NOVARTIS

Clinical Development

RAD001/Everolimus

CRAD001Y24135 / NCT01698918

An open-label, phase II, single-arm study of everolimus in combination with letrozole in the treatment of postmenopausal women with estrogen receptor positive HER2 negative metastatic or locally advanced breast cancer

Statistical Analysis Plan (SAP)
to support the Clinical Study Report (CSR) of the Extension

Author:	Lead Statistician, [redacted] Trial Statistician, [redacted]

Document type:	SAP Documentation
Document status:	Final 1.0
Release date:	21-JUN-2018
Number of pages:	25

Template Document History – Changes compared to previous version

Date	Version number	Summary of changes
21-JUN-2018	1.0	N/A - First version

Table of contents

List of Abbreviations

AE	Adverse Event
ATC	Anatomical Therapeutic Chemical (ATC) Classification System
AUC	Area Under the Curve
[redacted]	[redacted]
BMI	Body Mass Index
BOR	Best Overall Response
BP	Blood Pressure
CBR	Clinical Benefit Rate
CIs	Confidence Intervals
CR	Complete Response
CRF	Case Report/Record Form; the term CRF can be applied to either EDC or paper
CRO	Contract Research Organization
CSR	Clinical Study Report
CT	Computed tomography
CTC	Common Terminology Criteria
CTCAE	Common Terminology Criteria for Adverse Events
[redacted]	[redacted]
CYP3A4	Cytochrome P450 3A4
DI	Dose Intensity
ECG	Electrocardiogram
ECOG	Eastern Cooperative Oncology Group
eCRF	Electronic Case Report Form
ER	Estrogen Receptor
FAS-2L	Second line Efficacy Analyses a Subset of the FAS
FAS	Full Analysis Set
HER-2	Human Epidermal Growth Factor Receptor 2
IHC	Immunohistochemistry
MedDRA	Medical Dictionary for Regulatory Activities
MRI	Magnetic Resonance Imaging
mTOR	Mammalian Target Of Rapamycin
o.d.	omnia die/once a day
ORR	Overall Response Rate
OS	Overall Survival
OSDQ	Oral Stomatitis Daily Questionnaire
PD	Progressive Disease
PFS	Progression Free Survival
PgP	P-glycoprotein
[redacted]	[redacted]
PR	Partial Response
PS	Performance Status
PT_TXT	Preferred Term
Q1	the 25th Percentile
Q3	the 75th Percentile

RDI	Relative Dose Intensity
RAP	Report Analysis Preparation
RECIST	Response Evaluation Criteria in Solid Tumors
SD	Stable Disease
SD	Standard Deviation
SOC_TXT	Body System/Primary System Organ Class
UNK	Unknown
WHO	World Health Organization

1 Introduction

This statistical analysis plan (SAP) describes the detailed statistical methodology and analyses for the Extension Phase of the Clinical Study Report (CSR) of study CRAD001Y24135, an open-label, phase II, single-arm study of everolimus in combination with letrozole in the treatment of postmenopausal women with estrogen receptor positive HER2 negative metastatic or locally advanced breast cancer.

The content of this SAP is based on protocol CRAD001Y24135 Amended Protocol Version 5. All decisions regarding final analysis, as defined in the SAP document, have been made prior to database lock of the study data.

1.1 Study design

This is an open-label, phase II, multicenter, international, single-arm trial for patients with ER+, HER2- metastatic or locally advanced unresectable breast cancer. Enrolled patients will receive everolimus in combination with letrozole in the first line setting until disease progression or any other reason for which the patient may be discontinued.

The study counts 2 phases:

- **Core Phase:** from FPFV to 24 months after LPFV and upon approval of amendment 5; safety and efficacy data (including PROs in selected subset) will be collected.
- **Extension Phase:** from the end of the Core Phase (upon approval of amendment 5) to LPLV of the Extension Phase for patients that are still deriving benefit at the end of the Core Phase. Patients may be transitioned to the Extension Phase and continue to receive the drugs up to 3 years or until disease progression or for any other reason the patient may be discontinued. Only safety and clinical benefit as assessed by investigator will be collected.

During the Core Phase following disease progression in the first line setting, patients will be offered everolimus in combination with exemestane. Patients who discontinue treatment in the first line setting due to unacceptable toxicity or due to withdrawal of consent will not be offered everolimus plus exemestane. Those patients treated in the second line setting will continue treatment until disease progression or any other reason for which the patient may be discontinued .

All patients who discontinue study treatment will have a 28 day follow up visit for safety and will be followed every 3 months thereafter for overall survival after completing treatment in either the first or second line treatment setting.

Patients still benefiting from everolimus following the overall survival cutoff and following approval of Amendment 5 may continue treatment in the Extension Phase of the study. Patients may continue on their existing line treatment (1^{st} or 2^{nd}) in the extension up to 3 years or until the following: disease progression as deemed by physician judgement or any other reason for which the patient may be discontinued. Patients entering the Extension Phase on 1^{st} line

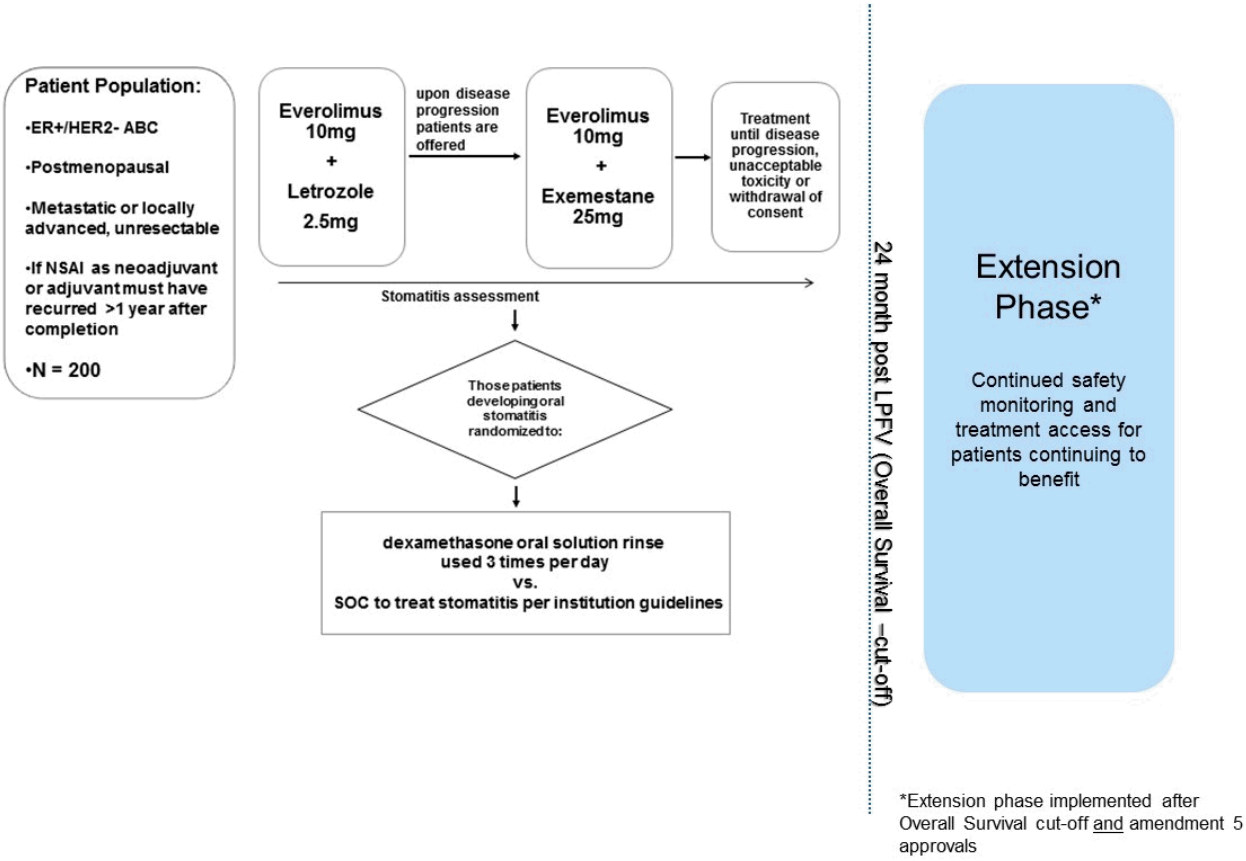
treatment and deemed to no longer be clinically benefiting will not be offered 2nd line treatment in the context of the study. Rather, these patients can start other standard treatments available or approved everolimus combination after exiting the study.

The extension will include required safety evaluations and standard of care, per investigator discretion and in consultation with the Sponsor.

The study design including Core and Extension Phases is illustrated in Figure 1.

The core phase has been reported previously (Primary analysis CSR: 31-Aug-2016; Overall Survival CSR: 23-Oct-2017), this is the SAP of the Extension phase of the study.

Figure 1 Study design



1.2 Study objectives and endpoints

The objectives and related endpoints are described in Table 1 below:

Table 1 Objectives and related endpoints

Objective	Endpoint	Phase
Primary: The primary objective is to estimate progression-free survival in patients treated with everolimus + letrozole in the first line setting.	PFS	Core
Secondary: Determine the overall response rate and clinical benefit rate of everolimus + letrozole in the first line setting	ORR, CBR	Core
Determine the progression free survival, overall response rate and clinical benefit rate of everolimus + exemestane in second line population	PFS, ORR, CBR	Core
Evaluate the safety of everolimus + letrozole in the first line setting	Incidence of Adverse events	Core
Evaluate the safety of everolimus + exemestane in the second line setting	Incidence of Adverse events	Core
Estimate the overall survival of patients treated with everolimus + letrozole in the first line setting	OS	Core
Evaluate a therapeutic intervention to reduce the severity and duration of stomatitis	Oral Stomatitis Daily Questionnaire (OSDQ) data	Core
Evaluate clinical benefit as assessed by the Investigator during the Extension Phase	**Proportion of patients with clinical benefit as assessed by the Investigator at scheduled visits**	**Extension**
Evaluate long term safety data	**Frequency and severity of AEs/SAEs**	**Extension**
[redacted]		

2 Statistical methods

2.1 Data analysis general information

The final analysis for the Extension Phase will be performed by [redacted] and reviewed by Novartis. SAS version 9.4 or later will be used to perform all data analyses and to generate tables and listings.

Data included in the analysis

During the Extension Phase, there will be no efficacy assessments other than the physician's determination of whether or not the patient is continuing to clinically benefit from the study treatment. Safety assessments will consist of monitoring and recording all adverse (AEs), including serious adverse events (SAE). All statistical analyses will be performed using all data collected in the database up to the data cut-off date. All data with an assessment date or event start date (e.g. vital sign assessment date or start date of an adverse event) prior to or on the cut-off date will be included in the analysis. Any data collected beyond the cut-off date will not be included in the analysis and will not be used for any derivations.

All events with start date before or on the cut-off date and end date after the cut-off date will be reported as 'ongoing'. The same rule will be applied to events starting before or on the cut-off date and not having documented end date. This approach applies, in particular, to adverse event and concomitant medication reports. For these cases, the end date will not be imputed and therefore will not appear in the listings.

General analysis conventions

Pooling of centers: Unless specified otherwise, data from all study centers will be pooled for the analysis. Due to expected small number of patients enrolled at centers, no center effect will be assessed.
Qualitative data (e.g., gender, race, etc.) will be summarized by means of contingency tables by first- and second-line therapy; a missing category will be included as applicable. Percentages will be calculated using the number of patients in the relevant population or subgroup as the denominator.
Quantitative data (e.g., age, body weight, etc.) will be summarized by appropriate descriptive statistics (i.e. mean, standard deviation, median, minimum, and maximum) by first- and second-line therapy.

2.1.1 General definitions

Investigational drug and study treatment

Investigational drug, will refer to everolimus only. Whereas, ***study treatment*** will refer to everolimus + letrozole and everolimus + exemestane.

The term investigational treatment may also be referred to as ***study treatment*** which is used throughout this document.

Date of first administration of investigational drug

The date of first administration of investigational drug is derived as the first date when a nonzero dose of investigational drug was administered and recorded on the dose administration DAR eCRF in the first line setting and second line setting respectively.

Date of last administration of investigational drug

The date of last administration of investigational drug is defined as the last date when a nonzero dose of investigational drug was administered and recorded on the DAR eCRF in the first line setting and the second line setting respectively.

Date of first administration of study treatment

The date of first administration of study treatment is derived as the first date when a nonzero dose of any component of study treatment was administered as per the Dosage Administration (e)CRF. (Example: if 1st dose of Everolimus (RAD001) is administered on 05-Jan-2015, and 1st dose of combination partner (Letrozole) is administered on 03-Jan-2015, then the date of first administration of study treatment is on 03-Jan-2015).

Date of last administration of study treatment

The date of last administration of study treatment is derived as the last date when a nonzero dose of any component of study treatment was administered as per Dose Administration (e)CRF. (Example: if the last Everolimus (RAD001) dose is administered on 15-Apr-2014, and the last dose of a combination partner (Letrozole) is administered on 17-Apr-2014, then the date of last administration of study treatment is on 17-Apr-2014).

Study day

The study day, describes the day of the event or assessment date, relative to the reference start date.

The study day is defined as:

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

The reference date for all assessments (safety, efficacy, pk, QoL/PRO, etc.) is the start of study treatment.

The study day will be displayed in the data listings. If an event starts before the reference start date, the study day displayed on the listing will be negative.

Time unit

A year length is defined as 365.25 days. A month length is 30.4375 days (365.25/12). If duration is reported in months, duration in days will be divided by 30.4375. If duration is reported in years, duration in days will be divided by 365.25.

Baseline

For safety and efficacy evaluations, the last available assessment on or before the date of start of study treatment is defined as "baseline" assessment.

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

On-treatment assessment/event and observation periods

For adverse event reporting the overall observation period will be divided into three mutually exclusive segments:

1. ***pre-treatment period***: from day of patient's informed consent to the day before first administration of study treatment
2. ***on-treatment period***: from date of first administration of study treatment to 28 days after date of last actual administration of any study treatment (including start and stop date)
3. ***post-treatment period***: starting at day 29 after last administration of study treatment.

Safety summaries (tables, figures) include only data from the on-treatment period with the exception of baseline data which will also be summarized where appropriate (e.g. change from baseline summaries). In addition, a separate summary for death including on treatment and post treatment deaths will be provided. In particular, summary tables for adverse events (AEs) will summarize only on-treatment events, with a start date during the on-treatment period (***treatment-emergent*** AEs).

However, all safety data (including those from the post-treatment period) will be listed and those collected during the pre-treatment and post-treatment period will be flagged.

Windows for multiple assessments

Not applicable.

Last contact date

Not applicable.

2.2 Analysis sets

Full Analysis Set

The Full Analysis Set (FAS) consists of all patients to whom the first line study treatment has been assigned. All primary efficacy analyses will be evaluated based on data from this population.

For second line efficacy analyses a subset of the FAS will be used and referred to as Full

Analysis Set 2nd line (FAS-2L), which consists of all patients in the FAS who received at least one dose of 2nd line study medication.

Safety

The Safety set consists of all patients who received at least one dose of first line study medication and had at least one post-baseline safety assessment.

The safety analysis for the second line will be performed on the subset of Safety set patients receiving at least one dose of second line medication. This subset will be referred to as Safety set 2nd line (Safety-2L).

Patients will be analyzed according to the study treatment actually received.

Patient Classification:

Patients may be excluded from the analysis populations defined above based on the protocol deviations entered in the database and/or on specific subject classification rules defined in

Table 2.

Table 2 Subject classification based on protocol deviations and non-PD criteria

Analysis sets	Protocol deviations leading to exclusion	Non protocol deviation leading to exclusion
FAS /FAS 2L	No written inform consent	Not applicable
Safety set/ Safety set 2L	No written inform consent	No dose of study medication

Withdrawal of Informed Consent

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

2.2.1 Subgroup of interest

For the extension phase no subgroup analyses will be performed. For the detail of the subgroup analysis on safety and efficacy on the core phase data please refer to SAP amendment 2 (RAP Module 3 Am. 2, 4-MAR-2016) section 2.9.1 and section 2.9.2 respectively.

2.3 Patient disposition, demographics and other baseline characteristics

The Full Analysis Set (FAS) will be used for all baseline and demographic summaries and listings. Summaries will be reported by first (FAS) and second-line (FAS 2L) therapy.

Basic demographic and background data

Baseline demographics and disease characteristics data will be listed and summarized for both the first line and the second line using the FAS and FAS-2L.

All demographic and background disease characteristics data will be listed in detail. Qualitative data (i.e., gender, race, ECOG performance status, etc.) will be summarized by means of contingency tables by the first line and the second line setting and quantitative data (i.e., age, body weight, etc.) will be summarized by appropriate descriptive statistics (frequency, mean, standard deviation, median, minimum, and maximum) by the first line and the second line.

Diagnosis and extent of cancer

Summary statistics will be tabulated for diagnosis and extent of cancer. According to the data collected on the eCRF, this analysis will include the following: primary site of cancer, details of tumor histology/cytology, histological grade, time since initial diagnosis, date of first recurrence/metastasis, date of most recent recurrence/metastasis, current disease status, number and type of metastatic sites in the current extent of disease, 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 the data collected on the radiology RECIST eCRFs, 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.

Medical history

Reported for Core Phase. Not applicable for the Extension Phase.

2.3.1 Patient disposition

The number (%) of patients in the FAS who are still on treatment, who discontinued the study phases and the reason for discontinuation will be presented:.

The following summaries will be provided (with % based on the total number of FAS patients):

- Number (%) of patients who are still on-treatment (based on the absence of the 'End of Treatment – first line of treatment' page);
- Number (%) of patients who discontinued the study treatment (based on the 'End of Treatment of first line of treatment' page)
- Reasons for study treatment discontinuation (based on 'End of Treatment of first line of treatment' page).
- Number (%) of patients who are still on-treatment (based on the absence of the 'End of Treatment – second line of treatment' page);

- Number (%) of patients who discontinued the study treatment (based on the 'End of Treatment of second line of treatment' page)
- Reasons for study treatment discontinuation (based on 'End of Treatment of second line of treatment' page).
- Number (%) of patients who completed from the post-treatment evaluations (based on 'Study Evaluation Completion' page);
- Number (%) of patients who discontinued from the post-treatment evaluations (based on 'Study Evaluation Completion' page);
- Reasons for discontinuation from the post-treatment evaluations phase (based on 'Study Evaluation Completion' page).

Protocol deviations

All protocol deviations will be listed.

Analysis sets

The number (%) of patients in each analysis set (defined in Section 2.2) will be summarized by the first line and the second line.

2.4 Treatments (study treatment, rescue medication, concomitant therapies, compliance)

2.4.1 Study treatment / compliance

Duration of study treatment exposure, cumulative dose, dose intensity (DI) and relative dose intensity (RDI) will be summarized separately for each study treatment both in first line (everolimus and letrozole) and second line treatment (everolimus and exemestane). 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 safety and safety-2L sets will be used for all summaries and listings of study treatment in the first line setting and the second line setting, respectively.

Duration of exposure to study treatment

The following algorithm will be used 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.

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.

The last date of exposure to study treatment is the latest of the last dates of exposure to investigational drug or any combination partner (see Section 2.1.1).

Summary of duration of exposure of study treatment in appropriate time units will include categorical summaries (<2 months, 2-<4 months, 4-<6 months etc.) and continuous summaries (i.e. mean, standard deviation etc.) using appropriate units of time.

Duration of exposure to investigational drug and combination partner

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

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

Summary of duration of exposure to investigational drug or combination partner in appropriate time units will include categorical summaries (<2 months, 2-<4 months, 4-<6 months etc.) and continuous summaries (i.e. mean, standard deviation etc.) using appropriate units of time.

Cumulative dose

Cumulative dose of a study treatment is defined as the total dose given during the study treatment exposure and will be summarized for each of the study treatment components (Everolimus or Letrozole for first line therapy, Everolimus or Exemestane for second line therapy).

The **planned cumulative dose** for a study treatment component refers to the total planned dose as per the protocol up to the last date of investigational drug administration.

The **actual cumulative dose** refers to the total actual dose administered, over the duration for which the subject is on the study treatment as documented in the Dose Administration eCRF.

For patients who did not take any drug the cumulative dose is by definition equal to zero.

For continuous dosing, the actual cumulative dose is the sum of the non-zero doses recorded over the dosing period and the planned cumulative dose is the planned starting dose summed over the same dosing period.

For intermittent dosing, the actual cumulative dose should be defined based on the days when the subject is assumed to have taken a non-zero dose during dosing periods.

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.

Relative dose intensity (RDI), expressed in %, is defined as follows:

RDI = [DI (dosing unit / unit of time) / PDI (dosing unit / unit of time)]*100.

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 each of the components.

2.4.2 Prior, concomitant and post therapies

Not applicable.

2.5 Analysis of the primary objective

The primary objective is to estimate the progression-free survival in patients treated with everolimus + letrozole in the first line setting.

2.5.1 Primary endpoint

The primary efficacy endpoint in this study is PFS, defined as the time from the date of enrollment to the date of first documented progression or death due to any cause. If a patient has not had an event, PFS will be censored at the date of the last adequate tumor assessment. The primary efficacy endpoint will be assessed as per local radiological assessment.

Please refer to section 2.6.1 of the SAP for the Core Phase (RAP Module 3 Am. 2, 4-MAR-2016).

2.5.2 Statistical hypothesis, model, and method of analysis

Please refer to section 2.6.1 of the SAP for the Core Phase (RAP Module 3 Am. 2, 4-MAR-2016).

2.5.3 Handling of missing values/censoring/discontinuations

Please refer to section 2.6.1 of the SAP for the Core Phase (RAP Module 3 Am. 2, 4-MAR-2016).

2.5.4 Supportive analyses

Not applicable.

2.6 Analysis of secondary efficacy objective(s)

Table 1 listed all secondary objectives and secondary efficacy endpoints. This SAP will only discuss of the secondary efficacy objectives related to the Extension Phase, which are described as follows:

- Evaluate clinical benefit as assessed by the investigator during the Extension Phase

All other secondary objectives and secondary endpoints have been detailed in Section 2.6.2 of the SAP for the Core Phase (RAP Module 3 Am. 2, 4-MAR-2016).

2.6.1 Secondary endpoints

Proportion of patients with clinical benefit

Proportion of patients with clinical benefit as assessed by the Investigator will be summarized at scheduled visits for the first line treatment on FAS population and the second line treatment on the FAS-2L population.

2.6.2 Statistical hypothesis, model, and method of analysis

Not applicable.

2.6.3 Handling of missing values/censoring/discontinuations

Not applicable.

2.7 Safety analyses

All safety analyses will be based on the safety set 1st and 2nd Line (Section 2.2).

2.7.1 Adverse events (AEs)

AE summaries will include all AEs occurring during on treatment period. All AEs collected in the AE (e)CRF page will be listed along with the information collected on those AEs e.g. AE relationship to study drug, AE outcome etc. AEs with start date outside of on-treatment period will be flagged in the listings.

AEs will be summarized by number and percentage of subjects having at least one AE, having at least one AE in each primary system organ class (SOC) and for each preferred term (PT) using MedDRA coding. A subject with multiple occurrences of an AE will be counted only once in the respective AE category. A subject with multiple CTCAE grades for the same preferred term will be summarized under the maximum CTCAE grade recorded for the event. AE with missing CTCAE grade will be included in the 'All grades' column of the summary tables.

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

The following adverse event summaries will be produced by first line or second line therapy; overview of adverse events and deaths (number and % of subjects who died, with any AE, any SAE, any AE leading to treatment discontinuation), AEs by SOC and PT, summarized by relationship (all AEs and AEs related to study treatment), seriousness (SAEs and non-SAEs), leading to treatment discontinuation..

First-line treatment-emergent adverse events are defined by all AEs reported with a start date within the first-line safety time window defined in Section 2.1.1. This will capture both new events as well as events which worsened since baseline during the first-line treatment period because any change in the severity or seriousness of an ongoing AE is recorded as a new AE in the case report form.

Second-line treatment-emergent adverse events are defined by all AEs reported with a start date within the second-line safety time window defined in Section 2.1.1. This will capture both new events as well as events which worsened since baseline during the second-line treatment period because any change in the severity or seriousness of an ongoing AE is recorded as a new AE in the case report form.

2.7.1.1 Adverse events of special interest / grouping of AEs

Not applicable.

2.7.2 Deaths

Separate summaries for on-treatment will be produced for the 1st and 2nd line, and all deaths (including post-treatment death) will be reported by system organ class and preferred term.

All deaths will be listed, post treatment deaths will be flagged. A separate listing of deaths prior to starting treatment will be provided for all enrolled subjects.

2.7.3 Laboratory data

Not applicable.

2.7.4 Other safety data

Not applicable.

2.8 Pharmacokinetic endpoints

Not applicable.

2.9 PD and PK/PD analyses

Not applicable.

2.10 Patient-reported outcomes

Not applicable.

[redacted]

2.13 Interim analysis

Not applicable.

3 Sample size calculation

The sample size is calculated for the core phase and based on an estimate of median PFS with reasonable accuracy (width of 95% confidence interval) for first line treatment with everolimus in combination with anastrozole or letrozole. The progression-free survival (PFS) for the population of patients treated with anastrozole or letrozole alone as first line therapy is approximately 9 months (Mouridsen et al, 2003, Bonneterre et al, 2001). Combining everolimus, the median PFS is expected to increase to 11 – 14 months. Considering a recruitment period of 18 months (1.5 years) and one year of follow up after the last patient is enrolled the expected 95% CIs for median PFS for 200 patients with 10% lost to follow-up, are provided in the table below for median PFSs of 11, 12, 13 and 14 months.

Median PFS	**Expected 95% CI**	**95% CI width**
11 months	9.32, 12.98	3.66
12 months	10.13, 14.22	4.09
13 months	10.93, 15.46	4.53
14 months	11.73, 16.71	4.98

4 Change to protocol specified analyses

No changes were made from protocol version 5 and the current SAP.

5 Appendix

5.1 Imputation rules

5.1.1 Study drug

The following rule should be used for the imputation of the dose end date for a given study treatment component:

If the dose end date is completely or partially missing and the EOT page is available:

Case 1: The dose end date is completely missing, and the EOT completion date is complete, then this latter date should be used.

Case 2: Only Year(yyyy) of the dose end date is available and yyyy < the year of EOT date:

Use Dec31yyyy

Case 3: Only Year(yyyy) of the dose end date is available and yyyy = the year of EOT date:

Use EOT date

Case 4: Both Year(yyyy) and Month (mm) are available for dose end date, and yyyy = the year of EOT date and mm < the month of EOT date:

Use last day of the Month (mm)

All other cases should be considered as a data issue and the statistician should contact the data manager of the study.

After imputation, compare the imputed date with start date of treatment, if the imputed date is < start date of treatment:

Use the treatment start date

Patients with missing start dates are to be considered missing for all study treatment component related calculations and no imputation will be made. If start date is missing then end-date should not be imputed.

5.1.2 AEdate imputation

Table 5-1 Imputation of start dates (AE)

Missing Element	Rule
day, month, and year	• No imputation will be done for completely missing dates

day, month	• If available year = year of study treatment start date then
	  ○ If stop date contains a full date and stop date is earlier than study treatment start date then set start date = 01JanYYYY
	  ○ Else set start date = study treatment start date.
	• If available year > year of study treatment start date then 01JanYYYY
	• If available year < year of study treatment start date then 01JulYYYY

Missing Element	Rule
day	• If available month and year = month and year of study treatment start date then o If stop date contains a full date and stop date is earlier than study treatment start date then set start date= 01MONYYYY. o Else set start date = study treatment start date. • If available month and year > month and year of study treatment start date then 01MONYYYY • If available month and year < month year of study treatment start date then 15MONYYYY

Table 5-2 Imputation of end dates (AE)

Missing Element	Rule (*=last treatment date plus <30> days not > (death date, cut-off date, withdrawal of consent date))
day, month, and year	• Completely missing end dates (incl. ongoing events) will be imputed by the end date of the on-treatment period*
day, month	• If partial end date contains year only, set end date = earliest of 31DecYYYY or end date of the on-treatment period *
day	• If partial end date contains month and year, set end date = earliest of last day of the month or end date of the on-treatment period*

Any AEs with partial/missing dates will be displayed as such in the data listings.

Any AEs which are continuing as per data cut-off will be shown as 'ongoing' rather than the end date provided.

5.2 Other imputations

Incomplete date of initial diagnosis of cancer

Missing day is defaulted to the 15th of the month and missing month and day is defaulted to 01-Jan.

5.3 AEs coding/grading

Adverse events are coded using the Medical dictionary for regulatory activities (MedDRA) terminology. The latest available MedDRA version at the time of the analyses would be used. The MedDRA version used would be specified in the footnote of relevant tables.

AEs will be assessed according to the Common Terminology Criteria for Adverse Events (CTCAE) version 4.0 (CTCAEv4.0).

The CTCAE represents a comprehensive grading system for reporting the acute and late effects of cancer treatments. CTCAE 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 is not necessarily twice as bad as a grade 1).

5.4 Laboratory parameters derivations

Not applicable.

5.5 Statistical models

5.5.1 Primary analysis

Not applicable.

5.5.2 Key secondary analysis

Not applicable.

6 Reference

1. Allison PD (2000). Survival analysis using the SAS system. Cary, NC, SAS Institute Inc.
2. Armitage P, McPherson CK, Rowe BC (1969): Repeated significance tests on accumulating data. J .R. Statist. Soc. A, 132, 232-244.
3. Brookmeyer, R. and Crowley, J. (1982): A Confidence Interval for the Median Survival Time, Biometrics 38, 29 - 41.
4. Clopper, C., Pearson, S (1934). The use of confidence or fiducial limits illustrated in the case of the binomial. Biometrika 26: 404-413.
5. Collet D (1994): Modeling survival data in medical research. London, Chapman & Hall
6. Dimitrienko A., Molenberghs G., Chuang-Stein Ch., Offen W. (2005). Analysis of clinical trials using SAS. SAS Institute Inc., SAS press.
7. Fleming TR, Rothmann MD, Lu HL (2009). Issues in Using Progression-Free Survival When Evaluating Oncology Products. Journal of Clinical Oncology 27, 2874-2880. Published by: European Organisation for Research and Treatment of Cancer, Brussels 2001.
8. FDA Guidelines (2005) Clinical Trial Endpoints for the Approval of Cancer Drugs
9. Biologics, April. Therasse P, Arbuck S, Eisenhauer E, et al (2000) New Guidelines to Evaluate the Response to Treatment in Solid Tumors, Journal of National Cancer

NOVARTIS

Clinical Development

RAD001/Everolimus

CRAD001Y24135 / NCT01698918

An open-label, phase II, single-arm study of everolimus in combination with letrozole in the treatment of postmenopausal women with estrogen receptor positive HER2 negative metastatic or locally advanced breast cancer

RAP Module 3 – Detailed Statistical Methodology Amendment 2

Author(s): [redacted], Trial Statistician
[redacted], CRO Statistician

Document type: RAP Documentation

Document status Final version number 1.0

Release date: 4 Mar 2016

Number of pages 46

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

Version	Date	Changes
Amendment 1	22-Jan-2016	Various edits were deemed necessary following discussion raised at Dry Run meeting. All changes are detailed below. **Section 2.2 Data Included in the analysis** Analysis of PFS second line timelines updated to match protocol Amendment 4 The following sentence was modified: The analysis of PFS for second line treatment will be performed 6 months later (18 months after the last patient's recruitment) provided that the number of patients in second line treatment is adequate for a reliable PFS analysis. Otherwise, this analysis will be performed 24 months after the last patient's recruitment. The analysis of the secondary objective, to evaluate a therapeutic intervention to reduce the severity and duration of stomatitis will be performed at the time of the primary analysis of 1st line PFS (12 months after the last patient's recruitment). In this analysis, the PRO questionnaire OSDQ data from patients who were randomized to receive either 0.5mg/5ml solution dexamethasone mouth rinse or standard of care up to the cutoff date for 1st line PFS analysis will be used. **Section 2.3 Definitions of analysis populations** The stomatitis analysis set (SAS) was removed as it was not used in the analysis. **Section 2.4 Major Protocol Deviations** The following PD was rephrase to be aligned with VAP M3. Previous wording: • Use of other anticancer agents before documented PD assessed by the investigator Current wording : • Administration of other antineoplastic medication/therapies other than the study drug during the study treatment **Section 2.6.1 Primary Efficacy** The following statement was introduced to clarify the date of enrollment: Date of enrollment is defined as the first day of study treatment. **Section 2.6.2.7.1 therapeutic intervention for stomatitis :** Following Dry Run 1, some revisions and clarifications of this analysis were deemed necessary. Categorical summaries replaced by shift tables : -shift tables will be presented to compare baseline value to the value at the end of the first episode of stomatitis. The definition of first incidence of stomatitis was clarified: -The analysis for this objective will be based on OSDQ data for the first incidence of stomatitis which will be identified using the answer

Version	Date	Changes
		provided to question 1b of the OSDQ: “When did you experience the first symptoms of mouth and throat soreness”. More information on the definition of the first incidence of stomatitis based on OSDQ data is provided in Section 3.1.5 of Appendix 16.1.9.” Censoring rules for duration of the first stomatitis incidence were updated to take into account the upper end of the relevant time window: -the patient discontinues from the study treatment with no resolution of the AE (up to the upper end of the relevant time window defined in Section 3.1.1.8 of Appendix 16.1.9) -In the absence of an end date of the stomatitis event, the censoring date applied will be the earliest from the following dates: upper end of the relevant time window defined in Section 3.1.1.8 of Appendix 16.1.9, analysis cut-off, new anticancer therapy start, death and last contact date. For duration of stomatitis based on AE CRF, the AE grouping Stomatitis/related events will be used: -As supportive analysis, the median duration will also be estimated using the duration calculated from start dates and end dates reported on the AE CRFs for the AE grouping stomatitis/related events (see Section 2.7.1.5). The same censoring rules as described above will be applied. Addition of the following sentence : -Definition of first and end dates of the first occurrence of stomatitis based on OSDQ and on AE CRFs is given in Section 3.1.5 of Appendix 16.1.9. Removal of the following sentences : -Furthermore, subgroup summaries by “Steroid” vs. “Non-steroid” use will be produced. **Section 2.7.1.4 AE summaries :** Death reporting updated to match the new BDM guidance on the reporting of deaths. Addition of the two following summaries : -All deaths during the study, by primary system organ class and preferred term (this summary will include deaths occurring during the study regardless of the treatment line and study phase) -On-treatment deaths by preferred term **Section 2.7.1.6 Time to first onset of specific AEs** Censoring rules were updated to take into account the upper end of the relevant time window: -The patient discontinues from the study drug with no event (up to the upper end of the relevant time window defined in Section 3.1.1.8 of Appendix 16.1.9) **Section 2.9.2 Efficacy** The definition of “visceral metastasis” has been clarified. Previous version: • presence of visceral metastasis (yes vs. no). Visceral refers to lung, liver, brain, pleural and peritoneal involvement

Version	Date	Changes
		Current version: • presence of visceral metastasis (yes vs. no). Visceral refers brain, pleura, pleural effusion, lung, liver, peritoneum and ascites. **Appendix 16.1.9 :** Creation of the following new sections and subsections to clarify the definition of the fist occurrence of stomatitis as well as its start and end dates: 3.1.5 Definition of the first occurrence of stomatitis 3.1.5.1 Based on OSDQ 3.1.5.2 Based on AE CRFs
Amendment 2	4-Mar-2016	Throughout the document, the terminology 'AE grouping stomatitis/related event' or 'stomatitis/related event' has been replaced with 'AESI grouping stomatitis' and 'stomatitis, respectively, as a result of the new CRS process. Stomatitis is considered as an AESI. **Section 2.6.2.7.1 Therapeutic intervention for stomatitis** Added 'last questionnaire date before' in paragraph: In the absence of an end date of the stomatitis event, the censoring date applied will be the last questionnaire date before the earliest from the following dates: upper end of the relevant time window defined in Section 3.1.1.8 of Appendix 16.1.9, analysis cut-off, new anticancer therapy start, death and last contact date. Text added: In the absence of an end date of the stomatitis event, the censoring date applied will be the earliest from the following dates: upper end of the relevant time window defined in Section 3.1.1.8 of Appendix 16.1.9, analysis cut-off, new anticancer therapy start, death and last contact date. Text removed: The same censoring rules as described above will be applied. **Section 2.7.1.4 AE Summaries** The terms 'clinically notable adverse events' was changed to 'Adverse events of special interest (AESI)'. **Section 2.7.1.5 title changed to: Adverse events of special interest** Text added: An adverse event of special interest is a grouping of adverse events that are of scientific and medical concern specific to RAD001. These groupings are defined using MedDRA terms, SMQs (standardized MedDRA queries), HGLTs (high level group terms), HLT (high level terms) and PTs (preferred terms). Customized SMQs (Novartis MedDRA queries, NMQ) may also be used. A NMQ is a customized

Version	Date	Changes
		group of search terms which defines a medical concept for which there is no official SMQ available or the available SMQ does not completely fit the need. It may include a combination of single terms and/or an existing SMQ, narrow or broad. For each specified AESI, number and percentage of subjects with at least one event of the AESI occurring during on-treatment period will be summarized. Summaries of these AESIs will be provided by treatment arm, (specifying grade, SAE, relationship, leading to treatment discontinuation, leading to dose adjustment/interruption, etc.). If sufficient number of events occurred, analysis of time to first occurrence will be applied. A listing of all grouping levels down to the MedDRA preferred terms used to define each AESI will be generated. Text deleted: Specific groupings of clinically notable adverse events will be considered and the number of patients with at least one event in each grouping will be reported. 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 groups are defined at the RAD001 project level and the latest groupings definition up-to-date at time of the analysis will be used. All notable adverse event groupings are defined through the use of Preferred Terms (PT), High Level Terms (HLT) or System Organ Classes (SOC) or through a combination of these three components. MAP Module 3 Post-text Supplement 2 v14.0 contains all the grouping terms to be used to map reported adverse events to the clinically notable adverse events groupings. This file may be updated (i.e., it is a living document) based on review of accumulating trial data. The latest version of this document available prior to database lock will be used. Note that certain adverse events may be reported within multiple groupings. **Section 2.7.1.6 title was changed to: Time to first onset of specific AESI** Text inserted: For the AE groupings stomatitis and non-infectious pneumonitis the following analysis of time to first occurrence will be considered. **Section 2.9.1 Safety** The terms ‘clinically notable adverse events’ was changed to ‘Adverse events of special interest (AESI)’.

Table of contents

List of tables

List of Abbreviations

AE	Adverse Event
ATC	Anatomical Therapeutic Chemical (ATC) Classification System
AUC	Area Under the Curve
[redacted]	[redacted]
BMI	Body Mass Index
BOR	Best Overall Response
BP	Blood Pressure
CBR	Clinical Benefit Rate
CIs	Confidence Intervals
CR	Complete Response
CRF	Case Report/Record Form; the term CRF can be applied to either EDC or paper
CRO	Contract Research Organization
CSR	Clinical Study Report
CT	Computed tomography
CTC	Common Terminology Criteria
CTCAE	Common Terminology Criteria for Adverse Events
[redacted]	[redacted]
CYP3A4	Cytochrome P450 3A4
DI	Dose Intensity
ECG	Electrocardiogram
ECOG	Eastern Cooperative Oncology Group
eCRF	Electronic Case Report Form
ER	Estrogen Receptor
FAS-2L	Second line Efficacy Analyses a Subset of the FAS
FAS	Full Analysis Set
HER-2	Human Epidermal Growth Factor Receptor 2
IHC	Immunohistochemistry
MedDRA	Medical Dictionary for Regulatory Activities
MRI	Magnetic Resonance Imaging
mTOR	Mammalian Target Of Rapamycin
o.d.	omnia die/once a day
ORR	Overall Response Rate
OS	Overall Survival
OSDQ	Oral Stomatitis Daily Questionnaire
PD	Progressive Disease
PFS	Progression Free Survival
PgP	P-glycoprotein
[redacted]	[redacted]
PR	Partial Response
PS	Performance Status
PT_TXT	Preferred Term
Q1	the 25th Percentile
Q3	the 75th Percentile

RDI	Relative Dose Intensity
RAP	Report Analysis Preparation
RECIST	Response Evaluation Criteria in Solid Tumors
SD	Stable Disease
SD	Standard Deviation
SOC_TXT	Body System/Primary System Organ Class
UNK	Unknown
WHO	World Health Organization

1 Introduction

1.1 Document content

This document has been developed according to Sponsor Protocol No. CRAD001Y24135 (Amended protocol version V02) and the corresponding case report forms (CRF).

This document provides the detailed statistical methodology of the Report Analysis Plan (RAP) of Study CRAD001Y24135: "An open-label, phase II, single-arm study of everolimus in combination with letrozole in the treatment of postmenopausal women with estrogen receptor positive metastatic breast cancer." It is structured as

- A draft of CSR Section 9.7 (Statistical methods planned in the protocol and determination of sample size)
- A draft of Appendix 16.1.9 (Documentation of statistical methods) of the CSR

The data analysis will be performed by [redacted] and reviewed by Novartis. The table/figure/listing shells of the statistical analysis plan can be found in Module 7.

1.2 References

Please refer to the following documents:

- CSR template (Full Clinical Study Report)
- The Master Analysis Plan for the project
- Guidelines for content of Statistical Appendices of the Clinical Study Report

2 CSR Section 9.7 – Statistical methods planned in the protocol and determination of sample size

2.1 CSR Section 9.7.1 – Statistical and analytical plans

The planned analysis is described in section 10 of the protocol (Appendix 16.1.1 of the CSR). Important information is given in the following sections and details are provided, as applicable, in Appendix 16.1.9 of the CSR.

Categorical data will be presented as frequencies and percentages. For continuous data, mean, standard deviation, median, minimum, and maximum will be presented unless specified in a different way.

It is planned that the data from all centers that participate in this protocol will be used.

2.2 Data Included in the analysis

An analysis cut-off date will be determined separately for the final efficacy and safety analyses in the first line setting and in the second line setting.

In this study the primary analysis of PFS for first line treatment will be performed 12 months after the last patient's recruitment. The analysis of PFS for second line treatment will be performed 6 months later (18 months after the last patient's recruitment) provided that the number of patients in second line treatment is adequate for a reliable PFS analysis. Otherwise, this analysis will be performed 24 months after the last patient's recruitment. The OS analysis will be performed 24 months after the last patient's recruitment.

The analysis of the secondary objective, to evaluate a therapeutic intervention to reduce the severity and duration of stomatitis will be performed at the time of the primary analysis of 1st line PFS (12 months after the last patient's recruitment). In this analysis, the PRO questionnaire OSDQ data from patients who were randomized to receive either 0.5mg/5ml solution dexamethasone mouth rinse or standard of care up to the cutoff date for 1st line PFS analysis will be used.

Only data with an assessment date or event start date (e.g. vital sign assessment date or start date of an adverse event) prior to or on the cut-off date will be included in all the analyses, including the final primary analysis. (Example: If the cut-off date is 15JUN2013 then an AE starting on 13JUN2013 will be reported, whereas an AE with start date on 17JUN2013 will not be reported.)

All events with start date before or on the cut-off date and end date after the cut-off date will be reported as 'continuing at the cut-off date'. The same rule will be applied to events starting before or on the cut-off date and not having documented end date. This approach applies, in particular, to adverse event and concomitant medication reports. For these events, the cut-off date will not be imputed and therefore will not appear in the listings.

If it is required to impute an end date to be able to perform a specific analysis (e.g. for a dose administration record with missing end date or end date after the cut-off date) the cut-off date needs to be imputed as an end date to allow for calculation of treatment exposure duration and dose intensity, the imputed date will be displayed and flagged in the listings.

Data included in efficacy analyses

Efficacy analyses will include all data observed in patients from FAS and FAS-2L population regardless whether it was observed on-treatment or after the study treatment discontinuation till the analysis cut-off date. In particular, the "28 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 and secondary efficacy analyses. Further details on censoring can be found in Section 2.6.1 and Appendix 16.1.9.

2.3 Definitions of analysis populations

The Full Analysis Set (FAS-population) consists of all patients to whom the first line study treatment has been assigned. All primary efficacy analyses will be evaluated based on data from this population.

For second line efficacy analyses a subset of the FAS will be used and referred to as Full Analysis Set 2nd line (FAS-2L), which consists of all patients in the FAS who received at least one dose of 2nd line study medication.

The Safety set consists of all patients who received at least one dose of first line study medication and had at least one post-baseline safety assessment.

The safety analysis for the second line will be performed on the subset of Safety set patients receiving at least one dose of second line medication. This subset will be referred to as Safety set 2nd line (Safety-2L).

The statement that a patient had no adverse events (AEs; on the adverse event eCRF), constitutes a safety assessment. Occurrence of a death constitutes a valid safety assessment as well. Patients who have received at least one dose of study drug but who have no post-treatment safety data of any kind will be excluded from the Safety Set and will be listed

The Stomatitis Analysis Set 1st line (SAS-1L) consists of patients in the FAS who filled in the OSDQ and who developed their first stomatitis event (based on OSDQ) in the first line. Similarly, the Stomatitis Analysis Set 2nd line (SAS-2L) includes all patients in the FAS-2L who filled in the OSDQ and who developed their first stomatitis event (based on OSDQ) in the second line. The first date reported in the OSDQ will be used to map the first stomatitis event either to first or second line treatment.

The **Stomatitis Randomized Set** consists of all patients in the FAS who were randomized to receive dexamethasone or standard of care to treat oral stomatitis.

2.4 Major Protocol Deviations

The following protocol deviations are considered for major deviations:

- Written informed consent not obtained
- ECOG performance status >2
- Not histologically or cytologically confirmed ER+ HER2- breast cancer
- No metastatic or locally advanced breast cancer
- Prior treatment for metastatic breast cancer
- Previous treatment with mTOR inhibitors
- Administration of other antineoplastic medication/therapies other than the study drug during the study treatment

Major protocol deviations will be summarized and tabulated. Other protocol deviations will also be identified, summarized and listed.

2.5 Subjects and treatments

2.5.1 Patient disposition

FAS and FAS-2L will be used for the patient disposition summaries. Based on the three eCRF pages ‘End of Treatment – first line of treatment’, ‘End of Treatment – second line of

treatment' and 'Study Evaluation Completion' there will be one combined summary for each treatment-line by showing:

1. Number (%) of patients who are still on-treatment (based on the absence of the 'End of Treatment – first line of treatment' page);
2. Number (%) of patients who discontinued the study treatment (based on the 'End of Treatment of first line of treatment' page)
3. Reasons for study treatment discontinuation (based on 'End of Treatment of first line of treatment' page).
4. Number (%) of patients who are still on-treatment (based on the absence of the 'End of Treatment – second line of treatment' page);
5. Number (%) of patients who discontinued the study treatment (based on the 'End of Treatment of second line of treatment' page)
6. Reasons for study treatment discontinuation (based on 'End of Treatment of second line of treatment' page).
7. Number (%) of patients who completed from the post-treatment evaluations (based on 'Study Evaluation Completion' page);
8. Number (%) of patients who discontinued from the post-treatment evaluations (based on 'Study Evaluation Completion' page);
9. Reasons for discontinuation from the post-treatment evaluations phase (based on 'Study Evaluation Completion' page).

2.5.2 Protocol deviation summaries

The number and percentage of patients in the FAS and FAS-2L with any protocol deviation will be tabulated by the deviation category (as specified in the VAP documents). All protocol deviations will be listed.

2.5.3 Groupings for analysis

The number and percentage of patients in each analysis population (definitions are provided in Section 2.3) will be summarized. The distribution of patients in selected analysis populations will also be summarized by country.

2.5.4 Background and demographic characteristics

Baseline demographics and disease characteristics data will be listed and summarized for both the first line and the second line using the FAS and FAS-2L.

2.5.4.1 Basic demographic and background data

All demographic and background disease characteristics data will be listed in detail. Qualitative data (i.e., gender, race, ECOG performance status, etc.) will be summarized by means of contingency tables by the first line and the second line setting and quantitative data (i.e., age, body weight, etc.) will be summarized by appropriate descriptive statistics (frequency, mean, standard deviation, median, minimum, and maximum) by the first line and the second line.

Baseline demographics and disease characteristics will also be summarized by treatment group for patients in the Stomatitis Randomized Set.

2.5.4.2 Diagnosis and extent of cancer

Summary statistics will be tabulated for diagnosis and extent of cancer. According to the data collected on the eCRF, this analysis will include the following: primary site of cancer, details of tumor histology/cytology, histological grade, time since initial diagnosis, date of first recurrence/metastasis, date of most recent recurrence/metastasis, current disease status, number and type of metastatic sites in the current extent of disease, 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 the data collected on the radiology RECIST eCRFs, 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.

2.5.4.3 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.

In addition, descriptive summaries of ongoing adverse events at second-line baseline by CTC maximum grade will be presented for the FAS-2L population.

2.5.4.4 Prior anti-neoplastic therapy

Prior anti-neoplastic therapy will be listed in three separate listings: (i) medications, (ii) radiotherapy, and (iii) surgery for the FAS.

The number and percentage of patients recording any prior anti-neoplastic medications, prior anti-neoplastic radiotherapy and prior anti-neoplastic surgery will be summarized.

The summary of prior anti-neoplastic medication will include the total number of regimens used, setting at last treatment, and the best response at last treatment.

All prior antineoplastic medications will also be tabulated by ATC code and preferred term.

2.5.4.5 Other

All data collected at baseline, including source of subject referral, child bearing potential [redacted] informed consent, will be listed for the FAS population.

2.5.5 Study treatment

Duration of study treatment exposure, cumulative dose, dose intensity (DI) and relative dose intensity (RDI) will be summarized separately for each study treatment both in first line (everolimus and letrozole) and second line treatment (everolimus and exemestane). 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 with dose reductions/interruptions will be presented, along with reasons for the dose changes.

The safety and safety-2L sets will be used for all summaries and listings of study treatment in the first line setting and the second line setting, respectively.

2.5.5.1 Duration of study treatment exposure

The following algorithm will be used 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.

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.

2.5.5.2 Cumulative dose

Cumulative dose is defined as the total dose given during the study treatment exposure 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.

2.5.5.3 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.

Relative dose intensity (RDI), expressed in %, is defined as follows:

RDI = [DI (dosing unit / unit of time) / PDI (dosing unit / unit of time)]*100.

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 each of the components.

2.5.5.4 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, by treatment line.

Interruption: An interruption is defined as a 0 mg / 0 tablets dose given on one or more days.

Reduction: A reduction is defined as a decrease in dose from the protocol planned dose or a decrease from the previous non-zero dose, even if this decrease has been directly preceded by an interruption. For example, in the sequence 10mg – 0mg – 5mg, the 5mg dose will be counted as a reduction. A decrease in frequency of administration which results in a lower cumulative dose is also counted as a reduction, e.g. in the sequence 10mg – 5mg – 5mg, two reductions will be counted.

If a patient moves from a higher than protocol planned dose down to the planned dose then this is not be counted as a reduction, however if they move directly from a higher than planned dose down to a lower than protocol planned dose or the planned dose on a less frequent regimen, then this is counted as a reduction.

If one of the study treatment component is permanently discontinued (before any protocol planned discontinuation date) while the other is ongoing it is counted as a reduction.

The number of dose reductions and interruptions per patient will be tabulated both separately and combined. The reasons for reductions/interruptions will also be summarized. Dose escalations should not be counted in these summaries.

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

2.5.6 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. Details regarding WHO Drug Reference List version will be included in the footnote in the corresponding tables and listings. 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 and preferred term by means of frequency counts and percentages both in first line (everolimus and letrozole) and second line treatment (everolimus and exemestane). 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.

Concomitant therapy will be descriptively analyzed by treatment period in the safety and safety-2L sets. Frequency distribution tables and listings will be prepared.

Concomitant medications which start during the first-line and are continuing in the second-line will be reported in both first-line and second-line tables.

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 (e.g. FAS and FAS-2L). Strong and moderate inhibitors, inducers, or substrates of the isoenzyme CYP3A will be identified as described in Appendix 16.1.9. The strong ones will be tabulated by ATC code. Both strong and moderate ones will be listed. Further on treatment anti-neoplastic therapies will be listed based on their identification (by the method given in Appendix 16.1.9) by the protocol deviation process.

Antineoplastic therapies administered since discontinuation of study drug will be listed and summarized by ATC class, preferred term and treatment period by means of frequency counts and percentages using the FAS and FAS-2L. Antineoplastic therapies administered after first-line treatment period in patients who did not start second-line study drug and in patients who received second-line treatment will be summarized separately.

2.6 Efficacy evaluation

The efficacy endpoints based on the tumor assessments will be derived according to the RECIST 1.0 guideline version 2 (see [Protocol Appendix 2] for details). The tumor endpoints derivation is based on the sequence of overall lesion responses at each assessment/time point. However, the overall lesion response at a given assessment/time point may be provided from different sources as illustrated in Table 2-1.

Table 2-1 Sources for overall lesion response

Source 1	Local (treating center) radiologist/investigator reported overall lesion response
Source 2	Calculated overall lesion response based on raw (i.e. individual lesion) measurements obtained from local (treating center) radiologist/Investigator

The primary efficacy analysis will be based on local (treating center) radiologist/Investigator tumor assessment data and will use Source 1 in Table 2-1 as a basis for endpoint derivation. In particular, the final local (treating center) radiologist/Investigator visit response for each assessment/time point collected on the RECIST 1.0 overall response category collection CRF page will be used to derive the primary efficacy endpoints. The assessment time point dates will be derived using the dates of the individual lesion measurements.

The overall response at each assessment will also be re-calculated using raw lesion measurements obtained from local (treating center) radiologist/investigator (Source 2, Table

2-1) and will be listed against Source 1, Table 2-1. Discrepancies between the calculated and the local radiologist/investigator reported overall responses will be identified and listed.

2.6.1 Primary Efficacy

Progression-free survival (PFS) derived from local radiological assessment will be used as the primary efficacy variable. The PFS is defined as the time from the date of enrollment to the date of the first documented disease progression or death due to any cause. Date of enrollment is defined as the first day of study treatment. If a patient has not progressed or died at the analysis cut-off date or when she receives any further anti-neoplastic therapy (including per protocol second-line treatment), PFS would be censored at the time of the last tumor assessment before the cut-off date or the anti-neoplastic therapy date. Further anti-neoplastic therapies taken during study treatment will be identified as protocol deviations (see Appendix 16.1.9), recognized from the data collected on 'Antineoplastic therapies since discontinuation of study drug' eCRF, and from the data on 'Dosage Administration Record'. Definitions and further details on PFS can be found in [Protocol Appendix 2].

Discontinuation due to disease progression (collected on the "End of treatment" and "Study Evaluation Completion" page) without supporting objective evidence (as defined in Section 3.1.3.2 of Appendix 16.1.9) satisfying progression criteria per RECIST 1.0 will not be considered a progressive disease.

Primary analysis

The primary analysis of PFS will be based on the data from investigator/local radiology review (Source 1 in Table 2-1). The analysis will be performed on the FAS and will use the default censoring and event date options from section 14.2.18 of [Protocol Appendix 2], i.e. A (1), B (1), C1 (1), C2 (1), D (1), E (1), and F (1). In particular, the PFS will be censored at the last adequate tumor assessment if one of the following occurs: absence of event before cut-off date for final primary analysis; the event occurred after a new anticancer therapy is given (including per protocol second-line treatment); the event occurred after two or more missing tumor assessments (see Section 3.1.3.5 of Appendix 16.1.9). See also Section 3.1.3.6 of Appendix 16.1.9 describing the special case of missing baseline tumor assessment. Discontinuation of study treatment (for any reason) will not be considered as a reason for censoring.

Kaplan-Meier estimates

The Kaplan-Meier estimate of the PFS survival distribution function will be computed. The results will be plotted graphically (Kaplan-Meier curves). The plots will display the number of patients at risk at equidistant time points.

The Median PFS will be provided along with the 95% confidence intervals. Additionally, the 25% and 75% percentiles will also be provided. The survival probabilities at 6, 10, 14 and 18 months, and the associate 95% confidence intervals will be summarized.

Sensitivity analyses of the primary endpoint PFS

The following sensitivity PFS analyses will be performed to address the impact of missing/unknown tumor assessments:

1. PFS using local radiology assessments (Source 1 in Table 2-1) on the FAS and using the following options from Table 14-5 of the [Protocol Appendix 2]: 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 Appendix 16.1.9). In the summary table, this approach will be referred as "Actual event PFS analysis".
2. PFS using local radiology assessments (Source 1 in Table 2-1) on the FAS and using the following options from Table 14-5 of the [Protocol Appendix 2]: 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 as "Backdating PFS analysis".

PFS censoring reasons

The number of patients with a PFS event and number of patients censored for the PFS analysis will be summarized for PFS based on local radiology assessments for the FAS. In addition, a summary of reasons for PFS censoring will be provided. The following categories will be used as appropriate:

- 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 or any other censoring reason)
- New cancer therapy added
- Event documented after two or more missing tumor assessments based on the distance D2 defined in Section 3.1.3.5 of Appendix 16.1.9

The distance D2 will also to be used to determine other PFS censoring reasons defined in Section 3.1.3.5 of Appendix 16.1.9.

For patients censored at the date of enrollment due to missing baseline tumor, the censoring reason will be "Adequate assessment no longer available".

2.6.2 Secondary Efficacy

The secondary efficacy objectives in this study are to determine the overall response rate (ORR) and clinical benefit rate (CBR) of everolimus + letrozole in the first line setting; determine the progression free survival, ORR and CBR of everolimus + exemestane in the second line setting; estimate the overall survival of patients treated with everolimus + letrozole in the first line setting; evaluate a therapeutic intervention to reduce the severity and duration of stomatitis.

2.6.2.1 Overall response rate (ORR) in first line

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 1.0 (see [Protocol Appendix 2]). ORR will be calculated based on the FAS, using local (treating center)

radiologist's/investigator's review of the tumor assessment data. However patients with only non-measurable disease at baseline will be included in the numerator if they achieve a complete response. Proportions of subjects with ORR will be presented along with exact 95% confidence intervals (Clopper and Pearson 1934).

The above analyses will also be repeated based on the data from subset of patients with only measurable disease at baseline.

2.6.2.2 Progression Free Survival in Second line

Progression-free survival in the second line (PFS-2L) is defined as the time interval between start of 2nd line treatment and documented disease progression or death due to any cause reported during or after second-line treatment. The median PFS in the second line will be estimated using the Kaplan-Meier method and presented along with 95% confidence intervals for the FAS-2L. The same censoring rules as for the PFS in the first line will be applied (see Section 2.6.1).

2.6.2.3 Overall response rate in second line

The analysis described in Section 2.6.2.1 will be repeated on FAS-2L population to estimate the ORR of the second line study treatment.

2.6.2.4 Overall survival (OS)

The overall survival (OS) following first line treatment with everolimus + letrozole is defined as the time from the date of enrollment to date of death due to any cause. The final OS analysis will be performed 24 months after the last patient's recruitment. If a death has not been observed by the date of analysis cut-off, then OS will be censored at the date of last contact (see Section 3.1.1.9 of Appendix 16.1.9).

The analysis of OS will be based on the data from the FAS population regardless of whether they were observed during the first line setting, during the second line setting, or during survival follow-up. The median OS as well as the 25^{th} and 75^{th} quartile will be estimated using the Kaplan-Meier method and presented along with 95% confidence intervals. In addition, Kaplan-Meier survival estimates with 95% confidence intervals will be summarized at 12, 18, 24 and 30 months.

2.6.2.5 Clinical benefit rate (CBR) in first line

CBR is defined as the proportion of patients with either a best overall response of CR, PR or SD lasting for 24 weeks or longer. A patient will be considered to have a SD for 24 weeks or longer if a SD response is recorded at 24 weeks or later. Taking into account the allowed visit window for tumor assessments, the SD response has to be recorded at 23 weeks or later to be included in the CBR calculation. CBR will be calculated on the FAS based on the local (treating center) radiologist's/Investigator's tumor assessment. Patients with only non-measurable disease at baseline will be included in the numerator if they achieve a complete response or stable disease lasting 24 weeks or longer. Proportions of subjects with CBR will be presented along with exact 95% confidence intervals (Clopper and Pearson 1934).

2.6.2.6 Clinical benefit rate (CBR) in second line

The analysis described in Section 2.6.2.5 will be repeated on FAS-2L population to estimate the CBR of the second-line study treatment.

2.6.2.7 Other secondary objective

2.6.2.7.1 Therapeutic intervention for stomatitis

The secondary objective, to evaluate a therapeutic intervention to reduce the severity and duration of stomatitis, will be performed using the patient reported outcome (PRO) Oral Stomatitis Daily Questionnaire (OSDQ) data from patients in the Stomatitis Randomized Set.

The analysis for this objective will be based on OSDQ data for the first incidence of stomatitis which will be identified using the answer provided to question 1b of the OSDQ: "When did you experience the first symptoms of mouth and throat soreness". More information on the definition of the first incidence of stomatitis based on OSDQ data is provided in Section 3.1.5 of Appendix 16.1.9.

The number of patients filling OSDQ and the number of patients missing OSDQ assessments as well as the reasons for missing OSQD assessments will be summarized by each treatment group for scheduled assessment time points.

Descriptive statistics (N, mean, SD, median, minimum, maximum) will be used to summarize the individual severity items of the OSDQ data at each scheduled assessment time point. Additionally, change from Day 1 at the time of each assessment will be summarized. Patients with an evaluable Day 1 score and at least one evaluable score after Day 1 during the treatment period will be included in the change from Day 1 analyses. A repeated measurements analysis model that includes terms for Day 1 score, time of visit, treatment group and time of visit by treatment group interaction will be used to compare the two treatment groups with respect to actual and change from Day 1 OSDQ severity scores longitudinally over time. To account for correlated repeated measures within patients a first-order autoregressive variance-covariance matrix will be used. Longitudinal plots at selected timepoints displaying the model adjusted, least square means and 95% confidence intervals will be produced.

For selected severity items, shift tables will be presented to compare Day 1 value to the value at the end of the first episode of stomatitis.

The time to first occurrence of stomatitis will be calculated using the dates reported in the OSDQ and is defined as time from first study treatment administration in the first line to start date of the AE, i.e. time in days is calculated as (start date of first occurrence of stomatitis) – (date of first administration of study treatment in the first line) +1. The duration of the first stomatitis incidence will also be calculated using the dates reported in the OSDQ. Duration of the first stomatitis incidence is defined as time from start date of the AE to end date of the AE, i.e. time in days is calculated as (end date of first occurrence of stomatitis) – start date of first occurrence of stomatitis) +1.

A patient will be censored for duration of the first stomatitis incidence if:

- the patient dies before resolution of the AE

- the patient receives a new anticancer therapy (excluding protocol specified second line with no event or before the event has occurred
- the patient discontinues from the study treatment with no resolution of the AE (up to the upper end of the relevant time window defined in Section 3.1.1.8 of Appendix 16.1.9)
- the AE is still on-going at the cut-off with no event

In the absence of an end date of the stomatitis event, the censoring date applied will be the last questionnaire date before the earliest from the following dates: upper end of the relevant time window defined in Section 3.1.1.8 of Appendix 16.1.9, analysis cut-off, new anticancer therapy start, death and last contact date.

Ascending Kaplan-Meier curves will be constructed. The median duration of the first stomatitis AE with 95% confidence interval will be estimated using Kaplan-Meier method for the two therapeutic interventions for stomatitis groups. Additionally, 25% and 75% percentiles will be also provided.

As supportive analysis, the median duration will also be estimated using the duration calculated from start dates and end dates reported on the AE CRFs for the AESI grouping stomatitis (see Section 2.7.1.5). In the absence of an end date of the stomatitis event, the censoring date applied will be the earliest from the following dates: upper end of the relevant time window defined in Section 3.1.1.8 of Appendix 16.1.9, analysis cut-off, new anticancer therapy start, death and last contact date.

Definition of first and end dates of the first occurrence of stomatitis based on OSDQ and on AE CRFs is given in Section 3.1.5 of Appendix 16.1.9.

These analyses will be repeated by treatment line for the SAS-1L and SAS-2L respectively.

2.6.3 Waterfall graphs

Waterfall graphs will be used to depict the anti-tumor activity. These plots will display the best percentage change from the treatment-line specific baseline in the sum of the longest diameter of all target lesions for each patient. The proportions of patients with various degrees of tumor shrinkage or growth which can be read directly from the graph can then re present a useful efficacy metric. Only patients with measurable disease at baseline and valid post baseline assessments will be included in the waterfall graphs. Construction of waterfall graphs is detailed in Appendix 16.1.9.

2.6.4 ECOG performance status

ECOG Performance Status (PS) scale will be used to assess physical health of patients, ranging from 0 (most active) to 5 (Dead):

Table 2-2 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

Score	Description
	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 screening/baseline, on Treatment Day 1 (prior to administration of the study drug, unless PS in screening was done less than 7 days before Day 1), and Day 1 of every cycle thereafter until progression.

Time based intervals will be used to group the ECOG PS data over time for each treatment line. Definition of the time intervals is provided in Appendix 16.1.9. Descriptive statistics (N, mean, SD, median, minimum and maximum) will be used to summarize the ECOG PS data at each scheduled assessment time point. Additionally, change from baseline at the time of each assessment will be summarized. Patients with an evaluable baseline score and at least one evaluable post baseline score during the treatment period will be included in the change from baseline analyses.

An analysis of the time to definitive deterioration of the ECOG PS by one category of the score from baseline (see section 3.1.4 Time interval for ECOG performance status in Appendix 16.1.9) will be performed in FAS set for the whole treatment period, i.e., not separately for the first-line and second-line settings. A deterioration is considered definitive if no improvements in the ECOG PS status is observed at a subsequent time of measurement during the treatment period following the time point where 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 is considered as worsening of 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. 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 are censored at the date of their last available questionnaire. For example, if the last assessment is at week 4 and the patient dies on week 10, the definitive deterioration date will be week 10. On the other hand, if the last assessment is at week 4 and the patient dies on week 16, which is after more than twice the planned period between two assessments (8 weeks) since the last assessment (week 4), then the patient is censored for definitive deterioration and the censoring date will be week 4.

Patients receiving any further anti-neoplastic therapy (excluding per protocol second-line treatment) prior to definitive worsening will be censored at their date of last available assessment prior to start of therapy. Patients that have not worsened at the data cut-off point will be censored at the date of last available assessment prior to data cut-off. Kaplan-Meier method will be used to estimate the survival distribution function of time to definitive worsening. The median time to definitive worsening along with a 95% confidence interval

will be presented.

2.7 Safety evaluation

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

All safety outputs will use the defined Safety sets (Section 2.3). The analyses will be performed per treatment line in the Safety and Safety-2L sets except for time to first onset of specific AEs which will be calculated for the whole treatment period.

The safety summary tables will only include "treatment-emergent" events/assessments occurring in the windows defined in Section 3.1.1.8 of Appendix 16.1.9, however, safety events/assessments outside of these windows will be listed and flagged.

2.7.1 Adverse events data

2.7.1.1 Coding of AEs

Adverse events are coded using the Medical dictionary for regulatory activities (MedDRA) terminology. Details regarding MedDRA version will be included in the footnote in the corresponding tables and listings.

2.7.1.2 Grading of AEs

AEs will be assessed according to the Common Terminology Criteria for Adverse Events (CTCAE) version 4.0 (http://evs.nci.nih.gov/ftp1/CTCAE/CTCAE_4.03_2010-06-14_QuickReference_5x7.pdf).

The CTCAE represents a comprehensive grading system for reporting the acute and late effects of cancer treatments. CTCAE v4.0 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 is not necessarily twice as bad as a Grade 1).

If CTCAE grading does not exist for an adverse event, grades 1 – 4 corresponding to the severity of mild, moderate, severe, and life-threatening will be used. CTCAE grade 5 (death) will not be used in this project; rather, this information will be collected on the "End of Treatment", "Study Completion Evaluation" or "Survival Information" eCRF pages.

2.7.1.3 General rules for AE Reporting

First-line treatment-emergent adverse events are defined by all AEs reported with a start date within the first-line safety time window defined in Section 3.1.1.8 of Appendix 16.1.9. This will capture both new events as well as events which worsened since baseline during the first-line treatment period because any change in the severity or seriousness of an ongoing AE is recorded as a new AE in the case report form.

Second-line treatment-emergent adverse events are defined by all AEs reported with a start date within the second-line safety time window defined in Section 3.1.1.8 of Appendix 16.1.9.

This will capture both new events as well as events which worsened since baseline during the second-line treatment period because any change in the severity or seriousness of an ongoing AE is recorded as a new AE in the case report form.

In the following sections, the term AE refers to treatment-emergent adverse events.

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.

2.7.1.4 AE summaries

Unless otherwise specified, the following adverse event summaries will be produced by treatment line, on the Safety set for the first-line treatment (or Safety-2L set for the second-line treatment):

- 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, regardless of study drug relationship by primary system organ class, preferred term and maximum CTC grade
- All deaths during the study, by primary system organ class and preferred term (this summary will include deaths occurring during the study regardless of the treatment line and study phase)
- On-treatment deaths by 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
- Adverse events of special interest (AESI) (Section 2.7.1.5) regardless of study drug relationship, by grouping and preferred term
- Adverse events of special interest (AESI) (Section 2.7.1.5) regardless of study drug relationship, by grouping, preferred term and maximum CTC grade
- Time to first occurrence of the adverse events of special interest (AESI): stomatitis
- Time to first occurrence of the adverse events of special interest (AESI): non-infectious pneumonitis

2.7.1.5 Adverse events of special interest

An adverse event of special interest is a grouping of adverse events that are of scientific and medical concern specific to RAD001. These groupings are defined using MedDRA terms, SMQs (standardized MedDRA queries), HGLTs (high level group terms), HLT (high level terms) and PTs (preferred terms). Customized SMQs (Novartis MedDRA queries, NMQ) may also be used. A NMQ is a customized group of search terms which defines a medical concept for which there is no official SMQ available or the available SMQ does not completely fit the need. It may include a combination of single terms and/or an existing SMQ, narrow or broad.

For each specified AESI, number and percentage of subjects with at least one event of the AESI occurring during on-treatment period will be summarized.

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

2.7.1.6 Time to first onset of specific AESI

For the AE groupings stomatitis and non-infectious pneumonitis the following analysis of time to first occurrence will be considered if sufficient number of events occurred.

All the adverse events of interest occurring within the windows defined in Section 3.1.1.8 of Appendix 16.1.9 (whole sequence of treatment) will be taken into account.

The time to the first occurrence of an AE is defined as the time from the start of study treatment to the date of the first occurrence of an adverse event, and it is calculated in days as:

(Start date of first occurrence of AE) - (date of first dose of study drug) + 1.

A patient will be censored for the time to first occurrence if:

- the patient dies with no event
- the patient receives a new anticancer therapy (excluding second-line treatment) with no event or before the event has occurred
- the patient discontinues from the study drug with no event (up to the upper end of the relevant time window defined in Section 3.1.1.8 of Appendix 16.1.9)
- the patient is still on-going at the cut-off date with no event

The censoring date applied will be the earliest among the following dates:

- upper end of the relevant time window defined in Section 3.1.1.8 of Appendix 16.1.9,
- analysis cut-off date,
- new anti-cancer therapy start date (excluding second-line treatment),
- death date or, in patients with no death data, last contact date (Section 3.1.1.9 of Appendix 16.1.9).

Kaplan-Meier curves for time to first onset of the AE will be estimated. Estimates of the median and 25th and 75th percentiles will be presented with 95% confidence intervals.

2.7.2 Laboratory data

Data from both central and local laboratories will be combined when analyzing laboratory data and will be assigned to the first or second line treatment period using the windows defined in Section 3.1.1.8 of Appendix 16.1.9. Baseline assessments in each treatment line period are defined in Section 3.1.1.7 of Appendix 16.1.9. All laboratory assessments will be listed and those collected outside of the windows defined in Appendix 16.1.9 will be flagged.

Laboratory data will be classified programmatically into CTC grades according to the NCI Common Terminology Criteria for Adverse Events (CTCAE) v4.0. 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 summaries will be produced for the laboratory data by laboratory parameter and for each treatment line, on the Safety (respectively Safety-2L) set for the first-line (respectively second-line treatment) with reference to the appropriate baseline:

- 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.

2.7.3 Vital signs

Vital sign assessments are performed in order to characterize basic body function. The parameters expected to be collected include: height (cm) (only collected in screening/baseline),

weight (kg), body temperature (°C), sitting blood pressure systolic, diastolic (mmHg) and sitting pulse (bpm).

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 of ≥15 bpm

Clinically notable below normal values

- Systolic BP: ≤ 90 mmHg and a decrease ≥ 20 mmHg from baseline
- Diastolic BP: ≤ 50 mmHg and a decrease ≥ 15 mmHg from baseline
- Body temperature: ≤ 35°C
- Weight: decrease from baseline of ≥ 10%
- Pulse rate: ≤ 50 bpm with decrease from baseline of ≥ 15 bpm

Vital signs shift table based on notable values will be produced for each vital sign parameter for each treatment line, on the Safety (respectively Safety-2L) set for the first-line (respectively second-line treatment) with reference to the appropriate baseline.

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

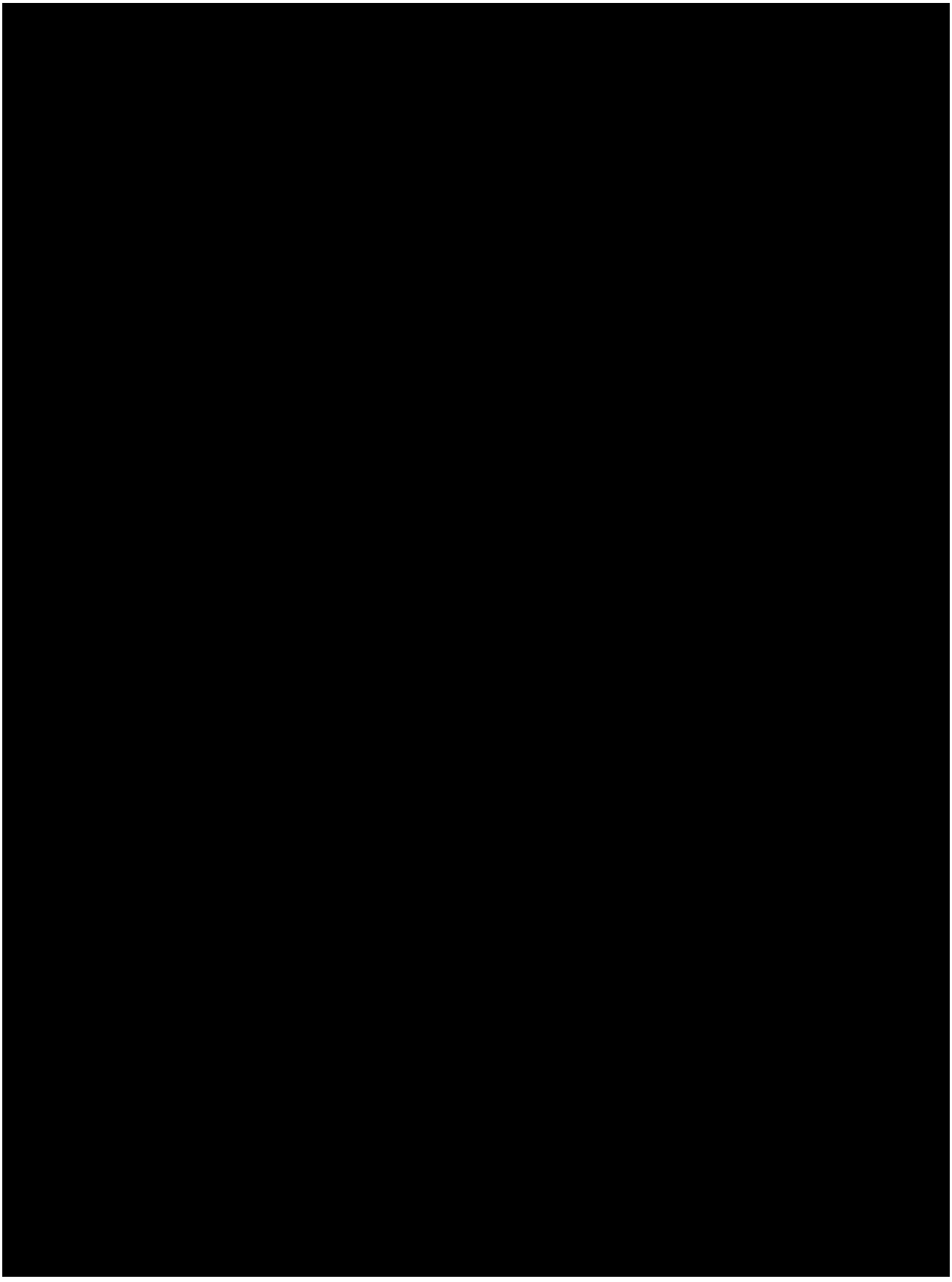
- Patients with clinically notable vital sign abnormalities.
- All vital sign assessments will be listed by patient and vital sign parameter.

In both listings, the clinically notable values will be flagged and assessments outside of the windows defined in Section 3.1.1.8 of Appendix 16.1.9 will be flagged.

2.7.4 Other safety data

Data from other tests (e.g. electrocardiogram, pulmonary function tests) will be listed, notable values will be flagged, and any other information collected will be listed as appropriate.

All assessments collected later than 28 days after the last treatment/exposure date (see Appendix 16.1.9) will be flagged in the listings.



2.9 Subgroup analyses

2.9.1 Safety

Selected first-line safety analyses will be repeated on the Safety set in the following subgroups:

- age (< 65 years and ≥ 65 years)
- race (Asian vs. non-Asian)

Selected safety analyses include:

- Adverse events, regardless of study drug relationship by primary system organ class and preferred term
- Serious adverse events, regardless of study drug relationship, by primary system organ class and preferred term
- Adverse events of special interest (AESI) regardless of study drug relationship, by grouping and preferred term

The objective for carrying out these subgroup analyses is to identify potential safety issues that are limited to a subgroup of patients, or safety issues that are more commonly observed in a subgroup of patients.

2.9.2 Efficacy

The primary PFS analysis in the first line will be repeated in the following subgroups for the FAS:

- age (< 65 years and ≥ 65 years)
- race (Asian vs. non-Asian)bone only lesions at baseline (yes vs. no)

- presence of visceral metastasis (yes vs. no).
- Visceral refers brain, pleura, pleural effusion, lung, liver, peritoneum and ascites.

2.10 Interim analyses

No interim analysis is planned.

2.11 CSR Section 9.7.2 – Sample size calculation

The sample size is calculated based on an estimate of median PFS with reasonable accuracy (width of 95% confidence interval) for first line treatment with everolimus in combination with anastrozole or letrozole. The progression-free survival (PFS) for the population of patients treated with anastrozole or letrozole alone as first line therapy is approximately 9 months (Mouridsen et al, 2003, Bonneterre et al, 2001). Combining everolimus, the median PFS is expected to increase to 11 – 14 months. Considering a recruitment period of 18 months (1.5 years) and one year of follow up after the last patient is enrolled the expected 95% CIs for median PFS for 200 patients with 10% lost to follow-up, are provided in the table below for median PFSs of 11, 12, 13 and 14 months.

Median PFS	**Expected 95% CI**	**95% CI width**
11 months	9.32, 12.98	3.66
12 months	10.13, 14.22	4.09
13 months	10.93, 15.46	4.53
14 months	11.73, 16.71	4.98

3 CSR Appendix 16.1.9 – Documentation of statistical methods

The statistical methods used to perform the analyses presented in the clinical study report will be described in Section 9.7. Section 16.1.9 will provide further details of the statistical methods not already provided in Section 9.7.

3.1 Definitions and general methodology

3.1.1 Definitions

3.1.1.1 Study drug and study treatment

Study drug is defined as everolimus.
Study treatment is defined as everolimus + letrozole and everolimus + exemestane

3.1.1.2 Date of first administration of study drug

The date of first administration of study drug is derived as the first date when a nonzero dose of study drug was administered and recorded on the dose administration DAR eCRF in the first line setting and second line setting respectively.

3.1.1.3 Date of last administration of study drug

The date of last administration of study drug is defined as the last date when a nonzero dose of study drug was administered and recorded on the DAR eCRF in the first line setting and the second line setting respectively.

3.1.1.4 Date of first administration of study treatment

The date of first administration of study treatment is derived as the first date when a nonzero dose of any component of study treatment was administered and recorded on the dose administration record (DAR) eCRF in the first line setting and the second line setting respectively.

The dates of first administration of study treatment are calculated separately for each treatment line and, for simplicity, are referred to as DFirstTrt1L and DFirstTrt2L, respectively.

3.1.1.5 Date of last administration of study treatment

The date of last administration of study treatment is derived as the last date when a nonzero dose of any component of study treatment was administered and recorded on the DAR eCRF in the first line setting and the second line setting respectively.

The dates of last administration of study drug are calculated separately for each treatment line and, for simplicity, are referred to as $D_{LastTrt}1L$ and $D_{LastTrt}2L$, respectively.

3.1.1.6 Study day

The study day for ***all assessments*** (i.e. safety and non-safety assessments) will be calculated as the difference between the date of the assessment and the start of study treatment of the

first line setting and the second line setting respectively (plus 1 if assessment occurred on or after start of study treatment). Note that the study day for OSDQ assessments will also be based on this rule. Note: if an adverse event starts before the start of study treatment the study day displayed on the listing will be negative.

3.1.1.7 Baseline

Baseline value is the result of an investigation describing the "true" uninfluenced state of the subject.

For ***all evaluations*** (i.e. safety and non-safety), the last available assessment before or at date of start of study treatment of first line setting and second line setting respectively is taken as 'baseline' assessment.

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

3.1.1.8 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:

- First-line treatment:
  - For patients who subsequently begin second-line study treatment: From the date of first administration of the first-line study treatment until the minimum between the date of last administration of the first-line study treatment plus 28 days and the date of first administration of second-line study treatment minus one day, that is, $D_{FirstTrt}1L$ until min ($D_{LastTrt}1L+28$ days, $D_{FirstTrt}2L-1$).
  - For patients who do not receive second-line study treatment: From the date of first administration of the first-line study treatment until the date of last administration of the study treatment in first-line treatment plus 28 days, that is, $D_{FirstTrt}1L$ until $D_{LastTrt}1L+28$ days.
- Second-line treatment: From the date of first administration of second-line study treatment until the date of last administration of the study treatment in second-line treatment plus 28 days, that is, $D_{FirstTrt}2L$ until $D_{LastTrt}2L+28$.
- Whole sequence of treatment: From the date of first administration of the first-line study treatment until the date of last administration of study treatment (first-line study treatment for patients not entering second-line and second-line study treatment otherwise) + 28 days.

3.1.1.9 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 3-1 below.

Table 3-1 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.

Source data	Condition
	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'.
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.

3.1.1.10 Month

For the analysis of time related endpoints, a month is defined as being equal to:

365.25 / 12 = 30.4375 days

For all time related efficacy endpoints, time will be calculated in days, but analyzed and reported in months.

3.1.2 Concomitant medications with specific impact on the analysis

3.1.2.1 Inhibitors of CYP3A4 and/or PgP

The following rules are consistent with Internal Clinical Pharmacology Drug-drug interaction (DDI) memo, which was updated Dec 2, 2009. The memo summarizes DDI data from three sources including the FDA's "Guidance for Industry, Drug Interaction Studies, the University of Washington's Drug Interaction Database, and Indiana University School of Medicine's Drug Interaction Table."

Co-administration with ***strong inhibitors of CYP3A4*** (e.g., ketoconazole, itraconazole, ritonavir) ***or P-glycoprotein (PgP)*** should be avoided.

Co-administration with ***moderate CYP3A4 inhibitors*** (e.g., erythromycin, fluconazole) ***or PgP inhibitors*** should be used with caution. If patient requires co-administration of moderate CYP3A4 inhibitors or PgP inhibitors, reduce the dose of everolimus to half the currently used dose. Additional dose reductions to every other day may be required to manage toxicities. If the inhibitor is discontinued the everolimus dose should be returned to the dose used prior to initiation of the moderate CYP3A4/PgP inhibitor.

Seville orange, star fruit, grapefruit and their juices affect P450 and PgP activity. Concomitant use should be avoided.

Avoid the use of ***strong CYP3A4 inducers***. If patient requires co-administration of strong CYP3A4 inducers (i.e., phenytoin, carbamazepine, rifampin, rifabutin, phenobarbital, St. John's wort), an increase in the dose of everolimus up to twice the currently used daily dose should be considered, using 2.5mg - 5mg increments. Enzyme induction usually occurs within 7-10 days; therefore everolimus dose should be increased by one increment 7 days after the start of the inducer therapy. If no safety concerns are seen within the next 7 days, the dose can be increased again one additional increment up to a maximum of twice the daily dose used prior to initiation of the strong CYP3A4 inducer.

This dose adjustment of everolimus is intended to achieve similar AUC to the range observed without inducers. However, there are no clinical data with this dose adjustment in patients receiving strong CYP3A4 inducers. If the strong inducer is discontinued the everolimus dose should be returned to the dose used prior to initiation of the strong CYP3A4/PgP inducer.

The following will be tabulated and summarized:

- Clinically relevant drug interactions: substrates, inducers, and inhibitors of isoenzyme CYP3A (table
- List of clinically relevant drug interactions mediated by PgP substrates
- List of clinically relevant drug interactions mediated by PgP inhibitors
- 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 will be patients who took these drugs during the study and therefore these 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.

3.1.2.2 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 enrollment but before end of treatment may be identified as protocol deviations. In addition, their efficacy data (other than overall survival) will be censored so that the tumor assessments made after the intake of anti-neoplastic drugs are not included in efficacy analyses. For details on the censoring rules see the [Protocol Appendix 2].

Clinical review of individual study data will be performed in order to identify those anti-neoplastic medications which are considered disallowed. At the project level a list of ATC classes and medications is maintained which includes:

- those ATC classes/medications which have been identified as being disallowed at the project level
- additional ATC classes/medications which are disallowed for specific indications
- those ATC classes/medications which need to be subjected to clinical review at the study level.

3.1.3 Implementation of RECIST

Response and progression evaluation will be performed according to the RECIST 1.0 guideline version 2 (as described in detail in [Protocol Appendix 2]) whenever it applies. The text below gives more detailed instructions and rules to provide further details needed for programming.

3.1.3.1 Overall lesions response for patients with only bone lesions at baseline

For patients with only bone lesions at baseline, the RECIST 1.0 will be extended to include the evaluation of overall lesion response, which will be based solely on non-target lesion responses or an occurrence of a new lesion. Bone lesions will be entered as non-target lesions. Specifically, in absence of new lesions, the overall lesion response at each assessment will be one of the following: complete response, stable disease, unknown, or progressive disease based on non-target lesion responses. Stable disease would include all assessments not qualifying for complete response, progressive disease or unknown. In presence of any new lesion, the overall lesion response will be progressive disease.

3.1.3.2 Disease progression

For patients with measurable disease at baseline, disease progression will only be assigned if it is documented as per RECIST 1.0 by an objective assessment method (e.g. CT scan, MRI, X-rays, photos for skin lesions, etc.).

Patients with bone only lesions, lytic or mixed (lytic+sclerotic) will be allowed to enter the study and the following will be considered disease progression among these patients:

- The appearance of one or more new 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.

In particular, to be considered as an event, disease progression should be demonstrated with objective evidence.

3.1.3.3 Best overall response

The best overall response (BOR) will be derived in agreement with RECIST 1.0 criteria. The definitions and the details on the derivation are given in Novartis internal guidelines version 2 [Protocol Appendix 2].

For each treatment line, only tumor assessments performed before the start of any further antineoplastic therapies (Section 3.1.2.2) will be considered in the assessment of BOR.
The best overall response is defined with regard to the tumor assessment schedule adopted in this study (i.e., every 8 weeks, ± 1 week). The study day within each line (as defined in Section 3.1.1.6) will be used for the BOR derivation (i.e., Study Day 1 is $D_{FirstTrt}1L$ for first-line treatment, and $D_{FirstTrt}2L$ for second-line treatment).
The best overall response for each patient is determined from the sequence of overall (lesion) responses according to the following rules:

- CR = at least two determinations of CR at least 4 weeks apart before progression.
- PR = at least two determinations of PR or better at least 4 weeks apart before progression (and not qualifying for a CR).
- SD = at least one SD assessment (or better) > 6 weeks after start of study treatment (and not qualifying for CR or PR).
- PD = progression ≤ 12 weeks after start of study treatment (and not qualifying for CR, PR or SD).
- UNK = all other cases (i.e. not qualifying for confirmed CR or PR and without SD for more than 6 weeks or early progression within the first 12 weeks).

For each treatment line, patients with best overall response "unknown" will be summarized by reason for having unknown status. The following reasons will be used:

- No valid post-baseline assessment
- All post-baseline assessments have overall response UNK
- New anti-neoplastic therapy(*) started before first post-baseline assessment
- SD too early
- PD too late

(*) in the analysis of 1st treatment line, this includes also the starting of the per protocol second-line treatment before any post baseline tumor assessment is performed

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

3.1.3.4 Change in imaging modality

Per RECIST 1.0, the imaging method used at baseline should be matched at all subsequent assessment. A strict implementation of RECIST 1.0 would mean that any change in imaging method compared to one used at baseline will lead to unknown overall response at given assessment. However, it is considered that a change in the use of contrast does not necessarily represent a change in the imaging method (for many tumor types, the assessment can be done despite the fact that contrast has changed).

Therefore, in the calculation of overall lesion response the following groups of methods/modalities as listed under the same bullet point will be considered the same:

- 'CT with contrast' and 'CT without contrast'
- 'Spiral CT with contrast' and 'spiral CT without contrast'

- 'MRI with contrast', 'MRI without contrast', 'Dynamic contrast – enhanced MRI' and 'GD-MRI'.

3.1.3.5 Determination of missing adequate tumor assessments

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

As detailed in [Protocol Appendix 2], the PFS censoring and event date options depend on the presence and the number of missing tumor assessments. For example:

- in the primary analysis of PFS, an event occurring after two or more missing assessments is censored at the last adequate assessment.
- in one of the sensitivity analyses of PFS, an event occurring after one or more missing assessments is back-dated to the date of 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 = 8+2(window) = 10 weeks then the analysis will assume one missing tumor assessment. If the distance is greater than D2 = (2*8) +2 =18 weeks then the analysis will assume two missing tumor assessments. The threshold D1 is formed based on the protocol specified interval between the tumor assessments plus the protocol allowed window around the protocol assessments. Similarly, the threshold D2 is formed based on two times the protocol specified interval between the 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 then the patient will be censored and the censoring reason will be 'Event documented after one 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 then the censoring reason will be 1.'Ongoing'; 2. 'New cancer therapy added'; 3. 'Withdrew consent'; 4. 'Loss to follow-up', respectively, depending on the case. However, if this distance is larger than D2 then the censoring reason will always default to 'Adequate assessment no longer available'.

3.1.3.6 No baseline tumor assessments

If there is no baseline tumor assessment, all post-baseline overall lesion responses will be unknown.

Since the timing of disease progression cannot be determined for patients with missing baseline tumor assessment, these patients are censored in the PFS analysis at date of enrollment in the first-line treatment period or at $D_{FirstTrt}2L$ in the analyses of the second-line treatment period.

Patients without baseline tumor assessment in the first-line who die within D2 distance (Section 3.1.3.5) of enrollment will still be counted as having an event in the analysis of PFS during/after first-line treatment. Likewise in the analyses of the second-line treatment period, patients without baseline tumor assessment for the second-line who die within 18 weeks of $D_{FirstTrt}2L$ will be counted as having an event in the analysis of PFS during/after second-line treatment.

All deaths will be counted in the overall survival analysis regardless of the presence or absence of baseline tumor assessment.

3.1.3.7 Construction of waterfall graphs

Waterfall graphs will be used to depict the anti-tumor activity. These plots will display the best percentage change from the treatment-line specific baseline in the sum of the longest diameter of all target lesions for each patient. The proportions of patients with various degrees of tumor shrinkage or growth which can be read directly from the graph can then represent a useful efficacy metric. Only patients with measurable disease at baseline and valid post baseline assessments will be included in the waterfall graphs.

However, caution needs to be applied to assessments where an occurrence of a new lesion or worsening in non-target lesions results in the overall tumor assessment being PD confounding the outcome obtained on the target lesions. These assessments will not be displayed as bars in the graph. If such a "contradicting" assessment represents the only post-baseline assessment for a patient then the patient will be represented by a special symbol (e.g. *) in the waterfall graph.

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

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

All possible assessment scenarios are described in Table 3-1.

Table 3-2 Inclusion/exclusion of assessments used in waterfall graph

	Criteria for inclusion/exclusion			Possible sources of contradictions	
case	Target response	Overall lesion response	Include in waterfall?	Non-target response	New lesion?

	Criteria for inclusion/exclusion			Possible sources of contradictions	
1	CR/PR/SD	PD	Yes but as * only	PD	Any
2	CR/PR/SD	PD	Yes but as * only	Any	Yes
3	UNK	UNK or PD	No	Any	Any
4	CR/PR/SD	UNK	No	UNK	No
5	CR/PR/SD	CR/PR/SD	Yes as a bar	SD/IR	No
6	PD	PD	Yes as a bar	Any	Any

Therefore, the following algorithm will be used to construct the graph:

Select "valid" post-baseline assessments to be included, i.e. for each patient and each assessment repeat the following four steps.

1. Check the target lesion response and overall lesion response at each assessment. If at least one of them is UNK then exclude the whole assessment. Otherwise, go to step 1.2.
2. Check the overall lesion response. If it is PD then go to step 1.3. Otherwise go to step 1.4.
3. Check target response. If it's PD then go to step 1.4. Otherwise flag the assessment with *.
4. Calculate the % change from baseline in target lesions.

For each patient, go through all valid assessments identified in step 1 and find the assessment with best % change from baseline in target lesions. The "best" means best for the patient, i.e. the largest shrinkage or if a patient only has assessments with tumor growth take the assessment where the growth is minimal. (Example 1: Patient 1 has the following % changes from baseline at assessments 1, 2, 3, 4 and 5, respectively: -10%; -25%; -13%; -4% and +6%. His/her best % change is then -25%. Example 2: Patient 2 has the following % changes from baseline at assessments 1, 2 and 3, respectively: +5%; +18% and +35%. His/her best % change is then +5%.

Construct the waterfall graph displaying the best % change from baseline for each patient. Patients having only * flagged assessment(s) will be displayed separately.

The waterfall graph will use the investigator reported overall lesion responses (see Source 1 in Table 1-1).

The best overall response (BOR) will be shown above each of the displayed bars in the graph.

The following conventions will be used to display the waterfall graphs:

1. Bars under the horizontal axis representing tumor shrinkage
2. Bars above the horizontal axis representing tumor growth
3. "Zero" bars with * symbol representing patients with contradiction.

The rationale for this order is that the third category (patients with contradiction) logically follows after the second category (patients with tumor growth).

3.1.4 Time interval for ECOG performance status

The following time based intervals will be used to group the ECOG PS data over time for each treatment line. Day x is defined in Section 3.1.1.6.

	Time Interval
Baseline	on or before the first date of study treatment in first line, second line, respectively
Week 4	day 2 to day 42
Weeks 8, 12, …	+/- 2 weeks centered around the planned assessment date: i.e. days 43, 70 for Week 8 (3rd assessment) days 71, 98 for Week 12 (4th assessment) days [(k-1)*28-13; (k-1)*28+14] (kth assessment, k>2)

If more than one assessment is done within the same time window, the assessment performed closest to the target date will be used.

3.1.5 Definition of the first occurrence of stomatitis

3.1.5.1 Based on OSDQ

Multiple occurrences of stomatitis can be recorded in the OSDQ questionnaire. To define the first occurrence of stomatitis based on the OSDQ data, the following rules will be applied.

Answer to the question 1b of the OSDQ questionnaire will be used to define the start and end dates of the first occurrence of stomatitis:

Q1b: When did you experience the first symptoms of mouth and throat soreness (only ask this question on Day 1 of the questionnaire)?

Start date of the first occurrence of stomatitis is defined as the first date ever recorded for question Q1b.

End date of the first occurrence of stomatitis is defined as the date of the last OSDQ questionnaire that was completed before a second date is recorded for question Q1b.

An example is provided below:

Patient	Questionnaire date	Date recorded for question Q1b	Stomatitis occurrence number
101	19-Nov-14	19-Nov-14	1
101	20-Nov-14	.	1
101	21-Nov-14	.	1
101	04-Dec-14	03-Dec-14	2

In the table above, the 3 first records belong to the first occurrence of stomatitis. This first occurrence started on 19-Nov-2014 and ended on 21-Nov-14.

3.1.5.2 Based on AE CRFs

Multiple AE CRF records of the AESI stomatitis can correspond to the same occurrence (investigators could report an end date for stomatitis and add a new record in the CRF only to update the grade of the current stomatitis occurrence or to use a different verbatim). In order to identify the first occurrence of the AESI stomatitis, it is needed to identify the AE records which belong to the same occurrence.

The definition is as follows:
If the difference between the previous AE record end date and the following AE record start date is less or equal to 1 day then the two records will be considered to belong to the same occurrence.
If this difference is greater than 1 day then the new AE record will be considered to be a new occurrence of the AESI stomatitis.

In the analysis of duration of the first incidence of the AESI stomatitis, only the first occurrence will be considered.

Start date of the first occurrence of the AESI stomatitis is defined as the earliest start date of the AE records belonging to the first occurrence of the AESI stomatitis.

End date of the first occurrence of the AESI stomatitis is defined as the latest end date of AE records belonging to the first occurrence of the AESI stomatitis.

An example is provided below:

Patient	AE PT	AESI grouping	Start date	End date	Occurrence number
101	Stomatitis	Stomatitis	19-Nov-14	22-Nov-14	1
101	Stomatitis	Stomatitis	23-Nov-14	30-Nov-14	1
101	Mouth Ulceration	Stomatitis	30-Nov-14	07-Dec-14	1
101	Stomatitis	Stomatitis	15-Feb-15	15-Apr-15	2

In the table above, the 3 first records belong to the same occurrence of the AESI stomatitis (first occurrence). The 4th record is a new occurrence (second occurrence).
In this example, the start date of the first occurrence of the AESI stomatitis is 19-Nov-14 and the end date is 07-Dec-14.

3.2 Details of the statistical analysis

3.2.1 Baseline comparability

Appropriate descriptive summary 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 performed in first line and second line respectively, but no p-values will be provided.

3.2.2 Time-to-event analyses

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

- Progression-free survival
- Overall survival
- Time to definitive deterioration of the ECOG score by one category of the score from baseline
- Time to the first incidence of AE
- [redacted]

3.2.2.1 Kaplan-Meier estimates

The survival function will be estimated using the Kaplan-Meier (product-limit) method as implemented in PROC LIFETEST with method=KM option in SAS. Median survival will be obtained along with 95% confidence intervals calculated from PROC LIFETEST output using the loglog option available within PROC LIFETEST. Kaplan-Meier estimates with 95% confidence intervals at specific time points will be summarized.

The Kaplan-Meier graphs as well as the statistics (hazard ratio, etc.) will be obtained from the SAS software.

Example code fragment

Here the input dataset, DSET, contains two variables, SURVTIME is the survival time (in months) for each subject, and the variable CENS indicates censoring with a value of 1 indicating a censored time, and 0 otherwise. The '(1)' following CENS in the time statement is used to indicate the state that denotes censoring. The output dataset quart gives estimates for the quartiles (with 95% confidence intervals) of the quartiles of the survival distribution. The quartile with percent label 50 corresponds to the median. The output dataset OUTDATA contains survival distribution function estimates with 95% confidence intervals constructed using Greenwood's estimate of the standard error of the Kaplan-Meier estimate. To identify the Kaplan-Meier estimate (and corresponding 95% confidence interval) at a fixed time point, identify the value of SURVTIME at which there was a failure that is either at, or the last one before the time point of interest. The estimate and confidence interval at this time point is the correct one to display.

```
ods output Quartiles=quart;
proc lifetest data=dset method=KM outsurv=outdata CONFTYPE=LOGLOG;
  time survtime*cens(1);
run;
```

[redacted]

[redacted]

3.2.3 Exact confidence intervals for response rate

Exact confidence intervals for proportions of patients with ORR and CBR will be derived using SAS PROC FREQ with the binomial option. Note that PROC FREQ will produce the confidence interval for the lowest level of the analysis variable, therefore care must be taken that the variable be coded correctly in order to achieve the desired result, i.e. with responders coded as 0 and non-responders as 1. It should also be noted that if the data analyzed contains only non-responders that the default interval produced will be for non-responders and the interval for responders must be produced by subtracting both the upper and lower limits from 1.

Example code fragment

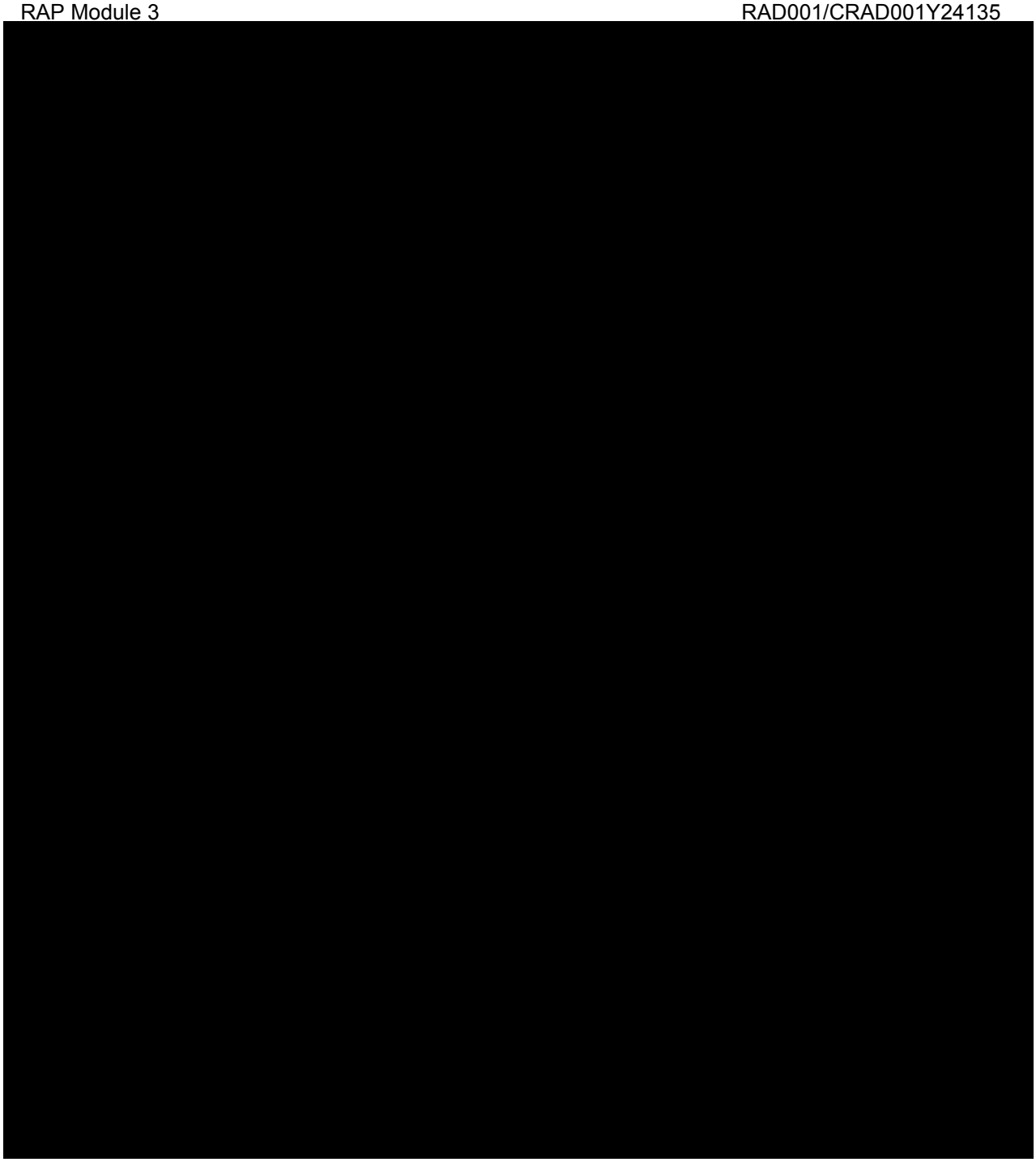
Here the input dataset, DSET, contains a line per patient with an indicator variable, VAR, that is equal to zero if a patient has the response of interest and one otherwise. The required exact limits are in the ods output dataset BINPROP, labeled XL_BIN and XU_BIN.

```
proc freq data=dset order=internal;
  table var / binomial;
run;
```

If the input dataset contained only non-responders, then the interval produced is for non-responders, and the required interval then needs to be derived with the following code:

```
data binprop2(keep=name1 nvalue1);
  set binprop(where=(name1 in ("XL_BIN", "XU_BIN")));
  if name1 eq "XL_BIN" then name1="XU_BIN";
                  else name1="XL_BIN";
  nvalue1 = 1-nvalue1;
run;
```

[redacted]



4 References

Allison PD (2000). Survival analysis using the SAS system. Cary, NC, SAS Institute Inc.

Armitage P, McPherson CK, Rowe BC (1969): Repeated significance tests on accumulating data. J .R. Statist. Soc. A, 132, 232-244.

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

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

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

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

Fleming TR, Rothmann MD, Lu HL (2009). Issues in Using Progression-Free Survival When Evaluating Oncology Products. Journal of Clinical Oncology 27, 2874-2880. Published by: European Organisation for Research and Treatment of Cancer, Brussels 2001.

FDA Guidelines (2005) Clinical Trial Endpoints for the Approval of Cancer Drugs and

Biologics, April. Therasse P, Arbuck S, Eisenhauer E, et al (2000) New Guidelines to Evaluate the Response to Treatment in Solid Tumors, Journal of National Cancer