Supplementary Appendix

This appendix has been provided by the authors to give readers additional information about their work.

Supplement to: Yao JC, Shah MH, Ito T, et al. Everolimus for advanced pancreatic neuroendocrine tumors. N Engl J Med 2011;364:514-23.

(PDF updated March 10, 2011.)

Acknowledgments

We thank the participating patients and their families, as well as the global network of research nurses, trial coordinators, and operations staffs, for their contributions, and we thank the investigators who enrolled patients in this trial, among them: **Belgium:** I. Borbath, J.-L. Van Laethem; **Brazil:** R. Montenegro; **Canada:** C. Card, W. Kocha, P. Metrakos, K. Mulder, L. Sideris; **France:** T. Andre, E. Baudin, C. Borg, G. Cadiot, S. Dominguez, B. Goichot, R. Guimbaud, C. Lepage, P. Rougier, P. Ruszniewski, J.F. Seitz, M. Ychou; **Germany:** C. Auernhammer, T. Gress, D. Hörsch M. Weber, B. Wiedenmann; **Greece:** G. Kaltsas; **Italy:** E. Bosi, V. Di Carlo, N. Fazio, G. Luppi, S. Ricci, S. Siena; **Japan:** M. Ikeda; **Slovakia:** R. Hyrdel; **South Korea:** Y.-J. Bang, H.-C. Chung, T.-W. Kim, Y.-S. Park; **Spain:** R. Salazar; **Sweden:** B. Eriksson; **Switzerland:** M. Fried; **Taiwan:** J.-S. Chen, L.-T. Chen, Y.-L. Lin, Y.-M. Shyr; **Thailand:** T. Ativitavas, P. Sunpaweravong; **United Kingdom:** A. Grossman, W. Mansoor, N. Reed, P. Ross; and **United States:** S. Agarwala, L. Anthony, T. Butler, E. Chiorean, J. Diener, P. Engstrom, R. Fine, K. Hartshorn, B. Koester, L. Kvols, B. Lembersky, A. Maniam, T. O'Dorisio, P. Philip, J. Pisegna, R. Pommier, S. Salmon, M. Scola, V. Sharma, and H. Xiong.

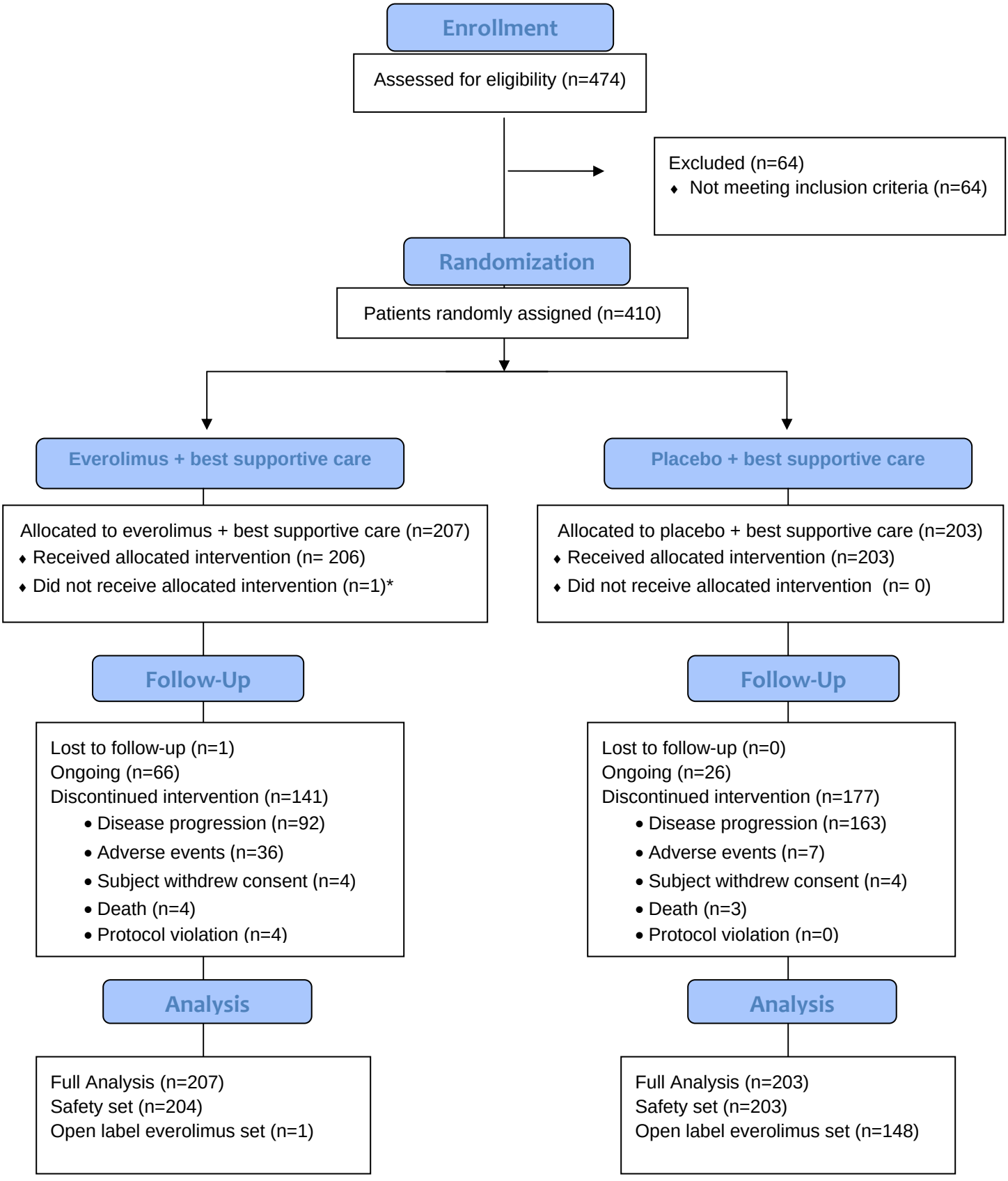
Response Criteria

Efficacy Assessment	Definition
Measurability of lesions at baseline	1. Measurable: unidimensional (LD only, size with conventional techniques ≥20 mm; spiral computed tomography ≥10 mm) 2. Nonmeasurable: all other lesions, including small lesions. (LD only, size with conventional techniques <20 mm; spiral computed tomography <10 mm)
Objective response	1. Target lesions: All measurable lesions up to a maximum of 5 per organ and up to 10 lesions total [more than one organ]) selected on the basis of their size and suitability for accurate repeated measurement CR: disappearance of all target lesions PR: ≥30% decrease in the sum of the LD of all target lesions taking the baseline sum of LD as reference PD: ≥20% increase in the sum of the LD of all measured target lesions, taking as reference the smallest sum of LD of all target lesions recorded at or after baseline SD: Neither sufficient shrinkage to qualify for PR or CR nor an increase in lesion which would quality for PD Unknown: Progression has not been documented and one or more target lesions have not been adequately assessed 2. Nontarget lesions: all other lesion that do not fulfill the criteria for target lesions at baseline CR: disappearance of all target lesions PD: Unequivocal progression of existing nontarget lesions SD: Neither sufficient shrinkage to qualify for CR nor PD Unknown: Progression has not been documented and one or more nontarget lesions have not been adequately assessed
Best Overall response	Best response recorded from treatment start to disease progression or recurrence CR: at least two determinations of CR at least 4 weeks apart before progression. PR: at least two determinations of PR or better at least 4 weeks apart before progression (and not qualifying for a CR). SD: at least one SD assessment (or better) > 6 weeks after randomization (and not qualifying for CR or PR) PD: any progression within 18 weeks after randomization (and not qualifying for CR, PR or SD) will lead to a best overall response evaluation of 'progressive disease' Unknown: all other cases
Duration of response	Applies only to patients with best overall response of CR or PR Start: date of first documented response (CR or PR) End: time of progression
Progression-Free Survival	PFS is defined as the duration of time from randomization to time of progression or death.

RECIST = Response Evaluation Criteria in Solid Tumors, LD = longest diameter, CR = complete response, PR = partial response, PD = progressive disease, SD = stable disease.

Adapated from Therasse P, Arbuck SG, Eisenhauer EA, Wanders J, Kaplan RS, Rubinstein L, Verweij J, Van Glabbeke M, van Oosterom AT, Christian MC, Gwyther SG. New guidelines to evaluate the response to treatment in solid tumors. European Organization for Research and Treatment of Cancer, National Cancer Institute of the United States, National Cancer Institute of Canada.J Natl Cancer Inst. 2000 Feb 2;92(3):205-16.

CONSORT Flow Diagram



* subject withdrew consent before receiving study medication

Appendix. Drug-related Adverse Events Occurring With At Least 1% Incidence of Grade 3 or 4 Events in Either Group.

Adverse event	**Everolimus (N = 204)**			**Placebo (N = 203)**		
	All grades n(%)	**Grade 3 n(%)**	**Grade 4 n(%)**	**All grades n(%)**	**Grade 3 n(%)**	**Grade 4 n(%)**
	no. of patients (percent)					
Anemia	35(17)	10(5)	2(1)	5(3)	0(0)	0(0)
Hyperglycemia	27(13)	11(5)	0(0)	9(4)	3(2)	1(<1)
Stomatitis	108(53)	10(5)	0(0)	23(11)	0(0)	0(0)
Thrombocytopenia	27(13)	7(3)	1(<1)	1(<1)	0(0)	0(0)
Diarrhea	69(34)	7(3)	0(0)	20(10)	0(0)	0(0)
Hypophosphatemia	15(7)	6(3)	0(0)	3(2)	1(<1)	0(0)
Neutropenia	13(6)	6(3)	0(0)	4(2)	4(2)	0(0)
Diabetes	17(8)	5(3)	0(0)	0(0)	0(0)	0(0)
Hemoglobin decreased	12(6)	5(3)	0(0)	2(1)	0(0)	0(0)
Lymphopenia	12(6)	5(3)	0(0)	3(2)	1(<1)	0(0)
Dyspnea	15(7)	3(2)	0(0)	6(3)	0(0)	0(0)
Fatigue	64(31)	3(2)	0(0)	29(14)	1 (<1)	0(0)
Mouth ulceration	12(6)	3(2)	0(0)	4(2)	0(0)	0(0)
Nausea	41(20)	3(2)	0(0)	37(18)	0(0)	0(0)
Pneumonitis	25(12)	3(2)	0(0)	0(0)	0(0)	0(0)
Pulmonary Embolism	3(2)	2(1)	1(<1)	0(0)	0(0)	0(0)

Abdominal pain	11(5)	2(1)	0(0)	9(4)	1(<1)	0(0)
Asthenia	26(13)	2(1)	0(0)	17(8)	2(1)	0(0)
Blood glucose decreased	2(1)	2(1)	0(0)	1(<1)	0(0)	0(0)
Dehydration	5(3)	2(1)	0(0)	0(0)	0(0)	0(0)
Hepatic Function abnormal	4(2)	2(1)	0(0)	1(<1)	1(<1)	0(0)
Interstitial lung disease	5(2)	2(1)	0(0)	0(0)	0(0)	0(0)
Leukopenia	12(6)	2(1)	0(0)	4(2)	1(<1)	0(0)
Palma-plantar erythrodysesthesia Syndrome	6(3)	2(1)	0(0)	0(0)	0(0)	0(0)
White blood cell count	4(2)	2(1)	0(0)	1(<1)	0(0)	0(0)
Alanine Aminotransferase increased	7(3)	1(<1)	0(0)	4(2)	2(1)	0(0)
Aspartate aminotransferase Increased	6(3)	1(<1)	0(0)	7(3)	2(1)	0(0)

NOVARTIS

Novartis Oncology

Clinical Development & Biostatistics

RAD001 (everolimus)

Study No. CRAD001C2324

Harmonization of Efficacy Analysis of Solid Tumor Studies

Protocol Post-text Supplement: 1
Guidelines for Response, Duration of Overall Response, TTF, TTP, Progression-Free Survival and Overall Survival (based on RECIST)

Authors (Version 2):	Capdeville R, Chen B, Gathmann I, Green S, Lebwohl D, Murawsky M, Wang L, Zuber E
Authors (Version 1):	Gathmann I, Mietlowski W, Kiese B, Mellars K, Capdeville R, Dugan M

Document type:	TA Specific Guideline
Document status:	Final Version 2, dated 18-Jan-2007 replacing Version 1, dated 13-Dec-2002
Release date:	17-Apr-2007
Number of pages:	21

Table of contents

List of tables

Glossary

CR	Complete response
CRF	Case Report Form
CSR	Clinical Study Report
CT	Computed tomography
DFS	Disease-free survival
eCRF	Electronic Case Report Form
FPFV	First patient first visit
MRI	Magnetic resonance imaging
LPLV	Last patient last visit
OS	Overall survival
PD	Progressive disease
PFS	Progression-free survival
PR	Partial response
RAP	Report Analysis Preparation
RECIST	Response Evaluation Criteria in Solid Tumors
SD	Stable disease
TTF	Time to treatment failure
TTP	Time to progression
UNK	Unknown

1 Introduction

The purpose of this document is to provide the working definitions and rules necessary for a consistent and efficient analysis of efficacy for oncology studies in solid tumors. This document is based on the RECIST criteria for tumor responses (Therasse, et al, 2000).

The efficacy assessments described in Section 2 and the definition of best response (Section 3.1) are based on the RECIST criteria but also give more detailed instructions and rules for determination of best response. Section 3.2 is summarizing the "time to event" variables and rules which are mainly derived from internal discussions, as the RECIST criteria do not define these variables in detail. Section 4 of this guideline describes data handling and programming rules. This section may be used in the analysis plan(s) to provide further details needed for programming.

As for a usual Novartis template, comments are written in *blue italic* font. The protocol authors must take these comments into consideration and provide project or study specific details in the protocol. Specifically, definitions highlighted in *red* must be discussed, defined and then documented in the study protocol. Any deviations to the guideline must be clearly specified in the protocol with justification.

2 Efficacy assessments

Tumor evaluations are made based on RECIST criteria (Therasse et al. (2000), New Guidelines to Evaluate the Response to Treatment in Solid Tumors, Journal of National Cancer Institute, Vol. 92; 205-16).

>> Indicate the assessment schedule for tumor assessments in the protocol. Frequency of tumor re-evaluation while on treatment should be adapted to the type and schedule of treatment, and the type of tumor treated. It should also be clearly stated how patients are followed for progression after discontinuation of study treatment.

>> It is assumed that all information which is considered for assessment of the tumor is captured in the RECIST eCRF, i.e. not merged from several sources.

2.1 Eligibility

Using RECIST criteria implies that only patients with measurable disease at baseline should be included in the study:

- **Measurable disease -** the presence of at least one measurable lesion. If the measurable disease is restricted to a solitary lesion, its neoplastic nature should be confirmed by cytology/histology.

>> If patients without measurable disease are allowed to be included in the study (when the primary endpoint is not the objective tumor response), please adapt the sentence accordingly.

- **Measurable lesions -** lesions that can be accurately measured in at least one dimension with longest diameter ≥ 20 mm using conventional techniques or ≥ 10 mm with spiral CT scan (with minimum lesion size no less than double the slice thickness).

- **Non-measurable lesions -** all other lesions, including small lesions (longest diameter <20 mm with conventional techniques or <10 mm with spiral CT scan), i.e., bone lesions, leptomeningeal disease, ascites, pleural/pericardial effusion, inflammatory breast disease, lymphangitis cutis/pulmonis, abdominal masses that are not confirmed and followed by imaging techniques, and cystic lesions.

>> If any of lesion should be handled differently, this must be clearly stated and justified in the protocol, e.g. tumor lesions that are situated in a previously irradiated area might or might not be considered measurable, and the conditions under which such lesions should be considered must be defined in the protocol when appropriate.

2.2 Methods of tumor measurement

The following considerations are to be made when evaluating the tumor:

- All measurements should be taken and recorded in metric notation, using a ruler or calipers. All baseline evaluations should be performed as closely as possible to the beginning of treatment and never more than 4 weeks before the beginning of the treatment.

>> If different window for baseline assessments is allowed in the protocol this must be justified in the protocol.

- The same method of assessment and the same technique should be used to characterize each identified and reported lesion at baseline and during follow-up.
- Imaging–based evaluation is preferred to evaluation by clinical examination when both methods have been used to assess the antitumor effect of a treatment.
- **CT and MRI:** CT and MRI are the best currently available and reproducible methods to measure target lesions selected for response assessment. Conventional CT and MRI should be performed with cuts of 7.5 mm or less in slice thickness contiguously. Spiral CT should be performed using a 5 mm contiguous reconstruction algorithm. This applies to tumors of the chest, abdomen and pelvis. To calculate the sum of all target lesions, their size must be entered throughout the study. Actual measurement of each lesion should be entered in the eCRF regardless of the size of the measurement. A size of 0 mm should only be entered if the lesion completely disappeared. If a very small lesion can not be reliably measured because of its size, it is recommended to enter the minimum lesion size (e.g. 5 mm for spiral CT). In other cases where the lesion cannot be reliably measured for reasons other than its size (e.g., borders of the lesion are confounded by neighboring anatomical structures), no measurement should be entered and the lesion cannot be evaluated.

>> If head and neck tumors and those of extremities are evaluated in the study, please specify the methods in detail in the protocol.

- **Chest x-ray:** Lesions on chest x-ray are acceptable as measurable lesions when they are clearly defined and surrounded by aerated lung. However, CT is preferable.
- **Ultrasound:** When the primary endpoint of the study is objective response evaluation, ultrasound (US) should not be used to measure tumor lesions. It is, however, a possible alternative to clinical measurements of superficial palpable lymph nodes, subcutaneous

lesions and thyroid nodules. US might also be useful to confirm the complete disappearance of superficial lesions usually assessed by clinical examination.

- **Endoscopy and laparoscopy:** The utilization of endoscopy and laparoscopy for objective tumor evaluation has not yet been fully and widely validated. Their uses in this specific context require sophisticated equipment and a high level of expertise that may only be available in some centers. Therefore, the utilization of such techniques for objective tumor response should be restricted to validation purposes in specialized centers. However, such techniques can be useful in confirming complete pathological response when biopsies are obtained.
- **Tumor markers:** Tumor markers alone cannot be used to assess response. If markers are initially above the upper normal limit, they must normalize for a patient to be considered in complete clinical response when all lesions have disappeared.

>> If tumor markers are used in the study for the response assessment, the criteria must be clearly stated in the protocol and the presence of abnormality in tumor markers be entered in the eCRF page for RECIST evaluations (see also section 2.4).

- **Cytology and histology:** Cytology and histology can be used to differentiate between PR and CR in rare cases (i.e., after treatment to differentiate between residual benign lesions and residual malignant lesions in tumor types such as germ cell tumors). Cytologic confirmation of neoplastic nature of any effusion that appears or worsens during treatment is required when the measurable tumor has met the criteria for response or stable disease. Under such circumstances, the cytologic examination of the fluid collected will permit differentiation between response and stable disease (an effusion may be a side effect of the treatment) or progressive disease (if the neoplastic origin of the fluid is confirmed).

>> When pathological response is being used, the protocol must clearly state details on how pathological responses are documented.

- **Clinical examination:** Clinical lesions will only be considered measurable when they are superficial (i.e., skin nodules and palpable lymph nodes). For the case of skin lesions, documentation by color photography, including a ruler to estimate the size of the lesion, is recommended.

>> If the protocol is considering specific symptoms as objective signs of clinical progression, e.g. bone pain or GI bleeding, then the criteria for clear worsening of these non-measurable 'lesions' indicative of PD should be clearly specified in the protocol. In that case, the protocol should clearly specify that additional criteria are used to complement RECIST criteria.

2.3 Baseline documentation of target and nontarget lesions

For the evaluation of lesions at baseline and throughout the study, the lesions are classified at baseline as either target or nontarget lesions:

- **Target lesions:** All measurable lesions up to a maximum of five lesions per organ and 10 lesions in total, representative of all involved organs should be identified as target lesions and recorded and measured at baseline. Target lesions should be selected on the basis of their size (lesions with the longest diameter) and their suitability for accurate repeated

measurements (either by imaging techniques or clinically).
A sum of the longest diameter for all target lesions will be calculated and reported as the baseline sum of the longest diameter. The baseline sum of the longest diameter will be used as reference by which to characterize the objective tumor response.

- **Nontarget lesions:** All other lesions are considered nontarget lesions, i.e. lesions not fulfilling the criteria for target lesions at baseline. Presence or absence or worsening of non-target lesions should be assessed throughout the study; measurements of these lesions are not required.

>> If the protocol is considering specific symptoms for assessment of the tumor, e.g. bone pain or GI bleeding, then these symptoms are to be entered as nontarget lesions with either presence or absence or as a new lesion (based on protocol specified criteria). In that case, the protocol should clearly specify that additional criteria are used to complement RECIST criteria.

>> For cancers which are known to metastasize in bone, the protocol should specify if and how bone lesions should be handled, e.g. if they are not identified only at baseline and end of study by scintigram but also followed throughout the study by bone X-ray or CT scan.

>> If no measurable lesions are identified at baseline, the patient is not evaluable for RECIST. Therefore, the Guideline leaves the decision to the study teams to determine how to handle time to event endpoints, e.g. in the metastatic setting nontarget lesions could be followed for PFS/ TTP.

2.4 Evaluation of target and nontarget lesions

To assess tumor response, the sum of the longest diameter for all target lesions will be calculated (at baseline and throughout the study). At each assessment response is evaluated first separately for the target (Table 2-1) and nontarget lesions (Table 2-2) identified at baseline. These evaluations are then used to calculate the overall lesion response considering both the target and nontarget lesions together (Table 2-3) as well as the presence or absence of new lesions.

The response for nontarget lesions is CR only if all nontarget lesions which were evaluated at baseline are now all absent. If any of the nontarget lesions is still present, the response can only be 'Incomplete response/Stable disease' unless any of the lesions was not assessed (in which case response is UNK) or there is unequivocal progression of the nontarget lesions (in which case response is PD).

>> If tumor markers are used as nontarget lesions to evaluate response, please specify criteria for CR, SD and PD in the protocol, e.g. CR='Normalization of tumor marker level', PD='Elevation of tumor markers to certain level', SD='Not qualifying for CR or PD'. These criteria are indication and study specific. In that case, the protocol should clearly specify that additional criteria are used to complement RECIST criteria.

Table 2-1 Response criteria for target lesions

Response Criteria	Evaluation of target lesions
Complete Response (CR):	Disappearance of all target lesions
Partial Response (PR):	At least a 30% decrease in the sum of the longest diameter of all target lesions, taking as reference the baseline sum of the longest diameters.
Progressive Disease (PD):	At least a 20% increase in the sum of the longest diameter of all measured target lesions, taking as reference the smallest sum of longest diameter of all target lesions recorded at or after baseline.
Stable Disease (SD):	Neither sufficient shrinkage to qualify for PR or CR nor an increase in lesions which would qualify for PD.
Unknown (UNK)	Progression has not been documented and one or more target lesions have not been assessed or have been assessed using a different method than baseline.

Table 2-2 Response criteria for nontarget lesions

Response Criteria	Evaluation of nontarget lesions
Complete Response (CR):	Disappearance of all nontarget lesions
Progressive Disease (PD):	Unequivocal progression of existing nontarget lesions.[1]
Incomplete Response/ Stable Disease (SD):	Neither CR nor PD
Unknown (UNK)	Progression has not been documented and one or more nontarget lesions have not been assessed or have been assessed using a different method than baseline.

[1] Although a clear progression of "nontarget" lesions only is exceptional, in such circumstances, the opinion of the treating physician does prevail and the progression status should be confirmed later on by the review panel (or study chair).

- The appearance of a new lesion is always associated with Progressive Disease (PD) and has to be recorded as a new lesion in the appropriate module of the CRF.
- If a lesion appears at the same anatomical location where a target lesion had previously disappeared, it is advised that the time point of lesion disappearance (i.e., the "0 mm" recording) be re-evaluated to make sure that the lesion was not actually present and/or not visualized for technical reasons in this previous assessment. If it has not been possible to change the 0 value, then the investigator/radiologist has to decide between the following two possibilities:
  - The lesion is a new lesion, in which case the overall tumor assessment will be considered as progressive disease
  - The lesion is clearly a re-appearance of a previously disappeared lesion, in which case the size of the lesion has to be entered in the CRF and the tumor assessment will remain based on the sum of tumor measurements as presented in table 1 above (i.e., a PD will be determined if there is at least 20% increase in the sum of the longest diameter of all measured target lesions, taking as reference the smallest sum of longest diameter of all target lesions recorded at or after baseline). Proper documentation should be available to support this decision.

- For those patients who have only one target lesion at baseline, the reappearance of the target lesion, which disappeared previously, is considered a PD.

2.5 Evaluation of overall lesion response

The evaluation of overall lesion response at each assessment is a composite of the target lesion response, nontarget lesion response and presence of new lesions as shown below in Table 2-3.

Table 2-3 Overall lesion response at each assessment

Target lesions	Nontarget lesions	New Lesions	Overall lesion response
CR	CR	No	CR[1]
CR	Incomplete response/SD[3]	No	PR
CR, PR, SD	UNK	No	UNK
PR	Non-PD and not UNK	No	PR[1]
SD	Non-PD and not UNK	No	SD[1, 2]
UNK	Non-PD or UNK	No	UNK[1]
PD	Any	Yes or No	PD
Any	PD	Yes or No	PD
Any	Any	Yes	PD

[1] This overall lesion response also applies when there are no nontarget lesions identified at baseline.
[2] Once confirmed PR was achieved, all these assessments are considered PR.
[3] As defined in section 2.4.

If the evaluation of any of the target or nontarget lesions identified at baseline could not be made during follow-up, the overall status must be 'unknown' unless progression was seen.

In some circumstances it may be difficult to distinguish residual disease from normal tissue. When the evaluation of complete response depends on this determination, it is recommended that the residual lesion be investigated (fine needle aspirate/biopsy) to confirm the CR.

3 Efficacy definitions

3.1 Best overall response

The best overall response is the best response recorded from the start of the treatment until disease progression/recurrence (taking as reference for PD the smallest measurements recorded since the treatment started). In general, the patient's best response assignment will depend on the achievement of both measurement and confirmation criteria.

To be assigned a status of PR or CR, changes in tumor measurements must be confirmed by repeat assessments that should be performed not less than 4 weeks after the criteria for response are first met.

>> Longer intervals may also be appropriate. However, this must be clearly stated in the protocol. The main goal of confirmation of objective response is to avoid overestimating the response rate observed. In cases where confirmation of response is not feasible, it should be made clear when reporting the outcome of such studies that the responses are not confirmed.

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 randomization/start of treatment (and not qualifying for CR or PR).

>> The protocol should state if randomization or start of treatment is used as start date (baseline). This is then used in all definitions.

>> If a different minimum follow-up period is required to classify for overall response= 'stable disease', this must be specified in the protocol.

- PD = progression ≤ 12 weeks after randomization/ start of treatment (and not qualifying for CR, PR or SD).

>> If PD in a different follow-up period is considered overall response='progressive disease', this must be specified in the protocol.

>> The protocol should state if discontinuation due to 'Disease progression' or death due to study indication is considered PD even if this was not accompanied by documentation of PD based on tumor measurements. This depends on Phase of the study and the primary endpoint (e.g. Phase III studies in which progression-free survival is primary endpoint should consider only documented PD, whereas Phase I and II studies may consider all clinical deteriorations PD). The following sentence therefore is only applicable if this is specified in the protocol:

Patients with symptoms of rapidly progressing disease without radiologic evidence will be classified as progression only when clear evidence of clinical deterioration is documented and/or patient discontinued due to 'Disease progression' or death due to study indication.

- UNK = all other cases (i.e. not qualifying for confirmed CR or PR and without SD after more than 6 weeks or early progression within the first 12 weeks)

Overall lesion responses of CR must stay the same until progression sets in, with the exception of a UNK status. A patient who had a CR can not subsequently have a lower status other than a PD, e.g. PR or SD, as this would imply a progression based on one or more lesions reappearing, in which case the status would become a PD.

Once confirmed overall lesion responses of PR must stay the same or improve over time until progression sets in, with the exception of a UNK status. However, if a patient has a PR (≥30% reduction of tumor burden compared to baseline) at one assessment, followed by a <30% reduction from baseline at the next assessment (but not ≥20% increase from previous smallest sum), the objective status at that assessment should be SD. Once a confirmed PR was seen, the overall lesion response should be considered PR (or UNK) until progression is documented.

>> Example: The sum of lesion diameters is 20 cm at baseline and then 14 cm – 15 cm – 14 cm – 16 cm – 16 cm at the subsequent visits. Assuming that nontarget lesions did not

progress, the overall lesion response would be PR – SD – PR – PR – PR. The second assessment with 14 cm confirms the PR for this patient. All subsequent assessments are considered PR even if tumor measurements decrease only by 20% compared to baseline (20 cm to 16 cm) at the following assessments.

If the patient progressed but continues study medication, further assessments are not considered for the determination of best overall response.

>> Note: these cases may be described as a separate finding in the CSR but not included in the best overall response rate.

The best overall response for a patient is always calculated, based on the sequence of overall lesion responses. However, the overall lesion response at a given assessment may be provided from different sources:

- Investigator overall lesion response
- Central Blinded Review overall lesion response
- Novartis calculated overall lesion response (based on measurements from either Investigator or Central Review)

The primary analysis of the best overall response will be based on the sequence of investigator/central blinded review/calculated (investigator)/calculated (central) overall lesion responses.

>> Specify which determination of best overall response will be considered primary (and delete the other terms in the text). If a central blinded review is used (e.g. in an open-label study in which response is the primary endpoint), the best overall response evaluated by the central blinded review will always be considered the primary response.

Based on the patients' best overall response during the study, the following rates are then calculated:

Overall response rate (ORR) is the proportion of patients with a best overall response of CR or PR. This is also referred to as 'Objective response rate' in some protocols or publications.

Disease control rate (DCR) is the proportion of patients with a best overall response of CR or PR or SD.

Another approach is to summarize the progression rate at a certain time point after baseline. In this case, the following definition is used:

Early progression rate (EPR) is the proportion of patients with progressive disease within 8 weeks of the start of treatment.

>> The protocol should define populations for which these will be calculated. Time duration for EPR is study specific. EPR is used for the multinomial designs of Dent and Zee and counts all patients who at the specified assessment (in this example the assessment would be at 8 weeks ± window) do not have an overall lesion response of SD, PR or CR. Patients with unknown assessment at that time point and no PD before, are also considered PD.

3.2. Time to event variables

>> The protocol should state which of the following variables is used in that study.

3.1.1 Progression-free survival

>> Usually in all Oncology studies, patients are followed for tumor progression after discontinuation of study medication for reasons other than progression or death. If this is not used, e.g. in Phase I or II studies, this should be clearly stated in the protocol.

Progression-free survival (PFS) is the time from date of randomization/start of treatment to the date of event defined as the first documented progression or death due to any cause. If a patient has not had an event, progression-free survival is censored at the date of last adequate tumor assessment.

3.1.2 Overall survival

>> All patients should be followed until death or until patient has had adequate follow-up time as specified in the protocol whichever comes first. The follow-up data should contain the date the patient was last seen alive / last contacted, the date of death and the reason of death ("Study indication" or "Other").

Overall survival (OS) is defined as the time from date of randomization/start of treatment to date of death due to any cause. If a patient is not known to have died, survival will be censored at the date of last contact.

3.1.3 Time to progression

>> Some studies might consider only death related to underlying cancer as an event which indicates progression. In this case the variable "Time to progression" might be used. TTP is defined as PFS except for death unrelated to underlying cancer.

Time to progression (TTP) is the time from date of randomization/start of treatment to the date of event defined as the first documented progression or death due to underlying cancer. If a patient has not had an event, time to progression is censored at the date of last adequate tumor assessment.

3.1.4 Time to treatment failure

>> This endpoint is often appropriate in studies of advanced disease where early discontinuation is typically related to intolerance of the study drug. In some protocols, time to treatment failure may be considered as a sensitivity analysis for time to progression. The list of discontinuation reasons to be considered or not as treatment failure may be adapted according to the specificities of the study or the disease.

Time to treatment failure (TTF) is the time from date of randomization/start of treatment to the earliest of date of progression, date of death due to any cause, or date of discontinuation due to reasons other than 'Protocol violation' or 'Administrative problems'. The time to treatment failure for patients who did not experience treatment failure will be censored at last adequate tumor assessment.

3.1.5 Duration of response

>> If the following variables are analyzed, it should be stated that this analysis might introduce a bias as it includes only responders. There have been reports where a treatment with a significantly higher response rate had a significantly shorter duration of response. The analysis of duration of responses should only be used as a descriptive analysis. If they are used as inferential comparison between treatments, clear justification must be given in the protocol. It should also be stated in the protocol if duration of response is to be calculated in addition for unconfirmed response.

Duration of overall response (CR or PR) applies only to patients whose best overall response was CR or PR. The start date is the date of first documented response (CR or PR) and the end date and censoring is defined the same as that for time to progression.

>> The following two durations might be calculated in addition for a large Phase III study in which a reasonable number of responders is seen. Justification must be given in the protocol when these endpoints are used for any comparison between treatments.

Duration of overall complete response (CR) applies only to patients whose best overall response was CR. The start date is the date of first documented CR and the end date and censoring is defined the same as that for time to progression.

Duration of stable disease (CR/PR/SD) applies only to patients whose best overall response was SD, PR or CR. The start and end date as well as censoring is defined the same as that for time to progression.

3.1.6 Time to response

Time to overall response (CR or PR) is the time between date of randomization/start of treatment until first documented response (CR or PR). This analysis will include all patients/responders. Patients who did not achieve a confirmed PR or CR will be censored at last adequate tumor assessment date when they did not progress (including deaths not due to underlying disease) or at maximum follow-up (i.e. FPFV to LPLV used for the analysis) when they had an event for progression-free survival.

Time to overall complete response (CR) is the time between date of randomization/start of treatment until first documented CR. This analysis will include all patients/responders. Patients who did not achieve a confirmed CR will be censored at last adequate tumor assessment date when they did not progress (including deaths not due to underlying disease) or at maximum follow-up (i.e. FPFV to LPLV used for the analysis) when they had an event for progression-free survival.

>> Indicate whether this analysis should include only the responder (in which case please delete 'patients/' and the sentence on patients who did not respond) or should estimate the time to response for the whole study population (in which case please delete '/responders'). If both methods should be used, please state that in the protocol.

3.1.7 Definition of start and end dates for time to event variables

Assessment date

For each assessment (i.e. evaluation number), the **assessment date** is calculated as the latest of all measurement dates (e.g. X-ray, CT-scan) if the overall lesion response at that assessment is CR/PR/SD/UNK. Otherwise – if overall lesion response is progression – the assessment date is calculated as the earliest date of all measurement dates at that evaluation number.

Start dates

>> State in the protocol if date of randomization or date of start of treatment is to be used for all definitions. For randomized studies specify exactly where the randomization date comes from, e.g. from IVRS, or if start of treatment is used as randomization date. For non-randomized studies please specify which treatment start date is taken if more than one treatment is to be given.

For all "time to event" variables, other than the duration of responses, the randomization/ date of treatment start will be used as the start date.

For the calculation of duration of responses the following start date should be used:

- **Date of first documented response** is the assessment date of the first overall lesion response of CR (for duration of overall complete response) or CR / PR (for duration of overall response) respectively, when this status is later confirmed.

End dates

The end dates which are used to calculate 'time to event' variables are defined as follows:

- **Date of death** (during treatment as recorded on the treatment completion page, or during follow-up as recorded on the study evaluation completion page or the survival follow-up page).
- **Date of progression** is the first assessment date at which the overall lesion response was recorded as progressive disease.

>> If applicable, if patients who discontinued due to 'Disease progression' are considered to be PD solely based on clinical deterioration, then add the following in the protocol:

When there is no documentation of radiologic evidence of progression, and the patient discontinued for 'Disease progression' due to documented clinical deterioration of disease, the date of discontinuation is used as date of progression.

- **Date of last adequate tumor assessment** is the date the last tumor assessment with overall lesion response of CR, PR or SD which was made before an event or a censoring reason occurred. In this case the last tumor evaluation date at that assessment is used. If no post-baseline assessments are available (before an event or a censoring reason occurred) the date of randomization/start of treatment is used.
- **Date of next scheduled assessment** is the date of the last adequate tumor assessment plus the protocol specified time interval for assessments. This date may be used if back-dating

is considered when the event occurred beyond the acceptable time window for the next tumor assessment as per protocol (see section 3.2.8).

Example (if protocol defined schedule of assessments is 3 months): tumor assessments at baseline – 3 months – 6 months – missing – missing – PD. Date of next scheduled assessment would then corresponds to 9 months.

- **Date of discontinuation** is the date of the end of treatment visit.
- **Date of last contact** is defined as the last date the patient was known to be alive. This corresponds to the latest date for either the visit date, lab sample date or tumor assessment date. If available, the last contact date from that survival follow-up page is used. If no survival follow-up is available, the date of discontinuation is used as last contact date.

>> In comparative studies with long follow-up period and therefore extended visit schedule, it may be useful to collect the survival status at a pre-specified cut-off within a limited timeframe for all patients with no documented death. In this case, this requires a contact to be made with the patient or with any reliable source of information on the patient's status, but not requiring a specific visit to be scheduled

- **Date of secondary anti-cancer therapy** is defined as the start date of any additional (secondary) antineoplastic therapy or surgery.

>> If this is applicable for the study, it should be specified in the protocol if new cancer therapy is considered an event or endpoints are censored.

3.1.8 Sensitivity analyses

This section outlines the possible event and censoring dates for progression (Table 3-4), as well is addressing the issues of missing tumor assessments during the study. For instance, if one or more assessment visits are missed prior to the progression event, to what date should the progression event be assigned? And should progression event be ignored if it occurred after a long period of a patient being lost to follow-up? It is important that the protocol and analysis plan specify the primary analysis in detail with respect to the definition of event and censoring dates and also include a description of one or more sensitivity analyses to be performed.

Based on definitions outlined in section 3.2.7, and using the draft FDA guideline on endpoints (*Clinical Trial Endpoints for the Approval of Cancer Drugs and Biologics, April 2005*) as a reference, the following analyses can be considered:

Table 3-1 Options for event dates used in PFS, TTP, Duration of response

Situation		Options for end-date (progression)[1] (1) = default unless specified differently in the protocol or analysis plan	Outcome
A	No baseline assessment	(1) Date of randomization/start of treatment	Censored
B	Progression at or before next scheduled assessment	(1) Date of progression (2) Date of next scheduled assessment[2]	Progressed Progressed
C1	Progression or death after exactly one missing assessment	(1) Date of progression (or death) (2) Date of next scheduled assessment[2]	Progressed Progressed
C2	Progression or death after two or more missing assessments	(1) Date of last adequate assessment[2] (2) Date of next scheduled assessment[2] (3) Date of progression (or death)	Censored Progressed Progressed
D	No progression	(1) Date of last adequate assessment	Censored
E	Treatment discontinuation due to 'Disease progression' without documented progression, i.e. clinical progression based on investigator claim	(1) N/A (2) Date of discontinuation (visit date at which clinical progression was determined)	Ignored Progressed
F	New anticancer therapy given	(1) Date of last adequate assessment (2) Date of secondary anti-cancer therapy	Censored Censored
G	Deaths due to reason other than deterioration of 'Study indication'	(1) Date of last adequate assessment	Censored (only TTP)

[1] = Definitions can be found in Section 3.2.7.

[2] = After the last adequate tumor assessment. "Date of next scheduled assessment" is defined in Section 3.2.7.

>> The primary analysis and the sensitivity analyses must be specified in the protocol. Clearly define if and why options (1) are not used for situations C, E and (if applicable) F.

Situations C (C1 and C2): Progression or death after one or more missing assessments: The primary analysis is usually using options (1) for situations C1 and C2, i.e.

- (C1) taking the actual progression or death date, in the case of only one missing assessment.
- (C2) censoring at the date of the last adequate assessment, in the case of two or more consecutive missing assessments.

In the case of two or missing assessments (situation C2), option (3) may be considered jointly with option (1) in situation C1 as sensitivity analysis. A variant of this sensitivity analysis consists of backdating the date of event to the next scheduled assessment as proposed with option (2) in situations C1 and C2.

Situation E: Treatment discontinuation due to 'Disease progression' without documented progression: By default, option (1) is used for situation E as patients without documented PD should be followed for progression after discontinuation of treatment. However, option (2) may be used as sensitivity analysis. If progression is claimed based on clinical deterioration instead of tumor assessment by e.g. CT-scan, option (2) may be used for

indications with high early progression rate or difficulties to assess the tumor due to clinical deterioration.

Situation F: New cancer therapy given: the handling of this situation must be specified in detail in the protocol. However, option (1), i.e. censoring at last adequate assessment may be used as a default in this case.

Additional suggestions for sensitivity analyses

Other suggestions for additional sensitivity analyses may include analyses to check for potential bias in follow-up schedules for tumor assessments, e.g. by assigning the dates for censoring and events only at scheduled visit dates. The latter could be handled by replacing in Table 4 the "Date of last adequate assessment" by the "Date of previous scheduled assessment (from baseline)", with the following definition:

- **Date of previous scheduled assessment (from baseline)** is the date when a tumor assessment would have taken place, if the protocol assessment scheme was strictly followed from baseline, immediately before or on the date of the last adequate tumor assessment.

In addition, analyses could be repeated using the Investigators' assessments of response rather than the calculated response. The need for these types of sensitivity analyses will depend on the individual requirements for the specific study and disease area and have to be specified in the protocol or RAP documentation.

4 Data handling and programming rules

The following section should be used as guidance for development of the protocol, data handling procedures or programming requirements (e.g. on incomplete dates).

4.1 Study/project specific decisions

For each study (or project) various issues need to be addressed and specified in the protocol or RAP documentation. Any deviations from protocol must be discussed and defined at the latest in the RAP documentation.

The proposed primary analysis and potential sensitivity analyses should be discussed and agreed with the health authorities and documented in the protocol (or at the latest in the RAP documentation before database lock).

4.2 Treatment and study completion CRFs

If study drug is discontinued, the **treatment completion page** is to be completed with a visit date reflecting the date the discontinuation decision was made, and with the 'Last known date subject took study drug' and one of the following reasons:

- Adverse event(s)
- Abnormal laboratory value(s)
- Abnormal test procedure results(s)
- Protocol violation
- Subject withdrew consent
- Lost to follow-up
- Administrative problems
- Death
- Disease progression
- Treatment duration completed as per protocol *(optional, to be used if only a fixed number of cycles is given)*

For reasons other than progression (and death) it should be checked if this was not in fact progression (especially reasons Adverse Events, Abnormal laboratory value (s), Abnormal test procedure result and subject withdrew consent). Also it should be checked if patient withdrew consent because of safety issues, in which case reasons Adverse Events, Abnormal laboratory value (s), Abnormal test procedure result should be used.

All patients who discontinued study drug for reasons other than documented progression, death or lost to follow-up will be followed for progression thereafter (patients who withdrew consent might not be followed with regular tumor assessments at the study site, but should ideally be followed until progression outside the study site). Ideally, all patients who discontinued study drug for progression without documented progression will still be followed with regular tumor assessments (e.g. in case of central radiology review). During that evaluation period, usually only tumor measurements (and/or response status) and survival data are collected. In some protocols, the subsequent anti-cancer therapies may also be recorded.

At the end of the study evaluation period, the **study evaluation completion page** is filled out with the following options:

- Subject withdrew consent
- Lost to follow-up
- Administrative problems *(when follow-up for progression has met protocol required events, e.g. follow-up stopped at certain number of events or certain time)*
- Death
- New cancer therapy *(optional, to be used when follow-up for progression is stopped in this case)*
- Disease progression

Thereafter, patients will be followed for survival using the survival follow-up pages. If information on death becomes available for patients who were lost to follow-up or withdrew consent, this may also be entered in the database. The reason for death must be documented (and will be coded using MedDRA); it must be also stated if death was due to 'Study indication' or 'Other' reason.

In comparative studies with long follow-up period and therefore extended visit schedule, it may be useful to collect the survival status at a pre-specified cut-off within a limited timeframe for all patients with no documented death. In this case, this requires a contact to be made with the patient or with any reliable source of information on the patient's status, but not requiring a specific visit to be scheduled

Until the specified cut-off point has been reached, the goal is to collect tumor assessments until disease progression for all patients regardless of whether the patients are still receiving study drug. If patients are not followed for progression, e.g. in a Phase I or II study mainly evaluating safety, the evaluation is completed when study drug is completed (in this case only the first completion page is used).

4.3 Medical validation of programmed overall lesion response

As RECIST is very strict regarding measurement methods (i.e. any assessment with more or less sensitive method than the one used to assess the lesion at baseline is considered UNK) and not available evaluations (i.e. if any target or nontarget lesion was not evaluated the whole overall lesion response is UNK unless remaining lesions qualified for PD), these UNK assessments may be re-evaluated by clinicians at Novartis or external experts. In addition, data review reports will be available to identify assessments for which the investigators' opinion does not match the programmed calculated response based on RECIST criteria. This may be queried for clarification. However, the investigator response assessment will never be overruled.

If Novartis elect to invalidate an evaluation of overall lesion response upon internal or external review of the data, the calculated overall lesion response at that specific assessment is to be kept in a dataset. This must be clearly documented in the RAP documentation and agreed before database lock. This dataset should be created and stored as part of the 'raw' data.

Any discontinuation due to 'Disease progression' without documentation of progression by RECIST criteria should be carefully reviewed. Only patients with documented deterioration of symptoms indicative of progression of disease should have this reason for discontinuation of treatment or study evaluation.

4.4 Programming rules

The following should be used for programming of efficacy results:

4.4.1.1 Calculation of 'time to event' variables

Time to event = enddate - startdate + 1 (in days)

When no post-baseline tumor assessments are available, the date of randomization/start of treatment will be used as enddate (duration = 1 day) when time is to be censored at last tumor assessment, i.e. time to event variables can never be negative.

4.4.1.2 Incomplete assessment dates

All investigation dates (e.g. X-ray, CT scan) must be completed with day, month and year.

If one ore more investigation dates are incomplete but other investigation dates are available, this/these incomplete date(s) are not considered for calculation of the assessment date (and assessment date is calculated as outlined in section 3.2.1). If all measurement dates have no day recorded, the 1st of the month is used.

If the month is not completed, for any of the investigations, the respective assessment will be considered to be at the date which is exactly between previous and following assessment. If a previous and following assessment is not available, this assessment will not be used for any calculation.

4.4.1.3 Incomplete dates for last contact or death

All dates must be completed with day, month and year. If the day is missing, the 15th of the month will be used for incomplete death dates or dates of last contact.

4.4.1.4 Nontarget lesion response

If no nontarget lesions are identified at baseline (and therefore not followed throughout the study), the nontarget lesion response at each assessment will be considered 'not applicable (NA)'.

4.4.1.5 Study/project specific programming

The standard analysis programs need to be adapted for each study/project.

4.4.1.6 Censoring reason

In order to summarize the various reasons for censoring, the following categories will be calculated for each time to event variable based on the treatment completion page, the study evaluation completion page and the survival page.

For survival the following censoring reasons are possible:

- Alive (treatment / study evaluation / survival)
- Lost to follow-up (during treatment* / study evaluation* / survival)

For PFS and TTP (and therefore duration of responses) the following censoring reasons are possible:

- Ongoing without event (treatment / study evaluation)
- Lost to follow-up (during treatment / study evaluation)
- Withdrew consent (during treatment* / study evaluation)

- Study evaluation stopped (*when follow-up for progression is stopped after certain number of events or at certain time, i.e. reason='Administrative problems' on study evaluation completion page, or when patients are not followed for progression after treatment completion*)
- Death due to reason other than underlying cancer *(only used for TTP)*
- New cancer therapy added *(optional; only if the protocol specified that PFS/TTP will be censored at that date)*

*Note = this category is to be used if no further information is available (as information on death may be available in patients who were originally lost to follow-up, and information on progression may be received in patients who withdrew consent to continue study drug).

5 References

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 Institute, Vol. 92; 205-16