MEASUREMENT OF EFFECT SOLID TUMOR EXAMPLES

Although response is not the primary endpoint of this trial, subjects with measurable disease will be assessed by standard criteria. For the purposes of this study, subjects should be re-evaluated every # of weeks weeks. In addition to a baseline scan, confirmatory scans will also be obtained # of weeks weeks following initial documentation of an objective response.

Antitumor Effect – Solid Tumors

For the purposes of this study, subjects should be re-evaluated for response every # of weeks weeks. In addition to a baseline scan, confirmatory scans should also be obtained # of weeks (not less than 4) weeks following initial documentation of objective response.

Response and progression will be evaluated in this study using the new international criteria proposed by the revised Response Evaluation Criteria in Solid Tumors (RECIST) guideline (version 1.1) [Eur J Ca 45:228-247, 2009]. Changes in the largest diameter (unidimensional measurement) of the tumor lesions and the shortest diameter in the case of malignant lymph nodes are used in the RECIST criteria. For primary brain tumors, response and progression will be evaluated using the RANO criteria [*J Clin Oncol* 28: 1963-1972.2010].

Definitions

Evaluable for toxicity: All subjects will be evaluable for toxicity from the time of their first treatment with Agent X**.**

Evaluable for objective response: Only those subjects who have measurable disease present at baseline, have received at least one cycle of therapy, and have had their disease re-evaluated will be considered evaluable for response. These subjects will have their response classified according to the definitions stated below.
Note: Subjects who exhibit objective disease progression prior to the end of cycle 1 will also be considered evaluable.

Evaluable Non-Target Disease Response: Subjects who have lesions present at baseline that are evaluable, but do not meet the definitions of measurable disease, have received at least one cycle of therapy, and have had their disease re-evaluated will be considered evaluable for non-target disease. The response assessment is based on the presence, absence, or unequivocal progression of the lesions.

Disease Parameters

Measurable Disease: Measurable lesions are defined as those that can be accurately measured in at least one dimension (longest diameter for non-nodal lesions and short axis for nodal lesions to be recorded) as $\geq$ 20 mm by chest x-ray, as $\geq$ 10 mm with CT scan, or

≥ 10 mm with calipers by clinical exam. All tumor measurements must be recorded in millimeters (or decimal fractions of centimeters).

Note: Tumor lesions that are situated in a previously irradiated area might or might not be considered measurable. ***If the investigator thinks it is appropriate to include them, the conditions under which such lesions should be considered must be defined in the protocol.***

Malignant lymph nodes: To be considered pathologically enlarged and measurable, a lymph node must be ≥ 15 mm in short axis when assessed by CT scan (CT scan slice thickness recommended to be no grater than 5 mm). At baseline and in follow-up, only the short axis will be measured and followed.

Non-measurable disease: All other lesions (or sites of disease), including small lesions (longest diameter < 10 mm or pathological lymph nodes with ≥ 10 to < 15 mm short axis) are considered non-measurable disease. Bone lesions, leptomeningeal disease, ascites, pleural/pericardial effusions, lymphangitis cutis/pulmonitis, inflammatory breast disease, and abdominal masses (not followed by CT or MRI) are considered as non-measurable.

Note: Cystic lesions that meet the criteria for radiographically defined simple cysts should not be considered as malignant lesions (neither measurable nor non-measurable) since they are, by definition, simple cysts.

'Cystic lesions' thought to represent cystic metastases can be considered as measurable lesions if they meet the definition of measurability described above. However, if non-cystic lesions are present in the same subject, these are preferred for selection as target lesions.

Target lesions: All measurable lesions up to a maximum of 2 lesions per organ and 5 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), be representative of all involved organs, but in addition should be those that lend themselves to reproducible repeated measurements. It may be the case that, on occasion, the largest lesion does not lend itself to reproducible measurement in which circumstance, the next largest lesion which can be measured reproducible should be selected. A sum of the diameters (longest for non-nodal lesions, short axis for nodal lesions) for all target lesions will be calculated and reported as the baseline sum diameters. If lymph nodes are to be included in the sum, then only the short axis is added into the sum. The baseline sum diameters will be used as reference to further characterize any objective tumor regression in the measurable dimension of the disease.

Non-target lesions: All other lesions (or sites of disease) including any measurable lesions over and above the 5 target lesions should be identified as **non-target lesions** and should also be recorded at baseline. Measurements of these lesions are not required, but

the presence, absence, or in rare cases unequivocal progression of each should be noted throughout follow-up.

Methods for Evaluation of Measurable Disease

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.

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 unless the lesion(s) being followed cannot be imaged, but are assessable by clinical exam.

Clinical lesions: Clinical lesions will only be considered measurable when they are superficial (e.g., skin nodules and palpable lymph nodes) and $\geq$ 10 mm diameter as assessed using calipers (e.g., skin nodules). In the case of skin lesions, documentation by color photography including a ruler to estimate the size of the lesion is recommended.

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.

Conventional CT and MRI: This guideline has defined measurability of lesions on CT scan based on the assumption that CT slice thickness is 5 mm or less. If CT scans have slice thickness greater than 5 mm, the minimum size for a measurable lesion should be twice the slice thickness. MRI is also acceptable in certain situations (e.g. for body scans).

Use of MRI remains a complex issue. MRI has excellent contrast, spatial, and temporal resolution; however, there are many image acquisition variables involved in MRI which greatly impact image quality, lesion conspicuity, and measurement. Furthermore, the availability of MRI is variable globally. As with CT, if an MRI is performed, the technical specifications of the scanning sequences used should be optimized for the evaluation of the type and site of disease. Furthermore, as with CT, the modality used at follow-up should be the same as was used at baseline and the lesions should be measured/assessed on the same pulse sequence. It is beyond the scope of the RECIST guidelines to prescribe specific MRI pulse sequence parameters for all scanners, body parts, and diseases. Ideally, the same type of scanner should be used and the image acquisition protocol should be followed as closely as possible to prior scans. Body scans should be performed with breath-holding techniques, if possible.

PET-CT: At present, the low dose or attenuation correction CT portion of a combined PET-CT is not always of optimal diagnostic CT quality for use with RECIST measurements. However, if the site can document that the CT performed as part of a PET-CT is of identical diagnostic quality to a diagnostic CT (with IV and oral contrast),

then the CT portion of the PET-CT can be used for RECIST measurements and can be used interchangeably with conventional CT in accurately measuring cancer lesions over time. Note, however, that the PET portion of the CT introduces additional data which may bias an investigator if is not routine or serially performed.

Ultrasound: Ultrasound is not useful in assessment of lesion size and should not be used as a method of measurement. Ultrasound examinations cannot be reproduced in their entirety for independent review at a later date and because they are operator dependent, it cannot be guaranteed that the same technique and measurements will be taken from one assessment to the next. If new lesions are identified by ultrasound in the course of the study, confirmation by CT or MRI is advised. If there is concern about radiation exposure at CT, MRI may be used instead of CT in selected instances.

Endoscopy, Laparoscopy: The utilization of these techniques for objective tumor evaluation is not advised. However, such techniques may be useful to confirm complete pathological response when biopsies are obtained or to determine relapse in trials where recurrence following complete response (CR) or surgical resection is an endpoint.

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 subject to be considered in complete clinical response. Specific guidelines for both CA-125 response (in recurrent ovarian cancer) and PSA response (in recurrent prostate cancer) have been published [*JNCI* 96:487-488, 2004; *J Clin Oncol* 17, 3461-3467, 1999; *J Clin Oncol* 26: 1148-1159, 2008]. In addition, the Gynecologic Cancer Intergroup has developed CA-125 progression criteria which are to be integrated with objective tumor assessment for use in first-line trials in ovarian cancer [*JNCI* 92:1534-1535, 2000].

Cytology, Histology: These techniques can be used to differentiate between partial responses (PR) and complete responses (CR) in rare cases (e.g. residual lesions in tumor types, such as germ cell tumors where known residual benign tumors can remain).

The cytological confirmation of the neoplastic origin of any effusion that appears or worsens during treatment when the measurable tumor has met criteria for response or stable disease is mandatory to differentiate between response or stable disease (an effusion may be a side effect of the treatment) and progressive disease.

FDG-PET: While FDG-PET response assessments need additional study, it is sometimes reasonable to incorporate the use of FDG-PET scanning to complement CT scanning in assessment of progression (particularly possible 'new' disease). New lesions on the basis of FDG-PET imaging can be identified according to the following algorithm:

a. Negative FDG-PET at baseline with a positive FDG-PET at follow-up is a sign of progressive disease (PD) based on a new lesion.

b. No FDG-PET at baseline and a positive FDG-PET at follow-up: If the positive FDG-PET at follow-up corresponds to a new site of disease

confirmed by CT, this is PD. If the positive FDG-PET at follow-up is not confirmed as a new site of disease on CT, additional follow-up CT scans are needed to determine if there is truly progression occurring at that site (if so, the date of PD will be the date of the initial abnormal FGD-PET scan). If the positive FDG-PET at follow-up corresponds to a pre-existing site of disease on CT that is not progressing on the basis of the anatomic images, this is not PD.

c. FDG-PET may be used to upgrade a response to a complete response (CR) in a manner similar to a biopsy in cases where a residual radiographic abnormality is thought to represent fibrosis or scarring. The use of FDG-PET in this circumstance should be prospectively described in the protocol and supported by disease-specific medial literature for the indication. However, it must be acknowledged that both approaches may lead to false positive CR due to limitations of FDG-PET and biopsy resolution/sensitivity.

Note: A 'positive' FDG-PET scan lesion means one which is FDG avid with an uptake greater than twice that of the surrounding tissue on the attenuation corrected image.

Response Criteria

Evaluation of Target lesions

Response	Evaluation of Target Lesions
Complete Response (CR)	Disappearance of all target lesions. Any pathological lymph nodes (whether target or non-target) must have reduction in short axis to < 10 mm.
Partial Response (PR)	At least a 30% decrease in the sum of diameters of target lesions, taking as reference the baseline sum of diameters.
Progressive Disease (PD)	At least a 20% increase in the sum of diameters of target lesions, taking as reference the *smallest sum on study* (this includes the baseline sum if that is the smallest on study). In addition to the relative increase of 20%, the sum must also demonstrate an absolute increase of at least 5 mm. ***Note:*** the appearance of one or more new lesions is also considered progression.
Stable Disease (SD)	Neither sufficient shrinkage to qualify for PR nor sufficient increase to qualify for PD, taking as reference the smallest sum diameters while on study.

Evaluation of Non-Target lesions

Response	Evaluation of Non-Target Lesions
Complete Response (CR)	Disappearance of all non-target lesions and normalization

	of tumor marker level. All lymph nodes must be non-pathological in size (< 10 mm short axis). ***Note:*** If tumor markers are initially above the upper normal limit, they must normalize for a subject to be considered in complete clinical response.
Non-CR/ Non-PD [Incomplete response/ Stable Disease (SD)]	Persistence of one or more non-target lesion(s) and/or maintenance of tumor marker level above the normal limits.
Progressive Disease (PD)	Appearance of one or more new lesions and/or *unequivocal progression* of existing non-target lesions. *Unequivocal progression* should not normally trump target lesion status. It must be representative of overall disease status change, not a single lesion increase. Although a clear progression of 'non-target' lesions only is exceptional, the opinion of the treating physician should prevail in such circumstances, and the progression status should be confirmed at a later time by the review panel (or Principal Investigator).

Evaluation of 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 progressive disease the smallest measurements recorded since the treatment started). The subject's best response assignment will depend on the achievement of both measurement and confirmation criteria.

For Subjects with Measurable Disease (e.g. Target Disease)

Target lesions	**Non-Target Lesions**	**New Lesions**	**Overall Response**	**Best Overall Response when Confirmation is Required***
CR	CR	No	CR	≥ 4 wks. Confirmation **
CR	Non-CR/Non-PD	No	PR	≥ 4 wks. Confirmation **
CR	Not evaluated	No	PR	
PR	Non-CR/Non-PD/not evaluated	No	PR	
SD	Non-CR/Non-PD/not evaluated	No	SD	Documented at least once ≥ 4 wks from baseline **
PD	Any	Yes or No	PD	No prior SD, PR or CR
Any	PD ***	Yes or No	PD	
Any	Any	Yes	PD	
* See RECIST 1.1 manuscript for further details on what is evidence of a new lesion. ** Only for non-randomized trials with response as primary endpoint.				

*** In exceptional circumstances, unequivocal progression in non-target lesions may be accepted as disease progression.

Note: Subjects with a global deterioration of health status requiring discontinuation of treatment without objective evidence of disease progression at that time should be reported as *"symptomatic deterioration."* Every effort should be made to document the objective progression even after discontinuation of treatment.

For Subjects with Non-Measurable Disease (e.g. Non-Target Disease)

Non-Target Lesion	New Lesions	Overall Response
CR	No	CR
Non-CR/non-PD	No	Non-CR/non-PD *
Not all evaluated	No	Not evaluated
Unequivocal PD	Yes or No	PD
Any	Yes	PD

* 'Non-CR/non-PD' is preferred over 'stable disease' for non-target disease since SD is increasingly used as an endpoint for assessment of efficacy in some trials so to assign this category when no lesions can be measured is not advised.

Duration of Response

Duration of overall response: The duration of overall response is measured from the time measurement criteria are met for CR or PR (whichever is first recorded) until the first date that recurrent or progressive disease is objectively documented (taking as reference for progressive disease the smallest measurements recorded since the treatment started)
The duration of overall CR is measured from the time measurement criteria are first met for CR until the first date that progressive disease is objectively documented.

Duration of stable disease: Stable disease is measured from the start of the treatment until the criteria for progression are met, taking as reference the smallest measurements recorded since the treatment started, including the baseline measurements.

Progression-Free Survival

Progression-Free Survival (PFS) is defined as the duration of time from start of treatment to time of progression or death, whichever occurs first.

Response Review

For trials where the response rate is the primary endpoint, it is strongly recommended that all responses be reviewed by an expert(s) independent of the study at the study's completion. Simultaneous review of the subjects' files and radiological images is the best approach.