

Statistical Analysis Plan I3Y-MC-JPBM

A Randomized, Double-Blind, Placebo-Controlled, Phase
3 Study of Nonsteroidal Aromatase Inhibitors
(Anastrozole or Letrozole) plus LY2835219, a CDK4/6
Inhibitor, or Placebo in Postmenopausal Women with
Hormone Receptor-Positive, HER2-Negative
Locoregionally Recurrent or Metastatic Breast Cancer
with No Prior Systemic Therapy in this Disease Setting

NCT02246621

Approval Date: 14-Apr-2017

**1. Statistical Analysis Plan:
I3Y-MC-JPBM: A Randomized, Double-Blind, Placebo-
Controlled, Phase 3 Study of Nonsteroidal Aromatase
Inhibitors (Anastrozole or Letrozole) plus LY2835219, a
CDK4/6 Inhibitor, or Placebo in Postmenopausal Women
with Hormone Receptor-Positive, HER2-Negative
Locoregionally Recurrent or Metastatic Breast Cancer
with No Prior Systemic Therapy in this Disease Setting**

Confidential Information

The information contained in this Statistical Analysis Plan (SAP) is confidential and the information contained within it may not be reproduced or otherwise disseminated without the approval of Eli Lilly and Company or its subsidiaries. This document and its associated attachments or appendices are subject to United States Freedom of Information Act Exemption 4.

Abemaciclib (LY2835219)

This study is a global, multicenter, double-blind, placebo-controlled, Phase 3 trial for women with hormone receptor positive, HER2 negative locoregionally recurrent or metastatic breast cancer randomized to receive nonsteroidal aromatase inhibitors with or without LY2835219.

Eli Lilly and Company
Indianapolis, Indiana USA 46285

Statistical Analysis Plan Version 1 electronically signed and approved by Lilly:
08-May-2015
Statistical Analysis Plan Version 2 electronically signed and approved by Lilly:
04-Dec-2015
Statistical Analysis Plan Version 3 electronically signed and approved by Lilly:
20-Dec-2016
Statistical Analysis Plan Version 4 electronically signed and approved by Lilly:
31-Mar-2017
Statistical Analysis Plan Version 5 electronically signed and approved by Lilly on date
provided below.

Approval Date: 14-Apr-2017 GMT

2. Table of Contents

Section	Page
1. Statistical Analysis Plan: I3Y-MC-JPBM: A Randomized, Double-Blind, Placebo-Controlled, Phase 3 Study of Nonsteroidal Aromatase Inhibitors (Anastrozole or Letrozole) plus LY2835219, a CDK4/6 Inhibitor, or Placebo in Postmenopausal Women with Hormone Receptor-Positive, HER2-Negative Locoregionally Recurrent or Metastatic Breast Cancer with No Prior Systemic Therapy in this Disease Setting	
2. Table of Contents	2
3. Revision History	8
4. Study Objectives	9
4.1. Primary Objective	9
4.2. Secondary Objectives	9
4.3. Exploratory Objectives	9
5. Study Design	10
5.1. Summary of Study Design	10
5.2. Determination of Sample Size	11
6. A Priori Statistical Methods	12
6.1. General Considerations	12
6.1.1. Populations	12
6.1.2. Definitions and Conventions	12
6.2. Handling of Dropouts or Missing Data	13
6.3. Patient Disposition	13
6.4. Patient Characteristics	13
6.4.1. Demographics and Performance Status	13
6.4.2. Baseline Disease Characteristics	14
6.4.3. Historical Illnesses	15
6.4.4. Prior Therapies	15
6.4.5. Post Study Treatment Discontinuation Therapies	15
6.5. Treatment Compliance	15
6.6. Concomitant Therapy	15
6.7. Efficacy Analyses	16
6.7.1. General Considerations	16
6.7.1.1. Population	16
6.7.1.2. Stratification Factors	16
6.7.1.3. Hypothesis Tests and Confidence Intervals for Efficacy Data	16

6.7.2. Primary Endpoint: Progression-Free Survival.....	16
6.7.2.1. Definition	16
6.7.2.2. Hypotheses and Analysis.....	17
6.7.2.3. Other Analyses	18
6.7.2.3.1. Progression-Free Survival Curves and Hazard Ratio	18
6.7.2.3.2. Restricted Mean Difference.....	18
6.7.3. Gated Secondary Endpoint: Overall Survival.....	19
6.7.3.1. Background	19
6.7.3.2. Definition	20
6.7.3.3. Hypotheses and Analysis.....	20
6.7.3.4. Other Analyses	22
6.7.3.5. Exploratory Pooled Overall Survival Analyses.....	22
6.7.4. Other Secondary Endpoints	23
6.7.4.1. Objective Response Rate, Disease Control Rate, and Clinical Benefit Rate	23
6.7.4.2. Duration of Response	23
6.7.5. Sensitivity Analyses.....	24
6.7.5.1. Progression-Free Survival	24
6.7.5.2. Overall Survival	24
6.8. Health Outcomes/Quality-of-Life Analyses.....	25
6.8.1. Instruments	25
6.8.2. Quality of Life	25
6.8.3. Health State Utility	25
6.8.4. Utilization	25
6.9. Bioanalytical and Pharmacokinetic/Pharmacodynamic Methods	26
6.10. Tailoring Biomarker Analyses.....	26
6.11. Safety Analyses	26
6.11.1. Extent of Exposure.....	26
6.11.2. Adverse Events	27
6.11.3. Deaths.....	28
6.11.4. Clinical Laboratory Evaluation	28
6.11.5. Vital Signs and Other Physical Findings	28
6.11.6. Electrocardiograms	29
6.12. Subgroup Analyses.....	29
6.13. Protocol Violations.....	29
6.14. Interim Analyses and Data Monitoring	30
6.14.1. Safety Interim Analyses	30
6.14.2. Efficacy Interim Analyses.....	30

6.14.3. Pharmacokinetic/Pharmacodynamic Interim Analyses..... 31

6.15. Analyses for the Japanese Regulatory Authority 31

6.16. Annual Report Analyses..... 31

6.17. Clinical Trial Registry Analyses..... 31

7. Unblinding Plan..... 31

8. Changes to Planned Analyses in Protocol..... 34

9. References 35

10. Appendices 37

Redacted Version

Table of Contents

Table	Page
Table JPBM.6.1. Rules for Determining Date of Progression or Censor for Progression-Free Survival	17
Table JPBM.6.2. Example Nominal Significance Levels	22

Redacted Version

Table of Contents

Figure	Page
Figure JPBM.5.1.Illustration of study design.	10
Figure JPBM.6.1.Initial graph.....	20

Redacted Version

Table of Contents

Appendix		Page
Appendix 1.	Abemaciclib Consolidated AE Terms	38

Redacted Version

3. Revision History

Statistical Analysis Plan (SAP) Version 1 was approved prior to the first internal blinded safety review and the first unblinded safety analysis by the Data Monitoring Committee (DMC).

Statistical Analysis Plan Version 2 was approved after the third unblinded safety analysis by the DMC, but prior to unblinding of the sponsor and prior to the first interim analysis for efficacy.

The overall changes and rationale for the changes incorporated in Version 2 are as follows:

- Updates to the interim analysis plan and the analysis of overall survival in Protocol Amendment (a).

Statistical Analysis Plan Version 3 was approved after unblinded safety analyses by the DMC but prior to unblinding of the sponsor and prior to the first interim analysis for efficacy. The overall changes and rationale for the changes incorporated in Version 3 are as follows:

- Timing of final PFS analysis changed to 240 events and second interim analysis which was added with protocol amendment (a) was removed per amendment (b). Alpha spend for the interim was changed from 0.0001 to 0.00025.
- Updates to the overall survival analysis plan. Specifically, the pooled (JPBL and JPBM) overall survival analysis was reclassified as an exploratory analysis.
- Modification to the definition for baseline value of an efficacy assessment to include an assessment prior to first dose if no assessment prior to the date of randomization is available.
- Clarification to the definition of a treatment emergent adverse event.
- Inclusion of list of consolidated MedDRA preferred terms used for adverse event reporting.

Statistical Analysis Plan Version 4 was approved after unblinded safety analyses by the DMC but prior to unblinding of the sponsor and prior to the first interim analysis for efficacy. The overall changes and rationale for the changes incorporated in Version 4 are as follows:

- The OS analysis after 236 OS events was replaced by analyses after 189 and 252 OS events.
- A statement was added to specify the key quality of life items and domains for comparative statistical analysis to meet registration requirements.
- The list of consolidated MedDRA preferred terms was updated in Appendix 1 following medical review of observed adverse events.

Statistical Analysis Plan Version 5 was approved after unblinded safety analyses by the DMC but prior to unblinding of the sponsor and prior to the first interim analysis for efficacy. Version 5 updates the gated analysis of OS to split alpha between the ITT population and the population of patients with visceral disease at baseline.

4. Study Objectives

4.1. Primary Objective

The primary objective of this study is to compare treatment with abemaciclib plus nonsteroidal aromatase inhibitor (NSAI) therapy versus placebo plus NSAI therapy with respect to progression-free survival (PFS) in postmenopausal women with hormone-receptor positive (HR+), human epidermal growth factor receptor 2-negative (HER2-) locoregionally recurrent or metastatic breast cancer who have not received prior systemic therapy in this disease setting.

4.2. Secondary Objectives

The secondary objectives of the study are to compare the combination treatment of abemaciclib and NSAI therapy versus placebo plus NSAI therapy with respect to the following:

- overall survival (OS)
- overall survival rate at 1, 2, and 3 years;
- objective response rate (clinical response + partial response [ORR = CR + PR]);
- duration of response (DoR) for CR + PR;
- disease control rate (DCR = CR + PR + stable disease [SD]);
- clinical benefit rate (CBR = CR + PR + SD \geq 6 months);
- the safety and tolerability;
- change in symptom burden from baseline using the European Organization for Research and Treatment of Cancer Quality of Life Questionnaire Core 30 (EORTC QLQ-C30), EORTC QLQ-BR23 (breast) questionnaire, and health status scores from the EuroQoL 5-Dimension 5 Level (EQ-5D 5L);
- the pharmacokinetics (PK) of abemaciclib, its metabolites, and NSAI therapy.

4.3. Exploratory Objectives

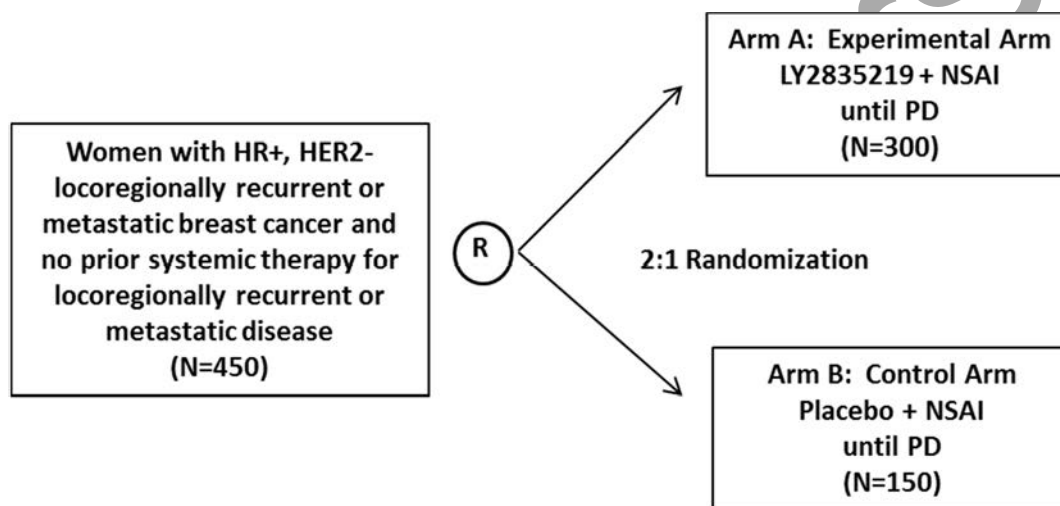
- To explore potential biomarkers related to the mechanism of action of abemaciclib, the cell cycle, and/or the pathogenesis of breast cancer.
- To explore change in tumor size.

5. Study Design

5.1. Summary of Study Design

Study I3Y-MC-JPBM (JPBM) is a multicenter, randomized, double-blind, placebo-controlled, Phase 3 study evaluating treatment of abemaciclib with NSAI or placebo with NSAI in postmenopausal women with hormone receptor-positive (HR+), HER2- locoregionally recurrent (not amenable to curative therapy) or metastatic breast cancer who have not received prior systemic therapy in this disease setting.

Figure JPBM.5.1 illustrates the study design.



Abbreviations: HER2- = human epidermal growth factor receptor 2-negative; HR+ = hormone receptor-positive; N = number; NSAI = nonsteroidal aromatase inhibitor; PD = progressive disease; R = randomized.

Figure JPBM.5.1. Illustration of study design.

Approximately 450 patients will be randomized 2:1 between the 2 arms:

- Experimental Arm A: abemaciclib 150 mg orally every 12 hours on Days 1 to 28 plus either anastrozole 1 mg or letrozole 2.5 mg once daily of a 28-day cycle
- Control (Placebo) Arm B: placebo orally every 12 hours on Days 1 to 28 plus either anastrozole 1 mg or letrozole 2.5 mg once daily of a 28-day cycle.

Patients will be randomized using the following stratification factors: nature of disease (visceral metastases versus bone-only metastases versus other) and prior (neo)adjuvant endocrine therapy (aromatase inhibitor therapy versus other versus no prior endocrine therapy). The presence of visceral metastases refers to lung, liver, pleural, peritoneal, or adrenal gland involvement at the time of randomization. Prior (neo)adjuvant endocrine therapy refers to aromatase inhibitor

therapy (e.g., anastrozole, exemestane and letrozole) versus other (e.g., tamoxifen and fulvestrant) versus no prior endocrine therapy.

5.2. Determination of Sample Size

The study will enroll approximately 450 patients in 2:1 randomization.

A 2-look group-sequential design of the primary endpoint of investigator-assessed PFS will be used to accommodate an event-driven plan for the interim and final PFS analyses (see Section 6.7.2 for details). There is 1 planned interim analysis and 1 final analysis for PFS in this study. The interim analysis is planned to take place after approximately 189 investigator-assessed PFS events have occurred. The primary PFS analysis will be performed after 240 PFS events have occurred. The cumulative 1-sided type I error rate of .025 will be maintained using the method described in Section 6.7.2.2. Assuming a hazard ratio (HR) of 0.67, this sample size yields more than 80% statistical power to detect superiority of the abemaciclib plus NSAI arm over the placebo plus NSAI arm with the use of a 1-sided log-rank test and a type I error of 0.025. If the true median PFS for the placebo plus NSAI arm is 10 months, then the HR of 0.67 amounts to an approximately 5-month (50%) improvement in median PFS for the abemaciclib plus NSAI arm under an additional assumption of exponential survival distribution.

6. A Priori Statistical Methods

6.1. General Considerations

6.1.1. Populations

The **entered** population includes all patients who sign the informed consent document.

The **enrolled** or **intent-to-treat (ITT)** population includes all randomized patients.

The **visceral disease (VIS)** population includes all randomized patients with visceral disease at baseline as per CRF.

The **randomized and treated (RT)** population includes all randomized patients who received at least one dose of abemaciclib, placebo, or NSAID.

Unless otherwise noted, all disposition analyses will be performed on the entered population, all patient characteristic and efficacy analyses will be performed on the ITT population, and all safety and exposure analyses will be performed on the RT population. An alpha controlled OS analysis will be performed on the VIS population.

All analyses will be performed by treatment arm. Unless otherwise noted, all analyses on the ITT population will be performed by assigned treatment arm and all analyses on the RT population will be performed by actual treatment received.

6.1.2. Definitions and Conventions

Study drug refers to abemaciclib or placebo.

Study treatment refers to abemaciclib + NSAID or placebo + NSAID.

The **date of randomization** is the date the patient was randomly assigned to abemaciclib + NSAID arm or placebo + NSAID arm using the interactive web response system (IWRS).

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

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

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 08JUN2014 and the date of first dose was 06JUN2014, 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 05JUN2014 and the date of first dose was 06JUN2014, 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.

6.2. 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.

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, rescreened after screen failure, enrolled in the study, treated in the study, reasons for discontinuation from study treatment (RT 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 and Performance Status

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

- race
- ethnicity
- age
- country

- height
- weight
- body mass index (BMI)
- baseline Eastern Cooperative Oncology Group (ECOG) performance status (PS).

6.4.2. Baseline Disease Characteristics

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

- study entry diagnosis
- disease extent at study entry (de novo metastatic disease vs metastatic recurrent disease vs locoregionally recurrent disease)
- prior (neo)adjuvant endocrine therapy (aromatase inhibitor vs other vs none)
- prior (neo)adjuvant chemotherapy (yes or no)
- Disease Free Interval (DFI)
- nature of disease (visceral metastases, bone only metastases, or other)
- measurable vs bone only non-measurable disease
- number of organs involved (1, 2, or 3+)
- metastatic site
- estrogen receptor status
- progesterone receptor status.

Prior (neo)adjuvant endocrine therapy will be reported directly from the 'Prior Neo(adjuvant) Endocrine Therapy' electronic case (clinical) report form (eCRF).

Prior (neo)adjuvant chemotherapy will be reported based on data reported on the 'Prior Systemic Therapy' eCRF. Chemotherapy includes both cytotoxic and targeted agents.

Disease free interval (DFI) is defined as the period of time between the completion of adjuvant endocrine therapy and disease recurrence for those patients who received adjuvant endocrine therapy. It will be calculated based on data reported on the 'Prior Systemic Therapy' eCRF. It will be calculated as the progressive disease date – end date of therapy + 1. If only the month and year of a treatment date or progression date is available, the day will be imputed to the 15th. The median and range will be reported.

Nature of disease will be reported directly from the 'Nature of Disease' eCRF. Disease measurability and number of organs involved will be derived from the 'Target Tumor Identification and Results' and 'Non-Target Tumor Identification and Results' eCRFs 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 non-target lesions.

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 (endocrine therapy vs chemotherapy) and specific therapy. Frequency of each specific therapy will be tabulated within each type of therapy.

6.4.5. 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.

6.5. Treatment Compliance

Treatment compliance of abemaciclib/placebo will be measured by pill counts and summarized by cycle. Within 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). The total assigned dose for a patient with no adjustments, omissions, or extensions for logistical reasons is $150 \text{ mg per dose} \times 2 \text{ doses per day} \times 28 \text{ days} = 8400 \text{ mg}$.

Treatment compliance of NSAID will be measured based on pill count data (for those sites using Lilly supplied NSAID) or data reported on 'Exposure Compliance: NSAID' form (for those sites using locally supplied NSAID), and will be summarized by cycle. Within each cycle, compliance will be calculated as the ratio of the number of doses taken to the total number of assigned doses (minus any dose adjustments and doses omitted/withheld for medical or logistical reasons). The total assigned dose for a patient with no adjustments, omissions, or extensions due to logistical reasons is $1 \text{ dose per day} \times 28 \text{ days} = 28 \text{ doses}$.

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 for the ITT population using the preferred name.

6.7. Efficacy Analyses

6.7.1. General Considerations

6.7.1.1. Population

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

6.7.1.2. Stratification Factors

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

- nature of disease (visceral metastases vs bone only metastases vs other)
- prior (neo)adjuvant endocrine therapy (aromatase inhibitor vs other vs none).

The stratification factors are captured in the IWRS and on electronic clinical (case) report forms (eCRFs). Unless otherwise specified, all stratified analyses will be based on the stratification factors per IWRS. A cross tabulation of the frequency of each level of each stratification factor per IWRS and eCRF will be produced.

6.7.1.3. Hypothesis Tests and Confidence Intervals for Efficacy Data

Unless otherwise noted, all hypothesis tests will be performed at the one-sided .025 level and all confidence intervals (CIs) will utilize a 95% confidence level.

6.7.2. Primary Endpoint: Progression-Free Survival

6.7.2.1. Definition

The primary efficacy measure is progression-free survival as defined by RECIST Version 1.1 and determined by the investigator. The PFS time is measured from the date of randomization to the date of objective progression or the date of death due to any cause, whichever is earlier.

If a patient is not known to have progressed or died at the time of analysis, PFS time will be censored at the last known progression-free assessment. The detailed censoring rules are described in the table below ([Table JPBM.6.1](#)).

Table JPBM.6.1. Rules for Determining Date of Progression or Censor for Progression-Free Survival

Rule	Situation	Date of Progression or Censor	Outcome
1	No baseline tumor assessments	Date of Randomization	Censored
2	No post baseline assessments and no death	Date of Randomization	Censored
3	No documented progression and no death (with a post-baseline tumor assessment)	Date of last adequate tumor assessment	Censored
4	Patient lost to follow-up (or withdrew consent from study participation) before documented progression or death	Date of last adequate tumor assessment	Censored
5	Documented progression	Date of documented progression. If a tumor assessment was done on multiple days, use the earliest date for that visit.	Progressed
6	Death without documented progression	Date of death	Progressed
7	Documented progression or death after missing ≥ 2 consecutive post-baseline tumor assessments	Date of last adequate tumor assessment before missed assessments or date of randomization, whichever is later	Censored

Note: Progression-free survival and associated outcome is determined by the earliest of the dates above, if more than 1 situation applies.

6.7.2.2. Hypotheses and Analysis

Letting $S_{LY}(t)$ and $S_P(t)$ denote the progression free survival functions of abemaciclib + NSAI and placebo + NSAI respectively, the null hypothesis

$$H_0: S_{LY}(t) = S_P(t)$$

will be tested against the one sided alternative hypothesis

$$H_1: S_{LY}(t) > S_P(t).$$

There is 1 planned interim analysis and 1 final analysis to test these hypotheses. At each analysis, the hypotheses above will be tested using a one sided stratified log rank test, stratified by nature of disease and prior (neo)adjuvant endocrine therapy.

The interim analysis is planned to take place after approximately 189 investigator-assessed PFS events have occurred. The cumulative 1-sided type I error rate of .025 will be maintained using the following method. At the interim analysis, the nominal alpha level will be .00025. The remaining alpha will be spent at the final analysis. The resulting boundary p-value for the final analysis is dependent on the exact number of events observed at each analysis and can be calculated using the method of Slud and Wei (1982). If the two analyses are performed at exactly 189 and 240 events, then the boundary p-value at the final analysis will be .0249993.

The actual alpha spent and p-value required to reject the null hypothesis will be calculated based on the actual number of events observed at the time of each analysis using software that implements the alpha-spending scheme noted above (for example, ADDPLAN 6.0 or SAS 9.2).

If statistical significance is not declared at the interim PFS analysis, the final PFS analysis will be performed after 240 PFS events have been observed based on investigator assessment. Once statistical significance is declared at either interim analysis or the final analysis, the study will be positive based on the primary endpoint of PFS, and testing of secondary endpoints will proceed.

The interim PFS analyses will be performed by the DMC. The requirements for unblinding the sponsor at the interim analyses are found in Protocol Section 12.2.14.

6.7.2.3. Other Analyses

6.7.2.3.1. Progression-Free Survival Curves and Hazard Ratio

The Kaplan-Meier (KM) method (Kaplan and Meier 1958) will be used to estimate the PFS curve for each treatment arm. Point estimates and confidence intervals for the first quartile, median, and third quartile for the PFS curve of each arm will be estimated. The PFS rates for each arm will be compared at 4 months intervals up to 24 months using a normal approximation for the difference between the rates.

A Cox proportional hazard model (Cox 1972) stratified by nature of disease and prior (neo)adjuvant endocrine therapy with treatment as a factor will be used to estimate the HR between the two treatment arms and the corresponding CI and Wald p-value (Agresti 2002).

6.7.2.3.2. Restricted Mean Difference

The common method for describing benefit on the time scale is to calculate the difference in median event time between the two treatment arms. An alternative method for describing benefit on the time scale is to estimate the average difference between the KM curves. This corresponds to calculating the difference in the average time to event for the two treatment arms (Irwin 1949; Karrison 1997; Meier et al. 2004). Similar to the HR, this method uses all of the available information across the KM curves, but has the additional advantage of assessing benefit on the time scale.

To estimate an improvement in PFS with abemaciclib, we will follow the method of Irwin (1949) detailed in Karrison (1997) and Meier (2004) for estimating the 'difference in average PFS', which we will refer to more formally as the restricted mean difference in PFS.

The area under each KM curve will be calculated using numerical integration (trapezium rule) per Karrison and implemented in SAS using PROC LIFETEST. The difference between treatment arms and a CI for the difference will be formed.

Since the KM curve may be ill-determined beyond a certain range, or even undefined (if the longest observation is censored), for evaluation and comparison of means, the area under each KM curve will be calculated between time 0 and restriction time T, which is why this is referred to as a "restricted mean". Following the suggestion of Karrison, the restriction time T will be chosen as largest time point t such that the standard error (SE) of the survival estimate at time t

in each treatment group is no more than 0.075. For this purpose, we will use the simple, albeit conservative, formula proposed by Peto et al. (1977) for calculating the SE of $S(t)$ as

$SE(S(t)) = S(t)\sqrt{(1-S(t))/n(t)}$, where $n(t)$ is the number of patients still at risk at time t .

6.7.3. Gated Secondary Endpoint: Overall Survival

6.7.3.1. Background

Overall survival is an important secondary endpoint for this study. Statistical tests will be conducted according to the graphical method of Maurer and Bretz (2013) so as to control the overall type I error rate at 0.025 (one-sided) or equivalently, 0.05 (two-sided). The graphical approach can be characterized by first defining a set of prespecified null hypotheses that are organized graphically by providing initial alphas for each hypothesis and weights for each edge of the graph that will determine the propagation of α through the entire hypothesis-testing scheme.

The initial graph allocates all alpha to the primary endpoint of PFS. That is, OS will be tested only if PFS is significant. Following a positive PFS result, the one-sided alpha of .025 will be allocated between the VIS and ITT populations using a graphical approach, with an initial allocation of .005 for the VIS and .020 for the ITT. If the test of OS on the VIS is positive, the test of OS on the ITT will receive all the alpha (Figure JPBM.6.1). Likewise, the test of OS on the ITT is positive, the test of OS on the VIS will receive all the alpha. More details concerning the graphical approach and alpha spending across multiple analyses of OS are provided in Section 6.7.3.3.

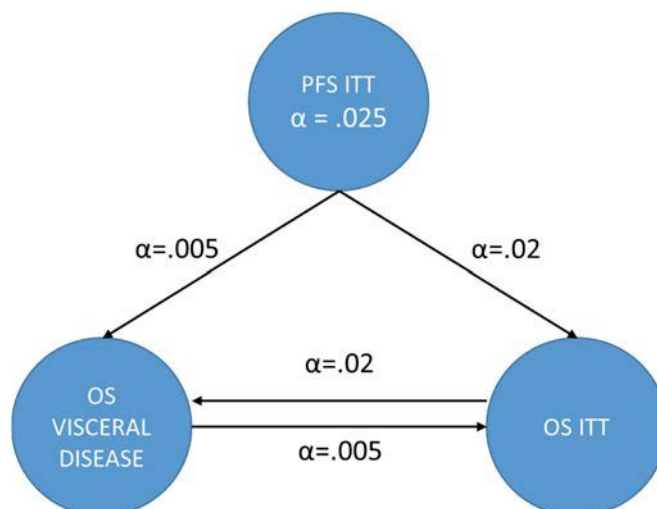


Figure JPBM.6.1. Initial graph.

6.7.3.2. Definition

The OS time is measured from the date of randomization to the date of death from any cause. For each patient who is not known to have died as of the data-inclusion cut-off date for a particular analysis, OS will be censored for that analysis at the date of last contact prior to the data inclusion cutoff date.

6.7.3.3. Hypotheses and Analysis

Letting $S_{LY}(t)$ and $S_P(t)$ denote the survival functions of abemaciclib + NSAI and placebo + NSAI respectively, the null hypothesis

$$H_{0i}: S_{LY}(t) = S_P(t), i=0,1,$$

will be tested against the one sided alternative hypothesis

$$H_{1i}: S_{LY}(t) \geq S_P(t), i=0,1,$$

where $i=0$ represents the hypothesis in the VIS and $i=1$ represents the hypotheses in the ITT. There are 4 planned interim analyses and 1 final analysis to test the null hypotheses which will occur at the following time points:

- The interim PFS analysis (189 PFS events)
- The final PFS analysis (240 PFS events)
- Approximately 189 OS events in the ITT, regardless of the number of events in the VIS
- Approximately 252 OS events in the ITT, regardless of the number of events in the VIS
- Final OS analysis: At least 315 OS events in the ITT and at least 189 events in the VIS

At the time of the final analysis, it is likely that either ITT or the VIS will have more than the prespecified number of events. To properly control Type I error, the analysis of the population with more than the prespecified events will only use those events that occurred through the time

of the prespecified number of events. For example, if at the time of the final analysis, 196 events have been observed in the VIS, the analysis of the VIS will use only data through the date of the 189th event.

At each analysis, the hypotheses above will be tested using a one sided stratified log rank test, stratified by nature of disease and prior (neo)adjuvant endocrine therapy for the ITT, and stratified by prior (neo)adjuvant endocrine therapy for the VIS.

Following a positive PFS result, the one-sided alpha of .025 will be allocated between the VIS and ITT populations using a graphical approach, with an initial allocation of .005 for the VIS and .020 for the ITT. The cumulative 1-sided type I error rate within each population will be maintained using the Lan-Demets method. Specifically, an α -spending function corresponding to the following O'Brien-Fleming type stopping boundary will be used for these interim analyses:

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

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 total alpha to be allocated by the alpha-spending function for each hypothesis is based on the alpha at each node of the graph at the time of analysis (taking into account of potential alpha re-allocation through rejection of other hypothesis). The planned testing levels at each analysis for example scenarios are illustrated in [Table JPBM.6.2](#), and the actual alpha levels will be adjusted according to the rejection history and the alpha-spending functions based on observed number of events. The table assumes the number of OS events in the VIS is 60% of that in the ITT at any given analysis, and that 47 and 95 OS events are observed in the ITT at the time of the interim and final PFS analyses (for example, ADDPLAN 6.0 or SAS 9.2).

Test	Node α at Analysis	Nominal Significance Level				
		Interim PFS Analysis	Final PFS Analysis	189 OS events	252 OS events	Final OS analysis
1. Initial Graph						
PFS	0.025	0.00025	.0249993	n/a	n/a	n/a
2. Graph following significant PFS result						
VIS OS	0.005	1×10^{-9}	3×10^{-7}	.0003	.0016	.0045
ITT OS	0.020	2×10^{-9}	2×10^{-5}	.0027	.0085	.0171
3. Graph following significant PFS and VIS OS result						
ITT OS	0.025	7×10^{-9}	4×10^{-5}	.0038	.0110	.0211
4. Graph following significant PFS and ITT OS result						
VISC OS	0.025	7×10^{-9}	4×10^{-5}	.0038	.0110	.0211

Table JPBM.6.2. Example Nominal Significance Levels

6.7.3.4. Other Analyses

The KM method will be used to estimate the OS curve for each treatment arm. Point estimates and CIs for the first quartile, median, and third quartile for the OS curve of each arm will be estimated. The OS rates at 1, 2, and 3 years for each arm will be estimated and compared using a normal approximation for the difference between the rates.

A stratified Cox proportional hazard model (stratified by nature of disease and prior (neo)adjuvant endocrine therapy) with treatment as a factor will be used to estimate the HR between the abemaciclib + NSAI group and the placebo + NSAI group, and the corresponding CI and Wald p-value.

A restricted mean difference analysis on OS will be conducted as described for PFS in Section [6.7.2.3.2.](#)

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.

6.7.3.5. Exploratory Pooled Overall Survival Analyses

The ITT populations of JPBL and JPBM will be combined to form a pooled population. Overall survival analyses will be performed on this population. The purpose of these analyses is to

evaluate whether adding abemaciclib to the appropriate endocrine therapy improves survival for patients with locally advanced or metastatic disease.

6.7.4. Other Secondary Endpoints

6.7.4.1. Objective Response Rate, Disease Control Rate, and Clinical Benefit Rate

Objective response rate (ORR), disease control rate (DCR), and clinical benefit rate (CBR) are summary measures of best overall response (BOR) as defined by RECIST v1.1. 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 post discontinuation therapy) will be included in the derivation. The one exception includes patients who receive surgery and/or radiotherapy for locally advanced breast cancer. For these patients, only those time point responses occurring prior to surgery/radiotherapy will be included in the derivation.

Each patient's BOR will be categorized as CR, PR, SD, progressive disease (PD), or not evaluable (NE). For patients with bone-only nonmeasurable disease (see Section 6.4.2), BOR will be limited to CR, SD, PD, and NE. Patients with SD will be further classified as SD \geq 6 months or SD <6 months. Stable disease \geq 6 months includes all patients with a best response of SD and a PFS time of \geq 6 months. A BOR of CR or PR will not require confirmation.

Objective response rate is the proportion of patients with a BOR of CR or PR. Clinical benefit rate is the proportion of patients with a BOR of CR or PR, or SD \geq 6 months. Disease control rate is the proportion of patients with a BOR of CR, PR, or SD. Patients with bone-only nonmeasurable disease cannot have a best response of PR, thus ORR will be reported for both the ITT population and the subset of patients with measurable disease.

For each of these rates, point estimates and confidence intervals (using the normal approximation to the binomial) will be calculated by treatment arm. Stratified tests comparing these rates between treatment arms will be conducted using a Cochran-Mantel-Haenszel (CMH) test.

6.7.4.2. Duration of Response

The DoR time is defined only for responders (patients with a BOR of CR or PR). It is measured from the date of first evidence of CR or PR to the date of objective progression or the date of death due to any cause, whichever is earlier. It is calculated as date of progression or death – date of first response evaluation of CR or PR + 1. The DoR will be censored according to the same rules as PFS, with the addition of the following rule: if a patient begins post discontinuation therapy, DoR will be censored on the day of the last response evaluation prior to the initiation of post discontinuation therapy.

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

6.7.5. Sensitivity Analyses

6.7.5.1. Progression-Free Survival

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:

Progression-Free Survival Sensitivity Analysis 1 (censoring for receiving subsequent systemic anticancer therapy): if a patient is initiated on another anticancer therapy prior to objective progression, including any postdiscontinuation treatment systemic therapy, radiotherapy, or surgical intervention, PFS will be censored at the date of the last complete objective progression-free disease assessment before initiation of the new therapy.

Progression-Free Survival Sensitivity Analysis 2 (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.

Progression-Free Survival Sensitivity Analysis 3 (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.

Progression-Free Survival Sensitivity Analysis 4: Progression-free survival 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 0.05 and exit p-value 0.01. The treatment factor will be kept out of the model throughout the covariate selection process and only added into the final model.

In addition, a PFS analysis based on independent central review data will be conducted. Details can be found in the Central Review SAP.

6.7.5.2. Overall Survival

One sensitivity analysis on OS is planned. Overall survival time for this analysis 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 disease related, survival time will be censored at the date of death.

Additional sensitivity analyses may be conducted to evaluate the impact of follow-up systemic therapy with another CDK4/6 inhibitor on OS. Survival times may be censored at the time of follow-up systemic therapy, and the time-dependent Cox model and inverse probability of treatment weighting method to adjust for the impact of follow-up therapies and time-dependent factors on survival might be used (Robins et al. 2000).

6.8. Health Outcomes/Quality-of-Life Analyses

6.8.1. Instruments

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

- the EORTC QLQ-C30
- the EORTC QLQ-BR23
- the EQ 5D 5L.

6.8.2. Quality of Life

Data from the EORTC QLQ-C30 instrument will be scored as described by Aaronson and colleagues (Aaronson et al. 1993). The EORTC QLQ-BR23 data will be scored as described by the EORTC scoring manual (Fayers et al. 2001). The key quality of life items and domains for comparative statistical analysis will be the pain item (item 9), physical functioning domain, and emotional functioning domain from the EORTC-QLQ-C30 questionnaire.

A mixed effects, repeated measures model will be applied to compare treatment arms by cycle with respect to each instrument. The model will include baseline score as a covariate and an unstructured covariance matrix will be utilized. For each instrument, the analysis will include all cycles for which at least 25% of patients in each arm have an assessment.

6.8.3. Health State Utility

The EQ-5D 5L data will be scored as described by van Hout et al. (van Hout et al. 2012). The index score is calculated from a set of item weights to derive a score of 0 to 1, with 1 representing the best health status. Geographic-specific weights will be used as appropriate and when available. The Visual Analog Scale (VAS) is scored from 0 (*worst imaginable health state*) through 100 (*best imaginable health state*) to represent the patient's self-report for each day. The EQ-5D 5L responses for each item will be summarized by frequency and corresponding percentages. Descriptive statistics for the index and VAS will be calculated by arm and cycle. Psychometric analyses, including calculation of reliability coefficients (Cronbach's alpha), will also be performed.

For both the index score and VAS, a mixed effects, repeated measures model will be applied to compare treatment arms by cycle. The model will include baseline score as a covariate and an unstructured covariance matrix will be utilized. The analysis will include all cycles for which at least 25% of patients in each arm have an assessment.

6.8.4. Utilization

Utilization data will be summarized by category across arms. The following categories will be described:

- analgesics (on study treatment and during short term follow up)
- transfusions (on study treatment and during short term follow up)

- surgery (on study treatment and during short term follow up)
- hospitalizations (on study treatment and during short term follow up)
- post discontinuation radiotherapy and systemic therapy.

For categorical variables, frequency and the corresponding proportions will be calculated and tests for differences in proportion between groups will be performed using a chi-squared test. Continuous variables will be described by the mean, median, and standard deviation. A t-test will be used to compare mean utilization.

6.9. Bioanalytical and Pharmacokinetic/Pharmacodynamic Methods

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

6.10. Tailoring 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.

Further analysis of biomarkers will be described in a separate biomarker SAP.

6.11. Safety Analyses

6.11.1. Extent of Exposure

For blinded study drug, extent of exposure will be measured by pill counts and summarized cumulatively. The summary will include total dosage taken and dose intensity. Dose intensity will be calculated as the ratio of total dose taken to the assigned cumulative dose. The assigned cumulative dose for each patient during each cycle is $150 \text{ mg per dose} \times 2 \text{ doses per day} \times 28 \text{ days} = 8400 \text{ mg}$. The assigned cumulative dose while on study is $2 \times 150 \text{ mg} \times \text{number of days on treatment}$.

For co-administered NSAID, extent of exposure will be measured based on pill count data (for those sites using Lilly supplied NSAID) or data reported on 'Exposure Compliance: NSAID' form (for those sites using locally supplied NSAID), and summarized cumulatively. The summary will include total doses taken and dose intensity. Dose intensity will be calculated as the ratio of total doses taken to the assigned number of doses. The assigned number of doses for each patient during each cycle is $1 \text{ dose per day} \times 28 \text{ days} = 28 \text{ doses}$. The assigned number of doses while on study is $1 \text{ dose per day} \times \text{number of days on treatment}$.

Dose adjustments and omissions, along with the reason for adjustment or omission, will be summarized for blinded study drug and NSAID.

6.11.2. Adverse Events

Adverse event (AE) terms and severity grades will be assigned by the investigator using Common Terminology Criteria for Adverse Events (CTCAE) Version 4. In addition, AE verbatim text will also be mapped by the sponsor or designee to corresponding terminology within MedDRA. Adverse events will be reported using the following reporting process:

- The CTCAE Version 4 term reported by the investigator will be mapped to the MedDRA PT and system organ class (SOC) of the corresponding MedDRA lower level term (LLT), unless the reported CTCAE term is ‘Other – specify’.
- If the reported CTCAE term is ‘Other – specify’ the MedDRA LLT, PT and SOC mapped from the verbatim AE term will be used.
- Preferred terms identified by Medical as clinically identical or synonymous will be grouped together under a single consolidated PT. For example, ‘Asthenia’ and ‘Fatigue’ will be reported as ‘Fatigue.’ See [Appendix 1](#) for a complete listing. This listing may be updated prior to database lock as new synonymous terms are observed.
- All listings and summaries will use the PT resulting from this process.

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

A treatment emergent adverse event (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). Comparisons of pre-existing conditions to on-treatment events at the LLT level 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.

To assess the relationship of the AE to the study treatment, the following terminologies are defined (in Protocol Section 8.1.2):

- **Related:** a direct cause and effect relationship between the study treatment and the AE is likely.
- **Possibly related:** a cause and effect relationship between the study treatment and the AE has not been demonstrated at this time and is not probable, but is also not impossible.
- **Unrelated:** without question, the AE is definitely not associated with the study treatment.

As per Lilly's standard operating procedures (SOPs), all "related" and "possibly related" AEs and SAEs will be defined as related to study treatment.

The following TEAE/SAE 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).

The four summaries will be produced for all TEAEs/SAEs and repeated for TEAEs/SAEs related to study treatment.

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 4. These calculated grades will be summarized by cycle and maximum post-baseline grade over the entire study.

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 PFS and OS will be performed for each of following potential prognostic subgroup variables:

- All baseline stratification factors
- nonsteroidal aromatase inhibitor received at cycle 1 (letrozole versus anastrozole)
- Disease setting (de novo metastatic versus recurrent metastatic versus locoregionally recurrent)
- Measurable disease at baseline (yes versus no)
- Number of organs involved (1 versus 2 versus 3+)
- Age (<65 years versus ≥65 years)
- Region (North America, Europe, Asia, and Other)
- Race (Caucasian, Asian, and Other)
- PgR status (positive versus negative)
- Baseline ECOG PS (0 versus 1).

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

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

6.14. Interim Analyses and Data Monitoring

6.14.1. Safety Interim Analyses

The Data Monitoring Committee (DMC) is responsible for providing external oversight of patient safety in Study JPBM independently of the Lilly study team and Lilly Global Product Safety (GPS).

During the study, safety interim analyses will be performed every 3 months. The first safety interim analysis will be triggered by the 90th patient enrolling, with the data cutoff for this analysis occurring 1 month after the trigger. The safety interim analyses will be conducted to evaluate the overall safety profile of abemaciclib when given in combination with NSAIs.

Each safety evaluation will be based, at least, on the following data reports:

- summary of treatment discontinuations and reasons for discontinuation
- summary of SAEs
- summary of TEAEs
- summary of CTCAE graded central laboratory results
- Lilly Safety System reports for all patients with SAEs
- summary of drug exposure and dose adjustments (including delays and reductions) and reasons for adjustment
- listing of treatment discontinuations due to adverse event or death.

Details pertaining to the conduct of these analyses are provided in the JPBM DMC Charter.

6.14.2. Efficacy Interim Analyses

One efficacy interim analysis of PFS and 3 interim analyses of OS are planned, as described in Sections 6.7.2.2 and 6.7.3.3. The interim PFS analysis (and corresponding OS analysis) will be conducted by the DMC. All other efficacy analyses will be conducted by the sponsor.

Rules for unblinding the sponsor at an interim analysis can be found in the protocol.

At the time of the interim PFS analysis, analyses of PFS and OS provided to the DMC will include:

- the boundary values for significance,
- the p-value for a stratified log rank test comparing the two treatment arms, stratified by the randomization factors,
- an estimate of the HR between the two arms based on a Cox proportional hazards model, stratified by the randomization factors, and
- a KM analysis by treatment arm.

Details pertaining to the conduct of these analyses are provided in the JPBM DMC Charter.

6.14.3. Pharmacokinetic/Pharmacodynamic Interim Analyses

A limited number of preidentified individuals independent of the study team may receive access to unblinded data, as specified in the unblinding plan, prior to the interim or final database lock, in order to initiate the final population PK model development processes for interim or final PK and PD analyses. Information that may unblind the study during the analyses will not be reported to study sites or blinded study team until the study has been unblinded. These analyses will be described in a separate document.

6.15. Analyses for the Japanese Regulatory Authority

Analyses conducted specifically for the Pharmaceuticals and Medical Devices Agency (PMDA) will be described in a separate SAP.

6.16. Annual Report Analyses

Annual report analyses, including Developmental Safety Update Report (DSUR) and Investigational Brochure (IB) analyses, are described in the LY2835219 Program Safety Analysis Plan.

6.17. 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, MedDRA PT.
- An AE is considered ‘Serious’ whether or not it is a treatment emergent adverse event (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.

Redacted Version

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 DMC, only the designed Statistical Analysis Center (SAC), who is independent of the sponsor, will perform analyses on unblinded data. For the interim PK analysis to occur prior to the interim/primary PFS analyses, 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. Changes to Planned Analyses in Protocol

Version 1 of the SAP does not contain any significant changes in planned analyses compared to the original protocol.

Version 2 of the SAP does not contain any significant changes in planned analyses compared to Protocol Amendment (a).

Version 3 of the SAP does not contain any significant changes in planned analyses compared to Protocol Amendment (b).

Redacted Version

9. References

- Aaronson NK, Ahmedzai S, Bergman B, Bullinger M, Cull A, Duez NJ, Filiberti A, Flechtner H, Fleishman SB, de Haes JCJM, Kaasa S, Klee MC, Osoba D, Razavi D, Rofe PB, Schraub S, Sneeuw KCA, Sullivan M, Takeda F. The European Organisation for Research and Treatment of Cancer QLQ-C30: A quality-of-life instrument for use in international clinical trials in oncology. *J Natl Cancer Inst.* 1993;85:365-376.
- Agresti A. Categorical Data Analysis. 2nd ed. New York: Wiley, 2002.
- Cox DR. Regression models and life-tables. *J Royal Stat Soc Ser B.* 1972;74(2):187-220.
- CPMP/EWP/2330/99. European Agency for the Evaluation of Medicinal Products; Committee for Proprietary Medicinal Products web site. Points to Consider on Application with 1. Meta-Analyses; 2. One Pivotal Study (adopted by CPMP May 2001). Available at: http://www.ema.europa.eu/docs/en_GB/document_library/Scientific_guideline/2009/09/WC500003657.pdf. Accessed April 09, 2015.
- FDA. Food and Drug Administration web site. Guidance for Industry: Integrated Summary of Effectiveness (2008). Available at: <http://www.fda.gov/downloads/Drugs/GuidanceComplianceRegulatoryInformation/Guidances/UCM079803.pdf>. Accessed April 09, 2015.
- Fayers PM, Aaronson NK, Bjordal K, Groenvold M, Curran D, Bottomley A, on behalf of the EORTC Quality of Life Group. *The EORTC QLQ-C30 Scoring Manual (3rd Edition)*. Published by: European Organisation for Research and Treatment of Cancer, Brussels 2001.
- Glimm E, Maurer W, Bretz F. Hierarchical testing of multiple endpoints in group-sequential trials. *Stat Med.* 2010;29(2):219-228.
- Irwin JO. The Standard Error of an Estimate of Expectational Life. *J Hygiene*, 1949;47:188-189.
- Kaplan EL, Meier P. Nonparametric estimation of incomplete observations. *J Amer Stat Assoc.* 1958;53:457-481.
- Karrison T. Use of Irwin's Restricted Mean as an Index for Comparing Survival in Different Treatment Groups-- Interpretation and Power Considerations, Controlled Clinical Trials. 1977;18:151-167.
- Maurer W, Bretz F. Multiple testing in group sequential trials using graphical approaches. *Stat Biopharmaceut Res.* 2013;5(4):311-320.
- Meier P, Karrison T, Chappell R, Xie H. The Price of Kaplan-Meier. *J Amer Stat Assoc.* 2004;99:467:890-896.
- Peto R, Pike MC, Armitage P, Breslow NE, Cox DR, Howard SV, Mantel N, McPherson K, Peto J, Smith PG. Design and analysis of randomized clinical trials requiring prolonged observation of each patient. II. Analysis and examples. *Br J Cancer.* 1977;35(1):1-39.
- Robins JM, Hernan MA, Brumback B. Marginal Structural Models and Causal Inference in Epidemiology. *Epidemiology.* 2000;11:550-560.
- Slud E, Wei LJ. Theory and method: two-sample repeated significance tests based on the modified Wilcoxon statistic. *J Amer Stat Assoc.* 1982;77(380):862-868.

van Hout B, Janssen MF, Feng YS, Kohlmann T, Busschbach J, Golicki D, Lloyd A, Scalone L, Kind P, Pickard AS. Interim scoring for the EQ-5D-5L: mapping the EQ-5D-5L to EQ-5D-3L value sets. *Value Health*. 2012;15(5):708-715.

Redacted Version

10. Appendices

Redacted Version

Appendix 1. Abemaciclib Consolidated AE Terms

Redacted Version

Consolidated Term / Consolidated Body System	CTCAE Preferred Terms
Abdominal pain / Gastrointestinal disorders	Abdominal discomfort, Abdominal pain, Abdominal pain lower, Abdominal pain upper, Gastrointestinal pain
Anaemia / Blood and lymphatic system disorders	Anaemia, Haematocrit decreased, Haemoglobin decreased, Red blood cell count decreased
Breast pain / Reproductive system and breast disorders	Breast discomfort, Breast pain
Cardiac failure / Cardiac disorders	Cardiac failure, Cardiac failure congestive
Conjunctival haemorrhage / Eye disorders	Conjunctival haemorrhage, Eye haemorrhage
Conjunctivitis / Infections and infestations	Conjunctivitis, Eye infection
Depression / Psychiatric disorders	Depressed mood, Depression
Ear infection / Infections and infestations	Ear infection, Otitis media
Ear pain / Ear and labyrinth disorders	Ear discomfort, Ear pain
Eye pain / Eye disorders	Eye pain, Ocular discomfort
Fatigue / General disorders and administration site conditions	Asthenia, Fatigue
Folate deficiency / Metabolism and nutrition disorders	Anaemia folate deficiency, Folate deficiency
Gastroenteritis / Infections and infestations	Gastric infection, Gastroenteritis, Gastroenteritis viral, Gastrointestinal infection
Haematuria / Renal and urinary disorders	Blood urine present, Haematuria
Hot flush / Vascular disorders	Flushing, Hot flush
Hypercalcaemia / Metabolism and nutrition disorders	Blood calcium increased, Hypercalcaemia
Hypercholesterolaemia / Metabolism and nutrition disorders	Blood cholesterol increased, Hypercholesterolaemia
Hyperglycaemia / Metabolism and nutrition disorders	Blood glucose increased, Hyperglycaemia
Hyperkalaemia / Metabolism and nutrition disorders	Blood potassium increased, Hyperkalaemia
Hypermagnesaemia / Metabolism and nutrition disorders	Blood magnesium increased, Hypermagnesaemia
Hypernatraemia / Metabolism and nutrition disorders	Blood sodium increased, Hypernatraemia
Hyperphosphataemia / Metabolism and nutrition disorders	Blood phosphorus increased, Hyperphosphataemia
Hypertension / Vascular disorders	Blood pressure diastolic increased, Blood pressure increased, Blood pressure systolic increased, Hypertension
Hypertriglyceridaemia / Metabolism and nutrition disorders	Blood triglycerides increased, Hypertriglyceridaemia
Hypoalbuminaemia / Metabolism and nutrition disorders	Blood albumin decreased, Hypoalbuminaemia
Hypocalcaemia / Metabolism and nutrition disorders	Blood calcium decreased, Hypocalcaemia, Calcium deficiency
Hypoglycaemia / Metabolism and nutrition disorders	Blood glucose decreased, Hypoglycaemia
Hypokalaemia / Metabolism and nutrition disorders	Blood potassium decreased, Hypokalaemia
Hypomagnesaemia / Metabolism and nutrition disorders	Blood magnesium decreased, Hypomagnesaemia
Hyponatraemia / Metabolism and nutrition disorders	Blood sodium decreased, Hyponatraemia
Hypophosphataemia / Metabolism and nutrition disorders	Blood phosphorus decreased, Hypophosphataemia

Consolidated Term / Consolidated Body System	CTCAE Preferred Terms
Hypotension / Vascular disorders	Blood pressure diastolic decreased, Blood pressure decreased, Blood pressure systolic decreased, Hypotension
Intestinal obstruction / Gastrointestinal disorders	Gastrointestinal obstruction, Intestinal obstruction, Large intestinal obstruction, Small intestinal obstruction
Iron deficiency / Metabolism and nutrition disorders	Blood iron decreased, Iron deficiency
Leukocytosis / Blood and lymphatic system disorders	Leukocytosis, White blood cell count increased
Leukopenia / Blood and lymphatic system disorders	Leukopenia, White blood cell count decreased
Lung infection / Infections and infestations	Lung infection, Pneumonia
Lymphocytosis / Blood and lymphatic system disorders	Lymphocytosis, Lymphocyte count decreased
Lymphopenia / Blood and lymphatic system disorders	Lymphopenia, Lymphocyte count decreased
Myocardial infarction / Cardiac disorders	Acute myocardial infarction, Myocardial infarction
Nephrolithiasis / Renal and urinary disorders	Nephrolithiasis, Ureterolithiasis
Neuropathy / Nervous system disorders	Acute polyneuropathy, Anaesthesia, Axonal neuropathy, Burning sensation, Dysaesthesia, Hypoaesthesia, Neuralgia, Neuritis, Neuropathy peripheral, Paraesthesia, Peripheral sensory neuropathy, Polyneuropathy, Sensory disturbance, Sensory loss, Skin burning sensation, Toxic neuropathy
Neutropenia / Blood and lymphatic system disorders	Neutropenia, Neutrophil count decreased
Oropharyngeal pain / Respiratory, thoracic and mediastinal disorders	Oropharyngeal discomfort, Oropharyngeal pain
Rash / Skin and subcutaneous tissue disorders	Exfoliative rash, Mucocutaneous rash, Rash, Rash erythematous, Rash follicular, Rash generalized, Rash macular, Rash maculo-papular, Rash maculovesicular, Rash morbilliform, Rash papular, Rash papulosquamous, Rash pruritic, Rash vesicular, Vulvovaginal rash
Thrombocytopenia / Blood and lymphatic system disorders	Platelet count decreased, Thrombocytopenia
Tracheitis / Infections and infestations	Tracheitis, Viral tracheitis
Urinary tract infection / Infections and infestations	Cystitis, Urinary tract infection
Urticaria / Skin and subcutaneous tissue disorders	Urticaria, Urticaria papular

Leo Document ID = 5420af9c-d790-4c2e-9dee-53f280cdb1c0

Approver: [REDACTED]
Approval Date & Time: 14-Apr-2017 16:09:31 GMT
Signature meaning: Approved

Redacted Version