NOVARTIS

Clinical Development

RAD001 (everolimus)

CRAD001Y2201 / NCT01783444

A three-arm, randomized, open label, phase II study of everolimus in combination with exemestane versus everolimus alone versus capecitabine in the treatment of postmenopausal women with estrogen receptor positive, locally advanced, recurrent, or metastatic breast cancer after recurrence or progression on prior letrozole or anastrozole

RAP Module 3 – Detailed Statistical Methodology – Addendum1

Author: [redacted] Trial Statistician

Document type: RAP Documentation

Document status: Final

Release date: 27-Oct-2017

Table of contents

Document History – Changes compared to previous version of RAP module 3.

Version	Date	Changes
Amendment 3	19Jul2017	• Section 3.8 updated: due to "Lack of PI oversight and no eCRF signatures at Database Lock (DBL) due to PI departure and no replacement", protocol deviation C17 was created and details were added on how to analyze efficacy patients with this protocol deviation C17 • Section 3.2.3: typo corrected '30.3475' should be '30.4375'
Amendment 2	13Jun2017	• Section 2.1.6: specification given in study day calculation for assessments before start of study treatment • Section 2.2: following protocol amendment, in sentence "The final PFS analysis is planned to be conducted when at least 150 PFS events per local tumor assessment have been documented...", 'at least' was replaced by 'approximately' • Section 3.6.4, specifications added on how to count dose reductions for Capecitabine • Section 3.8.2 the final OS analysis will be conducted at the same time as the final PFS analysis using the same cut-off date. • Sections 3.9.1.4 and 4.2: analysis on time to first occurrence of stomatitis removed • Section 3.9.1.6 added for the new process related to clinical trial safety disclosure. • Section 3.12; In the sentence "The primary objective of this study is to estimate the hazard ratio of PFS comparing everolimus + exemestane versus everolimus alone with 150 PFS events", 'approximately' was added. In addition, table 3-3 is updated to include approximate CI bounds for 146 observed PFS events.
Amendment 1	16Dec2016	• Section 2.2, reworded to include the conduction of interim analysis of PFS and the planned timeline for OS analysis to be consistent with protocol; • Section 2.3, definition of Safety set modified as per guidance • Section **Error! Reference source not found.** on major protocol deviations was removed as there is no per protocol set • Section 2.4.1, added information of CYP2C9 • Deleted Section 2.6.6 for waterfall plot • Section 3.6.1, clarification on capecitabine added • Section 3.6.3, clarification on capecitabine added • Section 3.6.4, clarification on counting interruption when last record is 0 mg and also if first dose is lower than the studied dose under the protocol • Section 3.8, added non-CR/non-PD in the definition of CBR • Section 3.8.1, deleted sensitivity analysis of PFS based on stratum information obtained from clinical database; deleted analyses for missing tumor assessments and subgroup analysis for PFS. Also deleted reference to log-rank test.

Version	Date	Changes
		• Section 3.8.2, included day 1 for baseline ECOG and QLQ; updated categories for questionnaire completion. Also definition of CBR is clarified. For QoL questionnaire, clarification made that baseline record used for time to deterioration should be a record when patient can still deteriorate. For overall survival, clarification on the timing of each analysis. • Section 3.9.1.2 and 3.9.2, updated to CTCAE v4.03 • Section 3.9.1.4, deleted analysis for time to first occurrence of G3/4 neutropenia and G3/4 thrombocytopenia, also deleted adverse event adusted for patient year exposure. • Section **Error! Reference source not found.**, CNAE section replaced by AESI section • Section 3.9.2, specific rules for reporting lab data was added • Section 3.9.3, respiration rate removed as not in CRF • [redacted] • Deleted Section 3.11 for subgroup analyses • Added new Section 3.11 for interim analysis • Section 4.2, Time to first occurrence of stomatitis and time to definitive deterioration from baseline in the global health status / QoL score of the EORTC QLQ-C30 questionnaire added • Section 4.2.3, reference to log-rank test removed • Section 4.2.4 removed • Section 5, added more reference • Typo corrections
Addendum 1	27Oct2017 (after DBL)	Duration of exposure for Capecitabine arm will be corrected in order to take the rest periods into account. Dose intensity as well as the planned dose intensity will be modified accordingly (section 3.6.1 and 3.6.3) [redacted] For efficacy endpoints : - 95% confidence interval for PFS and OS estimates will be provided to facilitate the comparison with Bolero 2 data and to show the potential increase in variability due to the low sample size. 90% CI were pre-planned to be aligned with the sample size calculation based on the precision of the estimate (width of the 90% CI of the HR) (section 3.8.1 and 3.8.2) - [redacted] - To get a more appropriate cox model adjusted on the known prognostic factors from the literature as well as to account

Version	Date	Changes
		for baseline imbalances between arms in some baseline characteristics, stratified adjusted cox model for both endpoints PFS and OS will be revised to include the following covariates : study treatment (eve+exe vs eve alone; eve+exe vs. cap.), performance status (0 vs. 1 or 2), presence of bone lesions only at baseline (yes vs. no), prior chemotherapy use (yes vs. no), number of organs involved (1 vs. 2 vs. >=3), age (<65 years vs. >=65 years old) and Race (Caucasian vs. non-caucasian), (section 3.8.1 and 3.8.2) - To further explore the longer PFS in Capecitabine arm versus the combination of everolimus and exemestane and given the observed imbalance in informative censoring between the two treatment arms, patient disposition (i.e. reasons for end of treatment) for patients censored due to "New cancer therapy started" reason in the main primary analysis will be presented (section 3.5). [redacted] Prior medications and further antineoplastic therapies after treatment discontinuation - To better describe the population at study entry, prior metastatic chemotherapy will be summarized using the regimen number to account for any possible combination of treatments. - To better understand the potential impact of the further antineoplastic therapy after treatment discontinuation on the overall survival (i.e. median OS estimate in the study different from the one observed in B2 study), the first therapy given after treatment discontinuation will be described taking into account the combination of treatments if any.

1 Introduction

This document describes the detailed statistical methodology of the Report Analysis Plan (RAP) of the study CRAD001Y2201: A three-arm, randomized, open-label, multi-center, international phase II study evaluating the combination treatment of everolimus (10 mg daily) with exemestane (25 mg daily) versus everolimus (10 mg daily) versus capecitabine (1250 mg/m^2 twice daily for 14 days in 3-week cycles) in patients with estrogen-receptor positive, HER2 negative, advanced breast cancer after recurrence or progression on letrozole or anastrozole. The data will be analyzed by Novartis. It is planned that data from all centers that participate in this study will be used.

2 Definitions and general methodology

2.1 Definitions

2.1.1 Study drug and study treatment

Study drug is defined as everolimus, exemestane or capecitabine. All study drugs are open-label.

Study treatment is defined as everolimus + exemestane, everolimus monotherapy or capecitabine monotherapy.

2.1.2 Date of first administration of study drug

The date of first administration of study drug is derived as the first date when a non-zero dose of study drug is administered and recorded on the dose administration record (DAR) eCRF. For the sake of simplicity, the date of first administration of study drug will also be referred to as the start of study drug.

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 non-zero dose of study drug is administered and recorded on the DAR eCRF.

2.1.4 Date of first administration of study treatment

The date of first administration of study treatment is derived as the first date when a non-zero dose of any component of study treatment is administered and recorded on the DAR eCRF. For the sake of simplicity, 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 non-zero dose of any component of study treatment was administered and recorded on the DAR eCRF.

2.1.6 Study day

The study day for ***safety assessments*** (e.g., adverse event onset, laboratory abnormality occurrence, vital sign measurement, dose interruption, etc.) will be calculated as the difference between the date of the assessment and the start of study treatment plus 1. *(Note: except in the case when the assessment is before start of study treatment in which case study day is calculated as the difference between the date of the assessment and the start of study treatment. In this particular case, the study day displayed on the listing will be negative.)*

The study day for ***all other, i.e., non-safety assessments*** (tumor assessment, death, disease progression, tumor response, ECOG performance status, QoL assessment) will be calculated as the difference between the date of the event and the randomization date plus 1. 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 02JAN2007, start of study drug is on 05JAN2007, and the date of death is 09JAN2007, then the study day when death occurred is 8.).

The study day will be displayed in data listings.

2.1.7 Baseline

Baseline value(s) is (are) the result of an investigation describing the "true" uninfluenced state of the subject.

For ***efficacy evaluations***, the last available assessment before or at the date of randomization is taken as 'baseline' value or 'baseline' assessment. In the context of the definition of baseline, efficacy evaluations also include the ECOG performance status, patient-reported outcome measures and clinical measurements included in the stratification.

For ***safety evaluations*** (i.e., laboratory and vital signs), the last available assessment before or at the date of the start of study treatment is taken as 'baseline' assessment.

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

2.1.8 On-treatment assessment/event

Safety summaries and selected summaries of deaths will summarize only on-treatment assessments/events. On-treatment assessment/event is defined as any assessment/event obtained in the 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: The calculation of study treatment duration, however, may use different rules as specified in Section 3.6.1.)

2.1.9 Last contact date

The last contact date will be derived for patients not known to have died at the analysis cut-off using the sources presented in Table 2-1 below:

Table 2-1 Last contact date data sources

Source data	Condition
Last contact date/last date patient was known to be alive from survival follow-up page	Patient status is reported to be alive. **Do not use if patient status is reported unknown.**
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.
End of treatment date from end of treatment page	No condition.
Tumor (RECIST) assessment date	Evaluation is marked as 'done'.
Laboratory/PK collection dates	Sample collection marked as done.
Vital signs date	At least one non-missing parameter value.
Performance status date	Non-missing performance status.
Start/End dates of AE	Non-missing verbatim term.

The last contact date on or before the data cut-off date should be used; the cut-off date should not be used as the censoring date (even in presence of post cut-off data) unless the patient was seen or contacted on the cut-off date.

Imputed dates (e.g., 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. Partially imputed dates (i.e., only day or day and month imputed) are allowed to be used for last contact date only if coming from Survival Follow-up page.

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

2.2 Data included in the analysis

An efficacy interim analysis of PFS was conducted to allow early termination of the everolimus monotherapy arm, in case the efficacy in the everolimus monotherapy arm was by far inferior compared to the everolimus + exemestane combination arm. This efficacy interim analysis was planned after 75 PFS events have been reached across the following 2 arms: everolimus monotherapy and everolimus + exemestane combination treatment.

The final PFS analysis is planned to be conducted when approximately 150 PFS events per local tumor assessment have been documented in each of the two following groups:

- the everolimus + exemestane combination arm combined with the everolimus monotherapy arm, and

- the everolimus + exemestane combination arm combined with the capecitabine monotherapy arm.

The final OS analysis was planned to be conducted 2 years after the randomization of the last patient. Following protocol amendment, the final OS analysis will be conducted at the time of the planned final PFS analysis.

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

2.3 Definitions of analysis populations

The **Full Analysis Set (FAS)** comprises all patients to whom study treatment has been assigned by randomization. All efficacy analyses will be conducted using data from this population according to the intent-to-treat (ITT) principle, i.e., patients will be analyzed according to the treatment and stratum they have been assigned to during the randomization procedure.

The **Safety Set** includes all patients who received any study treatment. In analyses of the safety set, patients will be analyzed according to the study treatment they actually received.

The actual treatment received corresponds to:

- the randomized treatment if patients took at least one dose of that treatment.
- the first treatment received if the randomized treatment was never received

2.4 Concomitant medications with specific impact on the analysis

2.4.1 Inducers, inhibitors and substrates of CYP3A

Everolimus is metabolized by CYP3A4 in the liver and, to some extent, in the intestinal wall.

Capecitabine and its metabolites do not inhibit the metabolism of CYP3A4 substrates but may inhibit the metabolism of CYP2C9 substrates. Therefore:

- Co-administration of CYP2C9 substrates (e.g. Irbesartan, Losartan, Phenytoin, Cyclophosphamide) should be exercised with caution and monitored closely per capecitabine local label.
- co-administration with strong inhibitors of CYP3A4 (e.g., ketoconazole, itraconazole, ritonavir) or P-glycoprotein (PgP) inhibitor is to be avoided;
- co-administration with moderate CYP3A4 inhibitors (e.g., erythromycin, fluconazole) or PgP inhibitors is to be used with caution;
- concomitant use of seville orange, star fruit, grapefruit and their juices that affect CYP3A4 and PgP activity is to be avoided;
- concomitant use of strong CYP3A4 inducers such as phenytoin, carbamazepine, rifampin, rifabutin, phenobarbital, St. John's wort is to be avoided.

The following will be tabulated and summarized:

1. clinically relevant drug interactions: substrates, inducers, and inhibitors of isoenzyme CYP3A;
2. list of clinically relevant drug interactions mediated by PgP substrates;
3. list of clinically relevant drug interactions mediated by PgP inhibitors;
4. list of clinically relevant drug interactions mediated by PgP inducers.

Despite the fact that some of these drugs should be avoided completely and some used with caution, there may be patients who took these drugs during the study. In that case, the concomitant medications need to be identified and classified (review to be performed by a Clinical Pharmacologist) and then tabulated and/or listed in the Clinical Study Report as appropriate.

2.4.2 Further anti-neoplastic therapy

Administration of anti-neoplastic drugs (apart from study treatment) and other investigational drugs is not allowed during study treatment. Patients who take such anti-neoplastic drugs after randomization but before the end of treatment will be identified as protocol deviations. In addition, their efficacy data (other than overall survival) will be censored so that tumor assessments made after the start of anti-neoplastic drugs will not be included in the primary efficacy analyses for PFS. The same rule will apply to efficacy analyses for best overall response (BOR), ECOG performance status and QoL assessments. For details on the censoring rules, see [Appendix 1 of the study protocol].

Clinical review of study data will be performed to identify anti-neoplastic medications that are not allowed during study treatment.

2.5 Implementation of RECIST 1.1

Response and progression evaluation will be performed according to the RECIST 1.1 guideline (as described in detail in [Appendix 1 of the study protocol]). The text below provides more detailed instructions and rules needed for programming purposes.

2.5.1 Overall lesion response for patients with bone lesions only at baseline

For patients with lytic or mixed lytic-blastic bone lesions only at baseline, RECIST 1.1 will be extended to allow the evaluation of overall lesion response in such patients, which will be based solely on non-target lesion responses and/or an occurrence of a new lesion. Bone lesions will be entered as non-target lesions. Specifically, in the absence of new lesions, the overall lesion response at each assessment will be one of the following: complete response, -non-CR/non-PD, unknown response, or progressive disease based on non-target lesion responses. -Non-CR/non-PD response -will include all assessments not qualifying for complete response, progressive disease or unknown response. In the presence of a new lesion, the overall lesion response will be progressive disease.

2.5.2 Disease progression

For patients with measurable disease at baseline, disease progression will only be assigned if it is documented as per RECIST 1.1 by an unbiased assessment method (e.g., CT scan, MRI,

X-ray, photography for skin lesions, etc.). A new lesion will be entered on the 'New lesion' RECIST eCRF with the corresponding measurement method (or method='Other').

Discontinuation due to disease progression without supporting objective evidence (as defined above) will not be attributed to progressive disease.

Patients with bone only lesions, lytic or mixed lytic-blastic, will be allowed to enter the study but such lesions will not be considered as measurable lesions in this study. The following criteria will be used to declare disease progression among these patients:

- the appearance of one or more new lytic lesions in bone,
- the appearance of one or more new lesions outside of bone,
- unequivocal progression of existing bone lesions.

Note: Pathologic fracture, new compression fracture, or complications of bone metastases will not be considered as evidence of disease progression unless one of the above-mentioned criteria is fulfilled.

2.5.3 Best overall response

The best overall tumor response will be assessed per RECIST 1.1. The definitions and details of the derivation are given in [Appendix 1 of the study protocol].

Only tumor assessments performed before the start of any further anti-neoplastic therapies (i.e., any additional anti-neoplastic medications or surgery) will be considered in the assessment of best overall response. These anti-neoplastic therapies will be identified from the data collected on 'Anti-neoplastic therapies since discontinuation of study drug' eCRF.

Since, in this study, tumor assessments are performed every 6 weeks, allowing for a ± 1-week deviation window, the standard definition of a best overall response evaluation of "stable disease", "progressive disease" or "unknown" given in [Appendix 1 of the study protocol] requires an adjustment.

The following definitions will be used:

- The determination of CR and PR remains the same as stated in [Appendix 1 of the study protocol].
- SD = at least one SD assessment (or better) after randomization (and not qualifying for CR or PR). Because the first assessment after randomization is scheduled at 6 weeks, allowing for a 1-week deviation window, SD is defined as at least one SD assessment (or better) > 5 weeks after randomization (and not qualifying for CR or PR).
- If progression is detected as the first evaluable assessment in ≤ 9 weeks after randomization, the best overall response evaluation will be "progressive disease."
- UNK = all other cases (i.e., not qualifying for confirmed CR or PR and without SD for more than 5 weeks or early progression within the first 9 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-randomization assessment,
- all post-randomization assessments have overall response UNK,

- new anti-neoplastic therapy started before first post-randomization assessment,
- SD in less than 5 weeks post-randomization ('SD too early'),
- PD after more than 9 weeks post-randomization ('PD too late').

Special (and rare) cases where best overall response is "unknown" due to both SD in less than 5 weeks and PD after more than 9 weeks post-randomization will be classified as 'SD too early.'

2.5.4 Change in imaging modality

Per RECIST 1.1, the imaging method used at baseline should be matched at all subsequent assessments. 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 change in technique (e.g. from CT to MRI, or vice-versa), or a change in any other imaging modality. A change in methodology will result by default in a UNK overall lesion response assessment. However, another response assessment than the Novartis calculated UNK response may be accepted from the investigator if a definitive response assessment can be justified, based on the available information.

2.5.5 Determination of missing adequate assessments

The term 'missing adequate assessment' is defined as a tumor assessment that is not done or for which the overall lesion response is 'unknown'. For the sake of simplicity, a 'missing adequate assessment' is also referred to as a 'missing assessment'.

As detailed in Appendix 1 to the study protocol, the PFS censoring and event date options depend on the presence and the number of missing tumor assessments. For example:

1. in the primary analysis of PFS, an event occurring after two or more consecutive missing assessments is censored at the last adequate assessment, i.e., the last assessment preceding the missing assessment(s);
2. in one of the sensitivity analyses of PFS, an event occurring after one or more missing assessments is back-dated to the date of the next scheduled assessment.

An exact rule to determine whether there is none, one or two missing assessments is therefore needed. This rule is based on the time interval (distance) between the last adequate tumor assessment date and the event date.

If the distance is greater than threshold D1 = 6+2 = 8 weeks, the analysis will assume one missing assessment. If the distance is greater than D2 = (2*6)+2 = 14 weeks, the analysis will assume two missing assessments. The threshold D1 is formed based on the protocol-specified interval between consecutive tumor assessments plus the protocol-allowed window around the assessments. Similarly, the threshold D2 is formed based on the two-fold of the protocol-specified interval between consecutive tumor assessments plus the protocol-allowed window around the assessments.

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, the patient will be censored and the censoring reason will be 'Event documented after two or more missing tumor assessments'.

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

1. analysis cut-off date

2. start date of further anti-neoplastic therapy

3. date of study treatment discontinuation due to consent withdrawal

4. date of study treatment discontinuation due to loss to follow-up

is smaller or equal to D2, the censoring reason will be 1.'Ongoing', 2. 'New cancer therapy added', 3. 'Withdrew consent' or 4. 'Loss to follow-up', respectively, depending on the case. However, if this distance is larger than D2, the censoring reason will always default to 'Adequate assessment no longer available'.

3 Statistical methods used in reporting

3.1 Enrollment status

Number of patients screened will be summarized by country, center and randomization stratum. Number of patients randomized will be summarized by country, center, randomization stratum and treatment group.

3.2 Background and demographic characteristics

The Full Analysis Set (FAS) will be used for all baseline disease characteristics, demographic summaries and data listings.

3.2.1 Basic demographic and background disease data

All demographic and background disease characteristics will be listed in detail. Qualitative data such as sex, race, disease stage, ECOG performance status, etc. will be presented by treatment arm using frequency tables (counts and proportions by category). Relevant descriptive statistics (mean, median, minimum, maximum and standard deviation in most cases) by treatment arm will be used to present quantitative data such as age, body weight, etc.

Number of patients in each randomization stratum (stratification information obtained from IRT) will be presented. Potential discrepancies between randomization stratification information (obtained from IRT) and strata formed based on baseline factors collected on eCRFs will be tabulated and listed.

3.2.2 Protocol eligibility criteria

Protocol eligibility criteria as per eCRFs will be summarized and listed.

3.2.3 Diagnosis and extent of cancer

Summary statistics will be tabulated for diagnosis and extent of cancer. According to data collected on the eCRF, this analysis will include the following variables: primary site of cancer, details of tumor histology/cytology, histological grade, time since initial diagnosis,

date of first metastatic recurrence, date of most recent recurrence/metastasis, presence/absence of target and non-target lesions, number and type of organs involved.

The numbers and percentages of patients in categories defined by the following variables 'presence/absence of target and non-target lesions', 'number of organs involved' and 'organ types involved' will be based on data collected on the radiology RECIST, and eCRFs for diagnosis and extent of cancer, in particular, on the individual target and non-target lesion codes.

Time since initial diagnosis, time since first recurrence/metastasis, as well as time between first diagnosis and first recurrence/metastasis will be summarized in months. A month is defined as 365.25 / 12 = 30.4375 days.

3.2.4 Medical history

Medical history and ongoing conditions, including cancer-related conditions and symptoms, 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 Medical dictionary for regulatory activities [MedDRA] terminology.)

3.2.5 Prior anti-neoplastic therapy

Prior anti-neoplastic therapy will be listed in three separate listings: 1. medications, 2. radiotherapy, 3. surgery.

The numbers and percentages of patients recording any prior anti-neoplastic medications, prior anti-neoplastic radiotherapy and prior anti-neoplastic surgery will be summarized by treatment.

Prior anti-neoplastic medications will be summarized by therapy type (radiotherapy, surgery, NSAI (defined as "Letrozole", "Anastrozole"), hormonal therapy other than NSAI, chemotherapy, immunotherapy, targeted therapy and others). Any indication-specific significant prior anti-neoplastic medications will be identified from the summaries mentioned above.

In addition, the type of prior chemotherapies received will be summarized by treatment arms (i.e. for monotherapy regimens, the drug received will be reported; for regimen comprised of a combination of several drugs the combination will be reported).

3.2.6 Other

All data collected at baseline, including source of subject referral and child bearing potential, will be listed.

3.3 Protocol deviation summaries

The number and percentage of patients in the Full Analysis Set with any protocol deviation will be tabulated by the deviation category (as specified in the VAP documents) and by treatment group. The protocol deviations will also be summarized by center.

Protocol deviations leading to the exclusion from the analysis populations will be tabulated separately by treatment group.

All protocol deviations will be listed.

3.4 Groupings for analysis

The number and percentage of patients in each analysis population (definitions are provided in Section 2.3) will be summarized by treatment group and randomization stratum. The distribution of patients in screening and in selected analysis populations will also be summarized by country, center, treatment group and randomization stratum.

3.5 Patient disposition

The Full Analysis Set will be used for patient disposition summaries. Based on the two eCRF pages 'End of Treatment' and 'Study Evaluation Completion' there will be one combined summary, stratified by treatment, showing:

- Number (%) of patients who are still on-treatment (based on the absence of the 'End of Treatment' page);
- Number (%) of patients who discontinued the study treatment (based on the 'End of Treatment' page);
- Reasons for study treatment discontinuation (based on the 'End of Treatment' page).
- Number (%) of patients who entered the post-treatment evaluation phase (based on the 'End of Treatment' page);
- Number (%) of patients who discontinued from the post-treatment evaluations (based on the 'Study Evaluation Completion' page);
- Reasons for discontinuation from the post-treatment evaluations phase (based on the 'Study Evaluation Completion' page).

The summary of reasons for End of treatment discontinuation will also be presented separately for patients who have been censored in the primary PFS analysis due to the reason of "Antineoplastic therapy started" reason.

3.6 Study treatment

Duration of study treatment exposure, cumulative dose, dose intensity (DI) and relative dose intensity (RDI) will be summarized by treatment. In addition, the duration of exposure to study treatment will be categorized into time intervals; frequency counts and percentages will be presented for the number of patients in each interval. The number of patients, who have dose reductions or interruptions, and the reasons, will be summarized by treatment.

Listings of all doses of the study treatment along with dose change reasons will be produced.

The Safety Set will be used for all summaries and listings of study treatment.

3.6.1 Duration of study treatment exposure

The following algorithm will be used for everolimus and exmestane to calculate the duration of study treatment exposure for patients who took at least one dose of any of the components of the study treatment:

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

For capecitabine, the duration of exposure will include the rest periods and will be then defined as:

Duration of exposure (days) = (date of last administration of study treatment) – (date of first administration of study treatment) + 7 days (corresponding to the theoretical rest period) + 1.

The duration includes the periods of temporary interruption (of any component of the study treatment for any reason).

Duration of exposure to each component of the study treatment will also be calculated.

3.6.2 Cumulative dose

Cumulative dose is defined as the total dose given during the study treatment exposure period and will be summarized for each of the study treatment components separately. For patients who do not receive any drug, the cumulative dose will be set to zero.

3.6.3 Dose intensity and relative dose intensity

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

DI (dosing unit / unit of time) = Cumulative dose (dosing unit) / Duration of exposure (unit of time).

For patients who do not receive any drug, the DI will be set to zero.

Planned dose intensity (PDI) is the assigned dose by unit of time planned to be given to patients as per protocol in the same dose unit and unit of time as that of the Dose Intensity.

For Capecitabine, the planned dose recommended as per protocol is 1250 mg/m2 twice daily for 14-days every 3 weeks (i.e. 3-week cycle). The planned dose intensity including the rest period of 7 days is calculated as:

PDI (Cap.) = 1250 mg/m2*2*14 days / 21 days = 1666.67 mg/m2/day.

Relative dose intensity (RDI) is defined as follows:

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

DI and RDI will be summarized separately for each of the study treatment components, but using the duration of the study treatment exposure, not the duration of exposure to each of the components.

For the calculation of the assigned body-surface-area-adjusted capecitabine dose at every cycle, the following formula will be used:

BSA [m2] = 234.94*(height[cm]**0.422)*(weight[kg]**0.515)/10000.

according to the height at baseline and the last available weight at or before the cycle start date.

3.6.4 Dose reductions or interruptions

The number of patients, who have dose reductions or interruptions, and the reasons for such reductions/interruptions, will be summarized separately for each of the study treatment components.

An interruption is defined as a 0mg dose given on one or more days. However, for last records, it will be counted as an interruption only if there are 2 or more last records with 0 mg dose.

If a patient moves from a dose level that is higher than the studied dose under the protocol to the dose level that is being studied in the protocol, such changes will not be counted as reductions. However, if any patient moves directly from a higher than studied dose down to a lower than protocol-studied dose, or to the dose level being studied under the protocol but on a less frequent regimen, such changes will be counted as reductions. If first dose is lower than the studied dose under the protocol, it will be counted as a dose reduction.

For everolimus and exemestane, reductions count should be based on the actual total daily dose (mg).

For capecitabine, reductions count should be based on the actual total daily dose adjusted to BSA (mg/m^2) with an acceptance windows of 10%: if the difference compared to the previous non-zero dose (mg/m^2) is greater than 10% then it is a reduction. If first dose is lower than the studied dose under the protocol (mg/m^2) by more than 10%, it will be counted as a dose reduction.

If one drug is permanently discontinued (before a protocol-planned discontinuation date) while the other is ongoing, such discontinuations will be classified as interruptions.

Dose reductions and interruptions will be tabulated both separately and in a combined fashion. In the combined summary, dose interruptions will be considered as dose reductions to 0 mg, and therefore all reductions/interruptions will be labeled as reductions and tabulated in one table. Dose escalations will be summarized by treatment and level of reduced dose (-1 versus -2) for all patients with dose reductions in a separate table.

3.7 Concomitant therapy

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

Concomitant medications entered into the database will be coded using the WHO Drug Reference List to allow for categorization by preferred term. In addition to categorizing medication data by preferred term, drugs will be classified according to their ATC classification in order to present and compare how they are being utilized.

Concomitant medications and significant non-drug therapies taken concurrently with the study drug(s) will be listed and summarized by ATC class, preferred term and treatment arm by

means of frequency counts and percentages. These summaries will include medications starting on or after the start of study treatment or medications starting prior to the start of study treatment and continuing after the start of study treatment.

Any prior concomitant medications or significant non-drug therapies starting and ending prior to the start of study treatment will be listed.

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., efficacy analyses) will be identified prior to database lock. Separate summaries of these concomitant medications will be produced and the corresponding analysis populations will be used. Strong and moderate inhibitors, inducers, or substrates of the isoenzyme CYP3A will be identified as described in Section 2.4.1. The strong ones will be tabulated by ATC code. Both strong and moderate ones will be listed. New anti-neoplastic therapies will be listed based on their identification (by the method given in Section 2.4.2) by the protocol deviation process.

Anti-neoplastic therapies since discontinuation of study drug will be listed and summarized by ATC class, preferred term and treatment arm by means of frequency counts and percentages in separate summaries using the Full Analysis Set. In addition, the first regimen with combination of several drugs if any after the study treatment discontinuation will be summarized by treatment arms, using the full analysis set population.

3.8 Efficacy evaluation

The efficacy endpoints based on tumor assessments will be derived according to RECIST 1.1 (see Appendix 1 of the study protocol for details). The tumor endpoint derivation is based on the sequence of overall lesion responses at each assessment/time point. The overall lesion response at a given assessment/time point will be provided by the local (treating center's) radiologist/investigator. In particular, the final local (treating center's) radiologist's/investigator's visit response for each assessment/time point collected on the RECIST 1.1 overall lesion response CRF page will be used to derive the primary efficacy endpoint. The tumor assessment dates will be derived by Novartis using the dates of the individual lesion measurements.

Data included in efficacy analyses

Efficacy analyses will include all data observed in patients from the FAS regardless of whether the data were observed on-treatment or after the study treatment discontinuation until the analysis cut-off date. In particular, the "30 days" rule applied to all safety analyses will NOT be used for efficacy analyses.

For patients who took other anti-neoplastic drugs, their efficacy data (other than overall survival) will be censored so that the tumor assessments made after the administration of the other anti-neoplastic drugs are not included in the primary efficacy analyses.

For patients with protocol deviation code C17 ("Lack of PI oversight and no eCRF signatures at Database Lock (DBL) due to PI departure and no replacement"), efficacy data will be analyzed as follows:

- Progression-Free Survival will be censored at randomization date with censoring reason "Adequate assessment no longer available"
- Overall Survival will be censored at randomization date with censoring reason "Lost to follow-up"
- Best Overall Response will be set to UNK with reason "No valid post-baseline assessment"
- Overall Response status as non-responder
- Clinical Benefit status as non-responder

3.8.1 Primary Efficacy

Progression-free survival (PFS) derived from an investigator's assessment of radiology data will be used as the primary efficacy variable. The 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, whichever occurs first. If a patient has not progressed or died at the analysis cut-off date, or if she receives any further anti-neoplastic therapy, PFS will be censored at the time of the last tumor assessment before the cut-off date or the anti-neoplastic therapy date whichever occurs first. Further anti-neoplastic therapies will be identified via protocol deviations and from data collected on 'Antineoplastic therapies since discontinuation of study drug' eCRF. Definitions and further details on PFS can be found in Appendix 1 of the protocol.

Discontinuation due to disease progression (collected on the "End of treatment" and "Study Evaluation Completion" page) without supporting objective evidence satisfying progression criteria per RECIST 1.1 will not be considered a progressive disease.

Primary analysis

The primary objective of this study is to estimate the hazard ratio of a progression-free survival event comparing the everolimus + exemestane combination therapy with the everolimus monotherapy in postmenopausal women with ER-positive, HER2-negative, advanced breast cancer (ABC) after recurrence or progression on letrozole or anastrozole.

The primary analysis of PFS will be based on data from investigator/local radiology review. The analysis will be performed on the FAS following the intention-to-treat (ITT) principle, i.e., patients will be analyzed according to the treatment group they were randomized and the stratum they were assigned to at baseline. The analysis will use the default censoring and event date definitions from Table 3-1 of Appendix 1 of the study protocol, i.e. A(1), B(1), C1(1), C2(1), D(1), E(1), and F(1). In particular, PFS will be censored at the last adequate tumor assessment if one of the following occurs: absence of a PFS event; a PFS event occurs after a new anticancer therapy is given; a PFS event occurs after two or more missing tumor assessments (see Section 2.5.5). Discontinuation of study treatment (for any reason) will not be considered as a reason for censoring.

Kaplan-Meier estimates

The Kaplan-Meier (KM) estimate of the progression-free survival function will be computed for each treatment group using PROC LIFETEST with method=KM option in SAS. The results will also be plotted graphically (Kaplan-Meier curves) by treatment and by randomization strata as assigned through IRT. The plots will display the number of patients at risk at equidistant time points.

The estimated median PFS for each treatment group will be provided along with the approximate 90% confidence intervals (Brookmeyer and Crowley, 1982). Additionally, the 25^{th} and 75^{th} percentiles will also be computed. The progression-free survival probabilities at 2, 4, 6 and 9 months (timepoints to be adapted depending on the length of the follow-up), along with 90% confidence intervals, will be presented by treatment group. The log-log transformation available within PROC LIFETEST will be used to compute the confidence intervals.

Hazard ratio estimate

The hazard ratio estimate of a PFS event comparing the everolimus + exemestane combination therapy with everolimus monotherapy, along with the two-sided 90% confidence interval, will be obtained from the stratified Cox proportional hazards model, fitted using SAS PHREG procedure with ties=EXACT option and the stratification information obtained through IRT. In the PHREG procedure, the MODEL statement will include the indicator of assignment to the everolimus + exemestane arm as the only covariate. The baseline hazard function will be allowed to vary across strata, i.e., the STRATA statement will include the stratification variable obtained through IRT.

Sensitivity and other supportive analyses of the primary endpoint PFS

To assess the impact of stratification, a sensitivity analysis will be performed to estimate the PFS treatment effect (everolimus + exemestane versus everolimus alone) using the unstratified Cox regression model yielding the hazard ratio estimate of a PFS event along with the 90% confidence interval.

The following sensitivity PFS analyses will be performed to assess the impact of missing/unknown tumor assessments (analyses 1 and 2 below) and to assess the impact of PFS censoring due to another anti-neoplastic therapy (analysis 3 below):

- PFS using local radiology assessments on the FAS and using the following options in Table 14-5 on page 121 of Appendix 1 in the study protocol: A(1), B(1), C1(1), C2(3), D(1), E(1), and F(1), i.e., taking the event whenever it occurs even after two or more missing tumor assessments (see Section 2.5.5). In the summary table, this approach will be referred to as the ‘actual event PFS analysis’.
- PFS using local radiology assessments on the FAS and using the following options in Table 14-5 on page 121 of Appendix 1 in the study protocol: A(1), B(1), C1(2), C2(2), D(1), E(1), and F(1), i.e., backdating of events occurring after missing tumor assessments. In the summary tables, this approach will be referred to as the ‘backdating PFS analysis’.
- PFS using local radiology assessments on the FAS and using the following options in Table 14-5 on page 121 of Appendix 1 in the study protocol: A(1), B(1), C1(1), C2(1),

D(1), E(1), and F(4), i.e., and not censoring patients at the start of new anti-neoplastic therapies (if tumor assessment data are available following the start of new anti-neoplastic therapies).

A multivariate stratified Cox regression model will be used to evaluate the effect of baseline demographic and disease characteristics of potential prognostic value on PFS. Robustness of the PFS hazard ratio estimate to the adjustment for various prognostic factors in the stratified Cox model including prior chemotherapy use (yes vs. no), performance status (0 vs. 1 or 2), presence of bone lesions only at baseline (yes vs. no), time since first diagnosis of metastasis/recurrence to randomization (< 6 months vs. >= 6 months), non-steroidal aromatase inhibitor (NSAI) (letrozole or anastrozole) use (adjuvant vs. metastatic setting), number of organs involved (1 vs. 2 vs. >=3 – to be adapted depending on the number of patients in each category), and PgR status (positive vs. negative) will be assessed. The strata will be based on stratification information obtained through IRT. A revised multivariate stratified cox model including the key prognostic factors as well as the baseline covariates where imbalances between arms have been observed will be provided: performance status (0 vs. 1 or 2), presence of bone lesions only at baseline (yes vs. no), number of organs involved (1 vs. 2 vs. >=3), prior chemotherapy use (yes vs. no), age (<65 years vs. >=65 years old) and Race (Caucasian vs. non-caucasian).

Further supportive analyses will include:

- 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. The following categories will be used as appropriate (based on the end of treatment page, the study evaluation completion page and the survival page and based on the distance D2 defined in Section 2.5.5):
  - Ongoing without event
  - Adequate assessment no longer available (when follow-up for progression is stopped at a certain time or interrupted for a certain time period before cut-off for any reason, e.g., due to loss to follow-up or consent withdrawal)
  - New cancer therapy added
  - Event documented after two or more missing tumor assessments (for primary analysis only)
- Timing of all tumor assessments will be depicted graphically as per local radiology review by treatment arm
- 95% confidence intervals will be provided for all estimates of the main primary analysis (i.e. median, survival probabilities and hazard ratio)

3.8.2 Secondary Efficacy

Key Secondary Objective

The key secondary objective of this study is to estimate the hazard ratio of a PFS event using local investigator's/radiologist's tumor assessments comparing the everolimus + exemestane

combination therapy with the capecitabine therapy in postmenopausal women with ER-positive, HER2-negative, advanced breast cancer after recurrence or progression on letrozole or anastrozole.

The stratified Cox regression model will be used to estimate the hazard ratio of a PFS event, along with the associated 90% confidence interval, comparing the everolimus + exemestane combination therapy with the capecitabine therapy where the stratification information will be obtained through IRT. The confidence interval for the hazard ratio will not be adjusted for multiple comparisons.

Distribution of PFS will be assessed using the Kaplan-Meier estimation method. The estimated median PFS for each treatment group will be provided along with the approximate 90% confidence intervals. Additionally, the 25^{th} and 75^{th} percentiles will also be computed. The progression-free survival probabilities at 2, 4, 6 and 9 months (timepoints to be adapted depending on the length of the follow-up), along with 90% confidence intervals, will be presented by treatment group.

The same statistical principles will be applied and supportive analyses conducted as described in Section 3.8.1 "Sensitivity and other supportive analyses of the primary endpoint PFS" for the key secondary treatment comparison of everolimus + exemestane combination therapy versus capecitabine therapy.

Other Secondary Objectives

Other secondary objectives of this study are to evaluate each of everolimus + exemestane versus everolimus monotherapy and everolimus + exemestane versus capecitabine monotherapy with respect to overall survival (OS), overall response rate (ORR), clinical benefit rate (CBR), deterioration in the ECOG performance status, changes in quality of life scores over time, and safety.

The analysis of all secondary efficacy endpoints (mentioned above) will be performed on the FAS.

Overall survival (OS)

Overall survival (OS) is defined as the time from the date of randomization to the date of death due to any cause. The cut-off date for the planned final OS analysis was specified 2 years after the randomization of the last patient. However, at that time, the number of PFS events needed for final PFS analysis was not reached. Therefore the final OS analysis will be conducted at the time of the planned final PFS analysis using the same cut-off date. If death has not been observed by the analysis cut-off date, then OS will be censored at the date of last contact.

The OS analysis will be based on data from the FAS on the ITT basis, i.e., according to the treatment group patients are randomized to at baseline. Distribution of OS in each of the three treatment arms will be assessed using the Kaplan-Meier (KM) estimation method, and the treatment-specific KM curves will be graphically displayed. The estimated median OS and probability of surviving at the estimated median OS, along with 90% confidence intervals, will be presented for the three treatment arms. Stratified Cox regression models will be used to estimate the hazard ratio (HR) of death from any cause, along with the associated 90%

confidence interval, comparing (i) the everolimus + exemestane combination therapy with everolimus monotherapy, and (ii) the everolimus + exemestane combination therapy with capecitabine therapy where the stratification information will be obtained through IRT and the baseline hazard functions will be allowed to vary across strata.

As a sensitivity analysis, a stratified multivariate Cox regression model will be fitted after additional adjustment for key potential prognostic factors including prior chemotherapy (yes vs. no), performance status (0 vs. 1, 2), presence of bone lesions only at baseline (yes vs. no), time since first diagnosis of metastasis/recurrence to randomization (<6 Months, >= 6 Months), non-steroidal aromatase inhibitor (NSAI) (letrozole or anastrozole) usage (adjuvant vs. metastatic), number of organs involved (1 vs. 2 vs. ≥ 3 – to be adapted depending on the number of patients in each category) and PgR status (positive vs. negative). A revised multivariate stratified cox model including the key prognostic factors as well as the baseline covariates where imbalances between arms have been observed will be added: performance status (0 vs. 1 or 2), presence of bone lesions only at baseline (yes vs. no), number of organs involved (1 vs. 2 vs. >=3), prior chemotherapy use (yes vs. no), age (<65 years vs. >=65 years old) and Race (Caucasian vs. non-caucasian). The strata will be based on stratification information obtained through IRT.

Survival status, reason for censoring and death cause will be listed. Patients not known to have died will be censored for 'Lost to follow-up' if the time between their last contact date and the analysis cut-off date is longer than 3 months and 2 weeks (104 days).

95% confidence intervals for the median, hazard ratio and survival probabilities of the overall survival will be also provided as supportive analysis.

Overall response rate (ORR)

Overall response rate (ORR) is defined as the proportion of patients with best overall response of complete response (CR) or partial response (PR) according to RECIST 1.1 (Appendix 1). ORR will be calculated based on the FAS according to the ITT principle, using local radiologist's/investigator's tumor assessment. Patients with bone lesions only at baseline will be included in the numerator if they achieve a complete response. ORR estimates will be presented by treatment group along with exact 90% confidence intervals (Clopper and Pearson 1934). The estimation procedure will be repeated based on data for a subset of patients in the FAS with measurable disease only at baseline.

Clinical benefit rate (CBR)

Clinical benefit rate (CBR) is defined as the proportion of patients with best overall response of CR, PR, or overall lesion response of stable disease (SD for measurable disease and non-CR/non-PD (NCRNPD) for non-measurable disease) with duration of 24 weeks or longer. A patient will be considered to have a SD/NCRNPD for 24 weeks or longer if SD/NCRNPD is recorded at 24 weeks or later after randomization. Taking into account the allowed time window for tumor assessment visits, the SD/NCRNPD response has to be recorded at 23 weeks or later after randomization to be included in the CBR calculation. Best overall

response of CR, PR and overall lesion response of SD/NCRNPD are defined according to RECIST 1.1 (see Appendix 1). CBR will be calculated based on the FAS according to the ITT principle, using local radiologist's/investigator's tumor assessment. Patients with non-measurable disease only at baseline will be included in the numerator if they achieve a complete response. CBR estimates will be presented by treatment group along with exact 90% confidence intervals.

ECOG performance status

ECOG performance status (PS) scale will be used to assess physical health of patients, ranging from 0 (most active) to 5 (least active):

Table 3-1 ECOG Performance Status 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

ECOG performance status will be assessed and recorded at baseline, on Treatment Day 1, at week 6 and every 6 weeks thereafter as well as at discontinuation from the study treatment (within one week after discontinuation). If patients discontinue study treatment for any reason other than progression, lost to follow-up or consent withdrawal, the ECOG performance status will continue to be performed every 6 weeks until progression, lost to follow-up, consent withdrawal or investigator decision in patient best interest.

Time windows will be defined for summaries of ECOG data. If more than one assessment is available in the same time window, the assessment closest to the planned date will be considered. If two assessments are obtained with the same time difference compared to the scheduled visit day, the assessment with the worst value will be considered. Data obtained at the end of treatment will be classified as other assessment in the corresponding time window. Note that only data collected under treatment (i.e. while the patient is treated and up to 7 days after last dose intake) will be included in the time to deterioration analysis. Post-treatment data will be summarized separately. The end of treatment assessment will be included if collected within 7 days of the last dose intake.

Time Window	Planned visit timing	Time Window Definition
Baseline	On or before Study Day 1	<= Study Day 1
Week 6	Study Day 43	Study Days 22-63

Week 12	Study Day 85	Study Days 64-105
Week 18	Study Day 127	Study Days 106-147
Every 6 weeks thereafter, Week y=18+6*k (with k=1,2 …)	Study Day (18+6*k)*7+1	Study Days (18+6*k)*7-20 to (18+6*k)*7+21 Note: EOT data visit are included if obtained within 7 days of last non-0 dose intake.
Study Day 1 = randomization date		

Frequencies of patients with ECOG PS values of 0, 1 or >=2 will be used to summarize the ECOG PS data at each time window. .

An analysis of the time to definitive deterioration of the ECOG PS by at least one category of the score from baseline will be performed. A deterioration will be considered definitive if no improvements in the ECOG PS are observed at subsequent measurement times during the treatment period following the time point at which the deterioration is observed. (Example 1: if the score is 1 at baseline and then 1, 2, 1, 2, 3 at study days 28, 57, 83, 115, 150, respectively, then the time to definitive worsening is 115 days. Example 2: if the score is 1 at baseline and then 1, 1, 2 at study days 28, 57, 83, respectively, with no assessment of the ECOG performance status after day 83, then the time to definitive worsening is 83 days.)

Death will be considered as worsening of the ECOG PS if it occurs close to the last available assessment where "close" is defined as being within twice the planned period between two assessments. Patients who die after more than twice the planned period between two assessments will be censored at the date of their last assessment before the cut-off. This avoids overestimating the time to definitive worsening in patients dying after an irregular assessment scheme. For example, if the last assessment is at week 6 and the patient dies at week 16, the definitive deterioration date will be week 16. On the other hand, if the last assessment is at week 6 and the patient dies at week 22, which is after more than twice the planned period between two assessments (6 weeks) since the last assessment (at week 6), then the patient is censored for definitive deterioration and the censoring date will be week 6.Patients receiving any further anti-neoplastic therapy prior to definitive worsening will be censored at their date of last assessment prior to the start of therapy. Patients that have not worsened at the data cut-off point will be censored at the date of last assessment prior to the cutoff.

The Kaplan-Meier estimation method will be used to assess the distribution of time to definitive worsening in the ECOG PS score, stratified by treatment. The estimated treatment-specific median times to definitive worsening will be presented along with 90% confidence intervals. The stratified Cox regression model will be used to estimate the hazard ratio of a definitive worsening in the ECOG PS score, along with the associated 90% confidence interval, comparing the everolimus + exemestane combination therapy with everolimus alone and comparing the everolimus + exemestane combination therapy with the capecitabine therapy.

Patient reported outcomes (PRO)

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

Quality of life questionnaire (QLQ)

The EORTC QLQ-C30 questionnaire along with the breast module (BR23) will be used to collect patients' quality of life (QoL) data. Raw QoL data will be scored according to the EORTC scoring manual. If at least half of the items comprising the scale have been answered, the score for this scale will be calculated. For single item scales with missing responses and scales where less than half of the items have been answered, the score for these scales will be set to missing. The global health status/QoL scale score is identified as the primary QoL variable of interest. Physical functioning, emotional functioning and social functioning scale scores in QLQ-C30, and functional and symptom scale scores in BR23 are identified as secondary QoL variables of interest.

QLQ-C30 and BR23 will be assessed and recorded at baseline, on Treatment Day 1, at week 6 and every 6 weeks thereafter as well as at discontinuation from the study treatment (within one week after discontinuation). If patients discontinue study treatment for any reason other than progression, lost to follow-up or consent withdrawal, the QLQ-C30 and BR23 performance status and will continue to be performed every 6 weeks until progression, lost to follow-up, consent withdrawal or investigator decision in patient best interest.

Time windows will be defined for descriptive summaries of PRO data. If more than one assessment is available in the same time window, the assessment closest to the planned date will be considered. If two assessments are obtained with the same time difference compared to the scheduled visit day, the assessment obtained prior to scheduled visit day will be considered. Data obtained at the end of treatment will be classified as other assessment in the corresponding time window. Note that only data collected under treatment (i.e. while the patient is treated and up to 7 days after last dose intake) will be included in the time to deterioration analysis. Post-treatment data will be summarized separately. The end of treatment assessment will be included if collected within 7 days of the last dose intake.

Time Window	Planned visit timing	Time Window Definition
Baseline	On or before Study Day 1	<= Study Day 1
Week 6	Study Day 43	Study Days 22-63
Week 12	Study Day 85	Study Days 64-105
Week 18	Study Day 127	Study Days 106-147
Every 6 weeks thereafter, Week y=18+6*k (with k=1,2 …)	Study Day (18+6*k)*7+1	Study Days (18+6*k)*7-20 to (18+6*k)*7+21 Note: EOT data visit are included if obtained within 7 days of last non-0 dose intake.
Study Day 1 = randomization date		

The number of patients filling QoL data and the number of patients missing/expected to have QoL assessments will be summarized by each treatment group for scheduled assessment time points. Furthermore, the amount and the pattern of missing data will be explored by treatment group and over time. The following categories will be used to describe whether the questionnaire was completed at a specific time point:

- yes, fully completed
- yes, partially completed
- no, patient (or subject) missed scheduled assessment visit
- no, patient refused due to poor health
- no, patient (or subject) refused (unrelated to health)- no, study staff felt patient was too ill
- no, questionnaire not available
- no, institutional error
- no, other

Other exploratory analyses may be performed to better understand the missingness pattern for QoL data.

Descriptive statistics (count, mean, median, standard deviation, first and third quartile) will be used to summarize domains at each scheduled assessment time. Patients will be included if they completed at least one questionnaire item. Additionally, change from baseline in the scale 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 analysis.

Time to definitive 10% deterioration in the global health status / QoL score, and in each of the three secondary scales, will be examined for the three treatment arms. In addition, time to definitive 5-point and 10-point deterioration in the global health status / QoL score will be explored for each treatment arm. The assessed distributions will be presented descriptively using Kaplan-Meier curves. Summary statistics based on Kaplan-Meier distributions will be presented, including the estimated median time to definitive 10% (5-, 10-point) deterioration and the probability of not experiencing definitive 10% (5-, 10-point) deterioration by 12 and 24 weeks. Both point estimates and 90% confidence intervals will be presented.

Definitive 10% (5- or 10-point) deterioration is defined as a decrease in score by at least 10% (5 or 10 points) compared to baseline, with no later increase above this threshold observed during the course of the study. Baseline is defined as the latest available assessment made on or before the date of randomization when patient can still deteriorate. Time to definitive deterioration is the number of days between the date of randomization and the date of the assessment at which definitive deterioration is seen.

Death will be considered as deterioration of symptoms/QoL if it occurs close to the last available assessment where “close” is defined as twice the planned period between two assessments. This avoids overestimating the time to definitive worsening in patients dying after an irregular assessment scheme. Patients who die after more than twice the planned

period between two assessments since the last assessment will be censored at the date of their last available questionnaire.

Patients receiving any further anti-neoplastic therapy before definitive worsening will be censored at the date of their last assessment before starting this therapy. Patients that have not worsened as of the cut-off date for the analysis will be censored at the date of their last assessment (questionnaire) before the cut-off.

Treatment satisfaction questionnaire for medication (TSQM)

Patients' self-reported satisfaction or dissatisfaction with the study treatment will be measured using the Treatment Satisfaction Questionnaire for Medication (TSQM) version 1.4. TSQM is planned to be administered at week 3, 6, 12, and the end-of-treatment (EOT) visit. The questionnaire will be administered in the patients' local language.

TSQM items will be divided into four scales: side effects, effectiveness, convenience, and global satisfaction. Raw TSQM data will be scored according to the scoring manual. Only scale scores with not more than one missing item will be computed.

The following time-based intervals will be used to group TSQM data over time. Day x is defined as date of TSQM administration – randomization date + 1. If more than one assessment is available in the same time window, the assessment closest to the planned date will be considered. If two assessments are obtained with the same time difference compared to the scheduled visit day, the assessment obtained prior to scheduled visit day will be considered.

Time window	Planned visit timing	Time Window Definition
Week 3	Study Day 22	Days 11 to 31
Week 6	Study Day 43	Days 32 to 63
Week 12	Study Day 85	Days 64 to 105
EOT	EOT Day	Days (EOT-20) to (EOT+7)

If EOT occurs at the same time of week 3, week 6 or week 12 time window, priority is given for reporting at EOT time window.

All summaries of TSQM data will be based on the FAS according to the ITT principle. The number of patients filling TSQM data and the number of patients missing/expected to have TSQM assessments will be summarized by each treatment group for scheduled assessment time points. Furthermore, the amount and the pattern of missing data will be explored by treatment group and over time. The following categories will be used to describe whether the questionnaire was completed at a specific time point:

- yes, fully completed
- yes, partially completed

- no

- no, patient (or subject) missed scheduled assessment visit

- no, patient refused due to poor health

- no, patient (or subject) refused (unrelated to health)

- no, study staff felt patient was too ill

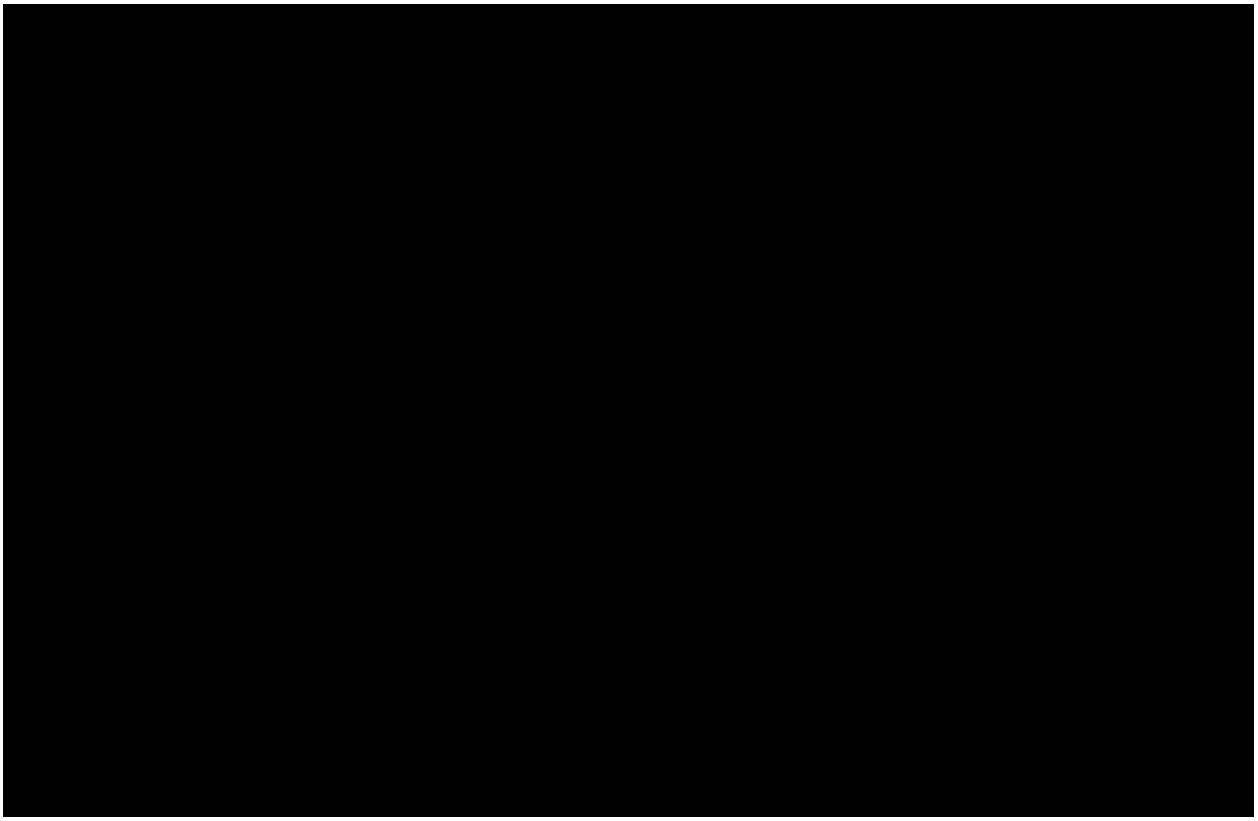
- no, questionnaire not available in the appropriate language

- no, institutional error

- no, other

Other exploratory analyses may be performed to better understand the missingness pattern for TSQM data.

Descriptive statistics (count, mean, median, standard deviation, first and third quartile) will be used to summarize individual item and multi-item scale scores by treatment group and assessment time point (week 3, week 6, week 12, and EOT). Differences in mean scale scores at weeks 3 and 12, and differences in mean change in scale scores between weeks 3 and 12 along with 90% confidence intervals comparing treatment satisfaction with everolimus + exemestane versus everolimus alone, and everolimus + exemestane versus capecitabine will be reported (no significance testing will be performed). The normal approximation will be used in the computation of confidence intervals.



3.9 Safety evaluation

For all safety analyses, the safety set will be used. The assessment of safety will be based mainly on the frequency of adverse events and on the number of laboratory values that fall outside of pre-determined ranges.

The overall observation period will be divided into three mutually exclusive segments:

- pre-treatment period: from day of patient's informed consent to the day before first dose of study treatment
- on-treatment period: from day of first dose of study medication to 30 days after last dose of study treatment
- post-treatment period: starting on day 31 after last dose of study treatment.

Safety summary tables will only include on-treatment events/assessments, i.e., those collected no later than 30 days after the date of the last study treatment administration. All safety events/assessments will be listed and those collected in the pre- and post-treatment period 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 graded 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. This grading system inherently places a value on the importance of an event, although there is not necessarily proportionality among grades (a grade 2 AE is not necessarily twice as severe as a grade 1 AE).

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 study; rather, this information will be collected on the "End of Treatment", "Study Evaluation Completion" or "Survival Information" eCRF pages.

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 treatment/exposure date. All AEs will be listed. AEs starting prior to study day 1 and AEs starting later than 30 days after the last treatment/exposure date will be flagged in the listings.

AEs will be summarized by presenting the number and percentage of patients having at least one AE, and having at least one AE in each body system/primary system organ class (SOC_TXT), and for each preferred term (PT_TXT) 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 CTC grade (AEVGRD1C). A patient with multiple CTC grades for an AE will be summarized under the maximum CTC grade recorded for the event. In the summaries presented by grade, all AEs will be pooled regardless of whether they are CTC gradable or not, i.e., regardless of whether the question "CTC AE" (variable CTIAEV1C) on the Adverse Events eCRF is answered 'Yes' or 'No'.

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

Any information collected (e.g., CTC grades, relatedness to study drug, action taken, etc.) will be listed as appropriate.

3.9.1.4 AE summaries

The following adverse event summaries will be produced:

- Adverse events, regardless of study drug relationship by primary system organ class and preferred term
- Adverse events with suspected relationship to study drug by primary system organ class, preferred term
- Adverse events with an overall incidence rate of 5% or more in either treatment arm, regardless of study drug relationship by primary system organ class and preferred term
- Adverse events, regardless of study drug relationship by primary system organ class, preferred term and maximum CTC
- Adverse events with suspected study drug relationship by primary system organ class, preferred term and maximum CTC
- CTC grade 3 or 4 adverse events, regardless of study drug relationship by primary system organ class and preferred term
- CTC grade 3 or 4 adverse events with suspected study drug relationship by primary system organ class and preferred term
- Deaths, by primary system organ class and preferred term
- On-treatment deaths, by primary system organ class and preferred term
- Serious adverse events, regardless of study drug relationship, by primary system organ class and preferred term
- Serious adverse events with suspected study drug relationship, by primary system organ class and preferred term
- Adverse events leading to study drug discontinuation, regardless of study drug relationship, by primary system organ class and preferred term
- Adverse events requiring dose adjustment or study-drug interruption, regardless of study drug relationship, by primary system organ class and preferred term
- Adverse events requiring additional therapy, regardless of study drug relationship, by primary system organ class and preferred term
- All grade 3/4 non-hematological adverse events

3.9.1.5 AEs of special interest (AESIs)

Specific groupings of adverse events (AESIs) will be considered and the number of patients with at least one event in each grouping will be reported for the entire population. Such groups consist of adverse events for which there is a specific clinical interest in connection with RAD001 treatment (i.e. where RAD001 may influence a common mechanism of action responsible for triggering them) or adverse events which are similar in nature (although not identical). The latest version of Case Retrieval Strategy (CRS) sheet should be used.

3.9.1.6 Clinical trial safety disclosure

For the legal requirements of ClinicalTrials.gov and EudraCT, two required tables on on-treatment adverse events which are not serious adverse events with an incidence greater than 5% and on on-treatment 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 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.

3.9.2 Laboratory data

All laboratory values will be converted into SI units and the severity grade calculated using appropriate common toxicity criteria (CTCAE).

A listing of laboratory values will be provided by laboratory parameter, patient, and treatment group. A separate listing will display notable laboratory abnormalities (i.e., newly occurring CTCAE grade 3 or 4 laboratory toxicities). The frequency of laboratory abnormalities will be displayed by parameter and treatment group.

Laboratory data summaries will include all laboratory assessments collected no later than 30 days after the last treatment/exposure date. All laboratory assessments will be listed and those collected later than 30 days after the last treatment/exposure date will be flagged in the listings.

Laboratory data will be classified (by Novartis biostatistics/SAS programming) into CTC grades according to the NCI Common Terminology Criteria for Adverse Events (CTCAE) v4.03. A severity grade of 0 will be assigned when the value is within normal limits. In the case when a local laboratory normal range overlaps into the higher (i.e., non-zero) CTC grade, the laboratory value will still be taken as within normal limits and assigned a CTC grade of zero.

The following rules will be applied:

- Conflict between normal range and grade definition: because many institutions have differences for normal ranges of metabolic laboratory and hematology values, the CTCAE often uses the terms 'Upper Limit of Normal (ULN)' and 'Lower Limit of Normal (LLN)' in lieu of actual numerical values. In some cases, an institution's LLN might be beyond the range specified for a grade. In this case, the institutional limits of normal should take precedence over the CTCAE values: the laboratory value will still be taken as within normal limits and assigned a CTC grade of zero.
- For the few parameters having comparison to baseline in CTCAE grading definition (Fibrinogen, INR, Hemoglobin, Creatinine), the highest grade will be retained. In other words, in the particular case when a value is grade x as per CTC grade definition based on threshold/ranges and also grade x+1 when comparing to baseline, grade x+1 is retained.

- Grade 5 will not be used. Grade values will be integers from -4 to 4.
- For calcium, CTCAE grading is based on Corrected Calcium and not on Calcium. Corrected Calcium (CALC) will be calculated from Albumin and Calcium: Corrected Calcium (mg/dL) = Calcium (mg/dL) ‑0.8 [Albumin (g/dL)-4].
- GFR will be calculated based on creatinine (CREA parameter) collected regularly during treatment and the CTC grade will be determined for GFR.
- For creatinine clearance (CRCLCG parameter), no calculation of CTC grade will be done.

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

- Number and percentage of patients with worst post-baseline CTC grade (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.

The following listings will be produced for the laboratory data:

- Listing of patients with laboratory values outside the laboratory reference ranges with values flagged to show the corresponding CTC grades and the classifications relative to the laboratory reference ranges.
- Listing of all laboratory data with values flagged to show the corresponding CTC grades and the classifications relative to the laboratory reference ranges.

3.9.3 Vital signs

Vital sign assessments will be performed in order to characterize basic body function. The parameters expected to be collected include: height (cm), weight (kg), body temperature (°C), pulse rate (beats per minute), systolic and diastolic blood pressure (mmHg).

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
- weight: increase from baseline of ≥ 10%
- pulse rate: ≥ 120 bpm with increase from baseline ≥ 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
- weight: decrease from baseline of ≥ 10%
- pulse rate: ≤ 50 bpm with decrease from baseline ≥ 15 bpm.

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

- summary statistics for change from baseline to the worst post-baseline value (in both directions, i.e., from baseline to the highest post-baseline and from baseline to the 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 group:

- 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/exposure date will be flagged.

3.9.4 Other safety data

Data from other tests (e.g., electrocardiogram, pulmonary function tests, LVEF) 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/exposure 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.

3.11 Interim analysis

One efficacy interim analysis of PFS was conducted to allow early termination of the everolimus monotherapy arm, in case of far inferior efficacy as compared to the everolimus + exemestane combination treatment arm. The efficacy interim analysis was planned after 75 PFS events have been observed as per local tumor assessment, across the following 2 arms: everolimus monotherapy and everolimus + exemestane combination arm.

Based on simulations, the probability of stopping the everolimus monotherapy arm was calculated.

Table 3-2 Probability (from 2000 simulated trials) of stopping a single agent arm under different truth

True hazard ratio	Probability of stopping everolimus monotherapy arm at 75 PFS events
0.1	99.30%
0.2	48.75%
0.3	6.45%
0.4	0.5%
0.5	0%

A general guidance was to stop the single agent arm if the observed hazard ratio was less than 0.20, (i.e., if the everolimus monotherapy arm was far inferior when compared to the everolimus + exemestane combination arm). The proposed decision guidance yielded high probability to stop the control arm if the combination arm is highly superior (HR $\leq$0.1), while keeping the probability low if the superiority was not so extreme (HR$\geq$0.3).

At the time of interim analysis, the observed hazard ratio along with the 90% confidence interval was provided for decision making. Simulation (L.J. Wei, 2007) was also carried out to predict the hazard ratio and 90% confidence interval at the final analysis, based on the data observed at interim.

The predicted hazard ratios and 90% confidence intervals were provided by the independent statistician to independent programmer as per following process:

- Trial statistician created R program that was used to calculate predicted hazard ratios and 90% confidence intervals and put this program into GPS under source control in the unrestricted area.
- Independent statistician imported this R program in the restricted area, put it under source control and used it to produce a csv file in the restricted area containing the predicted hazard ratios and 90% confidence intervals.
- Independent programmer used the csv file created by independent statistician to produce outputs displaying the predicted hazard ratios and 90% confidence intervals.

Based on the interim analysis results and DMC recommendation, the study continues without change.

3.12 Sample size calculation

The primary objective of this study is to estimate the hazard ratio of PFS comparing everolimus + exemestane versus everolimus alone with approximately 150 PFS events. For this number of PFS events, the precision of HR estimation is illustrated by tabulating the approximate 90% confidence intervals (Jennison and Turnbull 1999) for the hazard ratio (HR) (see Table 3-3) under different point estimates for the HR.

Table 3-3 Approximate* 90 percent CI bounds for HR

	Assuming 146 observed PFS events		Assuming 150 observed PFS events	
Estimated HR	**Lower bound of approximate 90% CI for HR**	**Upper bound of approximate 90% CI for HR**	**Lower bound of approximate 90% CI for HR**	**Upper bound of approximate 90% CI for HR**
0.55	0.419	0.722	0.420	0.719
0.60	0.457	0.788	0.459	0.785
0.65	0.495	0.853	0.497	0.850
0.70	0.533	0.919	0.535	0.916
0.75	0.571	0.985	0.573	0.981
* Jennison and Turnbull (1999)				

A total of 300 patients are planned to be recruited at a uniform rate over an 18-month enrollment period and randomized with equal allocation to one of the three treatment arms. Assuming the median PFS time to be 7 months in the everolimus + exemestane arm (Baselga et al 2012), 4 months in the everolimus monotherapy arm (NCI-Canada), and 6 months in the capecitabine arm (O'Shaughnessy et al 2012, Stocker et al 2007, Jäger et al 2010, Kaufmann et al 2010, Robert 2011), the expected time to observe 150 PFS events in each of the two pairwise treatment comparisons is about 28 months after the randomization date of the first patient in the study, assuming that about 10% of the patients will be lost to follow-up or withdraw consent.

3.13 Statistical outputs

Tables, figures and listings will be generated as described in [CRAD001Y2201 RAP module 7 – CSR deliverables].

4 Details of the statistical analysis

4.1 Baseline comparability

Appropriate descriptive statistics of baseline variables 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 presented by treatment group, 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 in the ECOG score by one category of the score from baseline
- Time to definitive 10% deterioration from baseline in the global health status / QoL score of the EORTC QLQ-C30 questionnaire
- Time to definitive 5-point deterioration from baseline in the global health status / QoL score of the EORTC QLQ-C30 questionnaire
- Time to definitive 10-point deterioration from baseline in the global health status / QoL score of the EORTC QLQ-C30 questionnaire
- Time to definitive 10% deterioration from baseline in each of the three secondary scales (physical functioning, emotional functioning, and social functioning) of the EORTC QLQ-C30 questionnaire

4.2.1 Analysis of time-to-event data with ties

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

4.2.2 Kaplan-Meier survival function estimation

The survival function in each treatment group will be estimated using the Kaplan-Meier (product-limit) method as implemented in PROC LIFETEST (see Figure 4-2 below). In each treatment group, the estimated median PFS time, along with the approximate 90% confidence interval, will be obtained from the PROC LIFETEST output. The log-log transformation option available within PROC LIFETEST will be used to compute the confidence intervals. The Kaplan-Meier graphs will be obtained from the SAS software. The hazard ratio point and interval estimates will be displayed in the figures.

4.2.3 Hazard ratio estimation

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 and key secondary analyses, i.e., the MODEL statement will include the indicator of assignment to the everolimus + exemestane arm as the only covariate, and the STRATA statement will include the stratification variable obtained through IRT.

General SAS code for the stratified Cox model

```
PROC PHREG data=dataset;
MODEL survtime*censor(1)=trt / TIES=EXACT;
STRATA stratum 1 .. <stratum k>;
```

```
RUN;
```

/* *survtime* represents variable containing event/censoring times;

censor represents censoring variable (1=censored, 0=event);

trt represents treatment group variable;

stratum1 to *stratumk* represent stratification variables */

The two-sided 90% asymptotic confidence interval for the hazard ratio will be based on ***the Wald test***.

4.3 Median follow-up of the study

Median study follow-up (in months) in this study will be calculated as

([analysis cut-off date] - [median randomization date] + 1)/30.4375, where 30.4375 is the average duration of a month in days: 365.25/12 ≈ 30.4375.

The median randomization date is obtained by first sorting all patients in the FAS by the randomization dates, respectively, and then taking the date of the median patient (i.e., the patient in the middle of the sorted list in case of an odd number of patients or the average between the two patients in the middle of the sorted list in case of an even number of patients).

The time from last contact date to data cut-off date will be summarized by time intervals in increments of 6 weeks.

The number of patients at risk of death and the number of deaths in intervals of time with increments of 6 weeks will be summarized during the course of the trial using the life table method. The Kaplan-Meier estimates of survival probabilities and associated standard errors will be provided.

5 References

Baselga J, Campone M, Piccart M, et al. (2012) Everolimus in postmenopausal hormone-receptor-positive advanced breast cancer. N Engl J Med. 366:520-9.

O'Shaughnessy JA, Kaufmann M, Siedentopf F, et al. (2012) Capecitabine monotherapy: review of studies in first-line HER-2-negative metastatic breast cancer. ncologist. 17:476-84.

Robert NJ, Diéras V, Glaspy J, et al (2011) RIBBON-1: randomized, double-blind, placebo-controlled, phase III trial of chemotherapy with or without bevacizumab for first-line treatment of human epidermal growth factor receptor 2-negative, locally recurrent or metastatic breast cancer. J Clin Oncol. 29(10):1252-60.

Jäger E, Al-Batran S, Saupe S, et al. (2010) A randomized phase III study evaluating pegylated liposomal doxorubicin (PLD) versus capecitabine (CAP) as first-line therapy for metastatic breast cancer (MBC): Results of the PELICAN study. J Clin Oncol 28:15s (suppl; abstr 1022).

Kaufmann M, Maass N, Costa SD, et al (2010) First-line therapy with moderate dose capecitabine in metastatic breast cancer is safe and active: results of the MONICA trial. Eur J Cancer. 46:3184-91.

Stockler M, Sourjina T, Grimison P, et al (2007) A randomized trial of capecitabine (C) given intermittently (IC) rather than continuously (CC) compared to classical CMF as first-line chemotherapy for advanced breast cancer (ABC). J Clin Oncol, 25:18s (2007 ASCO Annual Meeting Proceedings Part I., 2007: 1031)

Jenninson, C, Turnbull, BW: Group sequential methods with applications to clinical trials. Chapman and Hall/CRC. 2000

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