NOVARTIS

Clinical Development

LEE011 (ribociclib)

CLEE011F2301 (MONALEESA-3) / NCT02422615

MONALEESA-3: A randomized double-blind, placebo-controlled study of ribociclib in combination with fulvestrant for the treatment of men and postmenopausal women with hormone receptor positive, HER2-negative, advanced breast cancer who have received no or only one line of prior endocrine treatment

Statistical Analysis Plan (SAP) – Final CSR

Author:	Trial Statistician
Document type:	SAP Documentation
Document status:	Final
Release date:	09-Jan-2023
Number of pages:	29

Document History – Changes compared to previous version of SAP.

Version	Date	Changes
1.0	09-Jan-2023	First version

Table of contents

List of tables

List of figures

List of abbreviations

AE	Adverse event
AESI	Adverse events of special interest
ATC	Anatomical Therapeutic Chemical
BOR	Best overall response
CI	Confidence Interval
CBR	Clinical benefit rate
CR	Complete response
CRF	Case Report Form
CRS	Case retrieval sheet
CSR	Clinical study report
CT	Computed tomography
CTCAE	Common Terminology Criteria for Adverse Events
DAR	Dosage administration Record
DI	Dose Intensity
DOR	Duration of Response
ECG	Electrocardiogram
ECOG	Eastern Cooperative Oncology Group
EORTC QLQ-C30	European Organization for Research and Treatment of Cancer Core 30-item Quality of Life Questionnaire
EOT	End of treatment
ER	Estrogen receptor
FAS	Full analysis set
HER2	Human epidermal growth factor receptor 2
HR	Hazard Ratio
HR	Hormone receptor
MBC	Metastatic breast cancer
MedDRA	Medical Dictionary for Regulatory Activities
NCI	National Cancer Institute
NMQ	Novartis MedDRA queries
ORR	Overall response rate
OS	Overall survival
PD	Progressive disease
PDI	Planned dose intensity
PDS	Programming Datasets Specifications
PFS	Progression-free survival
PR	Partial response
PS	Performance Status
PT	Preferred term
QTcF	QT interval corrected by Fridericia method
RDI	Relative dose intensity
RECIST	Response Evaluation Criteria In Solid Tumors

SAE	Serious adverse event
SAP	Statistical Analysis Plan
SD	Stable disease
SD	Standard deviation
SMQ	Standardized MedDRA queries
SOC	System organ class
TFL	Tables Figures and Listings
TBIL	Total Bilirubin

1 Introduction

This document describes the detailed statistical methodology to be used for the final Clinical Study Report (CSR) of study CLEE011F2301, a phase III, randomized, double-blind, placebo-controlled study of ribociclib or placebo in combination with fulvestrant for the treatment of men and postmenopausal women with hormone receptor positive, HER2-negative, advanced breast cancer who have received no or only one line of prior endocrine treatment.

The content of this SAP is based on CLEE011F2301 protocol amendment 4 (Jan 29, 2020). All decisions regarding final analysis, as defined in the SAP document, have been made prior to database lock and unblinding of the study data.

CSR deliverables (shells for tables, figures, listings) and further programming specifications are described in Tables Figures and Listings (TFL) shells and Programming Datasets Specifications (PDS) documents, respectively.

1.1 Study design

This is a randomized, phase III, double blind, placebo controlled, international study to determine the efficacy and safety of treatment with fulvestrant with ribociclib versus fulvestrant with placebo in men and postmenopausal women with HR+, HER2-negative advanced breast cancer. This study will consist of 4 phases: screening (up to 28 days), randomized treatment, post-treatment efficacy follow-up, and post-treatment survival follow-up.

Approximately 660 patients will be randomly assigned to one of the following treatment arms in a 2:1 ratio.

- Experimental arm: fulvestrant (500 mg intramuscular [as two 5 mL injections] on Cycle 1 Days 1 and 15 (C1D1 and C1D15), and on CnD1 thereafter) + ribociclib (600 mg by mouth once daily for three weeks followed by one week break, in a 28-day cycle)
  OR
- Control arm: fulvestrant (500 mg intramuscular [as two 5 mL injections] on Cycle 1 Days 1 and 15 (C1D1 and C1D15), and on CnD1 thereafter) + ribociclib placebo (by mouth once daily for three weeks followed by one week break, in a 28-day cycle)

Randomization will be stratified by the following factors:

1. Lung or liver metastases: (yes vs no)
2. Previous endocrine therapy (A vs B) according to the following definition:

   A) Patients treatment naïve in the metastatic/advanced disease setting:
      i- whose disease relapsed >12 months after completion of (neo)adjuvant endocrine therapy with no subsequent treatment for advanced/metastatic disease,
      OR
      ii- with de-novo advanced/metastatic disease (no prior exposure to endocrine therapy).

   B) Patients who received up to 1 line of treatment for metastatic/advanced disease:
      i- whose disease relapsed on or within 12 months from completion of (neo) adjuvant endocrine therapy, with no subsequent treatment for advanced/metastatic disease,

OR

ii- whose disease relapsed > 12 months from completion of (neo) adjuvant endocrine therapy, and progressed on or after subsequent endocrine treatment for advanced/metastatic disease,

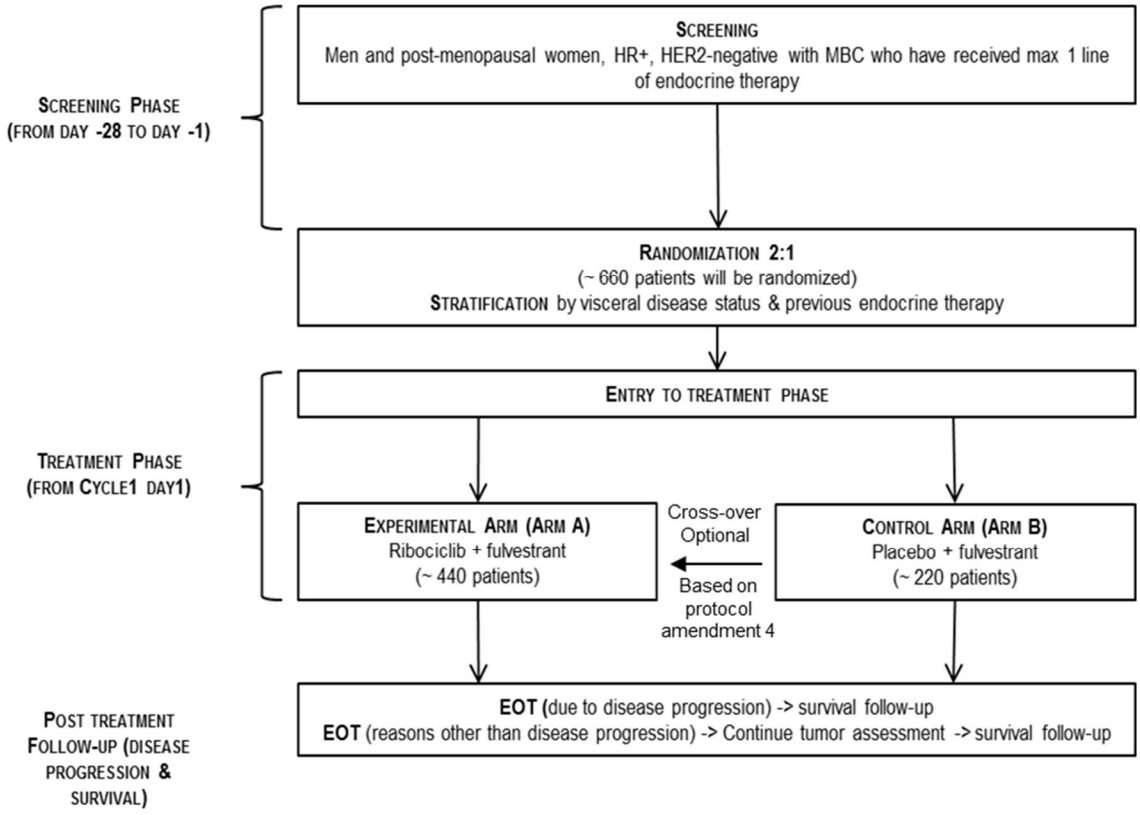
OR

iii- with advanced/metastatic disease at the time of diagnosis that progressed on or after endocrine therapy for advanced/metastatic disease with no prior (neo) adjuvant treatment for early disease.

Following the statistically significant PFS and OS benefit, with protocol amendment 4, study participants were unblinded, with an opportunity for those patients still on study treatment in the placebo combination arm to cross-over to the LEE011 combination arm. Crossover was optional and was conducted only upon documented consent of the study patient.

The study design is summarized in Figure 1-1.

Figure 1-1 Study Design



1.2 Objectives

The study objectives and corresponding endpoints as specified in the protocol are provided in Table 1-1. PFS, as assessed by the local radiologists/investigators and using RECIST 1.1 criteria is the primary endpoint. Overall survival is a secondary endpoint. Both PFS and OS have been tested statistically significant and results have been summarized by previous study reports.

Table 1-1 Study objectives

Objective	Endpoint
Primary	
To compare PFS between ribociclib in combination with fulvestrant to placebo in combination with fulvestrant among men and postmenopausal women with HR+, HER2-negative advanced breast cancer who received no or only one prior endocrine treatment for advanced disease	PFS based on local radiology assessments and using RECIST 1.1 criteria
Secondary	
To compare the two treatment arms with respect to overall survival.	Overall survival
To evaluate the two treatment arms with respect to overall response rate, clinical benefit rate, time to response and duration of response.	ORR as defined by RECIST 1.1. CBR, defined as percentage of patients with CR, PR or SD lasting 24 weeks or longer, TTR, DOR per RECIST 1.1
To evaluate the two treatment arms with respect to time to deterioration of ECOG performance status.	Time to definitive deterioration of ECOG performance status from baseline
To evaluate the safety and tolerability of ribociclib in combination with fulvestrant.	Frequency/severity of AEs, laboratory abnormalities
To evaluate patient reported outcomes for health-related quality of life in the two treatment arms.	Time to 10% deterioration in the global health status/QOL scale score of the EORTC QLQ-C30 Change from baseline in the global health status/QOL scale score of the EORTC QLQ-C30
To characterize the pharmacokinetics (PK) of ribociclib (and relevant metabolites such as LEQ803) when given in combination with fulvestrant.	Concentration by time point for ribociclib (and relevant metabolites such as LEQ803)

Objective	Endpoint
[redacted]	

2 Definitions and general methodology

2.1 Definitions

2.1.1 Study drug and study treatment

Study drug is defined as ribociclib or matching placebo.

Study treatment is defined as ribociclib + fulvestrant, or matching placebo + fulvestrant.

Cross-over open-label ribociclib treatment will refer to ribociclib drug administered to patients randomized to placebo plus fulvestrant arm who subsequently cross-over to ribociclib plus fulvestrant (cross-over period).

Cross-over open-label ribociclib plus fulvestrant treatment will refer to the combination of drugs ribociclib plus fulvestrant administered during the cross-over period.

2.1.2 Date of first administration of study drug

The date of first administration of study drug is defined as the first date when a nonzero dose of study drug is administered and recorded on the dose administration DAR CRF. The date of first administration of study drug will also be referred to as start of study drug. Similar definitions apply for the other components of study treatment.

For analyses related to the cross-over period, date of first administration of cross-over open-label ribociclib treatment is defined as the first date when a non-zero dose of cross-over open-label ribociclib treatment is administered and recorded on the DAR CRF, after the date of cross-over recorded on the cross-over Details CRF.

2.1.3 Date of last administration of study drug

The date of last administration of study drug is defined as the last date when a nonzero dose of study drug is administered and recorded on the DAR CRF. Similar definitions apply for the other components of study treatment.

For analyses related to the cross-over period, date of last administration of cross-over open-label ribociclib treatment is defined as the last date when a non-zero dose of cross-over open-label ribociclib treatment is administered and recorded on the DAR CRF, after the date of crossover recorded on the cross-over Details CRF.

2.1.4 Date of first administration of study treatment

The date of first administration of study treatment is defined as the first date when a nonzero dose of any component of study treatment is administered and recorded on the DAR CRF. The

date of first administration of study treatment will also be referred to as the start of study treatment.

For analyses related to the cross-over period, date of first administration of cross-over open-label ribociclib plus fulvestrant treatment is equal to the date of first administration of cross-over open-label ribociclib treatment.

2.1.5 Date of last administration of study treatment

The date of last administration of study treatment is defined as the last date when a nonzero dose of any component of study treatment was administered and recorded on the DAR CRF.

For analyses related to the cross-over period, date of last administration of cross-over open-label ribociclib plus fulvestrant treatment is defined as the last date when a non-zero dose of any component of the cross-over open-label ribociclib plus fulvestrant treatment was administered and recorded on the DAR CRF during the cross-over period.

2.1.6 Study day

The study day will be calculated as:

- The date of the event (visit date, onset date of an event, assessment date etc.) − reference start date + 1 if event is on or after the reference start date;
- The date of the event (visit date, onset date of an event, assessment date etc.) − reference start date if event precedes the reference start date.

The reference start date for safety assessments (e.g., adverse event onset, laboratory abnormality occurrence, vital sign measurement, dose interruption etc.) is the start of study treatment. (Note: if an adverse event starts before the start of study treatment the study day displayed on the listing will be negative).

The study day will be displayed in data listings.

For analyses related to the cross-over period, study day is defined as above, with the reference date being the date of first administration of cross-over open-label ribociclib treatment.

2.1.7 Baseline

For safety evaluations (e.g., laboratory assessments and ECG), the last available assessment before or at date of start of study treatment will be used as the ‘baseline’ assessment. Assessments specified to be collected post-dose on the first date of treatment are not considered as baseline values.

If patients have no value as defined above, the baseline results will be considered missing.

For cross-over analyses, the last available assessment before or on the date of first dose of cross-over open-label ribociclib treatment will be used as the ‘baseline’ assessment.

2.1.8 On-treatment assessment/event

Safety summaries and selected summaries of deaths will summarize only on-treatment assessments/events.

Double-blind on-treatment period:

An on-treatment assessment/event is defined as any assessment/event in the following time interval:

[date of first administration of study treatment, the earliest between the date of last administration of study treatment + 30 days and the date of first administration of cross-over open-label ribociclib plus fulvestrant treatment – 1 day], i.e., including the lower and upper limits.

The double-blind on-treatment period is the **default** on-treatment period used in the analyses unless specified otherwise.

Cross-over on-treatment period:

An on-treatment assessment/event is defined as any assessment/event in the following time interval: [date of first administration of the cross-over open-label ribociclib plus fulvestrant treatment, date of last administration of the cross-over open-label ribociclib plus fulvestrant treatment + 30 days], i.e., including the lower and upper limits.

(Note: However, the calculation of study treatment duration will use different rules as specified in Section 3.5.1).

An AE started in the screening phase and ongoing in the on-treatment phase will not be considered as an on-treatment AE unless it has worsened.

If the last date of study treatment is missing, any assessment/event occurring after the start of study treatment will be considered as on-treatment.

Data listings will include all assessments/events, flagging those which are not on-treatment.

Note: The date of first administration of study treatment and the date of last administration of study treatment are defined in Sections 2.1.4 and 2.1.5, respectively.

2.1.9 Last contact date

For patients not known to have died at the analysis cut-off, the last contact date will be derived using the last complete date among the following:

Table 2-1 Last contact date data sources

Source data	Conditions
Date of Randomization	No Condition
Last date patient was known to be alive from Survival Follow-up page	Patient status is reported to be alive, lost to follow-up or unknown.

Source data	Conditions
Start/End dates from further antineoplastic therapy	Non-missing medication/procedure term.
Start/End dates from drug administration record	Non-missing dose. Doses of 0 are allowed.
Date of discontinuation/study phase completion from end of treatment disposition page	No condition.
Date of discontinuation/Study phase completion from end of post treatment follow up phase disposition page	No condition
Date of ECG assessment	At least 1 non-missing parameter value
Date of PRO assessment	At least 1 non-missing answer to questionnaire
Tumor (RECIST) assessment date	For non-target lesion: non-missing lesion status For target lesion: non-missing lesion diameter For new lesion: “Is there a new lesion?” yes
Laboratory sample collection date	At least 1 non-missing parameter value
PK collection dates	At least 1 non-missing PK concentration
Vital signs date	At least one non-missing parameter value
Concomitant medication date	At least one non-missing name of medication
[redacted]	[redacted]
Body fluid/Tissue Collection date	Non-missing result (positive/negative tumor cells)
Hospitalization admission/discharge date	Non-missing type of facility admitted to
Cardiac imaging date	Non-missing LVEF or overall interpretation
Performance Status date	Non-missing performance status
Start/End dates of AE	Non-missing verbatim term

The last contact date is defined as the latest complete date from the above list on or before the data cut-off date. The cut-off date will not be used for last contact date, unless the patient was seen or contacted on that date. No dates post cut-off date will be used. Completely imputed dates (e.g., the analysis cut-off date programmatically imputed to replace the missing end date of a dose administration record) will not be used to derive the last contact date. Partial date imputation is allowed for event (death)/censoring coming from ‘Survival information’ eCRF.

2.2 Data included in the analysis

The statistical analyses will be performed using all data collected in the database up to the data cutoff date. Any data collected beyond the cutoff date will not be included in the analysis and will not be used for any derivations.

2.3 Analysis sets

2.3.1 Full analysis set (FAS)

The Full Analysis Set (FAS) consists of all randomized patients. Following the intent-to-treat principle, patients will be analyzed according to the treatment and stratum they were assigned to at randomization.

2.3.2 Safety Set

The Safety Set consists of all patients who received at least one dose of any component of study treatment and have at least one post-baseline safety assessment. Patients will be analyzed according to the treatment actually received. Treatment actually received refers to the treatment the patient was randomized to, unless the alternative treatment was received throughout the trial. If a patient takes at least one dose of the randomized treatment then the treatment actually received is the randomized treatment. The statement that a patient has no AE constitutes a safety assessment. Occurrence of a death constitutes a safety assessment as well.

2.3.3 Cross-over analysis set

The cross-over analysis set includes all patients in the placebo arm of the Safety set, who elected to cross-over to receive ribociclib combination therapy and received at least one dose of cross-over open-label ribociclib.

2.3.4 Patient classification

Patients may be excluded from the analysis sets defined above based on the protocol deviations entered in the database and/or on specific subject classification rules as defined in Table 2-1.

Table 2-1 Patient classification based on protocol deviations and non-protocol deviation criteria

Analysis set	Protocol deviations leading to exclusion	Non-protocol deviation criteria leading to exclusion
FAS	No written informed consent	NA
Safety set	No written informed consent	No post-baseline safety assessment or no dose of study treatment
Cross-over analysis set	No written informed consent	NA

Withdrawal of Informed Consent

Any data collected in the clinical database after a subject withdraws informed consent from all further participation in the trial, will not be included in the analysis data sets. The date on which a patient withdraws full consent is recorded in the eCRF.

3 Statistical methods used in reporting

3.1 Background and demographic characteristics

The FAS will be used for all baseline disease characteristics and demographic summaries and data listings.

3.1.1 Basic demographic and background data

Qualitative data (e.g., race, ECOG performance status, etc.) will be summarized by means of contingency tables by treatment arm and quantitative data (e.g., age, body weight, etc.) will be summarized by appropriate descriptive statistics (mean, standard deviation, median, minimum, and maximum) by treatment arm.

3.2 Protocol deviation summaries

The number and percentage of patients in the FAS with any protocol deviation will be tabulated by deviation category (as specified in the Study Specification Document) and by treatment arm. Additional protocol deviation summaries will be provided to address the potential impact of COVID-19 pandemic. The number and percentage of patients in the FAS with any protocol deviation with relationships to COVID-19 will be summarized by deviation category and by treatment arm.

All protocol deviations will be listed.

3.3 Groupings for analysis

The number and percentage of patients in each analysis set (definitions are provided in Section 2.3) will be summarized by treatment arm.

3.4 Patient disposition

Patient disposition for all randomized patients will be summarized based on FAS. There will be one combined by-treatment summary showing:

1. Number (%) of patients treated/untreated.
2. Number (%) of patients who are still on-treatment (based on the absence of the ‘End of treatment’ page)
3. Number (%) of patients who crossed over and are still on-treatment (based on the ‘Crossover details’ page and absence of the ‘End of treatment’ page)
4. Number (%) of patients who discontinued study treatment (based on the ‘End of Treatment’ page)
5. Number (%) of patients who crossed over and discontinued study treatment (based on the ‘Crossover details’ page and the ‘End of treatment’ page)
6. Reasons for study treatment discontinuation (based on ‘End of Treatment’ page)
7. Number (%) of patients who entered the post-treatment evaluations (based on ‘End of Treatment’ page)
8. Number (%) of patients who discontinued from the post-treatment evaluations (based on ‘End of post treatment follow up disposition’ page)

9. Reasons for discontinuation from the post-treatment evaluations phase (based on 'End of post treatment follow up disposition' page).
10. Number (%) of patients who entered survival follow-up.

3.5 Study treatment

Duration of study treatment exposure, cumulative dose, dose intensity (DI) and relative dose intensity (RDI) will be summarized by treatment. The number of patients with dose reductions/interruptions, and the reasons, will be summarized based on the safety set. Duration of exposure to cross-over open-label ribociclib plus fulvestrant treatment will be listed and the number of patients with dose reductions/interruptions and the reasons, will be listed based on the cross-over analysis set.

Details of the derivations and summaries are provided in the following sections.

3.5.1 Duration of study treatment exposure

The duration of exposure to study treatment will be calculated as

Duration of exposure to study treatment (days) = (last date of exposure to any study treatment component) – (date of first administration of study treatment) + 1.

Duration of exposure to ribociclib/placebo (days) = (last date of exposure to ribociclib/placebo) – (date of first administration to ribociclib/placebo) + 1.

Duration of exposure to fulvestrant (days) = (last date of exposure to fulvestrant) – (date of first administration to fulvestrant) + 1.

The last date of exposure is defined as follows for the study treatment components:

- For ribociclib /placebo: the last date of exposure is defined as the date of last administration of the corresponding medication;
- For fulvestrant, the last date of exposure is defined as following:
  1. If patient discontinues fulvestrant before C2D1 dose, then:
     a. Last date of exposure = last date of administration + 13 days.
     b. If patient died or lost to follow-up within last date of administration + 13 days, then last date of exposure is date of death or last contact date, respectively.
  2. If patient discontinues fulvestrant after C2D1 dose, then:
     a. Last date of exposure = last date of administration + 27 days.
     b. If patient died or lost to follow-up within last date of administration + 27 days, then last date of exposure is date of death or last contact date, respectively.

The duration of exposure includes the periods of temporary interruption (of any component of the study treatment for any reason). The duration of study treatment exposure will be summarized by treatment arm. In addition, the duration of exposure to study treatment will be categorized into time intervals (e.g., <3 months; 3-<6 months; 6-<9 months, etc.); frequency counts and percentages will be presented for the number of patients in each interval.

For analyses related to the cross-over on-treatment period, duration of exposure of cross-over open-label ribociclib plus fulvestrant treatment will be calculated as above considering the

combination of drugs ribociclib plus fulvestrant administered during the cross-over on-treatment period.

3.5.2 Cumulative dose and average daily dose

Cumulative dose for any component of study treatment is defined as the total dose of the medication given during the study treatment exposure.

Average daily dose is defined as [Cumulative dose (mg) / Number of dosing days]; drug free day(s) are not counted as dosing days.

Cumulative dose and average daily dose will be summarized using descriptive statistics by treatment arm for each component of study treatment. Patients with no exposure to the study treatment component will be excluded from the corresponding summary.

3.5.3 Dose intensity and relative dose intensity

Dose intensity (DI) for ribociclib/placebo for patients with non-zero duration of exposure to ribociclib/placebo is defined as follows:

DI (mg / day) = Cumulative dose (mg) / *adjusted* duration of exposure to ribociclib/placebo (day),

where *adjusted* duration of exposure (days) to ribociclib/placebo is the number of ribociclib /placebo dosing days a patient would be expected to have received per protocol, given their duration of exposure. Since ribociclib /Placebo follows a 3 weeks on, 1 week off schedule, the adjusted duration of exposure to ribociclib/placebo is the duration of exposure to ribociclib/placebo minus the planned off days for ribociclib /placebo. The adjusted duration of exposure to ribociclib/placebo (in days) is therefore 21 * (# completed 28 day period) + min(21, duration of last incomplete cycle).

For example, if the duration of exposure to ribociclib/placebo is 66 (corresponding to two cycles and 10 days) days, then the adjusted duration of exposure to ribociclib/placebo is 21*2+10=52 days. If the duration of exposure to ribociclib/placebo is 108 days (corresponding to three cycles and 24 days), then the adjusted duration of exposure is 21*3+21=84 days.

Specifically, let D1 represent the duration of exposure to ribociclib/placebo as defined above Then the adjusted duration of exposure is defined as
D=21*[D1/28] + min(21,D1-28*[D1/28]) days,

where [x] stands for the integer part of x. In this equation [D1/28] is the number of completed cycles, and D1-28*[D1/28] is the additional number of days in the last, incomplete cycle (if any). For example, if D1=30 then [D1/28]=1, D1-28*[D1/28]=2, and D=23. If D1=7 then D=7; if D1=22 then D=21; if D1=28 then D=21, etc.

Planned dose intensity (PDI) is defined as the assigned dose by unit of time planned to be given to patients as per protocol. The PDI for ribociclib/placebo is displayed in Table 3-1. Note that DI will also be calculated and DI for ribociclib/placebo will be reported in the units displayed in Table 3-1, whereas duration of exposure itself will be summarized in months.

Table 3-1 Planned dose intensity

Medication	PDI (dose unit/unit of time)
ribociclib /placebo	600 mg/day (3 weeks on 1 week off)

Ribociclib/placebo relative dose intensity (RDI) is defined as:

RDI = DI (dosing unit / unit of time) / PDI (dosing unit / unit of time).

Fulvestrant RDI is defined as:

RDI = actual cumulative dose / planned cumulative dose, where planned cumulative dose is defined as 500* (# completed D1 in a 28-days cycle) and added by 500 mg if patients completed C1D15.

DI for ribociclib/placebo and RDI for ribociclib/placebo and fulvestrant will be summarized separately for each of the study treatment.

3.5.4 Dose reductions, or interruptions

The number and percentage of patients with dose reductions, interruptions or delays, and the reasons, will be summarized by treatment arm as outlined below for double-blind on-treatment period. The dosage administration record of ribociclib will be listed by cross-over open-label ribociclib plus fulvestrant treatment for the cross-over on-treatment period.

Interruption: An interruption is defined as a 0 mg dose given on one or more days during the period where a patient is not on the "off" part of a treatment cycle, after which > 0mg dose resumes. For patients who had dose interruption checked but never resumed non-zero dose, the dose interruption will not be counted. For example, in the sequence of 600 mg – 0mg (dose break) -0mg (dose interruption) – 0 mg (dose permanently discontinuation) the 0mg (dose interruption) will not be counted as dose interruption. Interruptions will be summarized for each component of the study treatment.

Reduction: A reduction is defined as a decrease from the previous non-zero dose to another non-zero dose less than protocol planned dose, even if this decrease has been directly preceded by an interruption. For example, in the sequence of ribociclib 600mg – 0mg – 400mg, the 400mg dose will be counted as a reduction.

If due to dosing error, a patient took a dose during the dosing break with a dose that is lower than the previous dose, this will not be considered as dose reduction. For example, a patient took 600 mg from day 1-21, and mistakenly took 200 mg per day on day 22-28 which is supposed to be a dosing break. The patient resumed 600 mg dosing on day 29. This will not be considered as dosing reduction.

Dose reductions and interruptions will be tabulated separately. Dose escalations are not allowed according to the protocol and will not be counted in these summaries.

Missing data: If dose is recorded but frequency is missing or entered as 'none', it is assumed that the study drug was taken as per-protocol.

3.5.5 Discontinuation of study treatment components

The reasons for discontinuation of ribociclib/placebo will be summarized by treatment arm for the double-blind on-treatment period, based on the information on the ribociclib/placebo DAR CRF for patients who have the "dose permanently discontinued" box checked.

Partial discontinuation: A partial discontinuation is defined as the event when the last non-zero dose of ribociclib/placebo is more than 21 days before the last non-zero dose of fulvestrant when the permanent discontinuation checkbox is checked in the ribociclib/placebo DAR page. Partial discontinuation of ribociclib/placebo will be summarized by treatment for the double-blind on-treatment period.

The reaons for discontinuation and partial discontinuation will be listed using the cross-over analysis set separately.

3.6 Concomitant and post-treatment therapy

Concomitant therapies

Concomitant therapy is defined as all interventions (therapeutic treatments and procedures) besides the study treatment that were administered to a patient, coinciding with the study assessment period (even if started before the study assessment period).

Concomitant medications will be coded using the World Health Organization (WHO) Drug Reference Listing (DRL) dictionary that employs the WHO Anatomical Therapeutic Chemical (ATC) classification system.

Concomitant medications will be summarized by lowest ATC class, preferred term and treatment arm. The summary will include medications starting on or after the start of study treatment but no later than 30 days after last dose of study treatment or before the start of open-label ribociclib plus fulvestrant treatment.

The safety set will be used for all concomitant medication tables and listings.

Concomitant medications will be listed also for the cross-over period using the cross-over analysis set. Any concomitant therapies starting more than 30 days after the last date of cross-over open-label study treatment will be flagged in the listing.

3.7 Safety evaluation

The assessment of safety will be based mainly on the frequency of adverse events and on the number of laboratory/ECG values that fall outside of pre-determined ranges. Other safety data (e.g., vital signs and special tests) will be considered as appropriate.

All safety outputs will use the safety set. The safety summary tables will include 'double-blinded on-treatment' events/assessments, i.e., those collected on or after the first date of study treatment and collected no later than 30 days after the date of last study treatment administration or the date of the day prior to the start of cross-over open-label ribociclib plus fulvestrant treatment. The AEs started before the first dose but worsening during the treatment period are also considered as 'on-treatment' events.

Key safety analyses that include data from the cross-over on-treatment period will also be performed on the cross-over analysis set. See Section 2.1.8 for more details on the definition of 'on-treatment' event during the cross-over on-treatment period.

3.7.1 Adverse events (AEs)

3.7.1.1 Coding of AEs

Adverse events are coded using the Medical Dictionary for Regulatory Activities (MedDRA) terminology.

3.7.1.2 Grading of AEs

AEs will be assessed according to the Common Terminology Criteria for Adverse Events (CTCAE) version 4.03.

The CTCAE represents a comprehensive grading system for reporting the acute and late effects of cancer treatments. CTCAE v4.03 grading is by definition a 5-point scale generally corresponding to mild, moderate, severe, life threatening, and death.

If CTCAE grading does not exist for an adverse event, grades 1 – 4 corresponding to the severity of mild, moderate, severe, and life-threatening will be used. CTCAE grade 5 (death) will not be used in this project; if an AE results in death it will be documented in the outcome ("fatal"). Information on deaths will also be collected on the 'Death' CRF.

3.7.1.3 General rules for AE Reporting

AE summaries will include all AEs starting on or after study Day 1 (i.e., on or after the day of the first intake of study treatment) and starting no later than 30 days after the last administration of study treatment or the date of the day prior to the start of cross-over open-label ribociclib plus fulvestrant treatment (see Section 2.1.5). All AEs will be listed. AEs starting prior to study Day 1 and AEs starting later than 30 days after the last treatment date or the date of the day prior to the start of cross-over open-label ribociclib plus fulvestrant treatment will be flagged in the listings. AEs starting during the cross-over on-treatment period will be listed separately.

AEs will be summarized by presenting the number and percentage of patients having at least one AE, having at least one AE in each primary system organ class, and for each preferred term using MedDRA coding. A subject with multiple occurrences of an AE will be counted only once in the AE category.

Separate AE summaries will be presented by primary system organ class, preferred term, and maximum grade. A patient with multiple grades for an AE will be summarized under the maximum grade recorded for the event. AE with missing CTCAE grade will be included in the 'All grades' column of the summary tables.

In AE summaries the primary system organ class will be presented alphabetically and the preferred terms will be sorted within primary SOC in descending frequency. The sort order for the preferred term will be based on their frequency in the ribociclib arm.

The frequency of grade 3 and 4 AEs will be summarized separately.

Any information collected (e.g., grades, relationship to study treatment, action taken etc.) will be summarized and listed as appropriate.

3.7.1.4 AE summaries

The following adverse event summaries will be produced by treatment group:

- Summary of deaths and adverse events
- Adverse events, irrespective of causality, by primary system organ class, preferred term and maximum grade
- Adverse events with suspected relationship to study treatment by primary system organ class, preferred term and maximum grade
- Most common grade 3-4 adverse events, irrespective of causality, by preferred term and maximum grade (greater than x% in either arm)
- Adverse events, irrespective of causality, by primary system organ class and maximum grade
- Adverse events, irrespective of causality, by preferred term and maximum grade
- Adverse events with suspected relationship to study treatment by preferred term and maximum grade
- Grade 3 or 4 adverse events, irrespective of causality, by primary system organ class, preferred term and maximum grade
- Grade 3 or 4 adverse events with suspected relationship to study treatment by primary system organ class, preferred term and maximum grade
- On treatment deaths by preferred term
- Deaths, by primary system organ class and preferred term
- Serious adverse events, irrespective of causality, by primary system organ class and preferred term and maximum grade
- Serious adverse events with suspected relationship to study treatment, by primary system organ class, preferred term and maximum grade
- Adverse events leading to study drug discontinuation, irrespective of causality, by primary system organ class, preferred term and maximum grade
- Adverse events leading to study drug reductions, irrespective of causality, by primary system organ class, preferred term and maximum grade
- Adverse events leading to study drug interruptions, irrespective of causality, by primary system organ class, preferred term and maximum grade
- Adverse events requiring additional therapy, irrespective of causality, by primary system organ class, preferred term and maximum grade
- On-treatment deaths and SAEs with fatal outcome, by SOC and PT

AEs of interest will also be summarized. See Section 3.7.1.5 for the grouping details.

3.7.1.5 Grouping of adverse events of special interest

An adverse event of special interest is a grouping of adverse events that are of scientific and medical concern specific to ribociclib. These groupings are defined using MedDRA terms, SMQs (standardized MedDRA queries), HLGTs (high level group terms), HLTs (high level terms) and PTs (preferred terms). Customized SMQs (Novartis MedDRA queries, NMQ) may also be used. A NMQ is a customized group of search terms which defines a medical concept for which there is no official SMQ available or the available SMQ does not completely fit the need. It may include a combination of single terms and/or an existing SMQ, narrow or broad. For each specified AESI, the number and percentage of patients with at least one event of the AESI occurring during the double-blinded on-treatment period will be summarized.

Summaries of these AESIs will be provided by treatment arm, (specifying grade, SAE, relationship, leading to treatment discontinuation, leading to dose adjustment/interruption, etc.).

A Case Retrieval Sheet (CRS) with the exact composition of the AE groupings is to be used to map reported AEs to the AESI groupings. This file may be updated (i.e., it is a living document) based on review of accumulating trial data, and therefore the groupings are also subject to potential change. The most up-to-date version of the CRS will be used at the time of the analysis.

3.7.1.6 Clinical trial safety disclosure

For the legal requirements of ClinicalTrials.gov and EudraCT, two required tables adverse events which are not serious adverse events with an incidence greater than 5% and serious adverse events and SAE suspected to be related to study treatment will be provided by system organ class and preferred term on the safety set and the cross-over analysis set population.

If for a same patient, several consecutive AEs (irrespective of study treatment causality, seriousness and severity) occurred with the same SOC and PT:

- a single occurrence will be counted if there is ≤ 1 day gap between the end date of the preceding AE and the start date of the consecutive AE
- more than one occurrence will be counted if there is > 1 day gap between the end date of the preceding AE and the start date of the consecutive AE

For occurrence, the presence of at least one SAE / SAE suspected to be related to study treatment / non SAE has to be checked in a block e.g., among AE's in a ≤ 1 day gap block, if at least one SAE is occurring, then one occurrence is calculated for that SAE.

The number of deaths resulting from SAEs suspected to be related to study treatment and SAEs irrespective of study treatment relationship will be provided by SOC and PT.

3.7.2 Laboratory data

On analyzing laboratory data, data from all sources (central and local laboratories) will be combined. The summaries will include all laboratory assessments collected no later than 30 days after the last administration of study treatment or the date of the day prior to the start of open-label ribociclib plus fulvestrant treatment.

Laboratory data will be classified (by biostatistics/statistical programming) into CTC grades according to the NCI Common Terminology Criteria for Adverse Events (CTCAE) v4.03. The calculation of laboratory CTC grades will be based on the observed laboratory values only; clinical assessments will not be taken into account. The criteria to assign CTC grades in this study are given in Appendix 1.

For laboratory tests where grades are not defined by CTCAE, results will be graded by the low/normal/high classifications based on laboratory normal ranges.

The following summaries will be produced for the laboratory data (by laboratory parameter):

- Number and percentage of patients with each CTC grade as their worst post-baseline value (regardless of the baseline status). Each patient will be counted only for the worst grade observed post baseline.
- Shift tables using CTC grades to compare baseline to the worst post-baseline value will be produced for hematology and biochemistry laboratory parameters with CTC grades.
- For laboratory parameters where CTC grades are not defined, shift tables to the worst post-baseline value will be produced using the low/normal/high classifications based on laboratory reference ranges.

Number and percentage of patients meeting categorical liver function test criteria, including ALT, AST and ALT/AST (>3x, 5x, 8x, 10x, 20x ULN), Total Bilirubin (>1x, 2x ULN), ALP (>1.5x, 2x, 3x, 5x, 8x, 10x ULN), combined categories of ALT/AST and total bilirubin (e.g., ALT/AST>3x UNL & total bilirubin > ULN) as well as Hy's Law criteria (ALT or AST > 3 x ULN and TBIL >= 2 x ULN and ALP < 2 x ULN). For the combined categories, the assessments need not to be concurrent, i.e., patients are counted based on their most extreme value for each parameter (highest in the case of ALT, AST and TBIL; lowest in the case of ALP). Listing of patients with CTC grade 3 or 4 laboratory abnormalities will be produced and those assessments collected later than 30 days after the last double-blinded study treatment date and before the day of the start of cross-over open-label ribociclib plus fulvestrant treatment will be flagged in the listing.

Separate listing will be provided for the cross-over analysis set and those assessments collected later than 30 days after the last cross-over open-label study treatment date will be flagged in the listing.

3.7.3 Vital signs

Vital signs assessments are performed in order to characterize basic body function. The parameters expected to be collected include: height, weight, body temperature, heart rate, and systolic and diastolic blood pressure.

The criteria for clinically notable abnormalities are defined as follows:

Clinically notable elevated values

- Systolic BP: ≥ 180 mmHg and an increase ≥ 20 mmHg from baseline
- Diastolic BP: ≥ 105 mmHg and an increase ≥ 15 mmHg from baseline.
- Body temperature: ≥ 39.1°C

- Heart rate: ≥ 120 bpm with increase from baseline of ≥15 bpm

Clinically notable below normal values

- Systolic BP: ≤ 90 mmHg and a decrease ≥ 20 mmHg from baseline
- Diastolic BP: ≤ 50 mmHg and a decrease ≥ 15 mmHg from baseline
- Body temperature: ≤ 35°C
- Heart rate: ≤ 50 bpm with decrease from baseline of ≥ 15 bpm

The following summaries will be produced for each vital sign parameter:

- Summary statistic for change from baseline to the worst post-baseline value (in both directions, i.e., from baseline to highest post baseline and from baseline to lowest post baseline value).
- Number and percentage of patients with at least one post-baseline vital sign abnormality (in both directions, i.e., both elevated and below normal values).

In addition, patients with clinically notable vital sign abnormalities will be listed by treatment arm. Patients with notable values starting or worsening during the cross-over on-treatment period will be listed separately. The assessments collected later than 30 days after the last treatment date or after the start of cross-over open-label ribociclib plus fulvestrant treatment will be flagged.

3.7.4 ECG

All analyses of ECG data will be based on the average of all available replicate ECGs at each time point for each patient. For unscheduled assessments, 15-minute windows will be applied to group assessments for averaging.

The following parameters will be assessed: QT, QTcF, QTcB, PR, and QRS intervals in msec, heart rate (bpm), and the overall interpretation if clinically significant abnormalities are present.

- The number and percentage of patients with notable abnormalities will be summarized.
- Summary statistics and shift tables will be presented.
- ECG findings will be listed by treatment
- Notable ECG values starting or worsening during the cross-over on-treatment period will be listed separately.

Table 3-2 Clinically notable ECG values

ECG parameter (unit)	Clinically notable criteria
QT, QTcF, QTcB (ms)	New > 450
	New > 480
	New > 500
	Increase from Baseline > 30
	Increase from Baseline > 60

PR duration (ms)	Increase > 25% from baseline and to PR duration > 200, New >200
QRS duration (ms)	Increase > 25% from baseline and to QRS duration > 110, New >110
Heart Rate (bpm)	< 50 and decrease from Baseline of > 25% > 100 and increase from Baseline of > 25%

A newly occurring ECG abnormality is defined as an abnormal post-baseline ECG finding that is not present at Baseline. Baseline is defined as the average of the last ECG measurements (replicates taken on or before date of first dose of study treatment. The percentage of patients having notable ECG interval values is based on the number of patients at risk for the change with a value at baseline and post-baseline.

3.7.5 Cardiac imaging (MUGA / ECHO)

Shift tables comparing baseline to worst post-baseline cardiac imaging (MUGA or ECHO) overall interpretation will be provided. Percentages will be based on all patients in the Safety set.

Note: If there is any change in the methodology used throughout the study compared to baseline, the post-baseline values for which the methodology differs from baseline will be discarded in the tables presenting comparisons to baseline.

Descriptive statistics of the left ventricular ejection fraction (LVEF) at baseline, worst post-baseline value and change from baseline to worst post-baseline value will be provided.

A listing of patients with newly occurring clinically significant abnormality will be produced by treatment arm.

ECG data of patients with abnormal values starting or worsening during the cross-over on-treatment period will be listed separately.

3.7.6 Urinary Analysis

The following parameter will be summarized using shift table to compare baseline to the worst-post baseline values: urine bilirubin dipstick, urine blood dipstick, urine glucose dipstick, urine ketones dipstick, urine leukocyte dipstick, and urine nitrate dipstick. For all these parameters, both negative and trace are considered as normal and the more pluses the worse. The urine pH dipstick will be summarized using shift table with low/normal/high classifications based on laboratory reference ranges.

3.7.7 Other safety data

Data with notable values from other tests will be listed, and any other information collected will be listed as appropriate.

Notable values collected later than 30 days after the last study treatment date or after the start of cross-over open-label ribociclib plus fulvestrant treatment will be flagged in the listings.

Notable values starting or worsening during the cross-over on-treatment period will be listed separately.

4 Details of the statistical analysis

4.1 Duration of follow-up

Study follow-up will be summarized using the following methods:

- Summary of duration between randomization and cut-off date is defined as follows:
  - Randomization (recruitment) period = (Date of last patient randomized - Date of first patient randomized + 1) / 30.4375 (months)
  - Duration between randomization and data cut-off date = (Cut-off date – Date of randomization + 1) / 30.4375 (months). This item will be summarized overall.

All summaries will be reported in months.

Appendix 1 CTC grades for laboratory values in Novartis Oncology

CTC grades for laboratory values in Novartis Oncology (based on CTCAE v4.03 – June 2010) **Page 1**

				CTC Grades[1]				
Lab test (toxicity)	**SI unit**	**Lab test (NCDS)**	**Normal ranges (Merck manual, July 2015) and *conversion factors***	**0**	**1**	**2**	**3**	**4**
Hematology								
WBC ↓	10^9/L	WBC	**3.9 – 10.7 x 10^9/L**	≥ LLN	< LLN - 3.0 x 10^9/L	< 3.0 – 2.0 x 10^9/L	< 2.0 – 1.0 x 10^9/L	< 1.0 x 10^9/L
WBC [2] (Leukocytosis)	10^9/L	WBC			-	-	> 100 x 10^9/L	-
Hemoglobin [2] (Anemia)	g/L	HGB	**120 - 160 g/L** or 7.4 - 9.9 mmol/L (F) **140 - 170 g/L** or 8.7 – 10.6 mmol/L (M)	≥ LLN	< LLN - 100 g/L < LLN - 6.2 mmol/L	< 100 - 80 g/L < 6.2 - 4.9 mmol/L	< 80 g/L < 4.9 mmol/L	-
Hemoglobin ↑	g/L	HGB	*(16.113 x mmol/L = g/L)*		Increase >0-20 g/L above ULN	Increase >20-40 g/L above ULN	Increase >40 g/L above ULN	-
Platelets ↓	10^9/L	PLAT	**150 - 350 x 10^9/L**	≥ LLN	< LLN - 75.0 x 10^9/L	< 75.0 - 50.0 x 10^9 L	< 50.0 - 25.0 x 10^9/L	< 25.0 x 10^9/L
Neutrophils [3]↓	10^9/L	NEUT		≥2x10^9/L	< 2.0 - 1.5 x 10^9/L	< 1.5 - 1.0 x 10^9/L	< 1.0 - 0.5 x 10^9/L	< 0.5 x 10^9/L
Lymphocytes [3]↓	10^9/L	LYM		≥1.5x10^9/L	< 1.5 - 0.8 x 10^9/L	< 0.8 - 0.5 x 10^9/L	< 0.5 - 0.2 x 10^9/L	< 0.2 x 10^9/L
Lymphocytes ↑	10^9/L	LYM			-	> 4 - 20 x 10^9/L	> 20 x 10^9/L	-
Biochemistry								
AST ↑	U/L	AST	**0 - 35 U/L** or 0 – 0.58 ukat/L *(60 x ukat/L = U/L)*	≤ ULN	> ULN – 3.0 x ULN	> 3.0 - 5.0 x ULN	> 5.0 - 20.0 x ULN	> 20.0 x ULN
ALT ↑	U/L	ALT	**0 - 35 U/L** or 0 – 0.58 ukat/L *(60 x ukat/L = U/L)*	≤ ULN	> ULN – 3.0 x ULN	> 3.0 - 5.0 x ULN	> 5.0 - 20.0 x ULN	> 20.0 x ULN
Total bilirubin ↑	umol/L	BILI	**5.1 – 20.5 umol/L** or 0.3 – 1.2 mg/dL *(17.1 x mg/dL = umol/L)*	≤ ULN	> ULN - 1.5 x ULN	> 1.5 - 3.0 x ULN	> 3.0 - 10.0 x ULN	> 10.0 x ULN
Alk. Phosphatase ↑	U/L	ALP	**36 - 92 U/L** or 0.5 - 1.5 ukat/L *(60 x ukat/L = U/L)*	≤ ULN	> ULN - 2.5 x ULN	> 2.5 - 5.0 x ULN	> 5.0 - 20.0 x ULN	> 20.0 x ULN
Creatinine [4]↑	umol/L	CREAT	**61.9 - 115 umol/L** or 0.7 – 1.3 mg/dL *(88.4 x mg/dL = umol/L)*	≤ ULN	> ULN - 1.5 x ULN	> 1.5 - 3.0 x ULN	> 3.0 - 6.0 x ULN	> 6.0 x ULN
Creatinine kinase [4]↑	U/L	CK	**30 - 170 U/L** or 0.5 – 2.83 ukat/L *(60 x ukat/L = U/L)*	≤ ULN	> ULN - 2.5 x ULN	> 2.5 - 5.0 x ULN	> 5.0 - 10.0 x ULN	> 10.0 x ULN
Albumin [2] (Hypoalbuminemia)	g/L	ALB	**35 - 55 g/L** or 3.5 to 5.5 g/dL	≥ LLN	< LLN - 30 g/L	< 30 - 20 g/L	< 20 g/L	-
Total Cholesterol ↑	mmol/L	CHOL	**3.88 – 5.15 mmol/L** or 150 - 199 mg/dL *(38.67 x mg/dL = mmol/L)*	≤ ULN	> ULN - 7.75 mmol/L > ULN - 300 mg/dL	> 7.75 -10.34 mmol/L > 300 – 400 mg/dL	>10.34-12.92 mmol/L > 400 – 500 mg/dL	>12.92 mmol/L > 500 mg/dL
Lipase ↑	U/L	LIPASE	**<95 U/L** or <1.58 ukat/L *(60 x ukat/L = U/L)*	≤ ULN	> ULN - 1.5 x ULN	> 1.5 - 2.0 x ULN	> 2.0 - 5.0 x ULN	> 5.0 x ULN
Amylase ↑	U/L	AMYLASE	**0 - 130 U/L** or 0 – 2.17 ukat/L *(60 x ukat/L = U/L)*	≤ ULN	> ULN - 1.5 x ULN	> 1.5 - 2.0 x ULN	> 2.0 - 5.0 x ULN	> 5.0 x ULN
Uric acid [2] (Hyperuricemia)	umol/L	URATE	**150 - 470 umol/L** or 2.5 – 8 mg/dL *(59.48 x mg/dL = umol/L)*	≤ ULN	> ULN – 10 mg/dL > ULN – 595 umol/L	-	-	> 10 mg/dL > 595 umol/L

ULN = Upper Limit of Normal range; LLN = Lower Limit of Normal range

LAB - CTC grades in Novartis Oncology (26Oct15)

CTC grades for laboratory values in Novartis Oncology (based on CTCAE v4.03 – June 2010)

Page 2

				CTC Grades[1]				
Lab test (toxicity)	**SI unit**	**Lab test (NCDS)**	**Normal ranges (Merck manual, July 2015) and *conversion factors***	**0**	**1**	**2**	**3**	**4**
Phosphorus[2] (Hypophosphatemia)	mmol/L	PHOS	**0.97 – 1.45 mmol/L** or 3.0 - 4.5 mg/dL *(0.32 x mg/dL = mmol/L)*	≥ LLN	< LLN - 2.5 mg/dL < LLN - 0.8 mmol/L	< 2.5 - 2.0 mg/dL < 0.8 - 0.6 mmol/L	< 2.0 - 1.0 mg/dL < 0.6 - 0.3 mmol/L	< 1.0 mg/dL < 0.3 mmol/L
Calcium (corrected) [2] (Hypercalcemia)	mmol/L	CACALC	**2.2 - 2.6 mmol/L** or 9 - 10.5 mg/dL *(0.2495 x mg/dL = mmol/L)*	≤ ULN	> ULN - 11.5 mg/dL > ULN - 2.9 mmol/L	> 11.5 - 12.5 mg/dL > 2.9 - 3.1 mmol/L	> 12.5 - 13.5 mg/dL > 3.1 - 3.4 mmol/L	> 13.5 mg/dL > 3.4 mmol/L
Calcium (corrected) [2] (Hypocalcemia)	mmol/L	CACALC		≥ LLN	< LLN - 8.0 mg/dL < LLN - 2.0 mmol/L	< 8.0 - 7.0 mg/dL < 2.0 - 1.75 mmol/L	< 7.0 - 6.0 mg/dL < 1.75 - 1.5 mmol/L	< 6.0 mg/dL < 1.5 mmol/L
Magnesium [2] (Hypermagnesemia)	mmol/L	MG	**0.62 – 0.99 mmol/L** or 1.5 – 2.4 mg/dL *(0.4114 x mg/dL = mmol/L)*	≤ ULN	> ULN - 3.0 mg/dL > ULN - 1.23 mmol/L	-	> 3.0 – 8.0 mg/dL > 1.23 – 3.3 mmol/L	> 8.0 mg/dL > 3.3 mmol/L
Magnesium [2] (Hypomagnesemia)	mmol/L	MG		≥ LLN	< LLN - 1.2 mg/dL < LLN - 0.5 mmol/L	< 1.2 - 0.9 mg/dL < 0.5 - 0.4 mmol/L	< 0.9 - 0.7 mg/dL < 0.4 - 0.3 mmol/L	< 0.7 mg/dL < 0.3 mmol/L
Glucose (non-fasting) [2] (Hyperglycemia)	mmol/L	GLUCSN	**<7.8 mmol/L** or <140 mg/dL *(0.05551 x mg/dL = mmol/L)*	≤ ULN	-	> ULN - 250 mg/dL > ULN - 13.9 mmol/L	> 250 - 500 mg/dL > 13.9 - 27.8 mmol/L	> 500 mg/dL > 27.8 mmol/L
Glucose (fasting) [2] (Hyperglycemia)	mmol/L	GLUCSF	**3.9 – 5.8 mmol/L** or 70 - 105 mg/dL *(0.05551 x mg/dL = mmol/L)*	≤ ULN	> ULN - 160 mg/dL > ULN - 8.9 mmol/L	> 160 - 250 mg/dL > 8.9 - 13.9 mmol/L	> 250 - 500 mg/dL > 13.9 - 27.8 mmol/L	> 500 mg/dL > 27.8 mmol/L
Glucose [2] (Hypoglycemia)	mmol/L	GLUCSN/ GLUCSF		≥ LLN	< LLN - 55 mg/dL < LLN - 3.0 mmol/L	< 55 - 40 mg/dL < 3.0 - 2.2 mmol/L	< 40 - 30 mg/dL < 2.2 - 1.7 mmol/L	< 30 mg/dL < 1.7 mmol/L
Potassium [2] (Hyperkalemia)	mmol/L	K	**3.5 - 5.0 mmol/L** *(0.2558 x mg/dL = mEq/L = mmol/L)*	≤ ULN	> ULN - 5.5 mmol/L	> 5.5 - 6.0 mmol/L	> 6.0 - 7.0 mmol/L	> 7.0 mmol/L
Potassium [2] (Hypokalemia)	mmol/L	K		≥ LLN	< LLN - 3.0 mmol/L	-	< 3.0 - 2.5 mmol/L	< 2.5 mmol/L
Sodium [2] (Hypernatremia)	mmol/L	SODIUM	**136 - 145 mmol/L** *(0.435 x mg/dL = mEq/L = mmol/L)*	≤ ULN	> ULN - 150 mmol/L	> 150 - 155 mmol/L	> 155 - 160 mmol/L	> 160 mmol/L
Sodium [2] (Hyponatremia)	mmol/L	SODIUM		≥ LLN	< LLN - 130 mmol/L	-	< 130 - 120 mmol/L	< 120 mmol/L
Triglyceride [2] ↑	mmol/L	TRIG	**< 2.82 mmol/L** or < 250 mg/dL *(0.01129 x mg/dL = umol/L)*	< 150 < 1.71	≥ 150 - 300 mg/dL ≥ 1.71 – 3.42 mmol/L	> 300 - 500 mg/dL > 3.42 – 5.7 mmol/L	> 500 - 1000 mg/dL > 5.7 – 11.4 mmol/L	> 1000 mg/dL > 11.4 mmol/L
Coagulation								
INR [2]↑	1	INR	**0.8 – 1.2**	≤ ULN	> ULN - 1.5 x ULN	> 1.5 - 2.5 x ULN	> 2.5 x ULN	-
Activated partial thromboplastin time[2]↑	sec	APTT	**25 - 35 sec**	≤ ULN	> ULN - 1.5 x ULN	> 1.5 - 2.5 x ULN	> 2.5 x ULN	-
Fibrinogen [4]↓	g/L	FIBRINO	**1.5 – 3.5 g/L** or 150 – 350 mg/dL *(0.01 x mg/dL = g/L)*	≥ LLN	< LLN - 0.75 x LLN	< 0.75 - 0.5 x LLN	< 0.5 - 0.25 x LLN	< 0.25 x LLN

ULN = Upper Limit of Normal range; LLN = Lower Limit of Normal range

(1) = **LAB CTC grades 1, 2, 3, 4 overrule the study specific (central or local) normal range criteria**, e.g. if ULN of Sodium is 151 mmol/L and the value is 151 mmol/L, CTC grade 2 is assigned although the value is ≤ ULN.

(2) = **Life-threatening consequences and/or hospitalization are <u>not</u> considered** for determination of LAB CTC grades 3 and 4. Concomitant usage of anticoagulation therapy (for INR and Fibrinogen) is not considered either.

(3) = Values and LNRs for blood differentials can be given as %, absolute values should then be calculated using WBC. Generally, **≥ 1.5 x 10^9/L (lymphocytes) and ≥ 2 x 10^9/L (neutrophils) are considered as LAB CTC grade 0**

(4) = For Creatinine and Fibrinogen, the **comparison with baseline is <u>not</u> considered** for derivation of LAB CTC grades

NOVARTIS

Clinical Development

LEE011 (ribociclib)

CLEE011F2301 (MONALEESA-3) / NCT02422615

MONALEESA-3: A randomized double-blind, placebo-controlled study of ribociclib in combination with fulvestrant for the treatment of men and postmenopausal women with hormone receptor positive, HER2-negative, advanced breast cancer who have received no or only one line of prior endocrine treatment

Statistical Analysis Plan

Author: [redacted] Trial Statistician

Document type: SAP Documentation

Document status: Final Amendment v2.0

Release date: 11-Oct-2017

Number of pages: 77

Document History – Changes compared to previous version of SAP.

Version	Date	Changes
1.0	8/27/2015	Final version
2.0	6/06/2017	Amendment 1 Title: • Change the title to “statistical analysis plan” to align with new naming of analysis plan. • Update the name of the trial per protocol amendment. Section 1: • Add the protocol amendment the SAP is based on. • Update the naming of RAP M7 and M8. Section 1.1: • Clarify BIRC assessed PFS is a supporting end point for primary analysis. • Clarify the stratification factor definition, patient population, screening window per protocol amendment. Section 1.2: • Remove BIRC assessed PFS as a secondary end point per protocol amendment. • Remove PK parameters as secondary end points per protocol amendment. • [redacted] • [redacted] Section 2.1.7: clarify assessments done post dose on first day of treatment is not considered in safety baseline. Section 2.1.9: update last contact date definition per new SAP template and alignment with ML2 (CLEE011A2301) submission. Section 2.2: Remove PFS interim analysis and update OS interim analysis per protocol. Section 2.3.3, 2.3.5: Modify per protocol set definition to align with ML2 submission. Section 2.3.4: Clarify the definition of PAS. Section 2.4: Replace the prohibited medication table with a reference to the protocol and clarify that the most up to date list will be used for CSR purpose. Describe which analyses are affected by new anti-neoplastic therapies. Section 2.5.3: Add clarification about BOR analysis per protocol amendment and team discussion. Section 2.5.4: Update change in methodology per updated RECIST guidance. Section 2.5.5: • Remove the footnote of randomization as day 0 in Table 2-3 to avoid confusion. • Remove new anti-neoplastic therapy as a censoring reason since the primary analysis doesn’t censoring patients at new cancer therapy.

Version	Date	Changes
		Section 2.5.7, Remove the algorithm for constructing water graphs section per new SAP template. Section 3.2.1: Clarify when the stratification factors are per IRT. Section 3.2.2: Add disease free interval and number of de novo patients to align with ML2 submission. Section 3.2.4: Clarify how surgery, biopsy and radiotherapy will be reported. Section 3.3: add the PD list leading to exclusion from PPS. Section 3.5: • Add the summary for screening phase disposition. Section 3.6.1, • Add the definition of duration of exposure to LEE011/placebo, duration of exposure to fulvestrant to align with ML2 submission. • Update the definition of last date of exposure to fulvestrant to align with X2108 (CLEE011X2108). Section 3.6.3, • Remove dose intensity definition for fulvestrant since it is difficult to interpret. • Amend the definition of adjusted dose intensity for ribociclib/placebo to align with ML2 submission. • Remove the definition of RDI per patient since it is not used anywhere. • Update the definition of RDI for fulvestrant to align with X2108. Section 3.6.4, Align definition for dose interruption, dose delay, dose reduction with ML2 submission with the exception that partial discontinuation will not be considered as dose reduction. Section 3.6.5 Align the definition of partial discontinuation with ML2 submission. Section 3.7: add summary of anti-neoplastic medication after discontinuation by medication type. Section 3.8, Update the language to reflect the change of reading paradigm in BIRC assessment. Section 3.8.1.1, Update the PFS censoring option to match the updated RECIST guidance and add a table to summarize event/censoring cases for primary analysis. Section 3.8.1.4, • Add section audit-based BIRC assessment of PFS per PA (protocol amendment). Add a decision rule triggering full BIRC assessment. • Add derived local radiology assessment as one of the supportive analyses to align with ML2 submission. • Update “central radiology review” to “BIRC assessment” to align with BIRC read paradigm change. • Remove supportive analysis due to forced randomization since forced randomization is not allowed.

Version	Date	Changes
		• Remove IPCW analysis as a supportive analysis per team discussion. • Clarify how multivariate Cox regression will be done. • Clarify how the analysis using CRF based stratification factors will be done. Section 3.8.2, • Remove central assessed PFS as a secondary end point per protocol amendment. • Remove misplaced PK objective. Section 3.8.2.1: • Update the interim analysis for OS per protocol amendment. • Simplify the subgroup analysis by making a reference to the subgroup analysis section. • Clarify that the OS analysis will be done only after the primary analysis is statistically significant. • Clarify how the multivariate Cox regression should be done. Section 3.8.2.2: remove BIRC PFS section which has been replaced by Audit-based BIRC assessment of PFS in section 3.8.1.5. Section 3.8.2.5: Clarify censoring rule for DOR. Section 3.8.2.6: ECOG PS • Update the rule when 2 ECOG assessments fall in the same window to align with ML2 submission. • Clarify the definition of definitive deterioration. • Add estimation of HR to align with ML2 submission. Section 3.8.2.7: • Update the mixed model to align with ML2 submission. • Add EOT and efficacy FU to table 3-5. • Add linear model for selected time points. • Clarify time is a continuous variable in mixed model. Section 3.9.1.4: Add number of occurrences output per safety guidance. Section 3.9.1.5: Update CNAE to AESI per new SAP template. Section 3.9.2: • Add definition of Total ANC. • Update analysis for LFT to align with ML2 submission and add clarification. • Add analysis for time to onset and duration of AE to align with ML2 submission. • Add box plot. • Clarify that CTCAE calculation is based on observed data only. Section 3.9.4: • Add language to clarify how unscheduled assessment is summarized. • Remove parameter RR to align with the notable ECG table. • Update notable heart rate criteria to align with new SAP

Version	Date	Changes
		template. • Clarify the ECG baseline definition to align with ML2 submission. • Add the time to G2 or worse QT prolongation analysis. Section 3.9.6: add urinalysis section to align with ML2 submission. Section 3.10.1: Clarify what are evaluable concentrations. Section 3.10.2: Remove PK parameter section per protocol amendment. Clarify that the Asian subgroup is based on race. Section 3.10.3: • Add summary of PK concentration by dose. • Remove PK geometric mean graphic presentation per protocol amendment. • Add meal record listing. Section 3.10.4: • Delete the Exposure vs. tumor size analysis due to presence of non-measurable disease. • Add exposure vs PFS and TTR since PFS and TTR are more relevant end points in registration trial. • Delete the exposure vs. neutropenia analysis since the relationship is already clear based on ML2 data. • Remove scatter plot from exposure vs liver function analysis given that the boxplots will characterize the relationship between exposure and liver function. • Replace model based exposure vs ECG analysis by summary of ECG by dose group. Section 3.10.5: remove PK parameter imputation rule to align with protocol amendment. [redacted] Section 3.13 (old): update projected interim analysis timeline table per protocol amendment. Section 3.13.1: remove PFS interim analysis per protocol amendment. Move the projected timeline table to OS section. Section 3.13.2: update OS interim analysis per protocol amendment. Section 14: clarify that subgroups will be derived based on non-IRT data. Section 3.14: update subgroups per team discussion. [redacted]

Version	Date	Changes
		• Add audit size for BIRC assessed PFS section. Section 3.16: update numbers related to OS interim analysis per protocol amendment. Section 4.2: add other time to event end points mentioned in the document. Section 4.2.2: remove the SAS code per new SAP template Section 4.2.3: • Remove SAS code per new SAP template. • Align the log rank test with the test in section 3. Section 4.2.4: align the Kaplan-Meier estimates language with new SAP template. Section 4.3: remove PFS interim analysis per protocol amendment. Section 4.3.1: remove beta spending function section since futility interim analysis has been removed. Section 4.3.1.1: Remove predictive probability of success (PPOS) section since PFS interim analysis has been removed per protocol amendment. Section 4.3.2: remove technical detail about familywise type I error rate control about both PFS and OS tests due to the removal of PFS interim analysis. Section 4.4: • Remove typo about PFS calculation. • Add GAP analysis. Section 4.5: Remove IPCW per team discussion. Section 4.6: Remove the exposure vs ECG section. Section 4.7: Add audit-based BIRC PFS section Section 5: Add new references. Appendix 1: add lab CTC grades table.
2.0	10/11/2017	Amendment 2 Section 2.1.7 • Add clarification that a windows of 7 days is allowed for RECIST baseline if investigators overall response is within 7 days of treatment start date Section 3.2 • Clarify the definition of disease free interval • Add the subgroup definition by prior endocrine therapy status Section 3.6.3 • Clarify the definition of planned cumulative dose for fulvestrant Section 3.8 • For EQ-5D-5L PRO, change “index score” to “visual analog scale” to be consistent with the protocol Section 3.9 • Remove the definition of total ANC since some of the percentages of immature WBC is not available

Version	Date	Changes
		[redacted] Section 4.2 • Add time to first chemotherapy only • Correct the number of strata K in the stratified log-rank test • Add threshold of ±7 days for the agreement between local and central response

Table of contents

List of tables

List of figures

List of abbreviations

AE	Adverse event
AESI	Adverse events of special interest
ATC	Anatomical Therapeutic Chemical
AUC	Area Under the Curve
BIRC	Blinded independent review committee
BOR	Best overall response
BPI-SF	Brief Pain Inventory – Short form
CI	Confidence Interval
CBR	Clinical benefit rate
CR	Complete response
CRF	Case Report Form
CRO	Contract Research Organization
CRS	Case retrieval sheet
CSR	Clinical study report
CT	Computed tomography
CTCAE	Common Terminology Criteria for Adverse Events
CV	Coefficient of variation
ΔQTcF	Change from baseline in QTcF
DAR	Dosage administration Record
DI	Dose Intensity
DILI	Drug-Induced Liver Injury
DMC	Data Monitoring Committee
DOR	Duration of Response
ECG	Electrocardiogram
ECOG	Eastern Cooperative Oncology Group
EORTC QLQ-C30	European Organization for Research and Treatment of Cancer Core 30-item Quality of Life Questionnaire
EOT	End of treatment
EQ-5D-5L	EuroQol 5-level scale
ER	Estrogen receptor
FAS	Full analysis set
GGT	Gamma-glutamyl transferase
HER2	Human epidermal growth factor receptor 2
HR	Hazard Ratio
HR	Hormone receptor
HRQoL	Health-Related Quality of Life
IRT	Interactive Response Technology
LATA	Last adequate tumor assessment
LFT	Liver Function Test
LLOQ	Lower Limit of Quantification
MBC	Metastatic breast cancer

MedDRA	Medical Dictionary for Regulatory Activities
MRI	Magnetic resonance imaging
NCI	National Cancer Institute
NMQ	Novartis MedDRA queries
ORR	Overall response rate
OS	Overall survival
PAS	Pharmacokinetic analysis set
PD	Progressive disease
PDI	Planned dose intensity
PDS	Programming Datasets Specifications
PFS	Progression-free survival
PhRMA	Pharmaceutical Research and Manufacturers of America
PK	Pharmacokinetic
PPS	Per Protocol Set
PR	Partial response
PRO	Patient Reported Outcome
PS	Performance Status
PT	Preferred term
QTcF	QT interval corrected by Fridericia method
RDI	Relative dose intensity
RECIST	Response Evaluation Criteria In Solid Tumors
SAE	Serious adverse event
SAP	Statistical Analysis Plan
SD	Stable disease
SD	Standard deviation
SEC	Safety Event Categories
SMQ	Standardized MedDRA queries
SOC	System organ class
TA	Tumor assessment
TEAE	Treatment-emergent adverse event
TFL	Tables Figures and Listings
TBIL	Total Bilirubin
TTP	Time to progression
UNK	Unknown

1 Introduction

This document describes the detailed statistical methodology to be used for the Clinical Study Report (CSR) for the primary PFS analysis of study CLEE011F2301, a phase III, randomized, double-blind, placebo-controlled study of ribociclib or placebo in combination with fulvestrant for the treatment of men and postmenopausal women with hormone receptor positive, HER2-negative, advanced breast cancer who have received no or only one line of prior endocrine treatment.

The content of this SAP is based on CLEE011F2301 protocol amendment 2 (July 28, 2016). All decisions regarding final analysis, as defined in the SAP document, have been made prior to database lock and unblinding of the study data.

CSR deliverables (shells for tables, figures, listings) and further programming specifications are described in Tables Figures and Listings (TFL) shells and Programming Datasets Specifications (PDS) documents, respectively.

1.1 Study design

This is a randomized, phase III, double blind, placebo controlled, international study to determine the efficacy and safety of treatment with fulvestrant with ribociclib versus fulvestrant with placebo in men and postmenopausal women with HR+, HER2-negative advanced breast cancer. This study will consist of 4 phases: screening (up to 28 days), randomized treatment, post-treatment efficacy follow-up, and post-treatment survival follow-up.

Approximately 660 patients will be randomly assigned to one of the following treatment arms in a 2:1 ratio.

- Experimental arm: fulvestrant (500 mg intramuscular [as two 5 mL injections] on Cycle 1 Days 1 and 15 (C1D1 and C1D15), and on CnD1 thereafter) + ribociclib (600 mg by mouth once daily for three weeks followed by one week break, in a 28-day cycle)
  OR
- Control arm: fulvestrant (500 mg intramuscular [as two 5 mL injections] on Cycle 1 Days 1 and 15 (C1D1 and C1D15), and on CnD1 thereafter) + ribociclib placebo (by mouth once daily for three weeks followed by one week break, in a 28-day cycle)

Randomization will be stratified by the following factors:

1. Lung or liver metastases: (yes vs no)
2. Previous endocrine therapy (A vs B) according to the following definition:

   A) Patients treatment naïve in the metastatic/advanced disease setting:
      i- whose disease relapsed >12 months after completion of (neo)adjuvant endocrine therapy with no subsequent treatment for advanced/metastatic disease,
      OR
      ii- with de-novo advanced/metastatic disease (no prior exposure to endocrine therapy).

B) Patients who received up to 1 line of treatment for metastatic/advanced disease:

i- whose disease relapsed on or within 12 months from completion of (neo) adjuvant endocrine therapy, with no subsequent treatment for advanced/metastatic disease,

OR

ii- whose disease relapsed > 12 months from completion of (neo) adjuvant endocrine therapy, and progressed on or after subsequent endocrine treatment for advanced/metastatic disease,

OR

iii- with advanced/metastatic disease at the time of diagnosis that progressed on or after endocrine therapy for advanced/metastatic disease with no prior (neo) adjuvant treatment for early disease.

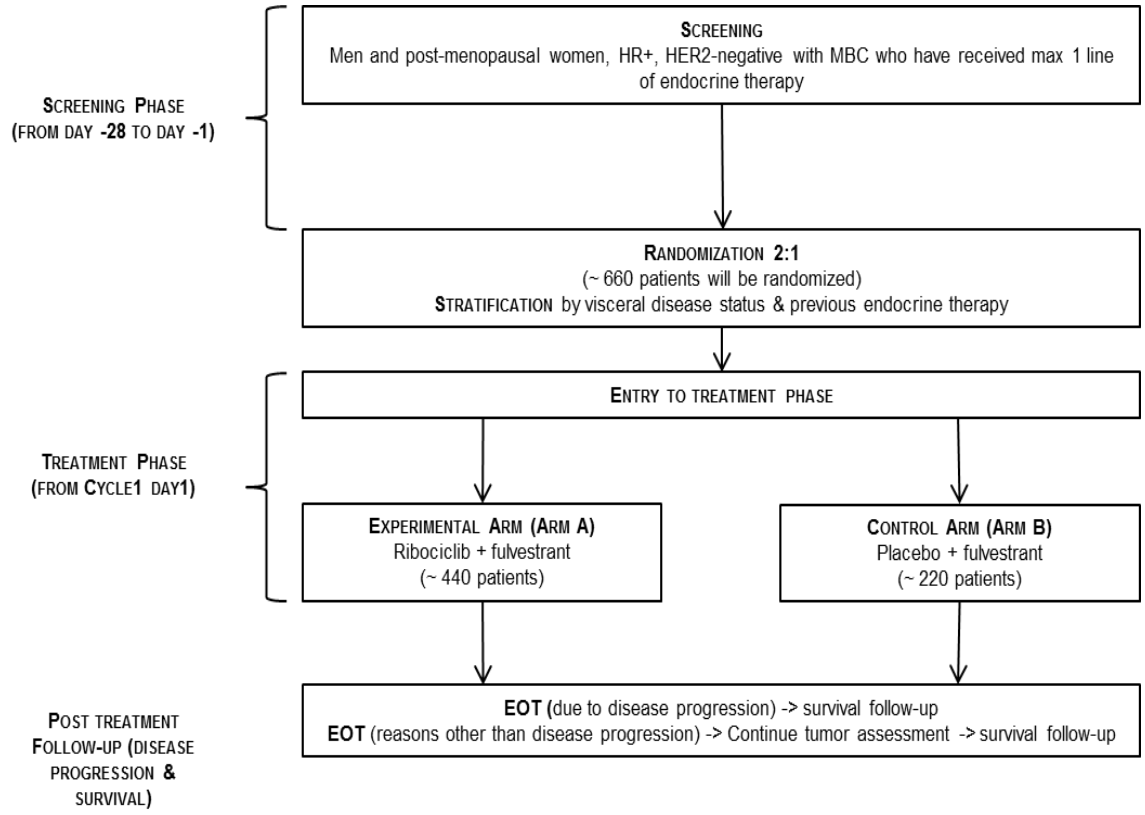
PFS, as assessed by the local radiologists/investigators and using RECIST 1.1 criteria will be the primary endpoint. PFS as assessed through audit-based blinded independent review committee (BIRC) will be a supporting end point for primary analysis.

An Independent Data Monitoring Committee (IDMC) will be constituted and will monitor safety and efficacy data. A Steering Committee (SC) will be established comprised of investigators and Novartis personnel participating in the trial to ensure transparent management of the study according to the protocol.

Overall survival is a secondary endpoint in this study and will be tested provided the primary endpoint PFS is statistically significant. A Lan-DeMets alpha spending function with O'Brien Fleming type stopping boundary (as implemented in East 6.3) will be used to maintain the overall type-I error rate for OS. The interim OS analyses will be performed by Novartis. Further details regarding the group sequential design are provided in Section 3.14.

The study design is summarized in Figure 1-1.

Figure 1-1 Study Design



1.2 Objectives

The study objectives and corresponding endpoints as specified in the protocol are provided in Table 1-1.

Table 1-1 Study objectives

Objective	Endpoint	Analysis
Primary		Refer to Section 3.8.1
To compare PFS between ribociclib in combination with fulvestrant to placebo in combination with fulvestrant among men and postmenopausal women with HR+, HER2-negative advanced breast cancer who received no or only one prior endocrine treatment for advanced disease	PFS based on local radiology assessments and using RECIST 1.1 criteria	
Secondary		
To compare the two treatment arms with respect to overall survival.	Overall survival	Refer to section 3.8
To evaluate the two treatment arms with respect to overall response rate, clinical benefit rate, time to response and duration of response.	ORR as defined by RECIST 1.1. CBR, defined as percentage of patients with CR, PR or SD lasting 24 weeks or longer, TTR, DOR per RECIST 1.1	Refer to section 3.8

Objective	Endpoint	Analysis
To evaluate the two treatment arms with respect to time to deterioration of ECOG performance status.	Time to definitive deterioration of ECOG performance status from baseline	Refer to section 3.8
To evaluate the safety and tolerability of ribociclib in combination with fulvestrant.	Frequency/severity of AEs, laboratory abnormalities	Refer to section 3.9
To evaluate patient reported outcomes for health-related quality of life in the two treatment arms.	Time to 10% deterioration in the global health status/QOL scale score of the EORTC QLQ-C30 Change from baseline in the global health status/QOL scale score of the EORTC QLQ-C30	Refer to section 3.8
To characterize the pharmacokinetics (PK) of ribociclib (and relevant metabolites such as LEQ803) when given in combination with fulvestrant.	Concentration by time point for ribociclib (and relevant metabolites such as LEQ803)	Refer to section 3.10
[redacted]		

2 Definitions and general methodology

2.1 Definitions

2.1.1 Study drug and study treatment

Study drug is defined as ribociclib or matching placebo.

Study treatment is defined as ribociclib + fulvestrant, or matching placebo + fulvestrant.

2.1.2 Date of first administration of study drug

The date of first administration of study drug is defined as the first date when a nonzero dose of study drug is administered and recorded on the dose administration DAR CRF. The date of first administration of study drug will also be referred to as start of study drug. Similar definitions apply for the other components of study treatment.

2.1.3 Date of last administration of study drug

The date of last administration of study drug is defined as the last date when a nonzero dose of study drug is administered and recorded on the DAR CRF. Similar definitions apply for the other components of study treatment.

2.1.4 Date of first administration of study treatment

The date of first administration of study treatment is defined as the first date when a nonzero dose of any component of study treatment is administered and recorded on the DAR CRF. The date of first administration of study treatment will also be referred to as the start of study treatment.

2.1.5 Date of last administration of study treatment

The date of last administration of study treatment is defined as the last date when a nonzero dose of any component of study treatment was administered and recorded on the DAR CRF.

2.1.6 Study day

The study day will be calculated as:

- The date of the event (visit date, onset date of an event, assessment date etc.) − reference start date + 1 if event is on or after the reference start date;
- The date of the event (visit date, onset date of an event, assessment date etc.) − reference start date if event precedes the reference start date.

The reference start date for safety assessments (e.g. adverse event onset, laboratory abnormality occurrence, vital sign measurement, dose interruption etc.) is the start of study treatment. (Note: if an adverse event starts before the start of study treatment the study day displayed on the listing will be negative).

The reference start date for all other, non-safety assessments (e.g., tumor assessment, death, disease progression, tumor response, ECOG performance status, and patient reported outcomes (PRO)) is the date of randomization. In other words, all efficacy time-to-event variables (e.g. progression-free survival, overall survival, time to response) will be calculated from date of randomization. (Example: if randomization date is 15DEC2014, start of study drug is on 18DEC2014, and the date of death is 28DEC2014 then the study day when the death occurred is 14).

The study day will be displayed in data listings.

2.1.7 Baseline

For efficacy evaluations, the last available assessment on or before the date of randomization will be used as the "baseline" value or "baseline" assessment. In the context of baseline definition, the efficacy evaluations also include ECOG performance status and patient-reported outcomes. For RECIST based endpoints including PFS and response rates, a window of 7 days from the start of study treatment will be allowed, i.e. the investigator/BIRC-reported responses will be maintained and baseline considered valid if the baseline assessment is within 7 days of treatment start date.

For safety evaluations (e.g. laboratory assessments and ECG), the last available assessment before or at date of start of study treatment will be used as the 'baseline' assessment. Assessments specified to be collected post-dose on the first date of treatment are not considered as baseline values.

If patients have no value as defined above, the baseline results will be considered missing.

2.1.8 On-treatment assessment/event

Safety summaries and selected summaries of deaths will summarize only on-treatment assessments/events. An on-treatment assessment/event is defined as any assessment/event in the following time interval:

[date of first administration of study treatment, date of last administration of study treatment + 30 days], i.e. including the lower and upper limits. (Note: However, the calculation of study treatment duration will use different rules as specified in Section 3.6.1).

An AE started in the screening phase and ongoing in the on-treatment phase will not be considered as an on-treatment AE unless it has worsened.

If the last date of study treatment is missing, any assessment/event occurring after the start of study treatment will be considered as on-treatment.

Data listings will include all assessments/events, flagging those which are not on-treatment.

Note: The date of first administration of study treatment and the date of last administration of study treatment are defined in Sections 2.1.4 and 2.1.5, respectively.

2.1.9 Last contact date

For patients not known to have died at the analysis cut-off, the last contact date will be derived using the last complete date among the following:

Table 2-1 Last contact date data sources

Source data	Conditions
Date of Randomization	No Condition
Last date patient was known to be alive from Survival Follow-up page	Patient status is reported to be alive, lost to follow-up or unknown.
Start/End dates from further antineoplastic	Non-missing medication/procedure term.

Source data	Conditions
therapy	
Start/End dates from drug administration record	Non-missing dose. Doses of 0 are allowed.
Date of discontinuation/study phase completion from end of treatment disposition page	No condition.
Date of discontinuation/Study phase completion from end of post treatment follow up phase disposition page	No condition
Date of ECG assessment	At least 1 non-missing parameter value
Date of PRO assessment	At least 1 non-missing answer to questionnaire
Tumor (RECIST) assessment date	For non-target lesion: non-missing lesion status For target lesion: non-missing lesion diameter For new lesion: “Is there a new lesion?” yes
Laboratory sample collection date	At least 1 non-missing parameter value
PK collection dates	At least 1 non-missing PK concentration
Vital signs date	At least one non-missing parameter value
Concomitant medication date	At least one non-missing name of medication
[redacted]	[redacted]
Body fluid/Tissue Collection date	Non-missing result (positive/negative tumor cells)
Hospitalization admission/discharge date	Non-missing type of facility admitted to
Cardiac imaging date	Non-missing LVEF or overall interpretation
Performance Status date	Non-missing performance status
Start/End dates of AE	Non-missing verbatim term

The last contact date is defined as the latest complete date from the above list on or before the data cut-off date. The cut-off date will not be used for last contact date, unless the patient was seen or contacted on that date. No dates post cut-off date will be used. Completely imputed dates (e.g. the analysis cut-off date programmatically imputed to replace the missing end date of a dose administration record) will not be used to derive the last contact date. Partial date imputation is allowed for event (death)/censoring coming from ‘Survival information’ eCRF.

The last contact date will be used for censoring of patients in the analysis of overall survival.

2.2 Data included in the analysis

The primary PFS analysis will be carried out after approximately 364 events have been documented based on local investigator assessment. Up to 3 analyses for OS may be performed as described in section 3.8. For each of the analysis time points, statistical analyses will be performed using all data collected in the database up to the data cutoff date. Any data

collected beyond the cutoff date will not be included in the analysis and will not be used for any derivations.

2.3 Analysis sets

2.3.1 Full analysis set (FAS)

The Full Analysis Set (FAS) consists of all randomized patients. Following the intent-to-treat principle, patients will be analyzed according to the treatment and stratum they were assigned to at randomization. The FAS will be the primary analysis set for efficacy analyses.

2.3.2 Safety set

The Safety Set consists of all patients who received at least one dose of any component of study treatment and have at least one post-baseline safety assessment. Patients will be analyzed according to the treatment actually received. Treatment actually received refers to the treatment the patient was randomized to, unless the alternative treatment was received throughout the trial. If a patient takes at least one dose of the randomized treatment then the treatment actually received is the randomized treatment. The statement that a patient has no AE constitutes a safety assessment. Occurrence of a death constitutes a safety assessment as well.

2.3.3 Per protocol set

The Per-Protocol Set (PPS) includes the subset of the patients in the FAS without major protocol deviations, and who took at least one dose of study treatment. Patients with any of the following protocol deviations will be excluded from the PPS.

- Written informed consent not obtained;
- Patient is not post-menopausal;
- Patient without HR+ and HER2- advanced breast cancer at baseline;
- Patient received prior fulvestrant treatment;
- Patient received prior CDK4/6 inhibitor;
- Patient with recurrence on or within 12 month of adjuvant therapy and subsequent advanced treatment.
- Baseline ECOG performance status>1;
- Neither measurable disease nor predominantly lytic bone lesion at baseline;
- Patient received more than 1 line of prior hormonal anti-cancer therapy for advanced breast cancer;
- Patient received prior chemotherapy for advanced breast cancer;
- Patient received different treatment throughout the study than the one randomized to.

A sensitivity analysis of the primary endpoint (i.e., PFS) will be performed using PPS if the primary efficacy analysis is significant and the FAS and PPS differ.

2.3.4 Pharmacokinetic analysis set

The Pharmacokinetic analysis set (PAS) consists of all patients who provide at least one evaluable PK concentration (see Section 3.10.1 for definition of evaluable PK concentration).

2.3.5 Patient classification

Patients may be excluded from the analysis sets defined above based on the protocol deviations entered in the database and/or on specific subject classification rules as defined in Table 2-1.

Table 2-1 Patient classification based on protocol deviations and non-protocol deviation criteria

Analysis set	Protocol deviations leading to exclusion	Non-protocol deviation criteria leading to exclusion
FAS	No written informed consent	NA
Safety set	No written informed consent	No post-baseline safety assessment or no dose of study treatment
Per Protocol set	Any major protocol deviation as listed in definition of per protocol set	No dose of study treatment.

Withdrawal of Informed Consent

Any data collected in the clinical database after a subject withdraws informed consent from all further participation in the trial, will not be included in the analysis data sets. The date on which a patient withdraws full consent is recorded in the eCRF.

Death events may be used in the analysis if captured from public records (registers), local law and patient informed consent permitting.

[redacted]

2.4 Concomitant medications with specific impact on the analysis

According to the study protocol, the following medications are prohibited during treatment in this study:

- Strong inhibitors or inducers of CYP3A4/5
- Substrates of CYP3A4/5 with a narrow therapeutic index
- Medications with a known risk of QT prolongation
- Other investigational and antineoplastic therapies with the exception of palliative radiotherapy
- Herbal medications preparations and dietary supplements (except for vitamins)

These substances are listed in Table 14-1 in the study protocol. A corresponding list for programming purposes will be saved in a separate document. If there is an update to the list of prohibited medications (e.g., in a protocol amendment), the most up-to-date list shall be used for the Clinical Study Report.

Some patients may take these substances during the treatment period so these concomitant medications will be selected via programming and tabulated and listed in the Clinical Study Report.

Treatment with the prohibited substances mentioned above will be identified in the database as protocol deviations.

With the exception of palliative radiotherapy, administration of anti-neoplastic drugs (apart from study treatment) and other investigational drugs is not allowed during study treatment. Patients who take anti-neoplastic drugs after randomization but before end of treatment will be identified through data review. Clinical review of individual study data will be performed in order to identify those anti-neoplastic medications which are considered disallowed. Tumor assessments (TAs) made after the start of anti-neoplastic therapies (whether on study treatment or afterwards) will not be included in the efficacy analyses based on best overall response i.e. Overall Response Rate (ORR), Clinical Benefit Rate (CBR), Time to response (TOR) and duration of response (DOR) (Section 3.8.2.2-3.8.2.5). In addition, a sensitivity analysis will be performed for the primary PFS analysis in which PFS will be censored at the last tumor assessment before the start of new anti-neoplastic therapy. Further details are provided in Section 3.8.1. Patient reported outcome (PRO) and ECOG based time to event end point will not consider assessments after the start of new anti-neoplastic therapy. Further details are provided in section 3.8.2.6 and 3.8.2.7. For these analyses, the following will not be considered as new antineoplastic therapies: (i) palliative radiotherapy; (ii) continuation of combination partner therapy alone after end of study treatment.

2.5 Implementation of RECIST

Response and progression evaluation will be performed according to the Novartis RECIST guideline (as described in detail in Appendix 2 of the Clinical Study Protocol), which is based on the RECIST 1.1 guidelines (Eisenhauer et al 2009). The text below gives instructions and rules to provide details needed for programming.

2.5.1 Overall lesion responses for patients with only non-measurable lesions at baseline

Patients with at least one predominantly lytic bone lesion but not having measurable disease per RECIST 1.1 are allowed to enter the study. For patients with non-measurable lesions only at baseline, the overall lesion response will be based solely on non-target lesion response or an occurrence of a new lesion. Non-measurable lesions will be entered as non-target lesions. Therefore, the best overall response is determined from non-target lesion response and presence of new lesions (refer to RECIST Novartis guidelines as described in detail in Appendix 2 of the Clinical Study Protocol).

Note: Pathologic fracture, new compression fracture, or complications of bone metastases will not be considered as evidence of disease progression, unless there is unequivocal progression of existing non-target lesions or a new lesion.

2.5.2 Disease progression

Progressive disease should only be assigned if it is confirmed by an assessment method as per RECIST 1.1 guidelines (e.g. CT scan photos for skin lesions, etc.). If a new lesion is detected using an objective assessment method other than radiologic scan then it should also be entered on the 'New lesion' RECIST CRF with appropriate method. Discontinuation due to disease progression or death due to study indication in death CRF page, without corresponding supportive data in the RECIST CRF (as defined above), will not be considered as progressive disease in the calculation of best overall response and in the analysis of PFS.

2.5.3 Best overall response (BOR)

The best overall tumor response will be assessed per RECIST 1.1 criteria. The definitions and the details on the derivation are given in Appendix 2 of the Clinical Study Protocol.

Only tumor assessments performed before the start of any further anti-neoplastic therapies (i.e. any additional anti-neoplastic therapy or surgery) and within 30 days after the last administration of study treatment will be considered in the assessment of best overall response. Further anti-neoplastic therapies will be identified from the data collected on 'Anti-neoplastic therapies since discontinuation of study treatment' CRFs. Palliative radiotherapy is the only setting of radiotherapy allowed during the study. Therefore palliative radiotherapy will not be considered as an anti-neoplastic therapy for assessment of BOR.

The standard definition of a best overall response evaluation of 'stable disease', 'disease progression' or 'unknown' given in the Appendix 2 of the Clinical Study Protocol will be used for this study. Best overall response (as reported by the investigator for local BOR, and as reported by BIRC for central BOR) for each patient is determined from the sequence of overall (lesion) responses according to the following rules:

- CR = at least two determinations of CR at least 4 weeks apart before progression.
- PR = at least two determinations of PR or better at least 4 weeks apart before progression (and not qualifying for a CR).
- SD = at least one SD assessment (or better) > 6 weeks after randomization (and not qualifying for CR or PR).
- Non-CR/non-PD = at least one non-CR/non-PD assessment (or better) > 6 weeks after randomization date (and not qualifying for CR). This applies only for patients with non-measurable disease alone at baseline.
- PD = progression ≤ 12 weeks after randomization (and not qualifying for CR, PR, SD or Non-CR/non-PD).
- UNK = all other cases (i.e. not qualifying for confirmed CR or PR and without SD or Non-CR/Non-PD after more than 6 weeks or progression within the first 12 weeks).

Patients with best overall response "unknown" will be summarized by reason for having unknown status. The following reasons will be used:

- No valid post-baseline assessment
- All post-baseline assessments have overall response UNK
- New anti-neoplastic therapy started before first post-baseline assessment
- SD too early (≤6 weeks after randomization)
- PD too late (>12 weeks after randomization and not qualifying for CR, PR, SD or NON-CR/NON-PD)

Special (and rare) cases where BOR is unknown due to both early SD and late PD will be classified as "SD too early".

2.5.4 Change in imaging modality

Per RECIST 1.1, a change in methodology can be defined as either a change in contrast use (e.g. keeping the same technique, like CT, but switching from with to without contrast use or vice-versa, regardless of the justification for the change) or a major change in technique (e.g. from CT to MRI, or vice-versa), or a change in any other imaging modality. A change from conventional to spiral CT or vice versa will not constitute a major "change in method" for the purposes of response assessment. A change in methodology will result by default in an UNK (unknown) overall lesion response based on the Novartis calculation as per Novartis calculated response. However, a response from the investigator or the central blinded reviewer that differs from the Novartis calculated UNK is acceptable if a definitive response assessment can be justified based on the available information.

Potential discrepancies between the modality used and overall lesion response (e.g. change in modality but response is different from 'Unknown') will be queried during the data validation process.

2.5.5 Determination of missing adequate assessments

The term 'missing adequate assessment' is defined as assessments that are not done or assessments for which the overall lesion response equals 'Unknown'. The 'missing adequate assessment' is also referred to as 'missing assessment'.

As detailed in Section 3.8.1 and in [Appendix 2 of the study protocol], the PFS censoring and event date options depend on the presence and the number of missing tumor assessments. An event occurring after two or more missing assessment is censored at the last adequate tumor assessment.

An exact rule to determine whether there are no, one or two missing TAs is therefore needed. This rule is based on the interval between the last adequate tumor assessment (LATA) date and the event date. The scheduled date of tumor assessments (in weeks from randomization), protocol specified windows for tumor assessments, and the thresholds for LATA to belong to a visit can be found in the following table.

Table 2-3 Schedule for tumor assessment and time windows

Assessment schedule	Scheduled date – 1 week	Scheduled date (weeks from randomization)	Scheduled date +1 week	Threshold*

Every 8 weeks for the first 18 months	Baseline	0	0	0	0
	C3D1	7	8	9	12
	C5D1	15	16	17	20
	C7D1	23	24	25	28
	C9D1	31	32	33	36
	C11D1	39	40	41	44
	C13D1	47	48	49	52
	C15D1	55	56	57	**60**
	C17D1	63	64	65	**68**
	C19D1	71	72	73	78
Every 12 weeks after 18 months	C22D1	83	84	85	90
	C25D1	95	96	97	102
	C28D1	107	108	109	114
	C31D1	119	120	121	126
* The mid-point between current and next visit (except for baseline) and the upper limit for LATA to be matched to a certain scheduled assessment, e.g. if LATA is at week 13, this is after threshold for C3D1 and before that for C5D1, so the matching scheduled assessment is C5D1. .					

To calculate the number of missing tumor assessments, the LATA before an event is matched with a scheduled tumor assessment using the time window in Table 2-3 (essentially whichever scheduled assessment it is closest to). Two thresholds, D1 and D2 are calculated for that scheduled assessment based on the protocol-specified schedule and windows

- An event after LATA+D1 will be considered as having >=1 missing assessment
- An event after LATA+D2 will be considered as having >=2 missing assessments

Since there is a change of schedule for tumor assessments in 18 months, D1 and D2 are defined differently depending on when LATA happens.

Rule 1: if LATA happens within 60 weeks from randomization (the matched scheduled tumor assessment is C15D1 or before)

- D1=8+2=10 weeks
- D2=2*8+2=18 weeks

Rule 2: if LATA happens after 60 weeks but within 68 weeks from randomization (the matched scheduled tumor assessment is C17D1)

- D1=8+2=10 weeks
- D2=8+12+2=22 weeks

Rule 3: if LATA happens after 68 weeks from randomization (the matched scheduled tumor assessment is C19D1 or later)

- D1=12+2=14 weeks
- D2=2*12+2=26 weeks

Therefore, using the D2 definition above, the censoring of an event occurring after ≥2 missing TAs (in primary PFS analysis) can be refined as follows: if the distance between the last

adequate TA date and the PFS event date is larger than D2 then the patient will be censored and the censoring reason will be 'Event documented after two or more missing tumor assessments'.

The same D2 will be used to determine the PFS censoring reason. If the distance between the last adequate tumor assessment date and the earliest of the following dates:

1. Analysis cut-off date
2. Date of consent withdrawal
3. Date of lost to follow-up

Is smaller or equal to D2 then the censoring reason will be 1. 'Ongoing without event', 2. 'Withdrew consent', or 3. 'Lost to follow-up', respectively. However, if this distance is larger than D2 then the censoring reason will be 'Adequate assessment no longer available'. If the event is documented after two missing assessments, then the censoring reason will be 'Event documented after two or more missing tumor assessments'.

2.5.6 No baseline tumor assessments

As specified in the [Appendix 2 of the Clinical Study Protocol], since the timing of disease progression cannot be determined for patients with missing baseline tumor assessment, these patients are censored in the PFS analysis at the date of randomization. This rule however only applies to the disease progression component of the PFS assessment, and not to the survival component.

Patients without baseline tumor assessments who die within D2 distance (see Section 2.5.5 for definition) of randomization will be counted as having an event in the primary analysis of PFS at the date of death (Note: all deaths will be counted in the overall survival analysis regardless of presence or absence of the baseline tumor assessment).

These patients will be excluded from the PPS for the analysis of PFS as defined in Section 2.3.

2.5.7 Construction of waterfall graphs

Waterfall graphs will be used to depict the anti-tumor activity. These plots will display the best percentage change from baseline in the sum of diameters of all target lesions for each patient. Only patients with measurable disease at baseline will be included in the waterfall graphs.

Special consideration is needed for assessments where the target lesion response is CR, PR or SD, but the appearance of a new lesion or a worsening of non-target lesions results in an overall lesion response of PD. As a conservative approach, such assessments will not be considered for display as bars in the graph, since the percentage change in the sum of diameters of target lesions reflects the non-PD target lesion response, but the overall lesion response is PD. A patient with such assessments will be represented by a special symbol (e.g. ✶) in the waterfall graph.

Assessments with "unknown" target lesion response and assessments with unknown overall response will be excluded from the waterfall plots. Patients without any valid assessments will be completely excluded from the graphs.

The total number of patients displayed in the graph will be shown and this number will be used as the denominator for calculating the percentages of patients with tumor shrinkage and tumor growth. Footnote will explain the reason for excluding some patients (due to absence of any valid assessment).

All possible assessment scenarios are described in Table 2-34.

Table 2-4 Assessments considered for calculation of best percentage change for waterfall graphs

Case	Target response	Overall lesion response	Calculate % change from baseline in sum of diameters?
1	UNK	Any	No, exclude assessment
2	Any	UNK	No, exclude assessment
3	CR/PR/SD	PD	No, flag assessment with ✶
4	PD	PD	Yes
5	CR/PR/SD	CR/PR/SD	Yes

3 Statistical methods used in reporting

3.1 Enrollment status

Enrollment by country and center will be summarized for all screened patients and also by treatment arm using the FAS. The reasons for screen failure will also be summarized.

3.2 Background and demographic characteristics

The FAS will be used for all baseline disease characteristics and demographic summaries and data listings.

3.2.1 Basic demographic and background data

Demographic and background disease characteristics data will be listed in detail. Qualitative data (e.g. race, ECOG performance status, etc.) will be summarized by means of contingency tables by treatment arm and quantitative data (e.g. age, body weight, etc) will be summarized by appropriate descriptive statistics (mean, standard deviation, median, minimum, and maximum) by treatment arm.

Discrepancies between randomization stratification information (obtained from the Interactive Response Technology (IRT) system) and stratum information based on data collected on CRFs will be tabulated and listed.

Unless otherwise specified, all stratification information (including stratified analyses, analyses by "stratum", strata as covariates) will be based on stratification data from IRT.

3.2.2 Diagnosis and extent of cancer

Summary statistics will be tabulated for diagnosis and extent of cancer. This analysis will include the following: primary site of cancer, histological grade, stage at initial diagnosis, time since initial diagnosis, time from initial diagnosis to first recurrence/progression, stage at time of study entry, presence/absence of target and non-target lesions, number and type of metastatic sites involved, HER-2 / estrogen / progesterone receptor status, number of de novo patients. disease free interval (DFI) for non-de novo patients and prior endocrine therapy status (ET). Estrogen and progesterone receptor status summary will be combined into 3 categories (ER+ PR+, ER+ PR-, ER- PR+).

De novo patients will be identified as those with no date of first recurrence/progression or the first recurrence/progression is within 90 days of initial diagnosis without prior antineoplastic medication.

DFI for non-de novo patients will be calculated as the time from initial diagnosis to first recurrence/progression and categorized as ≤12 months and > 12 months.

Patients will also be grouped as follows based on prior endocrine therapy:

- No prior ET, including two groups of patients:
  - De novo patients who diagnosed with advanced disease and never treated.
  - Patient diagnosed with early stages of disease, treated with surgery and/ or RT and/ or chemotherapy (but no endocrine therapy) for that early setting and relapsed afterwards with advanced disease.
- Prior ET therapy
  - 1st line ET
    - Progression while on or within 12 months of end of (neo-)adjuvant ET
    - Progression >12 months after end of (neo-)adjuvant ET
  - 2nd line ET including patients who had endocrine therapy under advanced/metastatic setting and then progressed.

Time since initial diagnosis and time from initial diagnosis to first recurrence/progression will be summarized in months. A month is defined as 365.25/12=30.4375 days.

3.2.3 Medical history

Medical history and ongoing conditions, including cancer-related conditions, will be summarized and listed. Separate summaries will be presented for ongoing and historical medical conditions. The summaries will be presented by primary system organ class and preferred term. Medical history/current medical conditions are coded using the latest version of Medical Dictionary for Regulatory Activities (MedDRA) terminology available at the time of the analyses. The MedDRA version used for reporting will be specified in the CSR and as a footnote in the applicable tables/listings.

3.2.4 Prior anti-neoplastic therapy

The number and percentage of patients recording any prior anti-neoplastic medications, radiotherapy or surgery (biopsy and non-biopsy separately) will be summarized by treatment arm both separately and in a combined fashion.

- Prior anti-neoplastic medications will be summarized by therapy type (e.g. chemotherapy, hormonal therapy etc), and also by lowest ATC class, preferred term and treatment. The total number of regimens along with the type (e.g. hormonal therapy), setting (e.g. adjuvant), best response and time from last treatment to progression for the last therapy will be summarized by treatment arm.
- For radiotherapy, the setting (e.g. adjuvant) for the last therapy will be summarized.

For surgery (excluding biopsies), the time since last surgery will be summarized by treatment arm. Separate listings will be produced for prior anti-neoplastic medications, radiotherapy, and surgery.

3.2.5 Other

All data collected at baseline, including source of patient referral, child bearing potential and [redacted]

3.3 Protocol deviation summaries

The number and percentage of patients in the FAS with any protocol deviation will be tabulated by deviation category (as specified in the Study Specification Document) and by treatment arm.

Protocol deviations leading to the exclusion from per protocol set will be tabulated separately by treatment arm.

All protocol deviations as well as protocol deviations leading to exclusion from per protocol set will be listed.

3.4 Groupings for analysis

The number and percentage of patients in each analysis set (definitions are provided in Section 2.3) will be summarized by treatment arm and randomization strata per IRT.

3.5 Patient disposition

Patient disposition for all randomized patients will be summarized based on FAS. There will be one combined by-treatment summary showing:

1. Number (%) of patients treated/untreated.
2. Number (%) of patients who are still on-treatment (based on the absence of the ‘End of treatment’ page)
3. Number (%) of patients who discontinued study treatment (based on the ‘End of Treatment’ page)
4. Reasons for study treatment discontinuation (based on ‘End of Treatment’ page)

5. Number (%) of patients who entered the post-treatment evaluations (based on 'End of Treatment' page)
6. Number (%) of patients who discontinued from the post-treatment evaluations (based on 'End of post treatment follow up disposition' page)
7. Reasons for discontinuation from the post-treatment evaluations phase (based on 'End of post treatment follow up disposition' page).
8. Number (%) of patients who entered survival follow-up .

In a separate summary, the reasons for patients not completing the screening phase will be presented based on "Screening Phase Disposition" eCRF page

3.6 Study treatment

Duration of study treatment exposure, cumulative dose, dose intensity (DI) and relative dose intensity (RDI) will be summarized by treatment. The number of patients with dose reductions/interruptions, and the reasons, will be summarized and listed. Details of the derivations and summaries are provided in the following sections.

The safety set will be used for all summaries and listings of study treatment.

3.6.1 Duration of study treatment exposure

The duration of exposure to study treatment will be calculated as

Duration of exposure to study treatment (days) = (last date of exposure to any study treatment component) – (date of first administration of study treatment) + 1.

Duration of exposure to ribociclib/placebo (days) = (last date of exposure to ribociclib/placebo) – (date of first administration to ribociclib/placebo) + 1.

Duration of exposure to fulvestrant (days) = (last date of exposure to fulvestrant) – (date of first administration to fulvestrant) + 1.

The last date of exposure is defined as follows for the study treatment components:

- For ribociclib /placebo: the last date of exposure is defined as the date of last administration of the corresponding medication;
- For fulvestrant, the last date of exposure is defined as following:
  1. If patient discontinues fulvestrant before C2D1 dose, then:
     a. Last date of exposure = last date of administration + 13 days.
     b. If patient died or lost to follow-up within last date of administration + 13 days, then last date of exposure is date of death or last contact date, respectively.
  2. If patient discontinues fulvestrant after C2D1 dose, then:
     a. Last date of exposure = last date of administration + 27 days.
     b. If patient died or lost to follow-up within last date of administration + 27 days, then last date of exposure is date of death or last contact date, respectively.

The duration of exposure includes the periods of temporary interruption (of any component of the study treatment for any reason). The duration of study treatment exposure will be summarized by treatment arm. In addition, the duration of exposure to study treatment will be

categorized into time intervals (e.g. <3 months; 3-<6 months; 6-<9 months, etc); frequency counts and percentages will be presented for the number of patients in each interval.

3.6.2 Cumulative dose and average daily dose

Cumulative dose for any component of study treatment is defined as the total dose of the medication given during the study treatment exposure.

Average daily dose is defined as [Cumulative dose (mg) / Number of dosing days]; drug free day(s) are not counted as dosing days.

Cumulative dose and average daily dose will be summarized using descriptive statistics by treatment arm for each component of study treatment. Patients with no exposure to the study treatment component will be excluded from the corresponding summary.

3.6.3 Dose intensity and relative dose intensity

Dose intensity (DI) for ribociclib/placebo for patients with non-zero duration of exposure to ribociclib/placebo is defined as follows:

DI (mg / day) = Cumulative dose (mg) / *adjusted* duration of exposure to ribociclib/placebo (day),

where *adjusted* duration of exposure (days) to ribociclib/placebo is the number of ribociclib /placebo dosing days a patient would be expected to have received per protocol, given their duration of exposure. Since ribociclib /Placebo follows a 3 weeks on, 1 week off schedule, the adjusted duration of exposure to ribociclib/placebo is the duration of exposure to ribociclib/placebo minus the planned off days for ribociclib /placebo. The adjusted duration of exposure to ribociclib/placebo (in days) is therefore 21 * (# completed 28 day period) + min(21, duration of last incomplete cycle).

For example, if the duration of exposure to ribociclib/placebo is 66 (corresponding to two cycles and 10 days) days, then the adjusted duration of exposure to ribociclib/placebo is 21*2+10=52 days. If the duration of exposure to ribociclib/placebo is 108 days (corresponding to three cycles and 24 days), then the adjusted duration of exposure is 21*3+21=84 days.

Specifically, let D1 represent the duration of exposure to ribociclib/placebo as defined above Then the adjusted duration of exposure is defined as
D=21*[D1/28] + min(21,D1-28*[D1/28]) days,

where [x] stands for the integer part of x. In this equation [D1/28] is the number of completed cycles, and D1-28*[D1/28] is the additional number of days in the last, incomplete cycle (if any). For example, if D1=30 then [D1/28]=1, D1-28*[D1/28]=2, and D=23. If D1=7 then D=7; if D1=22 then D=21; if D1=28 then D=21, etc.

Planned dose intensity (PDI) is defined as the assigned dose by unit of time planned to be given to patients as per protocol. The PDI for ribociclib/placebo is displayed in Table 3-1. Note that DI will also be calculated and DI for ribociclib/placebo will be reported in the units displayed in Table 3-1, whereas duration of exposure itself will be summarized in months.

Table 3-1 Planned dose intensity

Medication	PDI (dose unit/unit of time)
ribociclib /placebo	600 mg/day (3 weeks on 1 week off)

Ribociclib/placebo relative dose intensity (RDI) is defined as:

RDI = DI (dosing unit / unit of time) / PDI (dosing unit / unit of time).

Fulvestrant RDI is defined as:

RDI = actual cumulative dose / planned cumulative dose, where planned cumulative dose is defined as 500* (# completed D1 in a 28-days cycle) and added by 500 mg if patients completed C1D15.

DI for ribociclib/placebo and RDI for ribociclib/placebo and fulvestrant will be summarized separately for each of the study treatment component.

3.6.4 Dose reductions, interruptions and delays

The number and percentage of patients with dose reductions, interruptions or delays, and the reasons, will be summarized by treatment arm.

Interruption: An interruption is defined as a 0 mg dose given on one or more days during the period where a patient is not on the "off" part of a treatment cycle, after which > 0mg dose resumes. For patients who had dose interruption checked but never resumed non-zero dose, the dose interruption will not be counted. For example, in the sequence of 600 mg – 0mg (dose break) -0mg (dose interruption) – 0 mg (dose permanently discontinuation) the 0mg (dose interruption) will not be counted as dose interruption. Interruptions will be summarized for each component of the study treatment. Delay**:** A special case of ribociclib/placebo interruption occurs at the start of a new cycle, after a planned rest period. It will be determined based on ribociclib / placebo dose administration record where a planned dose break is followed by a dose interruption. Such instances will be identified as a subset of the interruptions and will be summarized separately as dose delays.

Reduction**:** A reduction is defined as a decrease from the previous non-zero dose to another non-zero dose less than protocol planned dose, even if this decrease has been directly preceded by an interruption. For example, in the sequence of ribociclib 600mg – 0mg – 400mg, the 400mg dose will be counted as a reduction.

If due to dosing error, a patient took a dose during the dosing break with a dose that is lower than the previous dose, this will not be considered as dose reduction. For example, a patient took 600 mg from day 1-21, and mistakenly took 200 mg per day on day 22-28 which is supposed to be a dosing break. The patient resumed 600 mg dosing on day 29. This will not be considered as dosing reduction.

Dose reductions and interruptions will be tabulated separately. Dose escalations are not allowed according to the protocol and will not be counted in these summaries.

Missing data: If dose is recorded but frequency is missing or entered as 'none', it is assumed that the study drug was taken as per-protocol.

3.6.5 Discontinuation of study treatment components

The reasons for discontinuation of ribociclib/placebo will be summarized by treatment arm, based on the information on the ribociclib/placebo DAR CRF for patients who have the "dose permanently discontinued" box checked.

Partial discontinuation: A partial discontinuation is defined as the event when the last non-zero dose of ribociclib/placebo is more than 21 days before the last non-zero dose of fulvestrant when the permanent discontinuation checkbox is checked in the ribociclib/placebo DAR page.

Partial Discontinuation of ribociclib/placebo will be summarized by treatment

3.7 Concomitant and post-treatment therapy

Concomitant therapies

Concomitant therapy is defined as all interventions (therapeutic treatments and procedures) besides the study treatment that were administered to a patient, coinciding with the study assessment period (even if started before the study assessment period).

Concomitant medications will be coded using the World Health Organization (WHO) Drug Reference Listing (DRL) dictionary that employs the WHO Anatomical Therapeutic Chemical (ATC) classification system.

Concomitant medications will be summarized by lowest ATC class, preferred term and treatment arm. These summaries will include:

1. medications starting on or after the start of study treatment but no later than 30 days after last dose of study treatment, and
2. medications starting prior to the start of study treatment and continuing after the start of study treatment.

All concomitant therapies will be listed. Any concomitant therapies starting and ending prior to the start of study treatment or starting more than 30 days after the last date of study treatment will be flagged in the listing.

The safety set will be used for all concomitant medication tables and listings.

Concomitant medications that have the potential to impact some specific analyses (e.g. PK, efficacy or safety analyses) will be identified prior to database lock. Separate summaries of these concomitant medications will be produced using the appropriate analysis set (e.g. FAS for those potentially affecting efficacy).

- Strong inhibitors or inducers of CYP3A4/5, substrates of CYP3A with a narrow therapeutic window, medications with a known risk of QT prolongation and other prohibited medications described in Section 2.4 will be identified. They will be tabulated by ATC class and preferred term.
- Any anti-neoplastic therapy administered concomitantly with study treatment will be listed based on their identification through the protocol deviation process.

Post treatment anti-neoplastic therapy

Anti-neoplastic therapies after discontinuation of study drug will be listed and tabulated by ATC class, preferred term and treatment arm by means of frequency counts and percentages using the FAS. In addition, 1st line and 2nd line anti-neoplastic medications after discontinuation of study drug will be summarized by medication type, e.g. hormonal therapy, chemotherapy, etc. The medication type will be determined by clinical review.

[redacted]

3.8 Efficacy evaluation

The efficacy endpoints based on the tumor assessments will be derived according to the RECIST guideline version 1.1 (see [Appendix 2 of the Clinical Study Protocol] for details). The tumor endpoint derivation is based on the sequence of overall lesion responses at each assessment/time point. However, the overall lesion response at a given assessment/time point may be provided from different sources as illustrated in Table 3-2.

Table 3-2 Sources for overall lesion response

Source 1	Investigator (local radiology) reported overall lesion response
Source 2	Novartis-calculated overall lesion response based on raw (i.e. individual lesion) measurements from investigator (local radiology)
Source 3	Final BIRC reported overall lesion response

The primary efficacy analysis will be based on the investigator/local radiology review. The investigator reported overall lesion response at each assessment/time point (Source 1 in Table 3-2) will be used to derive the efficacy endpoints.

The overall response at each assessment will also be calculated using the raw lesion measurements (Source 2 in Table 3-2). The calculated responses will be listed along with the responses given by the investigator. As a sensitivity analysis, PFS based on calculated overall lesion response (Source 2 in Table 3-2) will also be summarized. Tumor assessment data based on BIRC will be used for supportive efficacy analysis. The BIRC comprises of two independent radiologists and an adjudicator. The BIRC-reported overall tumor response data will be used to derive the supportive BIRC-based endpoints. Data from the two independent central readers will be listed together with the adjudication. Differences in overall responses between local radiology (Source 1) and central radiology (Source 3) will be listed

3.8.1 Primary efficacy

PFS based on local radiology assessment is the primary efficacy variable in this study. PFS is defined as the time from the date of randomization to the date of the first documented disease progression or death due to any cause. If a patient has not progressed or died at the analysis cutoff date, PFS will be censored at the time of the last adequate tumor assessment before the cut-off date. Definitions and further details on PFS can be found in [Appendix 2 of the study protocol].

Discontinuation due to disease progression (collected on the 'End of treatment' and 'End of post treatment follow up' disposition page) without supporting objective evidence (as defined in Section 2.5.2) satisfying progression criteria per RECIST will not be considered disease progression for PFS derivation.

3.8.1.1 Primary analysis

The primary analysis of PFS will be based on the local radiological assessments (Source 1 in Table 3-2). The analysis will be performed on the FAS and will use the default censoring and event date options from [Table 14-7 of study protocol Appendix 2], i.e. event/censoring rules will be based on options A(1), B(1), C1(1), C2(1), D(1), E(1), and F(1) (summarized in Table 3-3). In particular, PFS will be censored at the last adequate tumor assessment if a patient didn't have an event or the event occurred after two or more consecutive missing tumor assessments (see Section 2.5.5). For the primary analysis, in this study, PFS will not be censored if a new antineoplastic therapy is started; instead, an ITT approach will be used for the purposes of PFS derivation , i.e. option F(1) in [Table 14-7 of protocol Appendix 2] will be used. A sensitivity analysis will be performed censoring PFS at the last adequate tumor assessment prior to start of new antineoplastic therapy, i.e. using option F(2). Discontinuation of study treatment (for any reason) will not be considered as a reason for censoring.

Table 3-3 Outcome and event/censor dates for primary PFS analysis

Situation	Date	Outcome
No baseline assessment	Date of randomization *	Censored
Progression or death at or before next scheduled Assessment	Date of progression (or death)	Progressed
Progression or death after exactly one missing assessment	Date of progression (or death)	Progressed
Progression or death after two or more missing assessments	Date of last adequate assessment prior to missed assessment	Censored
No progression (or death)	Date of last adequate assessment	Censored
New anticancer therapy given prior to protocol defined progression	Ignore the new anticancer therapy and follow situations above	As per above situations
Treatment discontinuation due to 'Disease progression' without documented progression, i.e. clinical progression based on investigator claim	Ignore clinical progression and follow situations above	As per above situations

*The rare exception to this is if the patient dies no more than D2 days (see Section 2.5.5 for definition) after randomization, in which case this is a PFS event at the date of death.

3.8.1.2 Hypothesis and test statistic

The primary efficacy analysis will be the comparison of PFS between the two treatment arms using a stratified (strata based on IRT data) log-rank test at one-sided 2.5% level of significance.

Assuming proportional hazards model for PFS, the following statistical hypotheses will be tested to address the primary efficacy objective:

$$H_{01}: \theta_1 \geq 0 \text{ vs. } H_{a1}: \theta_1 < 0$$

where θ_1 is the log-hazard ratio (fulvestrant+ ribociclib arm vs. fulvestrant + placebo arm) of PFS.

3.8.1.3 Kaplan-Meier estimates

The survival distribution of PFS will be estimated using the Kaplan-Meier method. The results will be plotted graphically (Kaplan-Meier curves) by treatment arm. The plots will display the number of patients at risk at equidistant time points. The median, 25^{th} and 75^{th} percentiles for PFS for each treatment arm will be provided with associated 95% confidence intervals. The survival probabilities at 2- month intervals, and the associated 95% confidence intervals will be summarized by treatment arm. Kaplan-Meier estimates will be obtained using PROC LEFETEST with method=KM option in SAS. The loglog (double log) option available within PROC LIFETEST will be used to compute the confidence intervals.

Hazard Ratio

The PFS hazard ratio with two-sided 95% confidence interval will be derived from the stratified Cox proportional hazards model. In this analysis the baseline hazard function will be allowed to vary across strata. SAS PHREG procedure with ties=EXACT option will be used to carry out this analysis in which the model statement will include treatment arm variable as the only covariate and the STRATA statement will include the stratum information as obtained via IRT.

3.8.1.4 Sensitivity and supportive analyses of the primary endpoint PFS

Audit-based BIRC assessment of PFS

PFS assessed by Blinded Independent Review Committee (BIRC) will be used as a supportive analysis of the primary endpoint.

For studies with PFS based on local radiology assessment as the primary endpoint, PFS assessed centrally has generally been used as a secondary or supportive analysis of the treatment effect observed in the primary efficacy analysis. Although 100% central review of scans has been performed in many trials, there is a growing body of evidence that an audit based approach for central evaluation is sufficient (Zhang et al, 2012, FDA ODAC 2012).

An audit (sample) based approach will therefore be implemented for the BIRC assessment of PFS, whereby all assessments for a randomly selected subset of randomized patients will be assessed by BIRC. An independent random sampling process, implemented by the third party IRT vendor, will select approximately 40% of randomized patients. This random allocation

will be stratified by randomized treatment arm and the strata used for the randomization of patients to treatment arms.

The distribution of PFS based on the audit BIRC sample will be estimated using the Kaplan-Meier method. The median along with two-sided 95% confidence interval (CI) will be presented by treatment group. A Kaplan-Meier figure will also be displayed. A stratified Cox regression will be used to estimate the hazard ratio (HR), along with two-sided 95% CI based on the audit BIRC sample.

In order to determine whether a 100% BIRC review should be conducted, two additional methods will be used to summarize the data from the sample-based BIRC assessment.

1. NCI method: The NCI (National Cancer Institute) method (Dodd et al. 2011) uses an auxiliary variable estimator of the log-hazard ratio that combines information from patient-level investigator assessments from all patients (in the FAS) and the BIRC assessment of patients randomly selected for central review (see Section 4.6 for methodology details). This estimate and its one-sided 95% CI will be provided. The NCI method is used for the audit sample size determination (see Section 3.16) and summary of treatment effect for the supportive BIRC assessment.
2. PhRMA method: The data from the BIRC assessment generated following the sampling scheme as above will also be summarized using the method proposed by Amit et al. 2011, referred to as the PhRMA (Pharmaceutical Research and Manufacturers of America) method, based on the early discrepancy rate (EDR) and late discrepancy rate (LDR).

   The EDR quantifies the frequency with which the investigator declares progression early relative to BIRC within each arm as a proportion of the total number of investigator assessed PDs. The LDR quantifies the frequency that the investigator declares progression later than BIRC as a proportion of the total number of discrepancies within the arm (see Section 4.6 for further details). If the distribution of discrepancies is similar between the arms this suggests the absence of evaluation bias favoring a particular arm. With this approach, the differential discordance (DD) of the early discrepancy rate (EDR) and late discrepancy rate (LDR) between the two arms will be estimated as the rate on the ribociclib+fulvestrant arm minus the rate on the placebo+fulvestrant arm. The EDR and LDR results will also be summarized by treatment arm.

The following thresholds based on the NCI and PhRMA methods will be used to define the trigger for a full BIRC review:

- If the upper-bound of the one-sided 95% confidence interval for BIRC-based log-hazard ratio exceeds zero (i.e. HR>1) based on the NCI method

and/or

- If ≥ 15% differential discordance is observed in EDR or LDR based on the PhRMA method.

A full BIRC review will only be considered if the primary endpoint is statistically significant. In the event of a full BIRC read of all patients' imaging data, all analyses based on BIRC assessment will be repeated including all randomized patients, with the exception of the auxiliary variable estimate (NCI method) described above.

Sensitivity analyses and further supportive analyses

As a sensitivity analysis to assess the impact of stratification, the two treatment groups will be compared using an unstratified log-rank test. The hazard ratio together with associated 95% confidence interval obtained using the unstratified Cox regression model will also be presented.

The primary efficacy analyses; i.e. the stratified log-rank test, Kaplan-Meier estimates, estimate of the median PFS along with 95% confidence interval, and hazard ratio obtained using the Cox proportional hazards model, will be repeated (as appropriate) based on the data obtained:

1. Using the investigator/local assessments (Source 1 in Table 3-2) on the PPS and using the same conventions as for the primary analysis if the primary analysis is found to be statistically significant and if the PPS/FAS are different.
2. Using investigator/local assessment (Source 1 in Table 3-2) on the FAS and taking the event whenever it occurs - even after two or more consecutive missing tumor assessments. The following options from [Table 14-7 of Study Protocol Appendix 2] will be used: A(1), B(1), C1(1), C2(3), D(1), E(1) and F(1). In the summary table, this approach is referred as 'Actual event PFS analysis'
3. Using investigator/local assessment (Source 1 in Table 3-2) on the FAS and backdating events occurring after missing tumor assessments. The following options from [Table 14-7 in Section 3.2.9 of Study Protocol Appendix 2] will be used: A(1), B(1), C1(2), C2(2), D(1), E(1), and F(1). In the summary tables, this approach is referred as 'Backdating PFS analysis'.
4. Using investigator/local assessment (Source 1 in Table 3-2) on the FAS and censoring PFS at the last adequate tumor assessment before the start of any new antineoplastic therapy excluding palliative radio-therapy.

A stratified multivariate Cox regression model will be fitted to evaluate the effect of other baseline demographic or disease characteristic on the estimated hazard ratio. This model will include the treatment arm and the following key prognostic factors as covariates: age (>=65 vs <65), prior chemo therapy in (neo)adjuvant setting (yes vs no), ECOG performance status (0 vs. 1), and bone only lesion at baseline (yes or no), .In addition, stratified multivariate cox models will be fitted including the treatment arm and each of these prognostic factors individually, including treatment by factor interaction, to explore the relationship between each factor and treatment. Further supportive analyses will include:

- Stratified Cox regression using derived local radiology assessment (source 2 in Table 3-2). No p-value will be presented.
- Number of patients and number of events by treatment arm within each stratum will be presented along with the hazard ratio for treatment effect obtained using the Cox proportional hazards regression with corresponding confidence intervals. No p-values will be presented for this analysis. K-M plots of survival distributions will be presented by stratum.

- If there is a high rate of discrepancy between the strata classifications constructed using the eCRF data and those obtained from the IRT, a supportive analysis will be performed in which a stratified Cox regression model will be used to estimate the treatment hazard ratio and the associated 95% confidence interval based on the CRF-derived strata. No other inferential statistics will be provided.
- Number of patients with a PFS event and number of patients censored for the PFS analysis will be summarized. In addition, a summary of reasons for PFS censoring will be provided by treatment arm based on the reasons defined in Section 2.5.5. These summaries on censoring reasons will be produced for PFS as per local investigator radiology and BIRC assessment.
- Comparison of PFS event type/censor between local investigator radiology review and BIRC assessment
- If the primary analysis for PFS is statistically significant, the treatment effect for patients treatment naïve in the metastatic setting and for patients who received up to 1 line of treatment for advanced disease (groups A and B as defined in section 1.1) based on CRF data will be evaluated using unstratified Cox regression models. The hazard ratio with 95% CI will be displayed for each subgroup and Kaplan-Meier estimates will also be summarized. [redacted]

3.8.1.5 Handling missing month/day in date of death

For rare cases when either day is missing or both month and day are missing for the date of death, the imputation rules in section 3.8.2.1 will be implemented.

3.8.2 Secondary efficacy analyses

The secondary efficacy objectives in this study are to

- compare the two treatment arms with respect to overall survival
- evaluate the two treatment arms with respect to overall response rate and clinical benefit rate, time to response and duration of response;
- evaluate the two treatment arms with respect to time to deterioration of ECOG performance Status;
- evaluate patient reported outcomes for health-related quality of life in the two treatment arms.

The analysis of all secondary efficacy endpoints will be performed based on the FAS.

3.8.2.1 Overall survival

Comparing the overall survival (OS) between ribociclib +fulvestrant and placebo+fulvestrant is a secondary objective in this study. OS is defined as the time from date of randomization to date of death due to any cause. If a patient is not known to have died at the time of analysis

cut-off, then OS will be censored at the date of last contact. Assuming proportional hazards for OS, the following statistical hypotheses will be tested:

$$H_{02}: \theta_2 \geq 0 \quad \text{vs.} \quad H_{A2}: \theta_2 < 0$$

Where θ_2 is the log OS hazard ratio (fulvestrant+ ribociclib arm vs. fulvestrant + placebo arm). The analysis to test this hypothesis is a stratified log-rank test at an overall one-sided 2.5% level of significance. The stratification will be based on the randomization stratification factors.

A maximum of 3 analyses are planned for OS: at the time of the analysis for PFS (provided PFS is significant), at which point a total 161 deaths (46% of OS events) are expected, after 263 events (75% of OS events) have been documented, and a final analysis for OS when 351 deaths (100% of OS events) are expected (expected 56 months from date of first patient to be randomized).

OS tests will be carried out using a Lan-Demets (O'Brien-Fleming) alpha spending function. OS will only be tested if the primary efficacy analysis has been shown to be statistically significant. Analyses will be based on the FAS population according to the randomized treatment group and strata assigned at randomization. All deaths recorded up to the cut-off date will be included in the analysis. The survival distribution of OS will be estimated using the Kaplan-Meier method, and Kaplan-Meier curves, medians and 95% confidence intervals of the medians intervals along with the proportion of patients alive at 6 month interval starting from 12 month will be presented for each treatment group. The hazard ratio for OS will be calculated, along with its 95% confidence interval, using a stratified Cox model using the same stratification factors as the log-rank test.

If OS is statistically significant, a stratified multivariate Cox proportional hazard model adjusting for covariates will be fitted, including the treatment arm and the following key potential prognostic factors as covariates: age (>=65 vs <65), prior chemo therapy in (neo)adjuvant setting (yes vs no), ECOG performance status (0 vs. 1), and bone only lesion at baseline (yes or no). In addition, stratified multivariate Cox models will be fitted including treatment arm and each of these factors individually, including treatment by factor interaction, to explore the relationship between each factor and treatment.

[redacted] The pattern of censored data will be examined between the treatment arms: reasons for censoring ('Alive' or 'Lost to follow-up') and death cause will be summarized by treatment arm. In addition, survival status, reason for censoring and death cause will be listed. Patients not known to have died will be censored as 'Lost to follow-up' if the time between their last contact date and the analysis cut-off date is longer than the protocol specified interval between the survival follow-up assessments plus 2 weeks, i.e., 14 weeks (98 days) for this study.

Handling missing month/day in date of death

For rare cases when either day is missing or both month and day are missing for the date of death, the following imputation rules will be implemented:

- If only day is missing, then date of death is imputed as max [(1 mmm-yyyy), min(last contact date +1, cutoff date)].
- If both day and month are missing, then date of death is imputed as max [(1 Jan-yyyy), min (last contact date +1, cutoff date)].

3.8.2.2 Overall response rate

ORR is defined as the proportion of patients with best overall response of confirmed complete response (CR) or confirmed partial response (PR) according to RECIST 1.1 ([see Appendix 2 of the study protocol]). ORR will be calculated based on the FAS using investigators' review of tumor assessment data. Patients with only non-measurable disease at baseline will be part of the analysis and will be included in the numerator only if a complete response was observed. ORR will be presented by treatment arm along with standard Wald asymptotic (i.e. normal approximation) 95% confidence intervals. The Cochran-Mantel Haenszel chi-square test (strata based on IRT data) will be used to compare the two treatment arms with respect to the ORR at a one-sided 2.5% level of significance.

As a supportive analysis, ORR will also be summarized based on the BIRC review of tumor data. As a sensitivity analysis, ORR will be calculated and summarized for patients with only measurable disease at baseline.

3.8.2.3 Clinical benefit rate

CBR is defined as the proportion of patients with a best overall response of confirmed CR or PR, or SD lasting 24 weeks or longer, according to RECIST 1.1 criteria. A patient will be considered to have SD for 24 weeks or longer if a SD response is recorded at 24-1=23 weeks or later from randomization, allowing for the ±1 week visit window for tumor assessments. Patients with only non-measurable disease at baseline will be part of the analysis and will be included in the numerator only if they achieve a complete response or have a 'Non-CR/Non-PD' response 23 weeks or more after randomization. CBR will be calculated using the FAS based on the investigators' tumor assessments. CBR will be summarized for the two treatment arms using descriptive statistics. The Cochran-Mantel-Haenszel test (strata based on IRT data) at one sided 2.5% level of significance will be used to compare the two treatment arms with respect to CBR. CBR based on (1) BIRC review, and (2) only including patients with measurable disease at baseline, will be summarized for the two treatment arms using descriptive statistics.

3.8.2.4 Time to response

Time to response (CR or PR) is the time from date of randomization to first documented response (CR or PR, which must be confirmed subsequently) according to RECIST 1.1. All

patients in the FAS will be included in the time to response calculation. Patients without confirmed PR or CR will be censored at:

- the maximum follow-up time (i.e. FPFV to LPLV used for the analysis) for patients who had a PFS event (i.e. either progressed or died due to any cause);
- the last adequate tumor assessment date for all other patients.

Time to response data will be listed and summarized by treatment arm. The distribution of time to response will be estimated using the Kaplan-Meier method and the median time to response will be presented along with 95% confidence interval only if a sufficient number of responses are observed. A descriptive summary of time to response for the responders will also be presented.

3.8.2.5 Duration of response

DoR applies only to patients whose best overall response was CR or PR. The start date is the date of first documented response (CR or PR, which must be confirmed subsequently) and the end date is the date of event defined as the first documented progression or death due to underlying cancer. The start date will be determined using the time the response was first determined and not using the time the response was confirmed. If a patient has not had an event, duration will be censored at the date of last adequate tumor assessment using the same censoring rule described for primary PFS analysis. DoR will be summarized by treatment arm. The distribution of duration of response will be estimated using the Kaplan-Meier method and the median response duration will be presented along with 95% confidence interval only if a sufficient number of responses are observed.

3.8.2.6 ECOG performance status

The ECOG PS scale (Table 3-3) will be used to assess physical health of patients, ranging from 0 (most active) to 5 (least active):

Table 3-3 ECOG Performance Scale

Score	Description
0	Fully active, able to carry on all pre-disease performance without restriction
1	Restricted in physically strenuous activity but ambulatory and able to carry out work of a light or sedentary nature, e.g., light house work, office work
2	Ambulatory and capable of all self-care but unable to carry out any work activities. Up and about more than 50% of waking hours
3	Capable of only limited self-care, confined to bed or chair more than 50% of waking hours
4	Completely disabled. Cannot carry on any self-care. Totally confined to bed or chair
5	Dead

The following intervals will be used to group the ECOG PS data over time. Day in columns 2 and 3 is defined as date of ECOG PS assessment date – randomization date + 1. The corresponding Day in column 1 assumes that a patient is treated on the day of randomization; however the definition of Day in columns 2 and 3 still applies if this is not the case, i.e. randomization date is taken as the reference for the windows.

Table 3-4 Time windows for ECOG PS assessments

Assessment	Target day of assessment	Time Interval
Baseline		Day 1 (if not available use screening)
Cycle 2 Day 1	29	Day 2 to day 42
Cycle k Day 1 (k≥3)	d=(k-1)*28+1	Day d-14 to day d+13
End of Treatment		Assessment taken at the end of treatment visit

If more than one assessment is done within the same time window, the assessment performed closest to the target date will be used. If 2 assessments within a time window are equidistant from the target date, then the worst of the 2 assessments will be used.

Frequency counts and percentages of patients in each score category will be provided by treatment arm and time point.

An analysis of time to definitive deterioration of performance status will be performed. The time to definitive deterioration is calculated from the date of randomization. Definitive deterioration is defined as an increase in ECOG PS by at least one category from the baseline score or death due to any cause. Deterioration is considered definitive if no return to baseline or better in ECOG PS is observed subsequent to the deterioration during the treatment period.

If a definitive deterioration is observed after two or more consecutive missing assessments, time to deterioration will be censored at the date of the last ECOG PS assessment prior to the deterioration. The number of missing assessments is calculated based on the time window in Table 3-4 and a rule similar for tumor assessments (see detail in section 2.5.5).

Patients receiving any further anti-neoplastic therapy before definitive deterioration will be censored at the date of their last assessment before the start date of the therapy. Patients that have not worsened as of the cutoff date will be censored at the date of their last assessment before the cutoff.

Kaplan-Meier estimates will be constructed for each treatment arm. The median, 25th and 75th percentiles for time to definitive deterioration for each treatment group will be obtained along with 95% confidence intervals.

A stratified log-rank test at one-sided 2.5% level of significance will be used to test the difference in time to definitive worsening of performance status between treatment arms. A (stratified) Cox proportional hazards model will be used to estimate the hazard ratio (with 95% confidence interval).

3.8.2.7 Patient reported outcomes

The European Organization for Research and Treatment of Cancer's core quality of life questionnaire (EORTC-QLQ-C30, version 3.0) and the EuroQoL 5-level instrument (EQ-5D-5L, Version 4.0) will be used to evaluate patient-reported outcome measures of health-related quality-of-life, functioning, disease symptoms and treatment-related side effects. The BPI-SF will be used to assess patient's subjective assessment of pain.

The PRO instruments are planned to be administered during screening and every 8 weeks after randomization in the first 18 months, and every 12 weeks thereafter until the end of treatment. PRO assessments will continue to be collected during the efficacy follow-up after the end of treatment.

The following time based intervals will be used to group the PRO data over time. Day is defined as date of PRO assessment date – randomization date + 1.

Table 3-5 Time windows for patient reported outcomes

	Time Interval
Baseline	Screening assessment
Cycle 3, 5, 7, 9 until cycle 19 Day 1	+/- 4 weeks centered around the planned assessment date (except for the first window and the last window): i.e. days (1, 85] for Day 1 of cycle 3 (2^{th} assessment) days (85, 141] for Day 1 of cycle 5 (3^{th} assessment) days (k*56-27; k*56+29] for (k+1^{th} assessment) days (477, 547] for 10^{th} assessment on Day 1 cycle 19
Cycle 22, 25, 28, … Day 1	+/- 6 weeks centered around the planned assessment date: i.e. days (547, 631] for 11^{th} assessment days (631, 715] for 12^{th} assessment days (715, 799] for 13^{th} assessment
End of treatment	Assessment taken at the end of treatment visit
Efficacy follow-up	At each of the efficacy follow-up visits

If more than one assessment is done within the same time window, the assessment performed closest to the target date will be used. If 2 assessments within a time window are equidistant from the target date, then the assessment obtained prior to target date will be used. If the closest assessment to the target date has two questionnaires filled out on the same date, then the worst score of these assessments will be used for each subscale score.

The FAS will be used for all PRO summaries and listings.

The global health status/global QoL scale score of the EORTC QLQ-C30 is identified as a primary PRO variable of interest. The physical functioning, emotional functioning and social functioning sub-scale scores of the EORTC QLQ-C30, the visual analog scale (VAS) of the EQ-5D-5L, and worst pain item (based on question 3), pain severity index, and pain interference indices of the BPI-SF are identified as secondary PRO variables of interest. High scores in the EORTC QLQ-C30 represent a higher response level. Thus a high score for a functional scale represents a high / healthy level of functioning; a high score for the global health status / QoL represents a high QoL, but a high score for a symptom scale / item represents a high level of symptomatology /problems. Higher scores in the EQ-5D-5L also correspond to better health states. A higher score in BPI-SF corresponds to more pain.

The number of patients completing PRO questionnaires and the number of patients missing/expected to have PRO assessments will be summarized by treatment arm for scheduled assessment time points (the number of ongoing patients will be used as

denominator). Furthermore, the amount and the pattern of missing data may be explored by treatment arm and over time using summary statistics. The following categories will be used to describe whether the questionnaire was completed at a specific time point:

- yes, fully completed
- yes, partly completed
- no.

Scoring of raw data and methods for handling missing items or missing assessments will be handled according to scoring manuals for each respective patient questionnaire (Fayers 2001; Oemar and Janssen 2013; Cleeland 2009).

As the main analysis and to best utilize the repeated PRO assessments, a repeated measures model for longitudinal data will be used to compare the two treatment arms in sub-scales obtained from EORTC QLQ-C30 and the VAS score of EQ-5D-5L over time. This longitudinal model will include terms for treatment, the randomization stratification factors, time (duration in weeks counting from the time of baseline measurement to the time of a particular post baseline measurement), baseline value as main effects, and an interaction term for treatment by time. Time will be considered as a continuous variable in this analysis. As a first approach, an unstructured correlation matrix will be used to model the correlation within patients. The structure of the correlation matrix will be investigated and simplified using likelihood ratio tests if appropriate. Time is considered as a continuous variable in the analysis. The data for selected time points will be fitted using a linear model using treatment, stratification factors, baseline value as covariates. The differences in least square means between the experimental and control group, and the corresponding 2-sided 95% CI at selected time points will be presented.

For the mixed effects model, patients with baseline and at least one non-missing post-baseline assessments will be included. This analysis will only include assessments up to the time point where there are at least 50 patients on each of the treatments. Data collected up to EOT assessment will be included.

Analysis of the time to definitive 10% deterioration in the primary and the secondary PRO variables of interest will be performed. Definitive 10% deterioration is defined as a worsening in score by at least 10% compared to baseline, with no later improvement above this threshold observed during the treatment period, or death due to any cause. Time to deterioration is the number of days between the date of randomization and the date of the assessment at which deterioration is seen. If a patient has not had an event prior to analysis cut-off, start of new anti-neoplastic therapy, lost to follow-up, end of treatment or withdrawal of consent, the time to deterioration will be censored at the date of the last evaluation before the earliest of these times. Only assessments collected while the patient is on treatment and on or before the end of treatment visit will be included in the PRO time to deterioration analysis. If deterioration is observed after two or more missing assessments, time to deterioration will be censored at the date of the last assessment prior to the deterioration. The rules for calculating number of missing assessments are the same as those for tumor assessments (see details in Section 2.5.5). Time to 10% deterioration in the global health status/QOL scale, and secondary PRO scales of interest as listed above, will be compared between the two treatment arms using a

stratified log-rank test (strata based on IRT data) at 2 sided 5% level of significance. The survival distributions will be presented descriptively using Kaplan-Meier curves. Summary statistics from the Kaplan-Meier distributions will be determined, including the median time to 10% deterioration and the proportions of patients without10% deterioration at 2 month intervals. Both point estimates and 95% CIs will be presented. A stratified Cox regression model will be used to estimate the hazard ratio (HR) of time to deterioration, along with 95% confidence interval. Sensitivity analysis of time to definitive deterioration with different cut-off definitions (e.g. 5%, 15%) may also be considered if the number of events per arm is judged sufficient to draw relevant conclusions.

Descriptive statistics (n, mean, median, SD, min, max) will be used to summarize the individual item and subscale scores from the EORTC QLQ-C30, EQ-5D-5L, and worst pain item (question 3), pain severity index and pain interference indices of the BPI-SF at each scheduled assessment time point. Additionally, change from baseline in the subscale scores at the time of each assessment will be summarized. Patients with an evaluable baseline score and at least one evaluable post baseline score during the treatment period will be included in the change from baseline analyses.

3.9 Safety evaluation

The assessment of safety will be based mainly on the frequency of adverse events and on the number of laboratory/ECG values that fall outside of pre-determined ranges. Other safety data (e.g. vital signs and special tests) will be considered as appropriate.

All safety outputs will use the safety set. The safety summary tables will include ‘on-treatment’ events/assessments, i.e. those collected on or after the first date of study treatment and collected no later than 30 days after the date of last study treatment administration. The AEs started before the first dose but worsening during the treatment period are also considered as ‘on-treatment’ events. All safety events/assessments will be listed and those collected outside of the on-treatment window will be flagged.

3.9.1 Adverse events (AEs)

3.9.1.1 Coding of AEs

Adverse events are coded using the Medical Dictionary for Regulatory Activities (MedDRA) terminology.

3.9.1.2 Grading of AEs

AEs will be assessed according to the Common Terminology Criteria for Adverse Events (CTCAE) version 4.03.

The CTCAE represents a comprehensive grading system for reporting the acute and late effects of cancer treatments. CTCAE v4.03 grading is by definition a 5-point scale generally corresponding to mild, moderate, severe, life threatening, and death.

If CTCAE grading does not exist for an adverse event, grades 1 – 4 corresponding to the severity of mild, moderate, severe, and life-threatening will be used. CTCAE grade 5 (death)

will not be used in this project; if an AE results in death it will be documented in the outcome ("fatal"). Information on deaths will also be collected on the 'Death' CRF.

3.9.1.3 General rules for AE Reporting

AE summaries will include all AEs starting on or after study Day 1 (i.e. on or after the day of the first intake of study treatment) and starting no later than 30 days after the last administration of study treatment (see Section 2.1.5). All AEs will be listed. AEs starting prior to study Day 1 and AEs starting later than 30 days after the last treatment date will be flagged in the listings.

AEs will be summarized by presenting the number and percentage of patients having at least one AE, having at least one AE in each primary system organ class, and for each preferred term using MedDRA coding. A subject with multiple occurrences of an AE will be counted only once in the AE category.

Separate AE summaries will be presented by primary system organ class, preferred term, and maximum grade. A patient with multiple grades for an AE will be summarized under the maximum grade recorded for the event. AE with missing CTCAE grade will be included in the 'All grades' column of the summary tables.

In AE summaries the primary system organ class will be presented alphabetically and the preferred terms will be sorted within primary SOC in descending frequency. The sort order for the preferred term will be based on their frequency in the ribociclib arm.

The frequency of grade 3 and 4 AEs will be summarized separately.

Any information collected (e.g. grades, relationship to study treatment, action taken etc.) will be summarized and listed as appropriate.

3.9.1.4 AE summaries

The following adverse event summaries will be produced by treatment group:

- Summary of deaths and adverse events
- Adverse events, irrespective of causality, by primary system organ class, preferred term and maximum grade
- Adverse events with suspected relationship to study treatment by primary system organ class, preferred term and maximum grade
- Most common grade 3-4 adverse events, irrespective of causality, by preferred term and maximum grade (greater than x% in either arm)
- Adverse events, irrespective of causality, by primary system organ class and maximum grade
- Adverse events, irrespective of causality, by preferred term and maximum grade
- Adverse events with suspected relationship to study treatment by preferred term and maximum grade

- Grade 3 or 4 adverse events, irrespective of causality, by primary system organ class, preferred term and maximum grade
- Grade 3 or 4 adverse events with suspected relationship to study treatment by primary system organ class ,preferred term and maximum grade
- On treatment deaths by preferred term
- Deaths, by primary system organ class and preferred term
- Serious adverse events, irrespective of causality, by primary system organ class and preferred term and maximum grade
- Serious adverse events with suspected relationship to study treatment, by primary system organ class, preferred term and maximum grade
- Adverse events leading to study drug discontinuation, irrespective of causality, by primary system organ class, preferred term and maximum grade
- Adverse events leading to study drug reductions, irrespective of causality, by primary system organ class, preferred term and maximum grade
- Adverse events leading to study drug interruptions, irrespective of causality, by primary system organ class, preferred term and maximum grade
- Adverse events requiring additional therapy, irrespective of causality, by primary system organ class, preferred term and maximum grade
- Non-serious adverse events (at least x% incidence rate in either treatment arms) by primary system organ class and preferred term
- Serious and non-serious adverse events with number of occurrences (an occurrence is defined as >1 day between start and prior end date of record of same preferred term)
- On-treatment deaths and SAEs with fatal outcome, by SOC and PT

AEs of interest will also be summarized. See Section 3.9.1.5 for the grouping details.

3.9.1.5 Grouping of adverse events of special interest

An adverse event of special interest is a grouping of adverse events that are of scientific and medical concern specific to ribociclib. These groupings are defined using MedDRA terms, SMQs (standardized MedDRA queries), HLGTs (high level group terms), HLTs (high level terms) and PTs (preferred terms). Customized SMQs (Novartis MedDRA queries, NMQ) may also be used. A NMQ is a customized group of search terms which defines a medical concept for which there is no official SMQ available or the available SMQ does not completely fit the need. It may include a combination of single terms and/or an existing SMQ, narrow or broad. For each specified AESI, the number and percentage of patients with at least one event of the AESI occurring during the on-treatment period will be summarized.

Summaries of these AESIs will be provided by treatment arm, (specifying grade, SAE, relationship, leading to treatment discontinuation, leading to dose adjustment/interruption, etc.).

A Case Retrieval Sheet (CRS) with the exact composition of the AE groupings is to be used to map reported AEs to the AESI groupings. This file may be updated (i.e. it is a living document) based on review of accumulating trial data, and therefore the groupings are also subject to potential change. Table 3-6 provides the latest groupings at the time of the finalization of the SAP, from the CRS dated 13Dec2016. The most up-to-date version of the CRS will be used at the time of the analysis.

Table 3-6 AESI groupings

AESI grouping	MedDRA category
Anemia	SMQ
Diarrhea	SMQ
Hepatobiliary toxicity	SMQ
Infections	SMQ and SOC
Leukopenia	HLT
Nausea, emesis	HLT
Neutropenia	HLT and PT
Pneumonitis, interstitial lung disease	SMQ
Pulmonary embolism	SMQ and HLT
QTc prolongation	SMQ
Renal impairment	SMQ
Reproductive toxicity	SMQ
Thrombocytopenia	SMQ

3.9.2 Laboratory data

On analyzing laboratory data, data from all sources (central and local laboratories) will be combined. The summaries will include all laboratory assessments collected no later than 30 days after the last administration of study treatment. All laboratory assessments will be listed and those collected later than 30 days after the last treatment date will be flagged in the listings.

Laboratory data will be classified (by biostatistics/statistical programming) into CTC grades according to the NCI Common Terminology Criteria for Adverse Events (CTCAE) v4.03. The calculation of laboratory CTC grades will be based on the observed laboratory values only; clinical assessments will not be taken into account. The criteria to assign CTC grades in this study are given in Appendix 1.

For laboratory tests where grades are not defined by CTCAE, results will be graded by the low/normal/high classifications based on laboratory normal ranges.

The following summaries will be produced for the laboratory data (by laboratory parameter):

- Box plots of laboratory values by scheduled time point and treatment arm.

- Number and percentage of patients with each CTC grade as their worst post-baseline value (regardless of the baseline status). Each patient will be counted only for the worst grade observed post baseline.
- Shift tables using CTC grades to compare baseline to the worst post-baseline value will be produced for hematology and biochemistry laboratory parameters with CTC grades.
- For laboratory parameters where CTC grades are not defined, shift tables to the worst post-baseline value will be produced using the low/normal/high classifications based on laboratory reference ranges.

Number and percentage of patients meeting categorical liver function test criteria, including ALT, AST and ALT/AST (>3x, 5x, 8x, 10x, 20x ULN), Total Bilirubin (>1x, 2x ULN), ALP (>1.5x, 2x, 3x, 5x, 8x, 10x ULN), combined categories of ALT/AST and total bilirubin (e.g., ALT/AST>3x UNL & total bilirubin > ULN) as well as Hy's Law criteria (ALT or AST > 3 x ULN and TBIL >= 2 x ULN and ALP < 2 x ULN). For the combined categories, the assessments need not to be concurrent, i.e. patients are counted based on their most extreme value for each parameter (highest in the case of ALT, AST and TBIL; lowest in the case of ALP).The following listings will be produced for the laboratory data:

- Listing of patients with CTC grade 3 or 4 laboratory abnormalities;
- Listing of all laboratory data with values flagged to show the corresponding CTC grades and the classifications relative to the laboratory reference ranges.

Time to first occurrence of grade 2 or worse laboratory toxicity and time to first occurrence of grade 3 or worse laboratory toxicity will be summarized for neutrophil, and ALT/AST, using the Kaplan-Meier method. Median time to first occurrence and 95% CI will be provided. In addition, Kaplan-Meier plots will be generated. Simple descriptive statistics, median, min and max as well as 25th percentile and 75th percentile, will be presented.

Time to first occurrence of grade X or worse laboratory toxicity is defined as the time from the start of treatment to the start date of the first incidence of grade X or worse laboratory toxicity, i.e. time in days is calculated as (start date of first occurrence) – (date of first dose of study treatment) +1. A patient will be censored if:

- The patient did not report any post-baseline grade X or worse event on or before the analysis cut-off date.
- The patient discontinued treatment without reporting any grade X or worse event up to 30 days after study treatment discontinuation.
- The patient died without reporting any grade X or worse event.
- The patient received a new anticancer therapy before reporting any grade X or worse event.

The censoring date will be the earliest of the following dates: end of treatment + 30 days, analysis cut-off, new anti-cancer therapy start date, death date and last non-missing

assessment for the lab parameter. Note that patients who have grade X or worse toxicity at the baseline or missing baseline evaluation will be excluded from this analysis.

Duration of grade X or worse laboratory toxicity (for neutrophil, X=2; for ALT or AST, X=3) may also be summarized. Duration of grade X or worse event is calculated as:

(Date when the grade of the event decreases to below X) – (date of onset of grade X or worse event) + 1

For patients experiencing any grade X or worse event, the duration of the first such event will be summarized using the Kaplan-Meier method. Median duration and 95% CI will be presented. In addition, Kaplan-Meier plots will be generated.

A patient will be censored for the duration of grade X or worse event, if:

- The patient dies without reporting a decrease to below grade X
- The patient receives a new anticancer therapy before reporting a decrease to below grade X
- The patient discontinues from the study treatment without reporting a decrease to below grade X up to 30 days after study treatment discontinuation
- The patient is still ongoing at the analysis cut-off date without reporting a decrease to below grade X

The censoring date is the earliest of the following dates: end of treatment + 30 days, analysis cut-off, new anti-cancer therapy start date, death date and last non-missing assessment for the lab parameter.

In addition, failure curves (ascending Kaplan-Meier curves) for time to first occurrence of AE will be constructed by treatment arm. Median together with 95% confidence interval as well as 25th percentile and 75th percentile will be presented for each treatment arm.

3.9.3 Vital signs

Vital signs assessments are performed in order to characterize basic body function. The parameters expected to be collected include: height, weight, body temperature, heart rate, and systolic and diastolic blood pressure.

The criteria for clinically notable abnormalities are defined as follows:

Clinically notable elevated values

- Systolic BP: ≥ 180 mmHg and an increase ≥ 20 mmHg from baseline
- Diastolic BP: ≥ 105 mmHg and an increase ≥ 15 mmHg from baseline.
- Body temperature: ≥ 39.1°C
- Heart rate: ≥ 120 bpm with increase from baseline of ≥15 bpm

Clinically notable below normal values

- Systolic BP: ≤ 90 mmHg and a decrease ≥ 20 mmHg from baseline
- Diastolic BP: ≤ 50 mmHg and a decrease ≥ 15 mmHg from baseline
- Body temperature: ≤ 35°C
- Heart rate: ≤ 50 bpm with decrease from baseline of ≥ 15 bpm

The following summaries will be produced for each vital sign parameter:

- Summary statistic for change from baseline to the worst post-baseline value (in both directions, i.e. from baseline to highest post baseline and from baseline to lowest post baseline value).
- Number and percentage of patients with at least one post-baseline vital sign abnormality (in both directions, i.e. both elevated and below normal values).

In addition, the following two listings will be produced by treatment arm:

- Patients with clinically notable vital sign abnormalities.
- All vital sign assessments will be listed by patient and vital sign parameter.

In both listings, the clinically notable values will be flagged and also the assessments collected later than 30 days after the last treatment date will be flagged.

3.9.4 ECG

All analyses of ECG data will be based on the average of all available replicate ECGs at each time point for each patient. For unscheduled assessments, 15-minute windows will be applied to group assessments for averaging.

ECG data will be summarized by presenting summary statistics of the raw data and change from baseline by treatment arm and time point. The following parameters will be assessed: QT, QTcF, QTcB, PR, and QRS intervals in msec, heart rate (bpm), and the overall interpretation if clinically significant abnormalities are present.

- The number and percentage of patients with notable abnormalities will be summarized.
- Summary statistics and shift tables will be presented.
- Individual listings will be provided by subject.

Table 3-6 Clinically notable ECG values

ECG parameter (unit)	Clinically notable criteria
QT, QTcF, QTcB (ms)	New > 450 New > 480 New > 500 Increase from Baseline > 30 Increase from Baseline > 60
PR duration (ms)	Increase > 25% from baseline and to PR duration > 200, New >200
QRS duration (ms)	Increase > 25% from baseline and to QRS duration > 110, New >110
Heart Rate (bpm)	< 50 and decrease from Baseline of > 25%

> 100 and increase from Baseline of > 25%

A newly occurring ECG abnormality is defined as an abnormal post-baseline ECG finding that is not present at Baseline. Baseline is defined as the average of the last ECG measurements (replicates taken on or before date of first dose of study treatment. The percentage of patients having notable ECG interval values is based on the number of patients at risk for the change with a value at baseline and post-baseline.

Time to grade 2 or worse QTcF prolongation will be analyzed using the Kaplan-Meier method and the median time to grade 2 or worse QTcF prolongation will be presented along with a 95% confidence interval if there are sufficient events for each treatment group. Simple descriptive statistics, median, min and max as well as 25th percentile and 75th percentile, will be presented. Note that patients who have grade 2 or worse toxicity at the baseline or missing baseline evaluation will be excluded from this analysis.

3.9.5 Cardiac imaging (MUGA / ECHO)

Shift tables comparing baseline to worst post-baseline cardiac imaging (MUGA or ECHO) overall interpretation will be provided. Percentages will be based on all patients in the Safety set.

Note: If there is any change in the methodology used throughout the study compared to baseline, the post-baseline values for which the methodology differs from baseline will be discarded in the tables presenting comparisons to baseline.

Descriptive statistics of the left ventricular ejection fraction (LVEF) at baseline, worst post-baseline value and change from baseline to worst post-baseline value will be provided.

A listing of patients with newly occurring clinically significant abnormality will be produced by treatment arm.

3.9.6 Urinary Analysis

The following parameter will be summarized using shift table to compare baseline to the worst-post baseline values: urine bilirubin dipstick, urine blood dipstick, urine glucose dipstick, urine ketones dipstick, urine leukocyte dipstick, and urine nitrate dipstick. For all these parameters, both negative and trace are considered as normal and the more pluses the worse. The urine pH dipstick will be summarized using shift table with low/normal/high classifications based on laboratory reference ranges.

3.9.7 Other safety data

Data from other tests will be listed, notable values will be flagged, and any other information collected will be listed as appropriate.

All assessments collected later than 30 days after the last treatment date will be flagged in the listings.

Any statistical tests performed to explore the data will be used only to identify any interesting comparisons that may warrant further consideration.

Subgroup analyses will be explored as described in Section 3.14.1.

3.10 Pharmacokinetic (PK) analyses

3.10.1 General principle

All PK analyses will be based on PAS set, unless otherwise specified.

Evaluable pre-dose concentrations satisfy the below conditions:

- Sample was collected before the current dose
- In addition, for pre-dose concentrations collected on Day 15 of any cycle
    - Patient had at least 10 consecutive days of ribociclib dosing (either 10 doses at 600 mg or 10 doses at 400 mg or 10 doses at 200 mg) immediately prior to the pre-dose collection. Ten consecutive doses are expected to provide adequate time to reach steady state for ribociclib at 600 mg and after dose reduction.
    - The PK collection was done at 24 ± 2 hours after the previous dose on Day 14 of the corresponding cycle.
    - Patient did not vomit within the first 4 hours following the last dose (confirmed using the records of date/time of PK sample collection and dosing).
    - the concentration has not been flagged for exclusion by the pharmacokineticist

Evaluable post dose 2h, 4h, 6h concentrations on Day 15 of any cycle satisfy the below conditions:

- Patient had at least 10 consecutive days of ribociclib dosing (either 10 doses at 600 mg or 10 doses at 400 mg or 10 doses at 200 mg) immediately prior to the PK collection.
- The PK collection was done within ± 15 minutes window of 2 hours after the Day 15 dosing for the post 2h concentration
- The PK collection was done within ± 30 minutes window of 4, 6 hours after the Day 15 dosing
- Patient did not vomit within the first 4 hours following the current dose.
- the concentration has not been flagged for exclusion by the pharmacokineticist

Only evaluable PK concentrations which are not flagged for exclusion will be used for figures, summaries, and statistical analysis. Concentration listings will include all concentrations, with flags indicating those excluded from analyses.

3.10.2 PK concentrations

PK concentrations of ribociclib and LEQ803 at C2D1 will be summarized for fulvestrant+ ribociclib arm.

In addition, ribociclib and LEQ803 steady state concentrations (i.e., Cycle 1 Day 15 and Cycle 2 Day 15 concentrations) will be also summarized by ribociclib dose level, visit, and timepoint. Patients will be classified into the following dose groups at each visit (C1D15 and C2D15) and timepoint:

- RIBO600: consists of all patients who provided evaluable concentrations after receiving at least 10 consecutive daily ribociclib doses of 600 mg immediately prior to the PK collection without a dose change or interruption.
- RIBO400: consists of all patients who provided evaluable concentrations after receiving at least 10 consecutive daily ribociclib doses of 400 mg immediately prior to the PK collection without a dose change or interruption.
- RIBO200: consists of all patients who provided evaluable concentrations after receiving at least 10 consecutive daily ribociclib doses of 200 mg immediately prior to the PK collection without a dose change or interruption.

Additional separate summaries of ribociclib and LEQ803 will be produced for Asian and non-Asian patients (based on race) at each visit (C1D15 and C2D15) by ribociclib dose level and timepoint.

Descriptive statistics of concentrations will include n, m (number of non-zero concentrations), mean, CV%, SD, median, geometric mean, geometric CV%, minimum and maximum). Coefficient of variation CV% is calculated as below:

100*(SD/arithmetic mean)

Geometric CV% is calculated as follows from non-zero values:

$$CV(\%) = 100 \cdot \sqrt{\exp(\hat{\sigma}^2) - 1}$$

where $\hat{\sigma}^2$ denotes the variance of the log-transformed values.

All individual plasma ribociclib (and also LEQ803) concentration data will be listed for fulvestrant + ribociclib arm using the Full analysis set. The meal record will be listed using Full Analysis Set.

3.10.3 Analysis of relationship between efficacy or safety endpoints and exposure

3.10.3.1 Exposure vs. PFS and TTR

The relationship between ribociclib exposure and PFS/TTR will be explored using the Kaplan-Meier method. Kaplan-Meier plots of the distribution of PFS and TTR by category of average Ctrough at steady state (denoted as Ctrough_avg_ss) will be presented, along with the median time for each category if estimable. For each patient, Ctrough_avg_ss will be calculated as the geometric mean of evaluable Ctrough values from C1D15 and C2D15. For the patients with only one evaluable Ctrough, the Ctrough_avg_ss will be that evaluable Ctrough itself. The Ctrough_avg_ss categories are defined as follows: < 25%, 25% to < 50%, 50% to < 75%, and ≥ 75%. The data will be analyzed using PAS.

3.10.3.2 Exposure vs. QTcF

The relationship between ribociclib concentration and the mean change from baseline QTcF (i.e., ΔQTcF) will be explored graphically: the arithmetic mean ΔQTcF and geo-metric mean ribociclib concentration will be displayed together against nominal time (time in x-axis, ΔQTcF and concentration in y-axis) using the PAS. The mean will be taken with respect to all the measurements from all patients at the same given nominal time. If a patient does not have both QTcF and evaluable PK measurement at a certain time point, then this patient will not be included in the plot for that time point.

In addition, the relationship between ribociclib dose level and QTcF will also be explored using the safety set. Patients will be classified into the following dose groups at each scheduled QTcF collection time point on C1D15 and C2D15.

- RIBO600: consists of all patients who received at least 10 consecutive daily ribociclib doses of 600 mg without a dose change or interruption immediately prior to the corresponding QTcF collection timepoint.
- RIBO400: consists of all patients who received at least 10 consecutive daily ribociclib doses of 400 mg without a dose change or interruption immediately prior to the corresponding QTcF collection timepoint.
- RIBO200: consists of all patients who received at least 10 consecutive daily ribociclib doses of 200 mg without a dose change or interruption immediately prior to the corresponding QTcF collection timepoint.

Summary statistics of the raw QTcF data and change from baseline will be presented by visit, time point and dose group using the safety set. Only scheduled QTcF on C1D15 and C2D15 will be included in the analysis.

3.10.3.3 Exposure vs. liver function

Box plots for ribociclib Ctrough_avg_ss will be presented by worst post_baseline LFT lab category (for TBIL, AST, ALT, ALP separately) using the PAS set. The categories are defined as:

- TBIL ("≤ 2xULN" vs. ">2xULN"),
- AST ("≤ 5xULN" vs. ">5xULN"),
- ALT ("≤ 5xULN" vs. ">5xULN"),
- ALP ("≤ 1.5xULN" vs. ">1.5xULN")

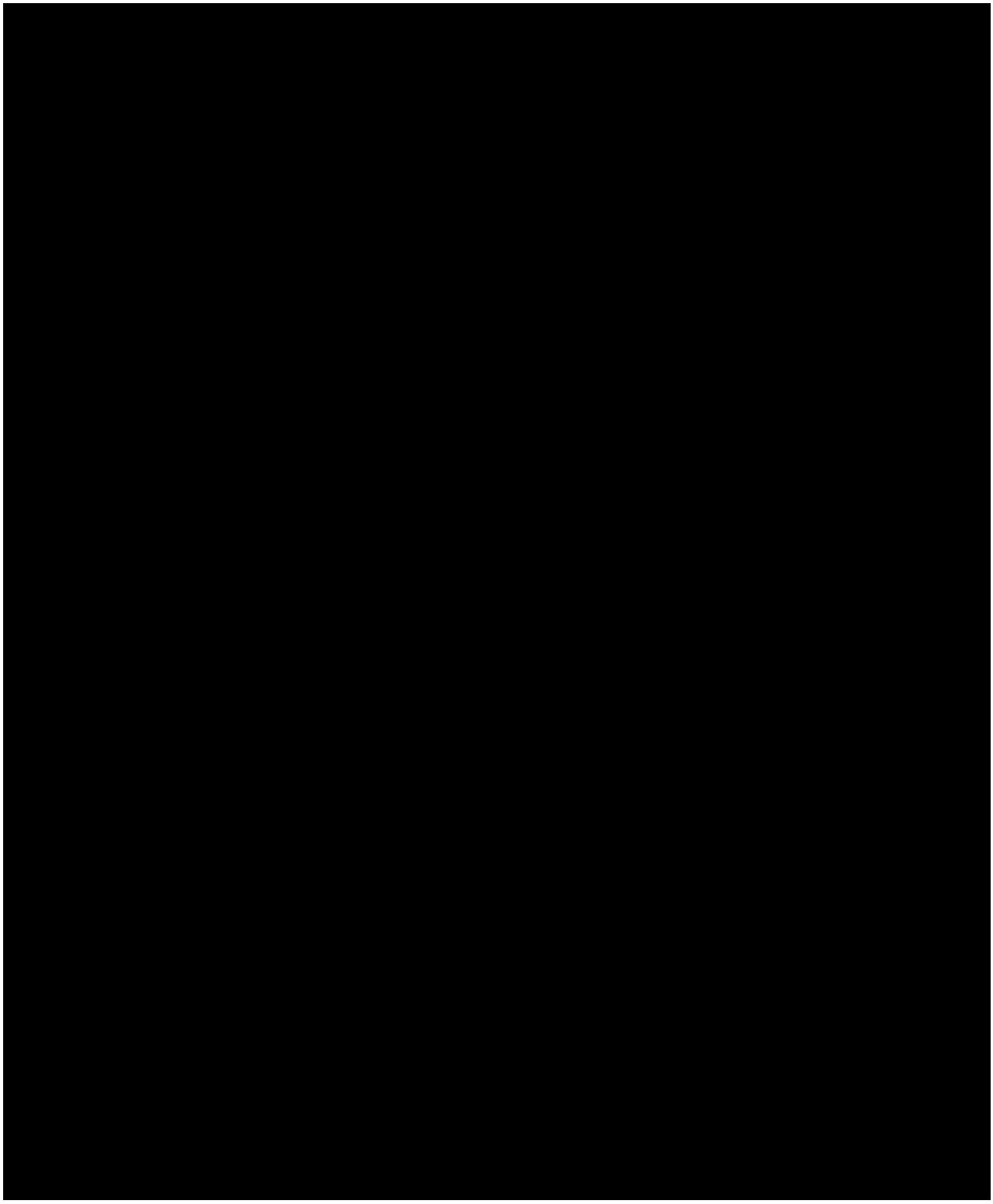
3.10.4 Handling missing and invalid values

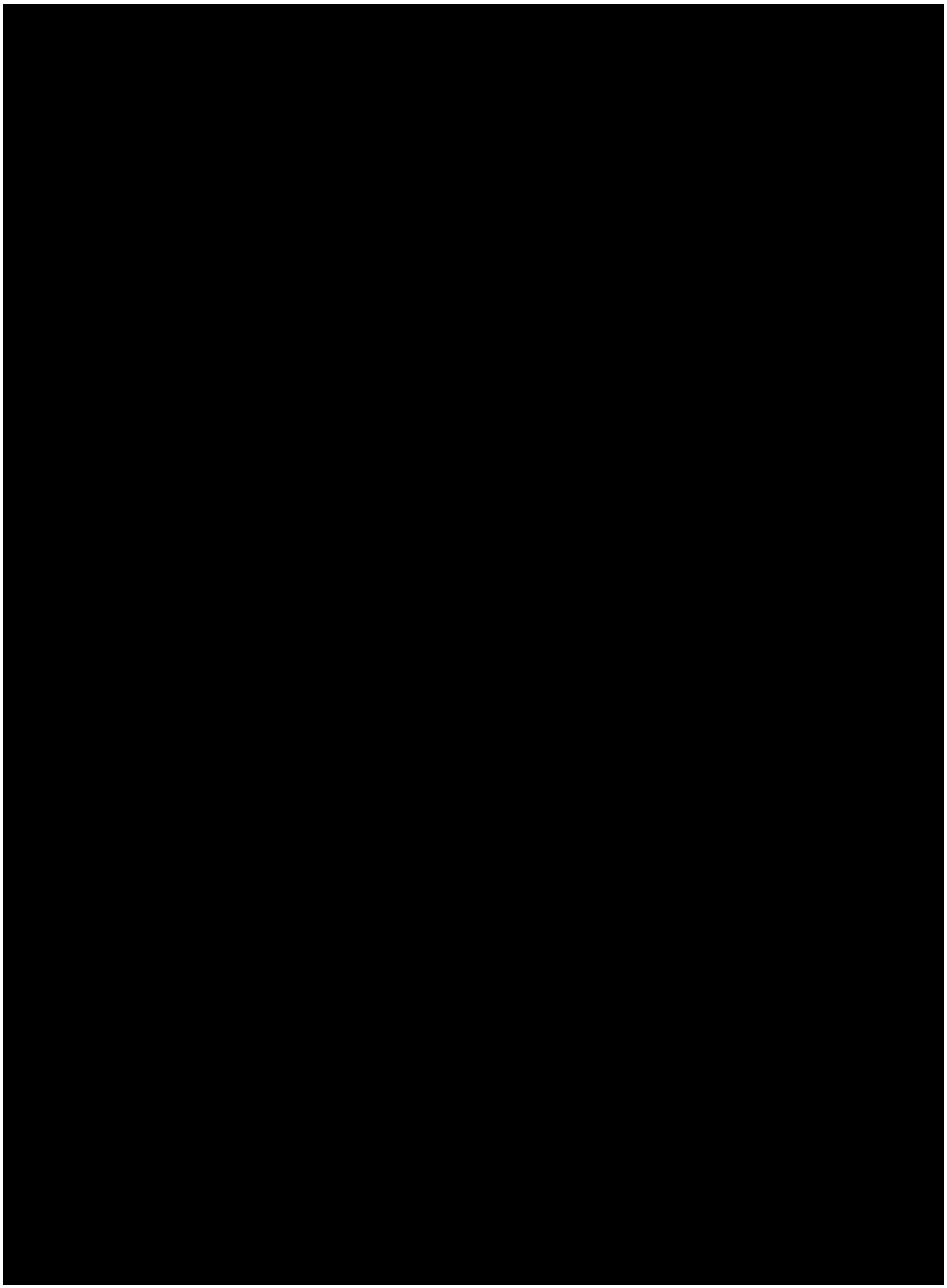
Plasma samples will be assayed for ribociclib concentrations by Novartis or Novartis designated laboratory using validated LC-MS/MS methods with a LLOQ of approximately 1.00 ng/mL for ribociclib.

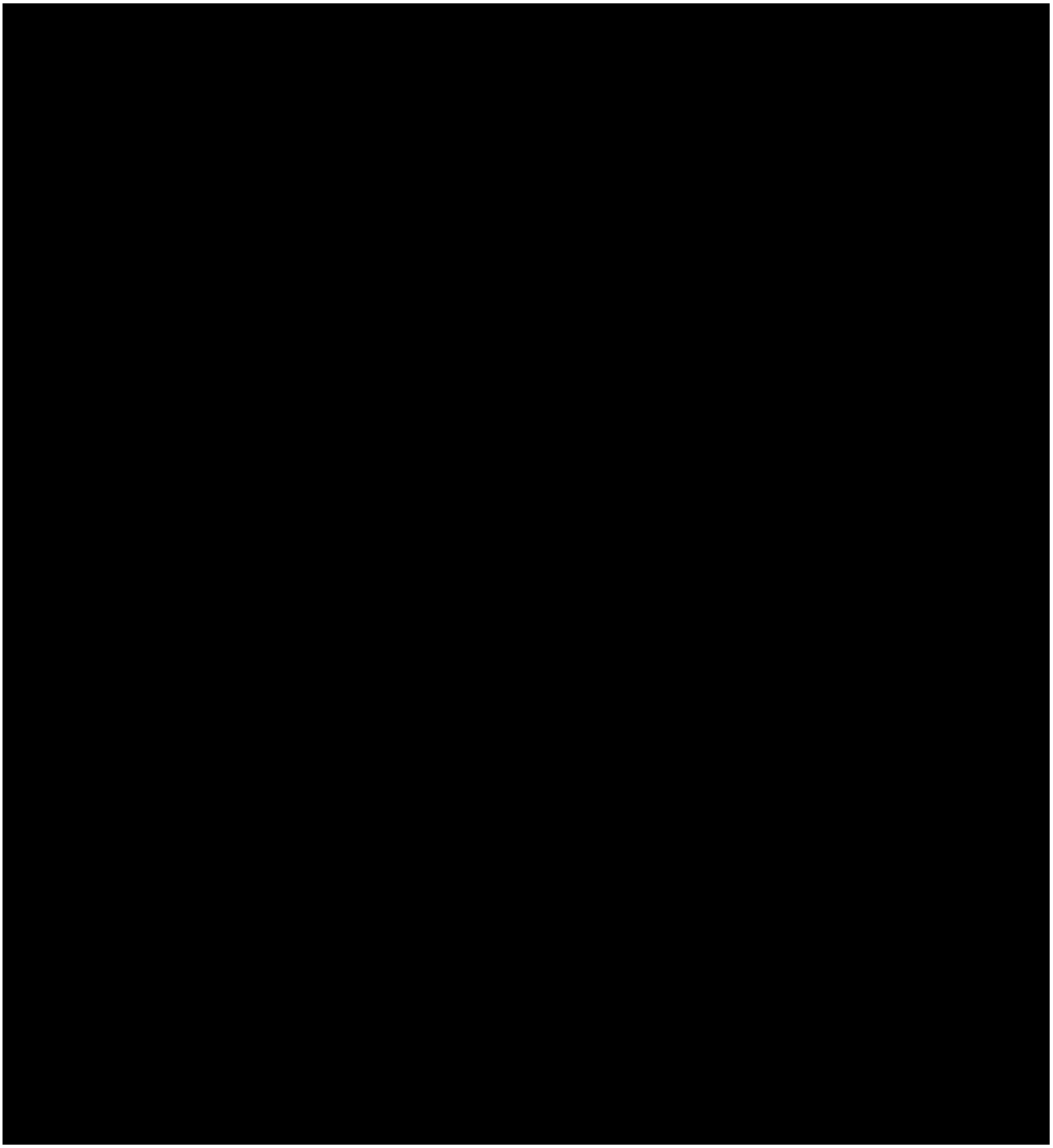
All concentrations below the LLOQ will be displayed in listings as zero with a flag and handled as zero in any calculations of summary statistics, but handled as missing for the

calculation of the geometric means and their CV%. The number of non-zero concentrations will also be reported in the summary statistics.

Missing values for any PK data will not be imputed and will be treated as missing.







[redacted]

3.14 Interim analyses

3.14.1 Primary Endpoint: PFS

No Planned interim analysis for PFS.

3.14.2 Overall Survival

Overall survival will be compared between the two treatment groups, provided the primary endpoint PFS is statistically significant favoring the test treatment arm (fulvestrant + ribociclib). A hierarchical testing procedure will be adopted in this study and the treatment effect on OS will be tested only if the primary efficacy endpoint PFS is statistically significant. A maximum of 3 analyses are planned for OS: at the time of analysis for PFS (provided PFS is significant), at which point a total of 161 deaths are expected, after 263 events have been documented, and a final analysis for OS when 361 deaths are expected (expected 56 months from date of first patient to be randomized).

An α-spending function according to Lan-DeMets (O'Brien-Fleming) , (as implemented in East 6.3) along with the testing strategy outlined below will be used to maintain the overall type I error probability (Lan and DeMets 1983). This guarantees the protection of the 2.5% overall level of significance across the two hypotheses and the repeated testing of the OS hypotheses in the interim and the final analysis (Glimm 2010). The trial allows for the stopping of the study for a superior OS result, provided the primary endpoint PFS has already been shown to be statistically significant favoring the test treatment arm (fulvestrant + ribociclib). Further, the exact nominal p-values that will need to be observed to declare statistical significance at the time of these analyses for OS will depend on the number of OS events that have been observed at the time of these analyses and the α for OS already spent at the time of earlier analyses.

The operating characteristics for OS are shown in Table 3-9 considering the hierarchical testing strategy of PFS and OS. The probabilities shown in Table 3-9 are conditional on PFS being statistically significant.

Statistical properties of the group sequential design are summarized in Table 3-9.

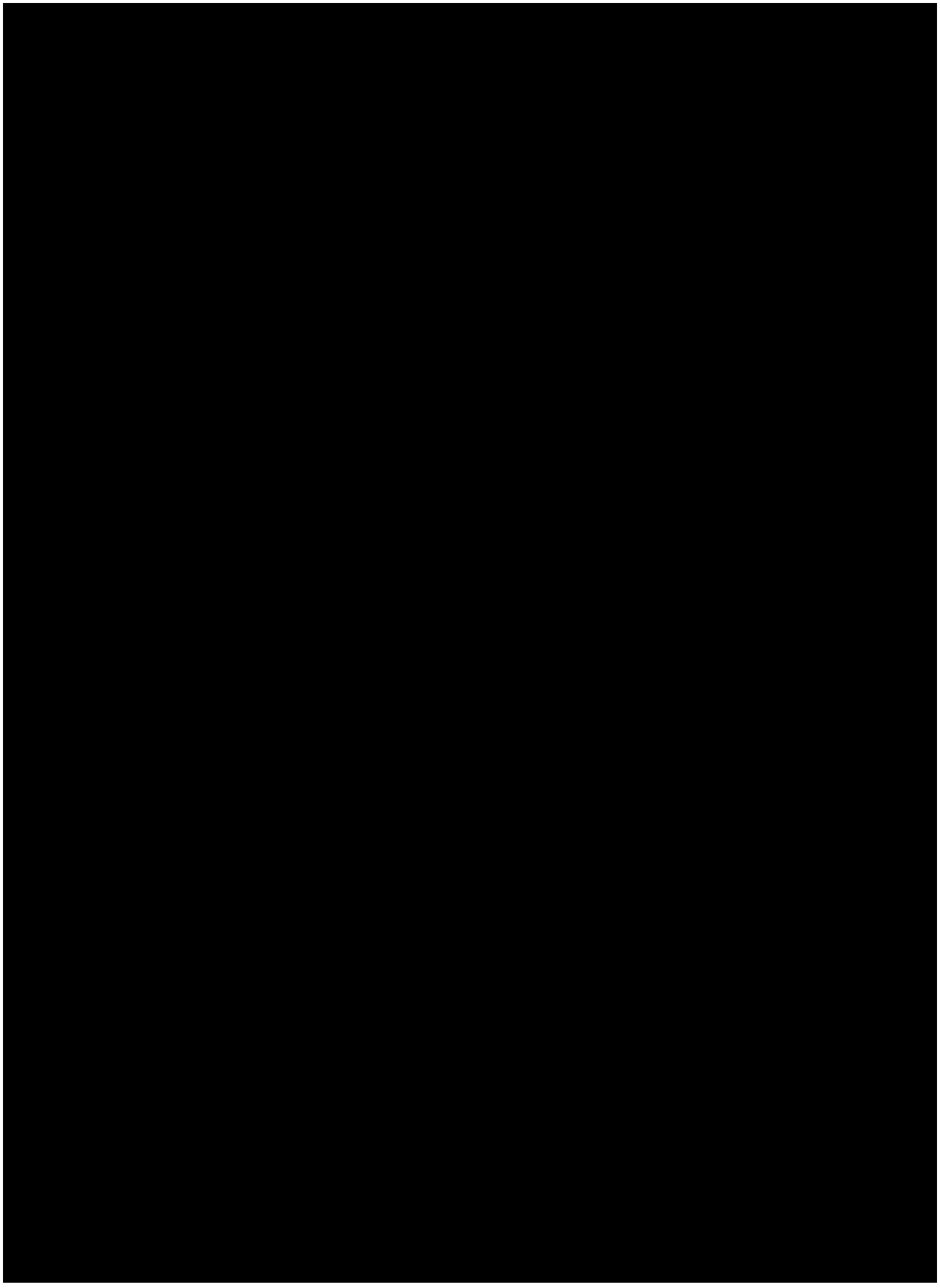
Table 3-9 Projected timelines for interim and final analyses

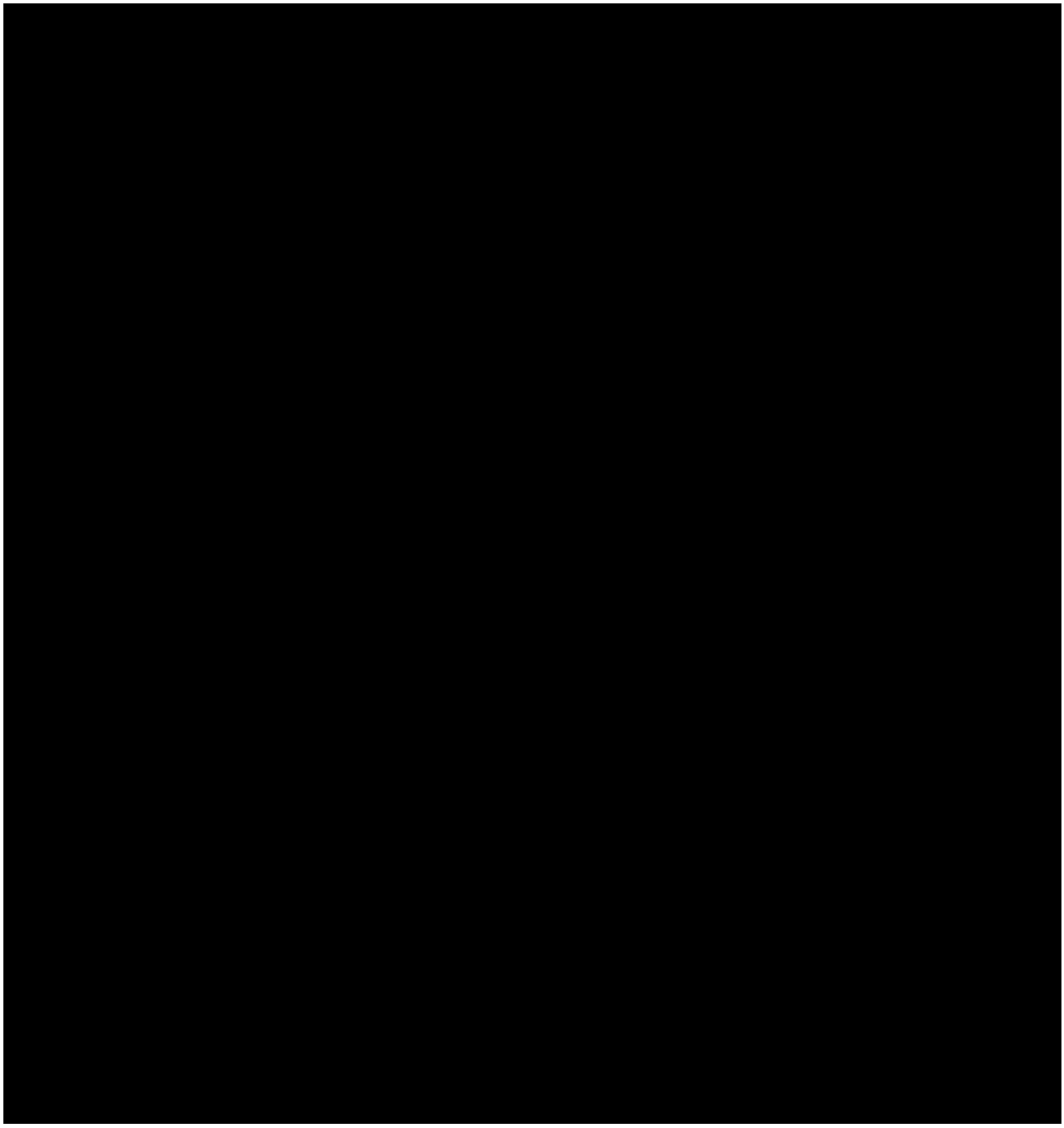
Months after randomization of the first patient	PFS events	Cumulative Power (%) against a hazard ratio of 0.67	OS events	Cumulative Conditional Power (%) against hazard ratio of 0.71
26	364 (100%)	95	161(46%)	14
39	--	--	263(75%)	60
56	--	--	351(100%)	85

Statistical significance of OS will only be declared if significance for primary PFS analysis has been declared.

3.14.3 Confidentiality of Interim OS results

At the time of primary PFS analysis, both PFS and interim OS analysis will be performed by the Sponsor's clinical team. Investigators and patients will remain blinded to study treatment and all patients will continue to be followed for OS until the final OS analysis (or earlier if OS reaches statistical significance at any of the interim analyses).





3.16 Sample size calculation

The median time to progression (TTP) for fulvestrant in first line post-menopausal advanced breast cancer patients is estimated to be between 8 month (Howell 2004) and 23 months in FIRST trial (Robertson et al 2012). For sample size calculation, the median PFS for 1st line patients is assumed to be 18 month. The median PFS for fulvestrant in relapsed advanced breast cancer is estimated to be between 4.8 months (SoFEA trial, Johnston et al 2013) and 6.5 months (CONFIRM trial, Di Leo et al 2010). Since the study population is closer to the population in CONFIRM trial, for sample size calculation, the median PFS for fulvestrant in second line is assumed to be 6.5 months. It is assumed that 40% and 60% of the patients will

be from first line and second line respectively. The median PFS in the control arm (fulvestrant + placebo) is estimated via simulation to be around 9 months. It is hypothesized that the addition of ribociclib will result in a clinically meaningful 33% reduction in the hazard rate of PFS, corresponding to an increase in median PFS to 13.4 months. A maximum of 364 PFS events will be required to detect a hazard ratio of 0.67 with 95% power using the log-rank test at one-sided cumulative 2.5% level of significance. Assuming that enrollment will continue for approximately 19 months at a uniform rate and 10% dropout rate by the time of PFS final analysis, a total of 660 patients should be randomized to the two treatment groups in 2:1 ratio to observe 364 events at approximately 7 months following the randomization of the last patient, i.e., 26 months from the randomization date of the first patient in this study. The sample size calculation was conducted with software package East 6.3. The sample size calculations are based on the estimates from data available from prior studies. A mid-study sample size re-assessment based on blinded pooled data may be performed prior to any efficacy analysis if data from the study indicate substantial deviation from the study assumptions.

The primary objective of the study is to compare investigator assessed PFS between arms in the overall population (both 1st line and 2nd line patients). If the primary analysis is statistically significant, treatment effect will be evaluated in 1st and 2nd line patients (i.e. previous endocrine therapy stratification factor A and B as defined in section 1.1) separately. Based on the estimated median PFS and proportion of patients in 1st line and 2nd line respectively, it is expected that the accumulation of events in 1st line will be slower than 2nd line patients. To ensure enough information coming from 1st line patients, the final analysis will be done after approximately 125 events in the first line patients or approximately 364 events in both arms whichever comes later.

Audit size for BIRC assessed PFS

The audit size of the sample-based BIRC assessment will be 40% of all randomized patients. Based on the audit size calculation approach proposed by Dodd, et. al (2011), assuming investigator and BIRC assessments are similar and the estimated log of investigator-based HR is -0.40 (i.e., HR=0.67), the audit size of 40% will ensure that the upper bound of a one-sided 95% CI for BIRC-based log-hazard ratio has 86% probability of being below 0 (i.e. HR < 1) if the correlation between investigator assessment and BIRC assessment is 0.7 (the estimated correlation based on data from the BELLE-2 [CBKM120F2302] study in metastatic breast cancer).

3.17 Power for analysis of secondary variable

OS will be compared between the two treatment arms, provided that the primary endpoint, PFS, is statistically significant. The power statements for overall survival below are therefore conditioned on significant primary PFS result. Estimate of first line median OS with fulvestrant alone is not published. However the median OS with fulvestrant alone in first line setting is expected to be between 34 months (letrozole alone, Mouridsen 2003) and 38 months (anastrozole alone, FACT trial, Bergh et al 2012) based on results from endocrine based monotherapy studies. For sample size calculation, we assume the median OS for fulvestrant

alone in first line to be 38 month. Median OS for fulvestrant in relapsed post-menopausal advanced breast cancer patients is estimated to be between 19 months (SoFEA trial, Johnston et al 2013) and 26 months (CONFIRM trial, Di Leo et al 2014). For sample size calculation, the median OS for fulvestrant in second line is assumed to be 26 months. Based on the expected split of first line and second line patients mentioned above, the median overall survival of control arm is estimated via simulation to be approximately 30 months. It is hypothesized that adding ribociclib to fulvestrant will result in a 29% reduction in the hazard rate for OS (corresponding to an increase in median survival to 42 months). To detect a hazard ratio of 0.71 with 85% cumulative power, a maximum of 351 deaths need to be observed, (using a log-rank test and a 3-look [superiority only] group sequential design at one-sided cumulative 2.5% level of significance). The sample size calculation was conducted with software package East 6.3.

Based on the number of patients planned to be randomized and 10% dropout rate by the time of OS final analysis, it is estimated that these 351 deaths will be observed at approximately 56 months after the randomization of the first patient. Therefore the estimated time for the final OS evaluation will be 30 months after the primary analysis of the PFS endpoint.

3.18 Sample size considerations for PK analysis

Sparse PK sampling will be performed on approximately 150 patients. Assuming a 2:1 ratio between the ribociclib and placebo arm, sparse plasma concentrations for ribociclib will be available in approximately 100 patients. Trough PK sampling will be performed in all remaining patients. PK data collected from this study may also be combined with data from other studies to support a population PK analysis of ribociclib using non-linear mixed effect modeling; details will be provided in a separate analysis plan. The choice of sampling time points was selected to ensure capture of Cmax. The aim of this analysis is to analyze the pharmacokinetics of ribociclib in a sample size as large as possible. This sample is anticipated to represent the covariate distributions in the targeted population. No specific hypothesis will be tested and therefore no specific sample size calculation has been performed. The decisions on number of sparse PK samples and number of patients needed are driven by feasibility.

4 Details of the statistical analysis

4.1 Baseline comparability

Appropriate descriptive summary statistics of baseline variables (see Section 3.2) will be provided as in-text tables in the core CSR and also in Section 14 in the post-text tables. The summaries will be grouped by treatment arms, but no p-values will be provided.

4.2 Time-to-event analyses

The following sections present a general methodology to be used to analyze the following time-to-event variables:

- Progression-free survival
- Overall survival

- Time to definitive deterioration of the ECOG score by one category of the score from baseline
- Time to response: defined as the time between date of randomization until first documented response (CR or PR) according to RECIST
- Time to definitive deterioration of PRO scores (e.g., global health status/QoL, physical functioning, emotional functioning, social functioning scales of the EORTC QLQ-C30, and the VAS of the EQ-5D-5L)
- Time to onset of safety events
- Duration of safety events
- Time to first chemotherapy or death
- Time to first chemotherapy
- Duration of response
- [redacted]

4.2.1 Analysis of time-to-event data with ties

The STRATA statement in LIFETEST procedure will be used to analyze time to event data with ties. The PHREG procedure in SAS with option TIES=EXACT will be used to fit the Cox proportional hazards model.

4.2.2 Hazard ratio

The hazard ratio as a measure of treatment effect will be derived from the Cox proportional hazards model using SAS procedure PHREG with TIES=EXACT option in the MODEL statement. The stratified unadjusted Cox model will be used (where the baseline hazard function is allowed to vary across strata) for the primary analysis, i.e. the MODEL statement will include only the treatment arm variable as a covariate and the STRATA statement will include stratification variable(s).

Hazard ratio with two-sided 95% confidence interval will be based on Wald test.

4.2.3 Hypothesis and test statistic

The primary efficacy analysis will be the comparison of the distribution of PFS between the two treatment arms using a stratified log-rank test at one-sided 2.5% level of significance, i.e.,

Assuming proportional hazards model for PFS, the following statistical hypotheses will be tested to address the primary efficacy objective:

$$H_{01}: \theta_1 \geq 0 \text{ vs. } H_{a1}: \theta_1 < 0$$

where θ_1 is the log-hazard ratio (fulvestrant+ ribociclib arm vs. fulvestrant + placebo arm) of PFS.

The ***stratified log-rank*** test (strata based on the randomization factor from IRT) will be implemented as follows: For each of the K=4 strata, the LIFETEST procedure will be run with the STRATA statement including only the treatment variable. The TIME statement will include the survival time and a (right) censoring variable. The rank statistic S_k and the corresponding variance $var(S_k)$ (k=1, 2, 3 and 4) will be estimated from this analysis.

The final test statistics will then be reconstructed using the formula: $Z=[S1+S2+S3+S4]/\sqrt{[var(S1)+var(S2)+var(S3)+var(S4)]}$. One-sided p-value will be computed using this Z statistic. Note: Under the null hypothesis, the asymptotic distribution of the test statistic Z is approximately normal (and correspondingly, Z^2 is approximately distributed as chi-square with one degree of freedom).

4.2.4 Kaplan-Meier estimates

An estimate of the survival function in each treatment group will be constructed using Kaplan-Meier (product-limit) method as implemented in PROC LIFETEST with METHOD=KM option. The PROC LIFETEST statement will use the option CONFTYPE=LOGLOG.

Median survival for each treatment group will be obtained along with 95% confidence intervals calculated from PROC LIFETEST output using the method of [Brookmeyer and Crowley 1982]. Kaplan-Meier estimates of the survival function with 95% confidence intervals at specific time points will be summarized. The standard error of the Kaplan-Meier estimate will be calculated using Greenwood's formula [Collett 1994].

4.2.5 Audit-based BIRC assessment of PFS

NCI method

The auxiliary variable estimator of the NCI audit-based method (Dodd et al. 2011) has the form

$$\tilde{\theta}_C = \hat{\theta}_{CA} + \hat{\lambda}(\hat{\theta}_{L\bar{A}} - \hat{\theta}_{LA})$$

where $\hat{\theta}_{CA}$ $\hat{\theta}_{L\bar{A}}$ and $\hat{\theta}_{LA}$ are estimators of the log-hazard ratio based on the central assessment in the audited subset of patients, the local assessment in the nonaudited subset of patients, and the local assessment in the audited subset, respectively. $\hat{\lambda}$ is defined as $\hat{\rho}\sqrt{\delta}(1-\delta)\sqrt{\hat{V}_{CA}/\hat{V}_L}$ and it depends on the variance estimators of $\hat{\theta}_{CA}$ ($\hat{V}_{CA}$)and $\hat{\theta}_L(\hat{V}_L)$ (the estimator of log-HR based on the local assessment in all patients), the proportion of patients in the audited subset(δ), and $\hat{\rho}$, an estimator of the correlation between $\hat{\theta}_{LA}$ and $\hat{\theta}_{CA}$. For the latter, a bootstrap approach will be used:

- Within the audited subset of size m, m patients will be sampled with replacement. Using this sample of m patients, the log-hazard ratio will be estimated based on the local and central assessments separately;
- This procedure will be repeated 10000 times, giving rise to 10000 pairs (local and central) of estimates of the log-HR;
- The sample correlation coefficient between these pairs of estimates will be used for $\hat{\rho}$.

The log-hazard ratio estimates contributing to the auxiliary variable estimate and corresponding variance estimates will be based on stratified Cox proportional hazards models, with stratification based on the randomization stratification factors. The upper bound of a 95% CI for θ_C will be calculated assuming asymptotic normality of $\tilde{\theta}_C$ and using the variance estimator for $\tilde{\theta}_C$ provided in Dodd et al., 2011, $\hat{V}_{CA}\{1-\hat{\rho}^2(1-\delta)\}$.

PhRMA method

The early discrepancy rate (EDR) and late discrepancy rate (LDR) will be calculated using the equations below together with information in Table 4-1.

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

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

Table 4-1 Local versus central disease progression assessments

		Central
Local	PD	No PD
PD	$a = a1 + a2 + a3$	b
No PD	c	d

$a1$: number of agreements on timing and occurrence of PD
$a2$: number of times local PD declared later than central PD
$a3$: number of times local PD declared earlier than central PD

The timing of local and central response assessment will be considered to agree if they occur within ±7 days of each other, aligned with the protocol-specified window for tumor assessments.

4.3 Group sequential design used in Phase III studies

The statistical methodology for the interim analyses of OS will be based on group sequential methodology with efficacy stopping boundaries defined by type I error spending functions.

This approach is flexible in dealing with any deviations from the targeted event totals, or unexpected changes to the plan.

If the exact number of events observed at the interim and final analyses deviates from the target numbers described in the protocol, the actual critical boundaries will be derived using the pre-specified error spending functions and the actual numbers of events observed.

- At interim analyses, information fractions will be computed as the ratio of the number of events observed at the considered interim analysis relative to the number targeted for the final analysis, as described in the sample size section of the protocol.
- At the final analysis, the critical value will be calculated using the exact number of observed events at the final cut-off date, considering the α-levels spent at interim analyses and considering the actual correlation among the test statistics, in order to achieve a cumulative type I error smaller than the desired significance level (i.e. smaller than 2.5% for a one-sided test and smaller than 5% for a two-sided test).

4.3.1 Alpha-spending function

The stopping boundaries to be used for OS test will be calculated using the α-spending function approach described in Lan and DeMets (Lan and DeMets, 1983). The spending function for one-sided test has the following functional form:

$$\alpha(t) = 2 - 2\Phi(Z_{\alpha/2} / \sqrt{t})$$

This function generates stopping boundaries that closely resemble the O'Brien-Fleming boundaries (O'Brien and Fleming, 1979).

4.4 Duration of follow-up

Study follow-up will be summarized using the following methods:

- Summary of duration between randomization and cut-off date, and follow-up times for primary PFS/OS, which are defined as follows:
  - Randomization (recruitment) period = (Date of last patient randomized - Date of first patient randomized + 1) / 30.4375 (months)
  - Duration between randomization and data cut-off date = (Cut-off date – Date of randomization + 1) / 30.4375 (months). This item will be summarized overall.
  - Follow-up time = (Date of event or censoring – Date of randomization + 1) / 30.4375 (months) regardless of censoring. Date of censoring is defined as the last adequate tumor assessment date for PFS or last contact date for OS. This item will be summarized by treatment arm.

All summaries will be reported in months. Date of censoring is the same as defined for the PFS and OS analysis.

In addition, the time from PFS/OS censoring date to data cut-off date will be summarized by time intervals in months: <3, 3 - < 6, 6 - < 12, 12 - < 18, 18 - <24 and by 12 month intervals thereafter if necessary. The gap time is calculated as ([analysis cut-off date] - [censoring date] + 1)/30.4375

4.5 Confidence intervals for response rate and clinical benefit rate

Response rate and clinical benefit rate will be summarized as percentages with 95% confidence intervals. A standard Wald asymptotic confidence interval, i.e., normal approximation, (implemented using SAS procedure FREQ for one-way tables) will be calculated.

5 References

Amit O, Mannino F, Stone AM, Bushnell W, Denne J, Helterbrand J and Burger HU. Blinded independent central review of progression in cancer clinical trials: results from a meta-analysis. European Journal of Cancer 2011; 47: 1772-1778.

Bergh J, Jönsson PE, Lidbrink EK, Trudeau M, Eiermann W, Brattström D, Lindemann JP, Wiklund F, Henriksson R. FACT: an open-label randomized phase III study of fulvestrant and anastrozole in combination compared with anastrozole alone as first-line therapy for patients with receptor-positive postmenopausal breast cancer. J Clin Oncol. 2012 Jun 1;30(16):1919-25. doi: 10.1200/JCO.2011.38.1095. Epub 2012 Feb 27.

Bretz F, Posch M, Glimm E, et al (2011). Graphical approaches for multiple comparison procedures using weighted Bonferroni, Simes or parametric tests. Biometrical Journal; 53, 894-913.

Bretz, F., Maurer, W., and Hommel, G. (2009) Test and power considerations for multiple endpoint analyses using sequentially rejective graphical procedures. Statistics in Medicine 30, 1489-1501.

Brookmeyer R and Crowley J (1982). A Confidence Interval for the Median Survival Time. Biometrics, 38, 29 - 41.

Buzdar AU, Jonat W, Howell A, et al: Anastrozole versus megestrol acetate in the treatment of postmenopausal women with advanced breast carcinoma: Results of a survival update based on a combined analysis of data from two mature phase III trials. Arimidex Study Group. Cancer 83:1142-1152, 1998

Cleeland CS (2009) The Brief Pain Inventory. User Guide. Houston, Texas.

Clopper, C.; Pearson, E.S. (1934). "The use of confidence or fiducial limits illustrated in the case of the binomial". Biometrika 26: 404–413

Collet D (1994). Modelling survival data in medical research. London, Chapman & Hall.

Di Leo A, Malorni L. Polyendocrine treatment in estrogen receptor-positive breast cancer: a "FACT" yet to be proven. J Clin Oncol. 2012 Jun 1;30(16):1897-900.

Dodd E L, Korn L E, Freidlin B, Gray R, Bhattacharya S. An audit strategy for

progression-free survival. Biometrics, Vol 67, Issue 3; 1092-99, 2011.

Dombernowsky P, Smith I, Falkson G, Leonard R, Panasci L, Bellmunt J, Bezwoda W, Gardin G, Gudgeon A, Morgan M, Fornasiero A, Hoffmann W, Michel J, Hatschek T, Tjabbes T, Chaudri HA, Hornberger U, Trunet PF. Letrozole, a new oral aromatase inhibitor for advanced breast cancer: double-blind randomized trial showing a dose effect and improved efficacy and tolerability compared with megestrol acetate. J Clin Oncol. 1998 Feb;16(2):453-61.

Dmitrienko A., Molenberghs G., Chuang-Stein Ch., Offen W. (2005). Analysis of clinical trials using SAS. SAS Institute Inc., SAS press

Dowsett M, Pfister C, Johnston RD, et al. Impact of tamoxifen on the pharmacokinetics and endocrine effects of the aromatase inhibitor letrozole in postmenopausal women with breast cancer. Clin Can Res 1999;5:2338-2343.

Dowsett M, Tobias JS, Howell A et al. (1999) 79(2):311-315. The effect of anastrozole on the pharmacokinetics of tamoxifen in post-menopausal women with early breast cancer. Br J Cancer. 1999;79(2):311-315

Eisenhauer E, et al (2009) New response evaluation criteria in solid tumors: revised RECIST guideline (version 1.1). European Journal of Cancer, Vol.45: 228-47

Fayers PM, Aaronson NK, Bjordal K, et al, on behalf of the EORTC Quality of Life Group (2001) The EORTC QLQ-C30 Scoring Manual (3rd Edition). Published by: European Organisation for Research and Treatment of Cancer, Brussels 2001

FDA briefing Document Oncologic Drugs Advisory Committee Meeting, July 24, 2012, Evaluation of radiologic review of progression-free survival in Non-hematologic malignancies.
http://www.fda.gov/downloads/advisorycommittees/committeesmeetingmaterials/drugs/oncologicdrugsadvisorycommittee/ucm311141.pdf

Finn, Rs ; Crown, JP ; Lang, I ; Boer, K ; Bondarenko, Im ; Kulyk, So ; Ettl, J ; Patel, R ; Pinter, T ; Schmidt, M ; Shparyk, Y ; Thummala, Ar ; Voytko, Nl ; Fowst, C ; Huang, X ; Kim, St ; Randolph, S ; Slamon, Dj. The cyclin-dependent kinase 4/6 inhibitor palbociclib in combination with letrozole versus letrozole alone as first-line treatment of oestrogen receptor-positive, HER2-negative, advanced breast cancer (PALOMA-1/TRIO-18): a randomised phase 2 study. Lancet Oncology, 2015 Jan, Vol.16(1), pp.25-35

Fleming TR, Harrington DP, O'Brien PC (1984). Designs for group sequential tests. Controlled Clinical Trials 5: 348-361.

Fleming TR, Rothmann MD, Lu HL (2009). Issues in Using Progression-Free Survival When Evaluating Oncology Products. Journal of Clinical Oncology 27, 2874-2880

Glimm, E., Maurer, W., and Bretz, F. (2010) Hierarchical testing of multiple endpoints in group-sequential trials. Statistics in Medicine 29, 219-228

Hernan M., Brumback B., Robins J. Marginal structural models to estimate the causal effect of zidovudine on the survival of HIV-positive men. *Epidemiology,* 11(5), 561-570 (2000).

Howell A, Robertson JF, Abram P, Lichinitser MR, Elledge R, Bajetta E, Watanabe T, Morris C, Webster A, Dimery I, Osborne CK. Comparison of fulvestrant versus tamoxifen for the treatment of advanced breast cancer in postmenopausal women previously untreated with endocrine therapy: a multinational, double-blind, randomized trial. J Clin Oncol. 2004 May 1;22(9):1605-13.

Hung HM, Wang SJ and O'Neil R (2007) Statistical considerations for testing multiple endpoints in group sequential or adaptive clinical trials. Journal of Biopharmaceutical Statistics 17: 1201-1210.

Hwang, I.K. Shih, W.J. and DeCani, J.S. (1990). Group sequential designs using a family of type I error probability spending functions. Statistics in Medicine 9, 1439-1445

Johnston S., Kilburn L.S., Ellis P. et al (2012) Fulvestrant alone or with concomitant anastrozole vs Exemestane following progression on non-steroidal aromatase inhibitor: first results of the SoFEa trial. EJC Volume 47, Supplement 2.

Johnston SR, Kilburn LS, Ellis P, Dodwell D, Cameron D, Hayward L, Im YH, Braybrooke JP, Brunt AM, Cheung KL, Jyothirmayi R, Robinson A, Wardley AM, Wheatley D, Howell A, Coombes G, Sergenson N, Sin HJ, Folkerd E, Dowsett M, Bliss JM; SoFEA investigators. Fulvestrant plus anastrozole or placebo versus exemestane alone after progression on non-steroidal aromatase inhibitors in postmenopausal patients with hormone-receptor-positive locally advanced or metastatic breast cancer (SoFEA): a composite, multicentre, phase 3 randomised trial. Lancet Oncol. 2013 Sep;14(10):989-98.

Jonat W, Kaufmann M, Blamey RW, Howell A, Collins JP, Coates A, Eiermann W, Janicke F, Njordenskold B, Forbes JF, Kolvenbag GJCM (1995), A randomized study to compare the effect of the Luteinising Hormone Releasing Hormone (LHRH) analogue Goserelin with or without Tamoxifen in pre- and perimenopausal patients with advanced breast cancer. European Journal of Cancer, 31A(2), 137-142

Klijn JG, Blamey RW, Boccardo F, Tominaga T, Duchateau L and Sylvester R (2001), Combined Tamoxifen and Luteinizaing Hormone-Releasing Hormone (LHRH) agonist versus LHRH agonist alone in premenopausal advanced breast cancer: a meta-analysis of four randomized trials. Journal of Clinical Oncology, 19: 343-353

Lan, K. K. G. & Demets, D. L. (1983). Discrete sequential boundaries for clinical trials. Biometrika 70, 659-63

Lien EA1, Anker G, Lønning PE, Solheim E, Ueland PM. Decreased serum concentrations of tamoxifen and its metabolites induced by aminoglutethimide. Cancer Res. 1990 Sep 15;50(18):5851-7

Mouridsen H, Gershanovich M, Sun Y, et al (2003) Phase III study of letrozole versus tamoxifen as first-line therapy of advanced breast cancer in postmenopausal women: analysis of survival and update of efficacy from the International Letrozole Breast Cancer Group. J Clin Oncol; 21: 2101-09

Mouridsen H, Gershanovich M, Sun Y, et al (2001) Superior efficacy of letrozole versus tamoxifen as first-line therapy for postmenopausal women with advanced breast cancer: results of a phase III study of the International Letrozole Breast Cancer Group. J Clin Oncol; 19: 2596-606

O'Brien PC, Fleming TR (1979). A multiple testing procedure for clinical trials, Biometrics. 35(3):549-56.

Oemar M, Jannsen B. (2013) EQ-5D-5L User Guide: Basic information on how to use the EQ-5D-5L instrument. Version 2.0, October 2013. © EuroQol Group 2013.

Paridaens RJ, Dirix LY, Beex LV, Nooij M, Cameron DA, Cufer T, Piccart MJ, Bogaerts J, Therasse P. Phase III study comparing exemestane with tamoxifen as first-line hormonal treatment of metastatic breast cancer in postmenopausal women: the European Organisation for Research and Treatment of Cancer Breast Cancer Cooperative Group. J Clin Oncol. 2008 Oct 20;26(30):4883-90. doi: 10.1200/JCO.2007.14.4659. Epub 2008 Sep 15.

Rabin R, Oemar M, Oppe M (2011) EQ-5D-5L User Guide: Basic information on how to use the EQ-5D-5L instrument. Version 1.0, April 2011. © EuroQol Group 2011

Reilly MC, Zbrozek AS, Dukes EM. The validity and reproducibility of a work productivity and activity impairment instrument. Pharmacoeconomics. 1993 Nov;4(5):353-65

Robertson JFR, Lindemann JPO, et al. Fulvestrant 500 mg versus anastrozole 1 mg for the first-line treatment of advanced breast cancer: follow-up analysis from the randomized 'FIRST' study. Breast Cancer Res Treat (2012) 136:503–511.

Robertson, Jfr ; Lindemann, J ; Garnett, S ; Anderson, E ; Nicholson, Ri ; Kuter, I ; Gee, Jmw. A Good Drug Made Better: The Fulvestrant Dose-Response Story. Clinical Breast Cancer, 2014 Dec, Vol.14(6), pp.381-389

Schuirmann DJ (1987). A comparison of the two one-sided tests procedure and the power approach for assessing the equivalence of average bioavailability. J Pharmacokinet Biopharm. 15(6):657-80.

Tang Z (2015). Optimal futility interim design: a predictive probability of success approach with time to event end point. Journal of Biopharmaceutical Statistics, 25(6), 1312-1319.
Zhang JJ, Chen H, He K, Tang S, Justice R, Keegan P, Pazdur R and Sridhara R. Evaluation of blinded independent central review of tumor progression in oncology clinical trials: a meta-analysis. Therapeutic innovation & regulatory science: 2012, 47(2): 167-174.

Appendix 1 CTC grades for laboratory values in Novartis Oncology

CTC grades for laboratory values in Novartis Oncology (based on CTCAE v4.03 – June 2010) **Page 1**

				CTC Grades[1]				
Lab test (toxicity)	**SI unit**	**Lab test (NCDS)**	**Normal ranges (Merck manual, July 2015) and *conversion factors***	**0**	**1**	**2**	**3**	**4**
Hematology								
WBC ↓	10^9/L	WBC	**3.9 – 10.7 x 10^9/L**	≥ LLN	< LLN - 3.0 x 10^9/L	< 3.0 – 2.0 x 10^9/L	< 2.0 – 1.0 x 10^9/L	< 1.0 x 10^9/L
WBC [2] (Leukocytosis)	10^9/L	WBC			-	-	> 100 x 10^9/L	-
Hemoglobin [2] (Anemia)	g/L	HGB	**120 - 160 g/L** or 7.4 - 9.9 mmol/L (F) **140 - 170 g/L** or 8.7 – 10.6 mmol/L (M)	≥ LLN	< LLN - 100 g/L < LLN - 6.2 mmol/L	< 100 - 80 g/L < 6.2 - 4.9 mmol/L	< 80 g/L < 4.9 mmol/L	-
Hemoglobin ↑	g/L	HGB	*(16.113 x mmol/L = g/L)*		Increase >0-20 g/L above ULN	Increase >20-40 g/L above ULN	Increase >40 g/L above ULN	-
Platelets ↓	10^9/L	PLAT	**150 - 350 x 10^9/L**	≥ LLN	< LLN - 75.0 x 10^9/L	< 75.0 - 50.0 x 10^9 L	< 50.0 - 25.0 x 10^9/L	< 25.0 x 10^9/L
Neutrophils [3]↓	10^9/L	NEUT		≥2x10^9/L	< 2.0 - 1.5 x 10^9/L	< 1.5 - 1.0 x 10^9/L	< 1.0 - 0.5 x 10^9/L	< 0.5 x 10^9/L
Lymphocytes [3]↓	10^9/L	LYM		≥1.5x10^9/L	< 1.5 - 0.8 x 10^9/L	< 0.8 - 0.5 x 10^9/L	< 0.5 - 0.2 x 10^9/L	< 0.2 x 10^9/L
Lymphocytes ↑	10^9/L	LYM			-	> 4 - 20 x 10^9/L	> 20 x 10^9/L	-
Biochemistry								
AST ↑	U/L	AST	**0 - 35 U/L** or 0 – 0.58 ukat/L *(60 x ukat/L = U/L)*	≤ ULN	> ULN – 3.0 x ULN	> 3.0 - 5.0 x ULN	> 5.0 - 20.0 x ULN	> 20.0 x ULN
ALT ↑	U/L	ALT	**0 - 35 U/L** or 0 – 0.58 ukat/L *(60 x ukat/L = U/L)*	≤ ULN	> ULN – 3.0 x ULN	> 3.0 - 5.0 x ULN	> 5.0 - 20.0 x ULN	> 20.0 x ULN
Total bilirubin ↑	umol/L	BILI	**5.1 – 20.5 umol/L** or 0.3 – 1.2 mg/dL *(17.1 x mg/dL = umol/L)*	≤ ULN	> ULN - 1.5 x ULN	> 1.5 - 3.0 x ULN	> 3.0 - 10.0 x ULN	> 10.0 x ULN
Alk. Phosphatase ↑	U/L	ALP	**36 - 92 U/L** or 0.5 - 1.5 ukat/L *(60 x ukat/L = U/L)*	≤ ULN	> ULN - 2.5 x ULN	> 2.5 - 5.0 x ULN	> 5.0 - 20.0 x ULN	> 20.0 x ULN
Creatinine [4]↑	umol/L	CREAT	**61.9 - 115 umol/L** or 0.7 – 1.3 mg/dL *(88.4 x mg/dL = umol/L)*	≤ ULN	> ULN - 1.5 x ULN	> 1.5 - 3.0 x ULN	> 3.0 - 6.0 x ULN	> 6.0 x ULN
Creatinine kinase [4]↑	U/L	CK	**30 - 170 U/L** or 0.5 – 2.83 ukat/L *(60 x ukat/L = U/L)*	≤ ULN	> ULN - 2.5 x ULN	> 2.5 - 5.0 x ULN	> 5.0 - 10.0 x ULN	> 10.0 x ULN
Albumin [2] (Hypoalbuminemia)	g/L	ALB	**35 - 55 g/L** or 3.5 to 5.5 g/dL	≥ LLN	< LLN - 30 g/L	< 30 - 20 g/L	< 20 g/L	-
Total Cholesterol ↑	mmol/L	CHOL	**3.88 – 5.15 mmol/L** or 150 - 199 mg/dL *(38.67 x mg/dL = mmol/L)*	≤ ULN	> ULN - 7.75 mmol/L > ULN - 300 mg/dL	> 7.75 -10.34 mmol/L > 300 – 400 mg/dL	>10.34-12.92 mmol/L > 400 – 500 mg/dL	>12.92 mmol/L > 500 mg/dL
Lipase ↑	U/L	LIPASE	**<95 U/L** or <1.58 ukat/L *(60 x ukat/L = U/L)*	≤ ULN	> ULN - 1.5 x ULN	> 1.5 - 2.0 x ULN	> 2.0 - 5.0 x ULN	> 5.0 x ULN
Amylase ↑	U/L	AMYLASE	**0 - 130 U/L** or 0 – 2.17 ukat/L *(60 x ukat/L = U/L)*	≤ ULN	> ULN - 1.5 x ULN	> 1.5 - 2.0 x ULN	> 2.0 - 5.0 x ULN	> 5.0 x ULN
Uric acid [2] (Hyperuricemia)	umol/L	URATE	**150 - 470 umol/L** or 2.5 – 8 mg/dL *(59.48 x mg/dL = umol/L)*	≤ ULN	> ULN – 10 mg/dL > ULN – 595 umol/L	-	-	> 10 mg/dL > 595 umol/L

ULN = Upper Limit of Normal range; LLN = Lower Limit of Normal range

LAB - CTC grades in Novartis Oncology (26Oct15)

CTC grades for laboratory values in Novartis Oncology (based on CTCAE v4.03 – June 2010)

Page 2

				CTC Grades[1]				
Lab test (toxicity)	**SI unit**	**Lab test (NCDS)**	**Normal ranges (Merck manual, July 2015) and *conversion factors***	**0**	**1**	**2**	**3**	**4**
Phosphorus[2] (Hypophosphatemia)	mmol/L	PHOS	**0.97 – 1.45 mmol/L** or 3.0 - 4.5 mg/dL *(0.32 x mg/dL = mmol/L)*	≥ LLN	< LLN - 2.5 mg/dL < LLN - 0.8 mmol/L	< 2.5 - 2.0 mg/dL < 0.8 - 0.6 mmol/L	< 2.0 - 1.0 mg/dL < 0.6 - 0.3 mmol/L	< 1.0 mg/dL < 0.3 mmol/L
Calcium (corrected) [2] (Hypercalcemia)	mmol/L	CACALC	**2.2 - 2.6 mmol/L** or 9 - 10.5 mg/dL *(0.2495 x mg/dL = mmol/L)*	≤ ULN	> ULN - 11.5 mg/dL > ULN - 2.9 mmol/L	> 11.5 - 12.5 mg/dL > 2.9 - 3.1 mmol/L	> 12.5 - 13.5 mg/dL > 3.1 - 3.4 mmol/L	> 13.5 mg/dL > 3.4 mmol/L
Calcium (corrected) [2] (Hypocalcemia)	mmol/L	CACALC		≥ LLN	< LLN - 8.0 mg/dL < LLN - 2.0 mmol/L	< 8.0 - 7.0 mg/dL < 2.0 - 1.75 mmol/L	< 7.0 - 6.0 mg/dL < 1.75 - 1.5 mmol/L	< 6.0 mg/dL < 1.5 mmol/L
Magnesium [2] (Hypermagnesemia)	mmol/L	MG	**0.62 – 0.99 mmol/L** or 1.5 – 2.4 mg/dL *(0.4114 x mg/dL = mmol/L)*	≤ ULN	> ULN - 3.0 mg/dL > ULN - 1.23 mmol/L	-	> 3.0 – 8.0 mg/dL > 1.23 – 3.3 mmol/L	> 8.0 mg/dL > 3.3 mmol/L
Magnesium [2] (Hypomagnesemia)	mmol/L	MG		≥ LLN	< LLN - 1.2 mg/dL < LLN - 0.5 mmol/L	< 1.2 - 0.9 mg/dL < 0.5 - 0.4 mmol/L	< 0.9 - 0.7 mg/dL < 0.4 - 0.3 mmol/L	< 0.7 mg/dL < 0.3 mmol/L
Glucose (non-fasting) [2] (Hyperglycemia)	mmol/L	GLUCSN	**<7.8 mmol/L** or <140 mg/dL *(0.05551 x mg/dL = mmol/L)*	≤ ULN	-	> ULN - 250 mg/dL > ULN - 13.9 mmol/L	> 250 - 500 mg/dL > 13.9 - 27.8 mmol/L	> 500 mg/dL > 27.8 mmol/L
Glucose (fasting) [2] (Hyperglycemia)	mmol/L	GLUCSF	**3.9 – 5.8 mmol/L** or 70 - 105 mg/dL *(0.05551 x mg/dL = mmol/L)*	≤ ULN	> ULN - 160 mg/dL > ULN - 8.9 mmol/L	> 160 - 250 mg/dL > 8.9 - 13.9 mmol/L	> 250 - 500 mg/dL > 13.9 - 27.8 mmol/L	> 500 mg/dL > 27.8 mmol/L
Glucose [2] (Hypoglycemia)	mmol/L	GLUCSN/ GLUCSF		≥ LLN	< LLN - 55 mg/dL < LLN - 3.0 mmol/L	< 55 - 40 mg/dL < 3.0 - 2.2 mmol/L	< 40 - 30 mg/dL < 2.2 - 1.7 mmol/L	< 30 mg/dL < 1.7 mmol/L
Potassium [2] (Hyperkalemia)	mmol/L	K	**3.5 - 5.0 mmol/L** *(0.2558 x mg/dL = mEq/L = mmol/L)*	≤ ULN	> ULN - 5.5 mmol/L	> 5.5 - 6.0 mmol/L	> 6.0 - 7.0 mmol/L	> 7.0 mmol/L
Potassium [2] (Hypokalemia)	mmol/L	K		≥ LLN	< LLN - 3.0 mmol/L	-	< 3.0 - 2.5 mmol/L	< 2.5 mmol/L
Sodium [2] (Hypernatremia)	mmol/L	SODIUM	**136 - 145 mmol/L** *(0.435 x mg/dL = mEq/L = mmol/L)*	≤ ULN	> ULN - 150 mmol/L	> 150 - 155 mmol/L	> 155 - 160 mmol/L	> 160 mmol/L
Sodium [2] (Hyponatremia)	mmol/L	SODIUM		≥ LLN	< LLN - 130 mmol/L	-	< 130 - 120 mmol/L	< 120 mmol/L
Triglyceride [2] ↑	mmol/L	TRIG	**< 2.82 mmol/L** or < 250 mg/dL *(0.01129 x mg/dL = umol/L)*	< 150 < 1.71	≥ 150 - 300 mg/dL ≥ 1.71 – 3.42 mmol/L	> 300 - 500 mg/dL > 3.42 – 5.7 mmol/L	> 500 - 1000 mg/dL > 5.7 – 11.4 mmol/L	> 1000 mg/dL > 11.4 mmol/L
Coagulation								
INR [2]↑	1	INR	**0.8 – 1.2**	≤ ULN	> ULN - 1.5 x ULN	> 1.5 - 2.5 x ULN	> 2.5 x ULN	-
Activated partial thromboplastin time[2]↑	sec	APTT	**25 - 35 sec**	≤ ULN	> ULN - 1.5 x ULN	> 1.5 - 2.5 x ULN	> 2.5 x ULN	-
Fibrinogen [4]↓	g/L	FIBRINO	**1.5 – 3.5 g/L** or 150 – 350 mg/dL *(0.01 x mg/dL = g/L)*	≥ LLN	< LLN - 0.75 x LLN	< 0.75 - 0.5 x LLN	< 0.5 - 0.25 x LLN	< 0.25 x LLN

ULN = Upper Limit of Normal range; LLN = Lower Limit of Normal range

(1) = **LAB CTC grades 1, 2, 3, 4 overrule the study specific (central or local) normal range criteria**, e.g. if ULN of Sodium is 151 mmol/L and the value is 151 mmol/L, CTC grade 2 is assigned although the value is ≤ ULN.

(2) = **Life-threatening consequences and/or hospitalization are <u>not</u> considered** for determination of LAB CTC grades 3 and 4. Concomitant usage of anticoagulation therapy (for INR and Fibrinogen) is not considered either.

(3) = Values and LNRs for blood differentials can be given as %, absolute values should then be calculated using WBC. Generally, **≥ 1.5 x 10^9 /L (lymphocytes) and ≥ 2 x 10^9 /L (neutrophils) are considered as LAB CTC grade 0**

(4) = For Creatinine and Fibrinogen, the **comparison with baseline is <u>not</u> considered** for derivation of LAB CTC grades