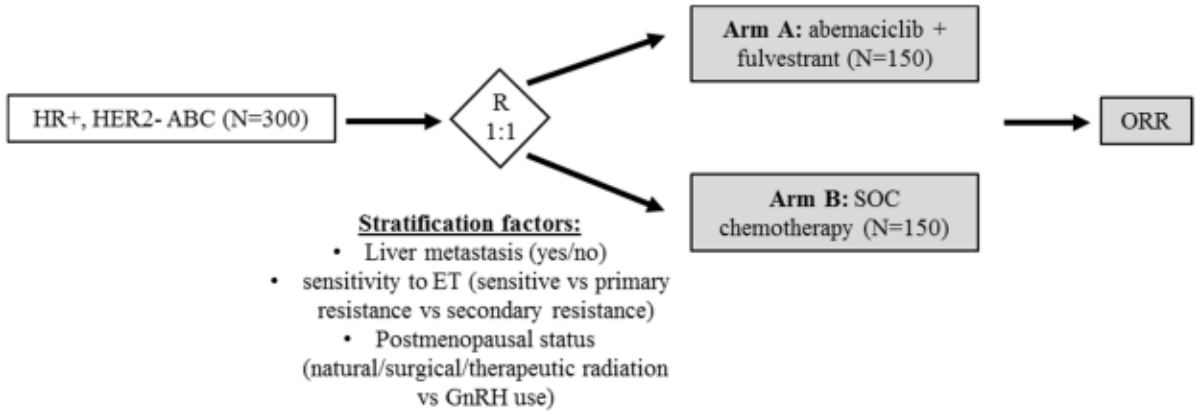
1. Primary resistance: a relapse while on the first 2 years of adjuvant ET, or progressive disease (PD) within first 6 months of first-line ET for MBC, while on ET
2. Secondary resistance: a relapse while on adjuvant ET but after the first 2 years, or a relapse within 12 months of completing adjuvant ET, or PD ≥6 months after initiating ET for MBC, while on ET
3. Sensitive: patients who do not meet the definition of primary resistance or secondary resistance will be considered to be sensitive

Patients who meet all criteria for enrollment will be randomly assigned 1:1 to:

1. **Arm A**: abemaciclib in combination with fulvestrant, or
2. **Arm B**: physician choice of standard chemotherapy (capecitabine, docetaxel, nab-paclitaxel, or paclitaxel).

Cross over between study treatments is not allowed in this study. Additionally, post-discontinuation therapy is neither specified nor restricted.

Taken together, these data will be used to characterize the efficacy, safety and tolerability of abemaciclib in combination with fulvestrant compared to chemotherapy. The study design is illustrated below.



Abbreviations: ABC= advanced breast cancer, ET = endocrine therapy; GnRH = gonadotropin-releasing hormone; HER2- = human epidermal growth factor receptor 2-negative; HR+ = hormone receptor positive; N = number of enrolled patients; ORR = objective response rate; R = randomization; SOC = standard of care.

5.2. Determination of Sample Size

Approximately 300 patients will be randomized in a 1:1 ratio to Arm A and Arm B for an estimated total of 150 patients per arm.

The primary analysis of ORR to assess ORR comparability will be performed approximately 6 months after last patient entering treatment; the final analysis of PFS to assess PFS comparability will be performed, if ORR comparability is declared, when approximately 180 PFS events in total have been observed.

ORR

Comparability of ORR will be declared if the following Bayesian criterion is achieved in the primary ORR analysis:

$$\text{Posterior P}(\text{true ORR}_{\text{Arm B}} - \text{true ORR}_{\text{Arm A}} < 10\%) > 80\%$$

Full detail regarding the Bayesian model for ORR (including prior specifications) is provided in Section 6.9.1.

CCI [redacted]

CCI [redacted]

[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]

[redacted]

CCI [redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]

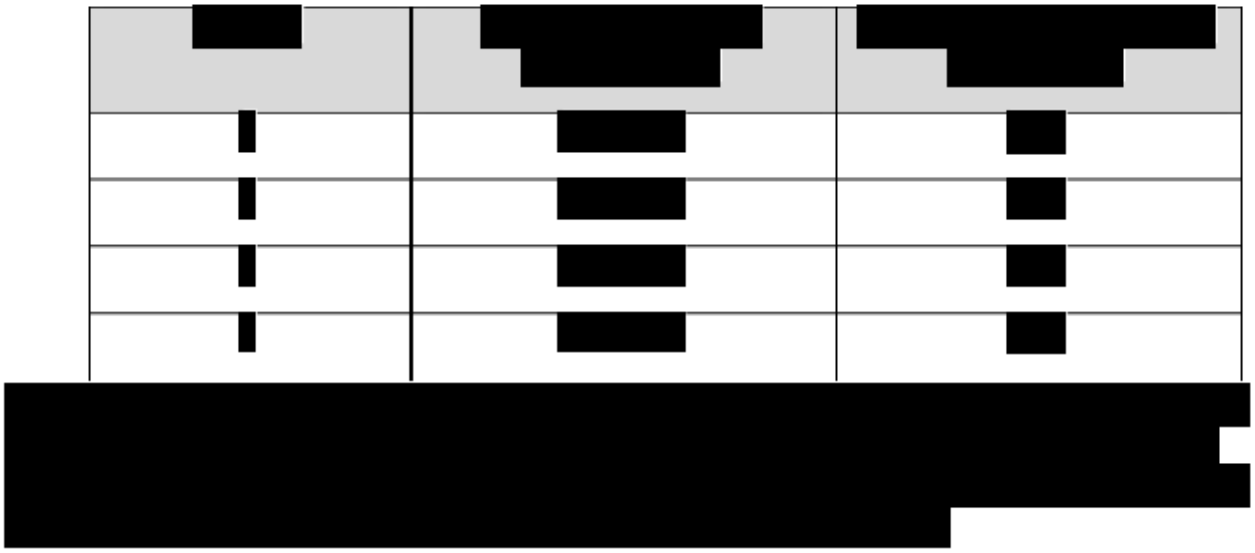
PFS

Progression-free survival comparability will be declared if ORR comparability is declared AND the following Bayesian criterion is achieved in the final PFS analysis:

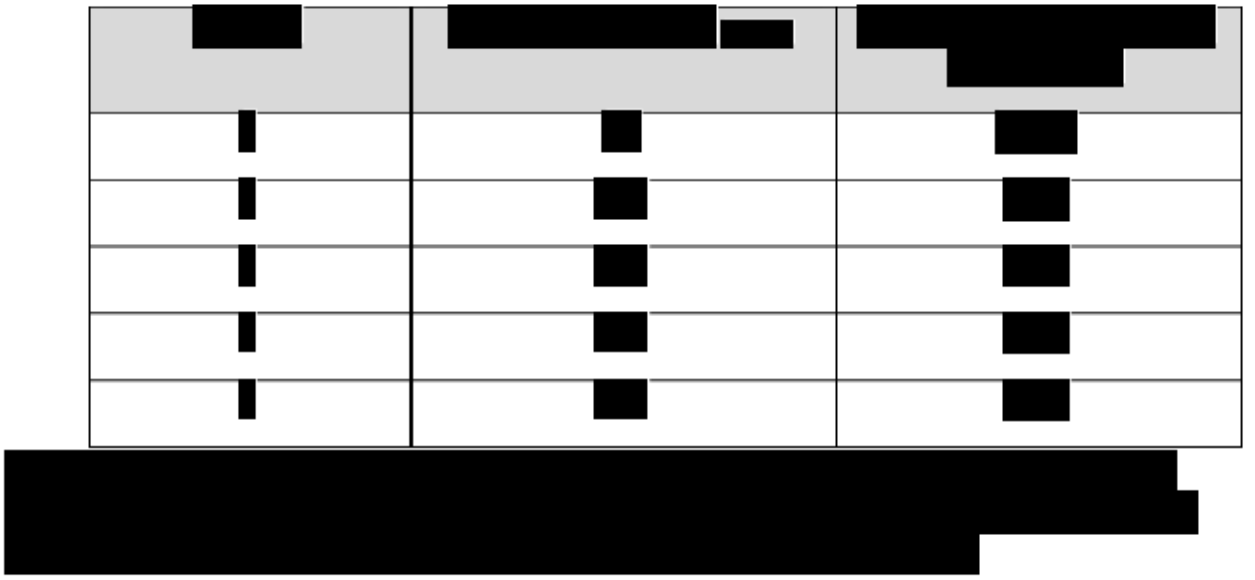
$$\text{Posterior P}\,(HR - 1 < 0.2) > 80\%, \text{ where HR} = \text{true hazard}_{\text{Arm A}} / \text{true hazard}_{\text{Arm B}}$$

CCI [redacted]

[redacted]



[redacted]



[redacted]

6. A Priori Statistical Methods

6.1. General Considerations

6.1.1. Populations

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

Disposition summaries will be performed on the **entered** population, defined as all patients who sign the informed consent form.

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. This analysis set will be used for all baseline and efficacy analyses.

Safety population: will include all randomized patients who received any quantity of study treatment, regardless of their eligibility for the study. The safety evaluation will be performed based on the first dose of study treatment a patient actually received, regardless of the arm to which he or she was randomized. The safety analysis set will be used for all dosing/exposure, safety analyses, and health economics analyses.

6.1.2. Definitions and Conventions

Study drug refers to abemaciclib (LY2835219).

Study treatment refers to abemaciclib plus fulvestrant or single-agent chemotherapy of physician's choice (capecitabine, docetaxel, nab-paclitaxel, or paclitaxel)

The **date of randomization** is the date the patient was randomly assigned to study treatment using the IWRS.

The **date of first dose** is the date of the first dose of study drug, fulvestrant, or single-agent chemotherapy.

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

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 08MAR2016 and the date of first dose was 06MAR2016, 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 05MAR2016 and the date of first dose was 06MAR2016, 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.

Additional exploratory analyses of the data will be conducted as deemed appropriate.

6.2. Adjustments for Covariates

The following baseline covariates will be the stratification factors for randomization: presence of liver metastases (Yes/No), sensitivity to prior ET (sensitive versus primary resistance versus secondary resistance), and postmenopausal status (natural/surgical/therapeutic radiation versus gonadotropin-releasing hormone [GnRH] agonist use). These stratification factors are thought to be associated with clinical outcomes. Balancing the 2 arms with respect to those factors will further reduce the potential for bias and improve the power of the analyses. For example, our internal data indicate that postmenopausal status is an important prognostic factor: in Study I3Y-MC-JPBL, the patients receiving GnRH agonist on abemaciclib arm (n=72) had a median PFS of 28.6 months versus 15.5 months for those patients who were truly postmenopausal (n=371).

6.3. 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. Rules for determining date of progression or censor for PFS are defined in Table JPCU.6.1.

6.4. Multiple Comparisons/Multiplicity

The primary endpoint is ORR comparability; PFS comparability is an important secondary endpoint, and will be statistically tested only if ORR comparability is declared, using a gatekeeping testing strategy. The family-wise error rate, i.e. the probability of making one or more false positive claims due to multiple endpoint-related multiplicity, will be well controlled at no more than 20%, per simulations.

6.5. 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, 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 pre-determined categories.

6.6. Patient Characteristics

6.6.1. Demographics and Performance Status

Patient demographics will be summarized. Patient demographics will include age, race, ethnicity, height, weight, and body mass index.

6.6.2. Baseline Disease Characteristics

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

- Initial 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
- Baseline Eastern Cooperative Oncology Group (ECOG) performance status (PS)
- HER2 status (negative/postive)
- Progesterone receptor status (negative/postive)
- Estrogen receptor status (negative/postive)
- Site of disease (liver, lung, etc.)

6.6.3. 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.

6.6.4. 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 (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.

6.6.5. Post-Study Treatment Discontinuation Therapies

Therapies received following study treatment discontinuation will be summarized by arm. Therapies will be summarized overall and by category: endocrine therapy or targeted/chemotherapy.

6.7. Treatment Compliance

Treatment compliance information for abemaciclib will be collected through pill counts at each cycle. Compliance will be calculated as the ratio of total dose taken to the total assigned dose (minus any dose adjustments and doses omitted/withheld for medical or logistical reasons). Treatment compliance information for capecitabine will be collected through couting number of doses taken at each cycle. Dosing information for fulvestrant, docetaxel, paclitaxel, and nab-paclitaxel will be collected at each cycle/visit.

6.8. 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 by number and percentage of patients for the safety population using the base name (without esters or salts).

6.9. Efficacy Analyses

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

The stratification factors for the primary and secondary analyses are

- Presence of liver metastases (Yes/No)
- Sensitivity to prior ET (sensitive versus primary resistance versus secondary resistance)
- Postmenopausal status (natural/surgical/therapeutic radiation versus GnRH agonist use).

The stratification factors will be derived based on CRF data.

Posterior medians and Bayesian credible intervals will be reported for ORR and PFS, for which Bayesian analyses have been specified.

All frequentist tests of treatment effects will be conducted at a 2-sided alpha level of 0.05, unless otherwise stated; and all confidence intervals (CIs) will be given at a 2 sided 95% level.

6.9.1. Primary Outcome and Methodology: Objective Response Rate

6.9.1.1. Definition

The primary endpoint of the study is **objective response rate (ORR)**, which is defined as the number of patients who achieve a best overall response (BOR) of complete response (CR) or

partial response (PR) divided by the total number of patients randomized to the corresponding treatment arm (ITT population), based on investigator-assessed tumor responses.

The BOR corresponds to the best response recorded from the start of the study treatment until the earliest of objective progression or start of new anticancer therapy (excluding maintenance endocrine therapy following study treatment); confirmations of CR and PR are not required.

6.9.1.2. Primary Analysis

Primary comparison between treatment arms will be based on the Bayesian decision rule as described in Section 5.2. The 80th percentile of the posterior distribution for the difference in true ORR (true $ORR_{Arm\ B}$ – true $ORR_{Arm\ A}$) will be calculated and compared to 10%, to determine if Posterior P(true $ORR_{Arm\ B}$ – true $ORR_{Arm\ A}$ < 10%) > 80%. In addition, the median of the posterior distribution for ORR, with 95% equal-tailed credible interval, will be summarized for each treatment arm.

An objective Bayesian approach will be implemented to derive the posterior distribution of the true (unknown) quantities of interest, $ORR_{Arm\ A}$ and $ORR_{Arm\ B}$, conditional on the observed data and prior distributions. For Arm A, the joint likelihood function is binomial(~150, $ORR_{Arm\ A}$) and the prior distribution for $ORR_{Arm\ A}$ is the (weakly informative) Jeffreys prior for the binomial class of likelihood functions, namely the beta distribution with both shape and scale parameters equal to 1/2. Because the beta priors are conjugate for the binomial likelihoods, the posterior distribution for $ORR_{Arm\ A}$ is a known beta distribution. Indeed, if r_A is the number of observed responders in N_A patients randomized to Arm A, the posterior distribution is:

$$[ORR_{Arm\ A} \mid r_A] \sim Beta(r_A + .5, N_A - r_A + .5).$$

Correspondingly for Arm B we have:

$$[ORR_{Arm\ B} \mid r_B] \sim Beta(r_B + .5, N_B - r_B + .5).$$

Note that because the two joint likelihoods and prior distributions for $ORR_{Arm\ A}$ and $ORR_{Arm\ B}$ are independent, the two posterior distributions are also independent. That is:

$$[ORR_{Arm\ A}, ORR_{Arm\ B} \mid r_A, r_B] = [ORR_{Arm\ A} \mid r_A] \cdot [ORR_{Arm\ B} \mid r_B]$$

Monte Carlo sampling (from the two independent beta posteriors) will be used to approximate the posterior distribution of the functional $ORR_{Arm\ B}$–$ORR_{Arm\ A}$.

The primary analysis of ORR to assess ORR comparability will be performed approximately 6 months after last patient entering treatment.

6.9.2. Gated Secondary Efficacy Analyses: Progression-free Survival

Progression-free survival (PFS) is an important secondary endpoint of the study, and will be assessed only if ORR comparability is declared, using a gatekeeping testing strategy.

PFS is defined as the time from randomization until the first occurrence of documented disease progression per RECIST 1.1 criteria, or death from any cause in the absence of progressive disease. PFS will be based on investigator-assessed tumor responses; there will not be an

independent central review of imaging data. Patients known to be alive and without disease progression will be censored at the time of the last adequate tumor assessment (a detailed PFS event/censoring scheme is provided in Table JPCU.6.1 below). A (stratified) Bayesian Cox regression model will be used to estimate the PFS HR comparing Arm A and Arm B. The 80th percentile of the posterior distribution of the HR comparing Arm A and Arm B will be calculated and compared to 1.2, to determine if Posterior P (HR – 1 < 0.2) > 80%. In addition, the posterior medians and associated 95% credible intervals for true median PFS by treatment arm will be provided.

The Bayesian model used to calculate the PFS HR posterior will combine the Breslow partial likelihood (Breslow, 1975) with a diffuse, weakly-informative prior distribution for the hazard ratio. Let $\theta = \log HR$. A priori, we will assume $\theta \sim N(0,100)$ where the second argument is the variance. The induced posterior distribution does not exist in closed form, hence Markov Chain Monte Carlo (MCMC) methods will be used to generate samples from the approximate posterior distribution for θ (and thus HR). Note that since the partial likelihood is used, the Bayesian model is complete without specification of a prior model for the underlying baseline hazard function.

Final analysis of PFS to assess PFS comparability will be performed, if ORR comparability is declared, when approximately 180 PFS events in total have been observed.

Table JPCU.6.1. PFS Event/Censoring Scheme

Situation	Date of Event or Censor	Event / Censor
Tumor progression or death	Earliest date of PD or death	Event
No tumor progression and no death	Date of last adequate radiological assessment or date of randomization (whichever is later)	Censored
Unless		
No baseline radiological tumor assessment available	Date of randomization	Censored
No adequate post baseline radiological tumor assessment available and death reported after 2 scan intervals following randomization	Date of randomization	Censored
New anticancer treatment started (excluding maintenance endocrine therapy following chemotherapy) and no tumor progression or death within 14 days	Date of adequate radiological assessment prior to (start of new therapy +14 days) or date of randomization (whichever is later)	Censored
Tumor progression or death documented immediately after 2 or more scan intervals following last adequate radiological tumor assessment or randomization (whichever is later)	Date of last adequate radiological assessment or date of randomization (whichever is later)	Censored

Notes:
(1) Symptomatic deteriorations (i.e. symptomatic progressions, which are not radiologically confirmed) will not be considered as progression
(2) Adequate radiological tumor assessment refers to an assessment with one of the following responses: CR, PR, SD or PD

6.9.3. Other Secondary Efficacy Analyses

6.9.3.1. Progression-free Survival 2

For patients who take maintenance endocrine therapy following study treatment, progression-free survival 2 (PFS2) is defined as the time from randomization to the disease progression date on first-line treatment post discontinuation of maintenance endocrine therapy, discontinuation date of first-line treatment post discontinuation of maintenance endocrine therapy, or starting date of the second-line treaetment post discontinuation of maintenance endocrine therapy, or death from any cause, whichever is earlier. For all other patients, PFS2 is defined as the time from randomization to the disease progression date on first-line treatment post discontinuation of study treatment, discontinuation date of first-line treatment post discontinuation of study treatment, or starting date of the second-line of post discontinuation of study treatment, or death from any cause, whichever is earlier.

If the patient is alive at the cutoff for analysis, and a PFS2 event has not been observed, PFS2 data will be censored on the last date the patient was known to be alive.

A Kaplan-Meier (KM) analysis of PFS2 will be performed to estimate the PFS2 curve for each arm. Point estimates, CIs for PFS2 quartiles, and PFS2 rates will be calculated every 3 months for the first 12 months.

6.9.3.2. Time to Response

Time to Response (TTR) is defined as the time from randomization until the date that measurement criteria for CR or PR (whichever is first recorded) are first met, per RECIST v1.1, and will be calculated for patients with CR or PR only.

A KM analysis of TTR will be performed to estimate the TTR curve for each arm. Point estimates and CIs for DoR quartiles will be calculated.

6.9.3.3. Duration of Response

Duration of response is defined as the time from the date that measurement criteria for CR or PR (whichever is first recorded) are first met until the first date that disease is recurrent or documented disease progression is observed, per RECIST 1.1 criteria, or the date of death from any cause in the absence of documented disease progression or recurrence. It will be calculated as date of progression or death – date of first response evaluation of CR or PR + 1 for patients with CR or PR only. The DoR will be censored according to the same rules as PFS.

A KM analysis of DoR will be performed to estimate the DoR curve for each arm. Point estimates, CIs for DoR quartiles, and DoR rates will be calculated every 3 months for the first 12 months.

6.9.4. Sensitivity Analyses

Progression-Free Survival Sensitivity Analysis 1 (nonobjective progression as a PFS event): if a patient is discontinued from study treatment due to investigator-determined non-objective progression (for example, symptomatic deterioration), then the patient's PFS time will be calculated using the date of non-objective progression as the progression date.

6.10. Health Outcomes/Quality-of-Life Analyses

The main analysis will be conducted in the safety population.

For each instrument, the compliance rate by visit will be calculated as the number of patients with completed assessments at each visit divided by the number of patients in safety population at each visit. Compliance rates and reasons for noncompliance for each instrument will be summarized by treatment arm.

6.10.1. Patient-Reported Outcome (PRO)

Patient-reported outcomes are measured through electronic versions of the following:

- Worst-pain NRS
- BM Frequency
- Bristol Stool Form Scale
- Loperamide doses
- FACT-GP5
- PRO CTCAE
- EORTC IL36

Worst-pain NRS, BM Frequency, Bristol Stool Form Scale, and Loperamide doses are single-item instruments and will be assessed daily. For such instrument with daily data collection, weekly average will be calculated for each patient; all other instruments will be assessed weekly. For each instrument, weekly average or weekly data collection will be used as the data at each visit for data analyses.

For each patient with data from baseline and at least 1 other visit post-baseline, the change from baseline score by visit, by all post-baseline visits, and maximum change from baseline score will be calculated and summarized using descriptive statistics between arms for each instrument, except for PRO CTCAE. Response to each item in PRO CTCAE will be treated as a categorical variable and will be summarized using descriptive statistics by visit between arms.

To describe and compare the Time to Deterioration (TTD) and also Time to Sustained Deterioration (TTSD) in the

- "worst pain" (2 point change)
- FACT-GP5 (one category change)
- PRO CTCAE (one category change)
- EORTC IL36 (European Organisation for the Research and Treatment of Cancer Quality of Life Questionnaire-Core30 (EORTC QLQ-C30) functional scales, and Global Health/Quality of Life (GHS) scale (10 point change)).

Pain Intensity:

Pain as assessed with the single "worst pain" NRS item, and will be characterized with descriptive non-inferential statistics based on a weekly average. Treatment arm comparisons will be conducted through use of mixed model repeated measures (MMRM) specified with an unstructured covariance matrix and estimated with restricted maximum likelihood (REML) to address the small sample negative bias in the variance component when using the maximum likelihood (ML) estimator (that is, the ML estimator usually has lower mean-squared error [MSE] than the REML estimator).

EORTC IL36:

EORTC IL36 instrument data will be scored as described by Aaronson et al. 1993. If not already addressed in the EORTC scoring manual (Fayers et al. 2001), descriptive statistics for each EORTC IL36 (functional scale and the Global Health Status/QOL scale) will be calculated and evaluated between arms.

6.10.2. Medical Resource Utilization

Utilization data will be summarized descriptively by healthcare resources used (HCRU) category (hospitalization days and reasons, emergency room visit reasons, growth factor use, analgesic use and transfusions), including a frequency table with tabular statistics. For categorical variables, frequency and the corresponding percentage will be derived and measures of central tendency and variability will be calculated for continuous variables by arm. Tests for differences in proportion between treatment groups and between response groups may be performed as

warranted (Chi-square for categorical variables, t-test for continuous variables and non-parametric tests for non-normally distributed data).

Time to first opioid use by treatment arm, responder type, age, baseline PRO score may be explored.

6.11. Safety Analyses

6.11.1. Extent of Exposure

Drug exposure, dose intensity, and drug adjustment (dose omissions, increases, reductions, interruptions, and delays) for abemaciclib, fulvestrant, and single-agent chemotherapy of physician's choice 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). The summary of dose adjustments will include reasons for adjustments.

For abemaciclib, extent of exposure will be measured by pill counts. Dose intensity will be expressed in mg/day. The assigned cumulative dose while on study is 2 × 150 mg × number of days on treatment.

For fulvestrant, extent of exposure will be measured using the fulvestrant administration on the eCRF. The assigned cumulative dose while on study is 500 mg × 2 + [500 mg × ceiling ((number of days on treatment – 28)/28)].

For single-agent chemotherapy of physician's choice, extent of exposure and assigned cumulative dose while on study will be measured based on data reported on the eCRF appropriate for the route of administration.

6.11.2. Adverse Events

Adverse event (AE) severity grades will be assigned by the investigator using Common Terminology Criteria for Adverse Events (CTCAE) Version 5.0. In addition, AE verbatim text will also be mapped by the sponsor or designee to corresponding terminology within MedDRA.

Preexisting conditions are defined as AEs that either is ongoing at informed consent or end on or after informed consent.

A treatment-emergent adverse event (TEAE) is defined as an AE that first occur or worsen in CTCAE grade after the first dose of study treatment. The MedDRA LLT will be used 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 summaries and listings will be produced:

- Overview of adverse events
- Summary of TEAEs by preferred term (PT) (any grade and Grade ≥ 3)
- Summary of TEAEs by system organ class (SOC) and PT (any grade and Grade ≥ 3)
- Summary of TEAEs by SOC and PT and maximum grade (1-5)
- Summary of SAEs by SOC and PT (any grade and Grade ≥ 3)
- Summary of AEs as reason for study treatment discontinuation by SOC and PT
- Listing of SAEs

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

6.11.3. Deaths and Other Serious Adverse Events

A summary of all deaths, including reasons for deaths, will be provided. All deaths, deaths on study treatment, deaths within 30 days of discontinuation of study treatment, deaths on study treatment or within 30 days of discontinuation of study treatment, and deaths after 30 days of discontinuation of study treatment will be summarized by reason for death. For deaths due to AE, the preferred term will be provided. In addition to the tabular summary, a by-patient listing of all deaths on study not attributed to study disease by the investigator will be provided.

6.11.4. Clinical Laboratory Evaluation

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

6.11.5. Vital Signs and Other Physical Findings

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

6.12. Subgroup Analyses

Subgroup analyses of ORR and PFS will be performed for each of following potential prognostic subgroup variables:

- presence of liver metastases (Yes/No)
- sensitivity to prior ET (sensitive versus primary resistance versus secondary resistance)
- postmenopausal status (natural/surgical/therapeutic radiation versus GnRH agonist use)
- number of organs involved (1 versus 2 versus 3+)
- age (<65 years versus ≥65 years)
- progesterone receptor status (positive versus negative)
- baseline ECOG PS (0 versus 1)

Subgroup analyses of ORR and PFS will also be performed for each of the physician choice of standard chemotherapy that patient are randomized to or would receive (capecitabine and taxane [docetaxel, nab-paclitaxel, paclitaxel]). Before a patient is randomized, the chemotherapy choice that the patient would receive if she was randomized to Arm B will be recorded in the Interactive Web Response System (IWRS). A chemotherapy subgroup will consist of patients who are randomized to Arm B to receive the chemotherapy and those who are randomized to Arm A and would receive the chemotherapy per IWRS.

If a level of a factor consists of fewer than 10% 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 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.

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.

6.13. Protocol Violations

Important protocol violations 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 which can only be identified by the clinical research associate during monitoring. Important protocol deviations are described in another document within the study Trial Master File.

6.14. Interim Analyses and Data Monitoring

One interim analysis of futility for ORR will be conducted approximately 6 months after the 100th patient has been randomized. The futility rule is described as follows:

$$\text{Posterior P(true ORR}_{\text{Arm B}} - \text{true ORR}_{\text{Arm A}} < 10\%) < 20\%$$

An assessment committee (AC) will be established to conduct interim analysis. The AC will include, at a minimum, Lilly Oncology Medical Director and statistician; Lilly study CRP/CRS will be excluded from the AC. Details on AC membership will be described in the Lilly Interim Analysis Authorization Form. Although the decision at the interim analysis will be made primarily based on the futility rule, AC members will also review and consider the totality of interim data to determine whether there are sufficient evidence to justify the termination of study treatment and/or enrollment due to futility.

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 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, by MedDRA preferred term.
- An adverse event is considered 'Serious' whether or not it is a treatment emergent adverse event (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 patients/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.

7. References

Cox DR. Regression models and life-tables. *Journal of Royal Statistical Society Ser B*. 1972; 74(2):187-220.

Leo Document ID = 500b77d0-0b15-49c9-8ad5-6da2238b15d8

PPD
Approval Date & Time: 26-Jul-2019 16:23:51 GMT
Signature meaning: Approved