Statistical Analysis Plan I3Y-MC-JPCW (1)

eMonarcHER: A Randomized, Double-Blind, Placebo-Controlled Phase 3 Study of Abemaciclib plus Standard Adjuvant Endocrine Therapy in Participants with High-Risk, Node-Positive, HR+, HER2+ Early Breast Cancer Who Have Completed Adjuvant HER2-Targeted Therapy

NCT04752332

Approval Date: 06-Dec-2021

Title Page

Protocol Title: eMonarcHER: A Randomized, Double-Blind, Placebo-Controlled Phase 3 Study of Abemaciclib plus Standard Adjuvant Endocrine Therapy in Participants with High-Risk, Node-Positive, HR+, HER2+ Early Breast Cancer Who Have Completed Adjuvant HER2-Targeted Therapy

Protocol Number: I3Y-MC-JPCW

Compound Number: LY2835219 (Abemaciclib)

Short Title: eMonarcHER: A Phase 3 Study of Abemaciclib plus Standard Adjuvant Endocrine Therapy in Participants with High-Risk, Node-Positive, HR+, HER2+ Breast Cancer

Sponsor Name: Eli Lilly and Company

Legal Registered Address: Indianapolis, Indiana USA 46285

Regulatory Agency Identifier Number(s)

Registry	ID
IND	106100
EudraCT	2020-004035-24

Approval Date: 06-Dec-2021 GMT

Table of Contents

Title Page **1**

Table of Contents **2**

Version history **4**

1. Introduction **5**
1.1. Objectives, Endpoints, and Estimands 5
1.2. Study Design 10

2. Statistical Hypotheses **11**
2.1. Multiplicity Adjustment 11

3. Analysis Sets **12**

4. Statistical Analyses **13**
4.1. General Considerations 13
4.1.1. Definitions 13
4.1.2. Handling of Dropouts or Missing Data 14
4.2. Participant Disposition 14
4.3. Primary Endpoint Analysis 14
4.3.1. Definition of Endpoint(s) 14
4.3.2. Main Analytical Approach 15
4.3.3. Sensitivity Analyses 16
4.3.4. Supplementary Analyses 17
4.4. Key Secondary Efficacy Endpoint Analyses 17
4.4.1. Distant Relapse-Free Survival (DRFS) 17
4.4.2. Definition of Endpoint 17
4.4.3. Main Analytical Approach 17
4.4.4. Another Key Secondary Endpoint, Overall Survival (OS) 18
4.4.5. Supportive Secondary Endpoint 19
CCI [redacted]
4.5.1. Efficacy Endpoint, Disease-Free Survival - Ductal Carcinoma In Situ (DFS-DCIS) 19
4.6. Safety Analyses 20
4.6.1. Extent of Exposure 21
4.6.2. Adverse Events (AEs) 21
4.6.3. Deaths 22
4.6.4. Clinical Laboratory Evaluation 22
4.6.5. Vital Signs and Other Physical Findings 22
4.6.6. Electrocardiograms 23
4.7. Other Analyses 23
4.7.1. Health Economics Including Patient-Reported Outcomes 23
4.7.2. Biomarkers 23
4.7.3. Pharmacokinetic (PK) 23
4.7.4. Patient Characteristics 23
4.7.5. Treatment Compliance 24
4.7.6. Concomitant Therapy 24

4.7.7. Medical Resource Utilization 24
4.7.8. Post-Study Treatment Therapy 24
4.7.9. Subgroup Analyses 24
4.7.10. Analyses of Pandemic Mitigations 25
4.8. Interim Analyses 25
4.8.1. Independent Data Monitoring Committee (DMC) 25
4.8.2. Interim Analyses for Efficacy/Futility 25
4.8.3. Interim Analysis for Safety 26
4.9. Changes to Protocol-Planned Analyses 26

5. Sample Size Determination 27

6. Supporting Documentation 28
6.1. Appendix 1: Clinical Trial Registry Analyses 28
6.2. Appendix 2: Development / Periodic Safety Update Report 29
6.3. Appendix 3: Protocol Deviations 29

7. References 30

Version history

This statistical analysis plan (SAP) for Study I3Y-MC-JPCW (JPCW) is based on the protocol amendment b dated 20 May 2021.

Table 1 SAP Version History Summary

SAP Version	Approval Date	Change	Rationale
1		Not Applicable	Original version

1. Introduction

There are no changes to the analyses described in the protocol. This SAP is an extension of the protocol. It contains additional details about the analysis plan for efficacy, safety, pharmacokinetics (PK), patient-reported outcomes (PROs), and exploratory endpoints.

Summaries of important protocol deviations will be described in an appendix in this document.

Tables, figures, and listings (TFLs) specifications will be contained in a separate document.

This statistical analysis plan for Study JPCW will describe analyses planned for efficacy (excluding PRO), safety and immunogenicity. Note that PROs, PK, and pharmacodynamics [PD]) will be specified in separate SAP addendums.

1.1. Objectives, Endpoints, and Estimands

Objectives	Endpoints
Primary	
• To compare the efficacy of abemaciclib plus physician's choice ET versus placebo plus physician's choice ET in the study population.	• IDFS as defined by the STEEP System (Hudis et al. 2007)
Secondary	
• To compare the efficacy of abemaciclib plus ET versus placebo plus physician's choice ET in the study population.	• OS • DRFS • Incidence of CNS metastases as the first site of disease recurrence
• To assess the safety profile of abemaciclib plus physician's choice ET versus placebo plus ET in the study population.	• Safety, including but not limited to TEAEs, SAEs, hospitalizations, clinical laboratory tests, vital signs, and physical examinations
• To evaluate participant-reported symptoms, function, and global health status/QOL (EORTC QLQ-C30).	• To compare EORTC QLQ-C30 scales between treatment arms
• To evaluate health status in the study population to inform decision modeling for health economic evaluation using the EQ-5D 5L.	• The EQ-5D 5L index score and the single-item health status measure
• To evaluate the PK of abemaciclib.	• Abemaciclib concentrations

Objectives	Endpoints
[redacted]	

Abbreviations: AE = adverse event; CNS = central nervous system; [redacted] DRFS = distant relapse-free survival; [redacted] EORTC QLQ-C30 = European Organization for Research and Treatment of Cancer quality of life questionnaire Core 30; EQ-5D 5L = EuroQOL 5 Dimension 5 Level; ET = endocrine therapy; [redacted] IDFS = invasive disease-free survival; OS = overall survival; PK = pharmacokinetics; [redacted] QOL = quality of life; SAE = serious adverse event; STEEP = Standardized Definitions for Efficacy End Points; TEAE = treatment-emergent adverse events.

[a] The 2 most common participant-felt symptoms from each monarchE treatment arm per Johnston et al. 2020.

Primary Estimand

The primary clinical question of interest is: What is the difference in invasive disease-free survival (IDFS) between Arms A (abemaciclib plus endocrine therapy [ET]) versus Arm B (placebo plus ET) in participants with hormone receptor-positive (HR+), human epidermal growth factor receptor 2-positive (HER2+) breast cancer at high risk of relapse following completion of standard adjuvant therapy?

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

- Population: adult participants with high-risk, node-positive, HR+, HER2+ early breast cancer who have completed standard adjuvant HER2-targeted therapy.
- Endpoint: IDFS, which is defined by the STEEP System (Hudis et al. 2007). IDFS is measured from the date of randomization to the date of first occurrence of:
    - ipsilateral invasive breast tumor recurrence
    - regional invasive breast cancer recurrence
    - distant recurrence
    - death attributable to any cause
    - contralateral invasive breast cancer
    - second primary non-breast invasive cancer

Treatment condition: randomized study intervention (abemaciclib plus ET or placebo plus ET) administered until evidence of disease recurrence or another protocol-defined reason for study intervention discontinuation.

- Intercurrent-event strategies (IES):
    - Study intervention discontinuation prior to disease recurrence is handled with treatment policy strategy, i.e., regardless of whether or not study intervention discontinuation had occurred.
    - Extended time without adequate assessment prior to disease recurrence is handled with treatment policy strategy, i.e. had extended time without adequate assessment.
- Population-level summary measure: hazard ratio of IDFS in abemaciclib plus ET versus placebo plus ET, estimated using a stratified Cox regression model (Cox 1972). Note that statistical significance will be assessed using a stratified log-rank test.

Rationale for IES: The interest lies in the treatment effect without the confounding effect of study intervention discontinuation prior to disease recurrence or extended time without adequate assessment.

- Study intervention discontinuation due to reasons other than disease recurrence/progression is handled with treatment policy as it reflects clinical practice. Time from randomization until disease recurrence/progression regardless of study intervention discontinuation will be considered in analysis.

- Disease recurrence/progression observed after an extended time and without adequate tumor assessment may have occurred much earlier but was not reported because the scheduled assessment was not done. This inadequate observation may introduce bias to IDFS estimates. If extended time without adequate assessment occurs, the participant will be censored and only the time up to the last adequate tumor assessment will be considered in analysis.

Secondary Estimands

Distant Relapse-Free Survival (DRFS)

A secondary research question is: What is the difference in DRFS between Arms A (abemaciclib plus ET) versus Arm B (placebo plus ET) in participants with HR+, HER2+ breast cancer at high risk of relapse following completion of standard adjuvant therapy?

The estimands for the secondary objectives are described by the following attributes:

- Population: participants with HR+, HER2+ node-positive early breast cancer with high risk-of disease recurrence who have completed adjuvant HER2-targeted therapy
- Endpoint: DRFS, which is defined as the time from randomization to distant recurrence or death from any cause, whichever occurs first. For participants who experienced an IDFS event other than distant recurrence or death, assessments will continue to be performed until an event of distant recurrence, death, or study completion, whichever occurs first.
- Treatment condition: randomized study intervention (abemaciclib plus ET or placebo plus ET) administered until evidence of disease recurrence or death or another discontinuation criterion is met (as defined in Protocol Section 7.1), whichever occurs first
- Intercurrent-event strategies (IES):
    - Study intervention discontinuation prior to disease recurrence is handled with treatment policy strategy, i.e., regardless of whether or not study intervention discontinuation had occurred.
    - Extended time without adequate assessment prior to disease recurrence is handled with treatment policy strategy; i.e. had extended time without adequate assessment.
- Population-level summary measure: hazard ratio of DRFS in abemaciclib plus ET versus placebo plus ET, estimated using a stratified Cox regression model (Cox 1972). Note that statistical significance will be assessed using a stratified log-rank test.

Rationale for IES: The interest lies in the treatment effect without the confounding effect of study intervention discontinuation prior to disease recurrence or extended time without adequate assessment.

- Study intervention discontinuation due to reasons other than disease recurrence/progression is handled with treatment policy as it reflects clinical practice. Time from randomization until disease recurrence/progression regardless of study intervention discontinuation will be considered in analysis.
- Disease recurrence/progression observed after an extended time without adequate tumor assessment may have occurred much earlier but is not reported because the scheduled assessment was not done. This inadequate observation may introduce bias to DRFS estimates. If extended time without adequate assessment occurs, the participant will be censored and only the time up to the last adequate tumor assessment will be considered in analysis.

Overall Survival (OS)

The additional secondary research question is: What is the difference in OS between Arms A (abemaciclib plus ET) versus Arm B (placebo plus ET) in participants with HR+, HER2+ breast cancer at high risk of relapse following completion of standard adjuvant therapy?

The estimands for the secondary objectives are described by the following attributes:

- Population: participants with HR+, HER2+ node-positive early breast cancer with high risk of disease recurrence who have completed adjuvant HER2-targeted therapy
- Endpoint: OS, which is defined as the time from randomization until death from any cause
- Treatment condition: randomized study intervention (abemaciclib plus ET or placebo plus ET) administered until evidence of disease recurrence or another discontinuation criterion is met (as defined in Protocol Section 7.1), whichever occurs first
- Intercurrent-event strategies (IES):
    - Study intervention discontinuation prior to disease recurrence is handled with treatment policy strategy, i.e., regardless of whether or not study intervention discontinuation had occurred
    - Post study intervention discontinuation of anticancer therapy prior to death is handled while on treatment strategy, i.e., consider the assessment of the endpoint up until the time that post study intervention discontinuation anticancer therapy is taken.
- Population-level summary measure: hazard ratio of OS in abemaciclib plus ET versus placebo plus ET estimated using a stratified Cox regression model (Cox 1972). Note that statistical significance will be assessed using a stratified log-rank test.

Rationale for IES: The interest lies in whether there is treatment benefit in OS when participants are treated with abemaciclib at an earlier stage and then with other anticancer therapies after disease recurrence on abemaciclib. Study intervention discontinuation due to reasons other than

death and new anticancer therapy taken prior to death are handled with treatment policy as it reflects clinical practice for participants with early breast cancer.

1.2. Study Design

The eMonarcHER Study is a Phase 3 global, randomized, double-blinded, placebo-controlled trial in participants with high-risk, node-positive, HR+, HER2+ early breast cancer who have completed adjuvant HER2-targeted therapy.

Approximately 2450 participants will be enrolled and randomized 1:1 to Arm A (abemaciclib plus ET) or Arm B (placebo plus ET), using the following stratification factors:

The stratification factors are captured in the Interactive Web Response System (IWRS) and are derived from information collected on electronic case report forms (eCRFs). Male patients are stratified as postmenopausal at the time of randomization. Unless otherwise specified, all stratified analyses will be based on the stratification factor per IWRS. A cross tabulation of the frequency of each level of the stratification factor per IWRS and eCRF will be produced.

Blinded study drug (abemaciclib or placebo) will be given continuously for up to 26 cycles (approximately 2 years, where a cycle = 28 days), or until evidence of disease recurrence or another discontinuation criterion is met, whichever occurs first.

Standard approved adjuvant ET is per physician's choice (such as tamoxifen or an aromatase inhibitor, with or without ovarian function suppression per standard practice). Adjuvant treatment with fulvestrant is not allowed at any time during the study.

Endocrine therapy will be taken as prescribed by the investigator during the on-study intervention period (up to 26 cycles, approximately 2 years) and then up to 10 years total duration, as medically indicated.

A detailed description of the study design is contained in the protocol.

2. Statistical Hypotheses

Primary Hypothesis (Arm A versus Arm B): Treatment of participants with high-risk, node-positive, HR+, HER2+ breast cancer (who have completed adjuvant HER2-targeted therapy) with abemaciclib plus ET will provide a clinically meaningful increase in IDFS over treatment with placebo plus ET.

2.1. Multiplicity Adjustment

A gated hypothesis testing procedure will be used to ensure control of the familywise error rate at 0.025 (one-sided) across the key endpoints: IDFS, DRFS, and OS, all in the intent-to-treat (ITT) population. The primary endpoint of IDFS will first be tested; the secondary DRFS will be tested only if statistical significance is achieved for IDFS. And finally, OS will be tested only if IDFS and DRFS achieve statistical significance. Other endpoints will not be error controlled.

3. Analysis Sets

The following populations are defined:

Population	Description	Endpoint
Intent to Treat (ITT)	All randomized participants, regardless of whether they took any doses of study intervention, or if they took the correct treatment. Participants will be analyzed according to the treatment group as randomized and not by actual treatment received.	Baseline, efficacy, and health economics including PRO analyses.
Safety	All randomized participants who received at least 1 dose of any study intervention. Participants will be analyzed according to the first dose of study intervention they actually received, regardless of the arm to which they were randomized.	Safety, including but not limited to TEAEs, SAEs, hospitalizations, clinical laboratory tests, vital signs, and physical examinations.
Pharmacokinetic (PK)	All randomized participants who received at least 1 dose of study intervention and have baseline and at least 1 postbaseline evaluable PK sample. Pharmacokinetic sampling is planned to occur in 20% of study participants.	Abemaciclib concentrations
Per Protocol	Includes all patients who are randomized and treated and do not have any important protocol deviations that could potentially affect the efficacy conclusions of the study.	IDFS, DRFS, and OS as sensitivity analyses

Abbreviations: DRFS = distant relapse-free survival; IDFS = invasive disease-free survival; ITT = intent-to-treat; OS = overall survival; PK = pharmacokinetics; SAE = serious adverse event; TEAE = treatment-emergent adverse event.

4. Statistical Analyses

4.1. General Considerations

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

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

Any change to the data analysis methods described in the protocol will require an amendment only if it changes a principal feature of the protocol. Before unblinding of the aggregate database, minor modifications or clarifications to the data analysis methods may be described and justified in the SAP. Any other change to the data analysis methods described in the protocol, and the justification for making the change, will be described in the SAP and the clinical study report (CSR). Additional exploratory analyses of the data will be conducted as deemed appropriate.

4.1.1. Definitions

Study drug refers to abemaciclib or placebo.

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

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

The **date of first dose** is the date of the first dose of study drug or ET. For patients receiving ET at the time of randomization, date of first dose is the date of randomization.

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

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

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

- the difference between the date of the event or assessment and the date of first dose plus 1 for all events or assessments occurring on or after the day of first dose. For example, if an event occurs on 08 March 2016 and the date of first dose was 06 March 2016, the study day of the event is 3.
- the difference between the date of the event or assessment and the date of first dose for all events or assessments occurring before the day of first dose. For example, if an event occurs on 05 March 2016 and the date of first dose was 06 March 2016, the study day of the event is -1.

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

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

- 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

4.1.2. Handling of Dropouts or Missing Data

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

4.2. Participant Disposition

A detailed description of participant disposition will be provided, including a summary of the number and percentage of participants entered into the study, enrolled in the study, and treated, as well as number and percentage of participants completing the study, as defined in the SAP, or discontinuing prior to study completion (overall and by reason for discontinuation).

4.3. Primary Endpoint Analysis

4.3.1. Definition of Endpoint(s)

The primary endpoint is IDFS as defined by the STEEP System (Hudis et al. 2007). Invasive disease-free survival time is measured from the date of randomization to the date of first occurrence of:

1. ipsilateral invasive breast tumor recurrence
2. regional invasive breast cancer recurrence
3. distant recurrence
4. death attributable to any cause
5. contralateral invasive breast cancer
6. second primary non-breast invasive cancer

Participants for whom no event has been observed will be censored on the day of their last assessment for recurrence or date of randomization if no postbaseline assessment for recurrence occurred. The detailed censoring rules are described in Table 2.

Tumor recurrence locations will be summarized as reported on the assessment for disease recurrence form. In accordance with the STEEP guidelines, events occurring in the same patient within 2 months of each other will be considered simultaneous, and will be classified as worst event, e.g., simultaneous local and distant events will be classified as distant.

In the analysis, the identification of IDFS events (that are not defined by death) is based on the tumor location of the recurrent disease as recorded by the investigator. If a patient has disease recurrence identified by the investigator (DRCRIND = “Yes”) and the corresponding location of tumor recurrence (DRCRLTLC) is local, regional, distant, contralateral, or second primary non-

breast neoplasm, that is defined as an invasive disease event (recurrence of noninvasive breast cancer is not counted as an event).

The date of each tumor recurrence per STEEP criteria is defined as the earliest date of the tumor assessment that confirms the recurrent tumor, using the method of radiographic examination or biopsy/FNA. If multiple tumor recurrences occur, IDFS events date will be the date of first tumor recurrence, or the date of death if no tumor recurrence is identified.

4.3.2. Main Analytical Approach

Table 2 defines the censoring rules to be applied to the IDFS primary analysis.

Table 2 Rules for Determining Date of Event or Censor for Invasive Disease-Free Survival

Situation	Date of Event or Censor	Event/Censor
IDFS event	Date of earliest IDFS event	Event
No IDFS event	Date of last assessment for disease recurrence	Censored
Unless		
IDFS event prior to the randomization date	Date of randomization	Censored
No post baseline disease recurrence assessment	Date of randomization	Censored
IDFS event documented after more than 12 months (+28 days)* following the last disease recurrence assessment or randomization (whichever is later)	Date of last assessment for recurrence prior to the documented IDFS event, or date of randomization (whichever is later)	Censored

Abbreviation: IDFS = invasive disease-free survival.

*12 months (+28 days) is the longest allowed interval between visits in long-term follow up after Year 5 defined by the schedule of activities

The primary IDFS analysis will be performed on the ITT population to test the superiority of Arm A (abemaciclib plus ET) over Arm B (placebo plus ET) and will use the log-rank test stratified by the randomization factors. The IDFS survival curves, and yearly IDFS rates with 95% CI for each treatment arm will be estimated using the Kaplan-Meier method (Kaplan and Meier 1958). Invasive disease-free survival events will be summarized as local, regional, contralateral, distant, second primary non-breast invasive cancer, and death.

One futility-only analysis, 1 efficacy interim analysis, and the final analysis are planned.

One futility-only analysis for IDFS will be conducted when approximately CCI have been observed in the ITT population. Futility should be declared if the observed IDFS hazard ratio is greater than 1.1.

In addition, there is 1 planned efficacy interim analysis to be performed after approximately CCI The final analysis for IDFS in this study will be performed after approximately CCI have been observed in the ITT population. CCI

CCI [redacted] will be maintained using the Lan-DeMets method (DeMets and Lan 1994) with the O'Brien-Fleming type alpha spending function. CCI [redacted] The actual boundaries for the interim and the final analyses will be calculated based on the actual number of events observed at the time of each analysis using the alpha-spending scheme noted above.

The efficacy and futility boundaries and properties of the design are described in the table below.

Table 3 Efficacy Information

CCI [redacted]

Abbreviations: IDFS=invasive disease-free survival; N/A= not applicable

[a] CCI [redacted]

[b] Dependent on the actual number of events observed at each analysis.

4.3.3. Sensitivity Analyses

The following sensitivity analyses will be conducted on the primary endpoint, IDFS:

- A log-rank test without stratification by randomization factors will be performed to test the superiority of abemaciclib plus ET versus placebo plus ET on the ITT population.
- An unstratified Cox proportional hazard model (Cox 1972) with treatment as a factor will be used to estimate the hazard ratio (HR) between the 2 treatment arms and the corresponding CI and Wald p-value.
- A multivariate Cox proportional hazard model constructed by selecting variables among all the potential variables (see specified factors in Section 4.7.9), using a stepwise

selection method with an entry p-value of .05 and an exit p-value of 0.1; the treatment factor will be kept out of the model throughout the covariate selection process and will only be added into the final model.

- Censoring for control arm patients receiving CDK4/6 Inhibitor: If a patient on control arm receives a CDK4/6 inhibitor prior to their first IDFS event, IDFS will be censored at the date of the last disease assessment prior to the CDK4/6 inhibitor start date.
- A stratified log rank test based on the per protocol population.
- A formal evaluation of the proportional hazard assumption for IDFS will be conducted. This will be done visually through inspection of the graph of log(–log[S(t)]) versus log(t) for the 2 treatment groups, as well as a test of the interaction between treatment and log(time) in the proportional hazards model.
- IDFS as defined by STEEP Version 2.0 in the ITT population will be compared between treatment arms using a stratified log-rank test, stratified by the randomization strata (based on IWRS data). IDFS per STEEP Version 2 (removes the second primary non-breast invasive cancer as an event) and is measured from the date of randomization to the date of first occurrence of:

    1. ipsilateral invasive breast tumor recurrence
    2. regional invasive breast cancer recurrence
    3. distant recurrence
    4. death attributable to any cause
    5. contralateral invasive breast cancer

4.3.4. Supplementary Analyses

The numbers and percentages of participants will be summarized by location and site(s) of recurrence as defined in Protocol Appendix 8 Breast Cancer Disease Recurrence for each treatment arm.

4.4. Key Secondary Efficacy Endpoint Analyses

4.4.1. Distant Relapse-Free Survival (DRFS)

4.4.2. Definition of Endpoint

DRFS is defined as the time from randomization to distant recurrence or death from any cause, whichever occurs first. For participants who experienced an IDFS event other than distant recurrence or death, assessments will continue to be performed until an event of distant recurrence, death, or study completion, whichever occurs first.

Assessments will also be performed for participants who discontinue treatment without an IDFS event per STEEP system or who are randomized and never received study intervention.

4.4.3. Main Analytical Approach

The DRFS in the ITT population will be compared between treatment arms using a stratified log-rank test, stratified by the randomization strata (based on IWRS data). The corresponding HR between treatment arms will be estimated using a stratified Cox regression model (Cox 1972), stratified by randomization strata. The DRFS survival curves, and yearly DRFS rates with 95%

CI for each treatment arm will be estimated using the Kaplan-Meier method (Kaplan and Meier 1958).

Hypothesis testing for DRFS will be conducted in a gated fashion. To maintain the experiment-wise type I error rate, DRFS will be hierarchically tested in the following way: only if the tests of IDFS in ITT is significant will DRFS also be tested inferentially for significance (Glimm et al. 2010).

The following sensitivity analyses will be conducted on the supportive secondary endpoint, DRFS:

- A log-rank test without stratification by randomization factors will be performed to test the superiority of abemaciclib plus ET versus placebo plus ET on the ITT population.
- An unstratified Cox proportional hazard model (Cox 1972) with treatment as a factor will be used to estimate the HR between the 2 treatment arms and the corresponding CI and Wald p-value.

4.4.4. Another Key Secondary Endpoint, Overall Survival (OS)

4.4.4.1. Definition of Endpoint

OS is defined as the time from randomization until 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 cut-off date.

4.4.4.2. Main Analytical Approach

OS in the ITT population will be compared between treatment arms using a stratified log-rank test, stratified by the randomization strata (based on interactive web-response system [IWRS] data). The corresponding HR between treatment arms will be estimated using a stratified Cox regression model (Cox 1972), stratified by randomization strata. The OS survival curves, and yearly OS rates with 95% CI for each treatment arm will be estimated using the Kaplan-Meier method (Kaplan and Meier 1958).

Hypothesis testing for OS will be conducted in a gated fashion. To maintain the experiment-wise type I error rate, OS will be hierarchically tested in the following way: only if the tests of IDFS and of DRFS in ITT is significant will OS also be tested inferentially for significance (Glimm et al. 2010).

At each analysis, the null hypothesis: treatment with abemaciclib plus ET is not different from placebo plus ET with respect to OS in ITT population will be tested using a 1-sided stratified log rank test, stratified by the randomization factors.

The cumulative 1-sided type I error rate of .025 will be maintained using the Lan-DeMets method (DeMets and Lan, 1994). Specifically, an α-spending function corresponding to the following O'Brien-Fleming type stopping boundary will be used for each interim efficacy analysis:

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

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

4.4.4.3. Sensitivity Analyses

The following sensitivity analyses will be conducted on the secondary endpoint, OS:

- A log-rank test without stratification by randomization factors will be performed to test the superiority of abemaciclib plus ET versus placebo plus ET on the ITT population.
- An unstratified Cox proportional hazard model (Cox 1972) with treatment as a factor will be used to estimate the HR between the 2 treatment arms and the corresponding CI and Wald p-value.
- Censoring for control arm patients receiving CDK4/6 Inhibitor: If a patient on control arm receives a CDK4/6 inhibitor prior to their first IDFS event, IDFS will be censored at the date of the last disease assessment prior to the CDK4/6 inhibitor start date.
- A stratified log rank test based on the per protocol population.

4.4.5. Supportive Secondary Endpoint

4.4.5.1. Incidence of CNS Metastases as the First Site of Disease Recurrence

Incidence of central nervous system (CNS) metastases as the first site of disease recurrence will be reported as the proportion of ITT participants with CNS metastases as the first invasive-disease event among ITT population by treatment arms. This incidence rate will be reported along with exact confidence intervals (CI: 95%) for each arm.

CCI [redacted]

4.6. Safety Analyses

All patients that receive at least 1 dose of any study treatment will be evaluated for safety and toxicity.

The Medical Dictionary for Regulatory Activities (MedDRA®) Version 23.1 (or higher) will be used when reporting adverse events (AEs) by MedDRA terms. The MedDRA Lower Level Term will be used in the treatment-emergent computation. Treatment-emergent adverse events (TEAEs) will be summarized by System Organ Class (SOC) and by decreasing frequency of Preferred Term (PT) within SOC. Preferred terms identified by Medical as clinically identical or synonymous will be grouped together under a single consolidated PT. For example, "Neutropenia" and "Neutrophil count decreased" will be reported as "Neutropenia."

Safety analyses will include summaries of the following:

- TEAEs, including severity and possible relationship to study intervention as determined by the investigator
- treatment-emergent serious adverse events (SAEs), including possible relationship to study intervention
- AEs leading to dose adjustments for abemaciclib or placebo
- discontinuations from study intervention due to AEs or death
- treatment-emergent abnormal changes in laboratory values
- treatment-emergent abnormal changes in vital signs

4.6.1. Extent of Exposure

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

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

For physician choice of ET, data are reported on Exposure Compliance Endocrine Study Treatment form and will be summarized cumulatively. The summary will include total doses missed since the previous visit and dose intensity. Dose intensity will be calculated as the ratio of total doses missed to the assigned number of doses. The assigned number of doses while on study is 1 dose per day × number of days on treatment.

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

4.6.2. Adverse Events (AEs)

All analyses of AEs will be conducted in the safety population.

The MedDRA PT derived from the verbatim term will be used when reporting AEs by MedDRA terms. The MedDRA Lower Level Term will be used in the treatment-emergent computation. Toxicity grades will be assigned by the investigator using National Cancer Institute - Common Terminology Criteria for Adverse Events (NCI-CTCAE) version 5.0.

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

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

An 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 of the investigator, they may jeopardize the patient and may require medical or surgical intervention to prevent one of the outcomes listed in this definition.

Adverse events of special interest (AESI): Neutropenia, Infections, Diarrhea, Hepatic events (including increases in aspartate aminotransferase [AST] and alanine aminotransferase [ALT]), venous thromboembolisms (VTEs), and interstitial lung disease (ILD)/pneumonitis. Categories of AESI may be modified as the understanding of the safety of abemaciclib increases. The final list of categories will be maintained at both compound and study level and reported in the CSR.

Consolidated AEs are composite AE terms consisting of synonymous PTs to allow meaningful interpretation of the AE data. The final list of consolidated AE categories and PTs will be maintained at both compound and study level and reported in the CSR.

Safety analyses will include summaries of the following:

- overview of TEAEs
- TEAE and consolidated AE by SOC and by decreasing frequency of PT within SOC (including severity and possible relationship to study intervention)
- treatment-emergent SAEs, including possible relationship to study intervention
- AESIs
- AEs leading to dose adjustments
- discontinuations from study intervention due to AEs or death
- treatment-emergent abnormal changes in laboratory values
- treatment-emergent abnormal changes in vital signs

4.6.3. 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 preferred term of the AE. A summary of deaths including reasons for death will be produced. All analyses of death will be conducted in the safety population.

4.6.4. Clinical Laboratory Evaluation

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

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

The analyses of laboratory abnormalities will be conducted in the safety population.

4.6.5. Vital Signs and Other Physical Findings

Temperature, blood pressure, pulse rate, respiration rate, oxygen saturation, weight and Eastern Cooperative Oncology Group Performance Status (ECOG PS) will be summarized by visit.

4.6.6. Electrocardiograms

Local electrocardiograms (ECGs) will be summarized. The summary will classify patients as having normal or abnormal ECG and summarize AEs identified by ECG.

4.7. Other Analyses

4.7.1. Health Economics Including Patient-Reported Outcomes

Health economics including PROs of health-related quality of life analysis details will be provided in an SAP addendum.

4.7.2. Biomarkers

Biomarkers assessed from blood or tissue samples and their relationship with clinical outcomes will be analyzed according to a separate translational research analysis plan

4.7.3. Pharmacokinetic (PK)

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

4.7.4. Patient Characteristics

4.7.4.1. Demographics

Patient demographics will be summarized for all randomized patients. Patient demographics will include gender, race, ethnicity, country, region, age, weight, height, and body mass index (BMI).

4.7.4.2. Baseline Disease Characteristics

Disease characteristics will be summarized and may include the following:

- Pathological diagnosis
- Primary tumor size by radiology prior to any systemic treatment
- Primary tumor size by pathology following definitive surgery
- Number of involved axillary lymph nodes at time of initial diagnosis and at definitive surgery
- Infraclavicular or Ipsilateral Internal Mammary Nodes at time of initial diagnosis and at definitive surgery
- Tumor stage (derived based on pathological results)
- Prior neoadjuvant therapy
- Surgery procedure (i.e., mastectomy or breast conserving surgery)
- Surgical margin status
- Menopausal status
- Hormone receptor status at biopsy and at definitive resection
- HER2 status at biopsy and at definitive resection
- ECOG PS

4.7.4.3. Historical Illnesses

Historical illnesses are clinically relevant events in the past that ended before the screening visit.

Historical illnesses (using PTs from the most current version of MedDRA) will be summarized.

4.7.4.4. Prior Therapies

Prior radiotherapy, surgery, and systemic therapy will be summarized by treatment arm. Prior radiotherapy and surgery will be categorized by reason for regimen. Prior systemic therapies will be categorized by reason for regimen and specific therapy. Frequency of each specific therapy will be tabulated within each reason for therapy.

4.7.5. Treatment Compliance

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

Compliance information for standard ET will be collected for each cycle to obtain the number of doses missed in the previous cycle. The estimate of percent compliance will be done using the same formula/calculation for blinded study drug percent compliance.

4.7.6. 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 using the preferred name by treatment arm.

4.7.7. Medical Resource Utilization

Summaries of hospitalizations, granulocyte-colony stimulating factor use, transfusions, and other supportive medication will be reported descriptively for each treatment arm. Duration of hospital stays and average number of hospital visits will be reported by treatment arm.

4.7.8. Post-Study Treatment Therapy

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

4.7.9. Subgroup Analyses

Subgroup analyses of the primary endpoint, IDFS, will be made to assess consistency of the intervention effect across the following subgroups:

CCI

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.

4.7.10. Analyses of Pandemic Mitigations

There may be times due to exceptional circumstances (for example, Covid-19 pandemic) where it may not be feasible for participants to come to investigator sites for study-required visits. To evaluate the impact of pandemic mitigations, missing visit/values will be summarized. Additional sensitivity analyses may be performed as deemed appropriate.

4.8. Interim Analyses

4.8.1. Independent Data Monitoring Committee (DMC)

Interim analyses for safety and efficacy will be conducted, using unblinded data, under the guidance of an independent DMC. The DMC will consist of at least 3 members, including at least 1 clinician and 1 statistician. The DMC will communicate any recommendations based on interim analysis to the Lilly senior management designee (SMD). If necessary, the SMD may form an Independent Review Committee to review and act upon the recommendations of the DMC. See Sections 4.8.2 and 4.8.3 for details on the conduct of interim analyses. Detailed information on the role of the DMC and frequency of meetings will be provided in the DMC charter separate from this protocol.

4.8.2. Interim Analyses for Efficacy/Futility

CCI

The efficacy interim analysis includes both an efficacy boundary and a futility boundary (see Table 3). The efficacy interim analysis will be conducted to provide early efficacy information and could potentially result in early communication with regulatory agencies. The DMC will be instructed to recommend to the SMD that the sponsor be unblinded if the analysis of IDFS is significant as described above and any additional criteria specified in the DMC charter are met.

If the study futility boundary is met at the futility-only analysis or the efficacy interim analysis, the DMC should recommend that the study be stopped for futility. If the analysis of IDFS is statistically significant or futile at the interim analysis, the DMC will be instructed to recommend to the SMD that the results be released to the sponsor. The SMD may convene an Internal Review Committee (IRC) to review the DMC's recommendation prior to releasing the results or unblinding the study team.

The sponsor has no intent to stop the study if the early efficacy boundary is crossed at the efficacy interim analysis, and all participants will continue follow-up for IDFS and OS until study close. Participants randomized to the control group will not be permitted to cross over to the experimental group, as this will confound the assessment of OS. In addition, participants will remain blinded for the duration of the study unless the criteria for unblinding are met. If the DMC makes a recommendation counter to this at an interim analysis, for example, the DMC recommends crossing all participants over to the experimental treatment, FDA will be consulted before any action is taken, as well as other regulatory agencies if deemed appropriate.

The unblinded analysis, including review of the efficacy along with the safety data, will be conducted by the DMC. Study sites will receive unblinded information about interim results ONLY if they need to know for the safety of their participants.

Unblinding details are specified in a separate unblinding plan document.

4.8.3. Interim Analysis for Safety

The DMC is responsible for providing external oversight of participant safety in eMonarcHER independently of the Lilly study team and Lilly Global Patient Safety (GPS). Safety interim analyses will be reviewed by the DMC at a frequency described in the DMC charter, but no less than approximately every 6 months while participants are still in the on-study intervention period. The safety interim analyses will be conducted to evaluate the overall safety profile of abemaciclib when given in combination with standard ET.

At each safety interim analysis, the DMC may recommend that the trial continue without modifications, continue with specific modifications, or be stopped for safety concerns. There will be no prespecified rules for stopping the trial due to safety concerns. The DMC members will review safety data at each safety interim analysis. If a significant safety signal is identified, the DMC may recommend a protocol amendment, termination of enrollment, and/or termination of study intervention. The recommendations of the DMC will be communicated to the Lilly SMD and, if necessary, an IRC.

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

4.9. Changes to Protocol-Planned Analyses

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

5. Sample Size Determination

Approximately 2450 participants will be randomized 1:1 to Arm A (abemaciclib plus ET) or Arm B (placebo plus ET) using the stratification factors described in Section 1.2.

A group-sequential design of the primary endpoint of IDFS will be used to accommodate an event-driven plan for the interim and final IDFS analyses.

Under additional assumptions below, a total sample size of approximately 2450 participants is required:

Sample size calculations were conducted using Cytel East 6.

6. Supporting Documentation

6.1. Appendix 1: 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 are summarized by treatment arms and by MedDRA PT.
- An AE is considered "Serious" whether or not it is a TEAE.
- An AE is considered in the "Other" category if it is both a TEAE and is not serious.
- For each SAE and "Other" AE, for each term and treatment arms, the following are provided:
    - the number of participants at risk of an event (if certain participants cannot be at risk for some reason, for example, gender-specific AEs, then the number will be adjusted to only include the participants at risk)
    - the number of participants who experienced each event term, and
    - the number of events experienced.
- For each SAE, for each term and treatment arms, the following are also provided for the EudraCT results submission:
    - the number of occurrences (events) causally related to treatment
    - the total number of deaths, and
    - the number of deaths causally related to treatment.
- Consistent with www.ClinicalTrials.gov requirements, a threshold for frequency of "Other" AEs can be implemented rather than presenting all "Other" AEs. For example, "Other" AEs that occur in fewer than 5% of participants in any treatment arms may not be included if a 5% threshold is chosen. The frequency threshold must be less than or equal to the allowed maximum of 5%.
- AE reporting is consistent with other document disclosures for example, the CSR, manuscripts, and so forth.
- A participant flow will be created that will describe
    - Number of participants per treatment arm. Screen failures do not need to be included. Number of participants who did not complete the study per treatment arm. This analysis will be based on study discontinuation, not treatment discontinuation.
    - Reasons participants did not complete the study.

6.2. Appendix 2: Development / Periodic Safety Update Report

The following reports are needed for the Development Safety Update Report (DSUR) or the Periodic Safety Update Report (PSUR):

- Estimated cumulative subject exposure
- Cumulative exposure to investigational intervention by demographic characteristics
- Listing of subjects who died during the DSUR/PSUR period, and
- Discontinuations due to AEs during the DSUR/PSUR period.

6.3. Appendix 3: Protocol Deviations

Important protocol deviations (IPD) that potentially compromise the data integrity and participants' safety will be summarized for the ITT population. These deviations will include deviations that can be identified programmatically and those that can only be identified by the clinical research associate during monitoring. Important protocol deviations are described in the Trial Issue Management Plan (TIMP) within the study Trial Master File.

The list of IPD that could potentially affect the efficacy conclusions of the study will be defined and documented in the TIMP prior to the final database lock in order to identify participants to be excluded from the per protocol population. A summary of all important protocol deviations will be provided.

The per-protocol (PP) population includes all patients who are randomized and treated and do not have any important protocol deviations that could potentially affect the efficacy conclusions of the study. The PP population is a subset of the ITT population that will be defined and finalized prior to database lock for the primary IDFS analysis. The PP population will be used for sensitivity analyses of IDFS, DRFS, and OS.

7. References

Cox DR. Regression models and life-tables. *J Royal Stat Soc Ser B*. 1972;34(2):187-202. https://doi.org/10.1111/j.2517-6161.1972.tb00899.x

DeMets DL, Lan KK. Interim analysis: the alpha spending function approach. *Stat Med*. 1994;13(13-14):1341-1352; discussion 1353-1356. https://doi.org/ 10.1002/sim.4780131308

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

Hudis CA, Barlow WE, Costantino JP, et al. Proposal for standardized definitions for efficacy end points in adjuvant breast cancer trials: the STEEP system. *J Clin Oncol*. 2007;25(15):2127-2132. https://doi.org/10.1200/JCO.2006.10.3523

Johnston SRD, Harbeck N, Hegg R, et al.; monarchE Committee Members and Investigators. Abemaciclib combined with endocrine therapy for the adjuvant treatment of HR+, HER2-, node-positive, high-risk, early breast cancer (monarchE). *J Clin Oncol*. 2020;38(34):3987-3998. https://doi.org/10.1200/JCO.20.02514

Kaplan EL, Meier P. Nonparametric estimation of incomplete observations. *J Amer Stat Assoc*. 1958;53(282):457-481. http://dx.doi.org/10.1080/01621459.1958.10501452

Leo Document ID = cfba2939-ba0e-4f46-bec3-62e814344aa0

Approver: PPD

Approval Date & Time: 06-Dec-2021 19:01:48 GMT

Signature meaning: Approved