# iRECIST: Guidelines for response criteria for use in trials testing immunotherapeutics


## Authors

Prof Lesley Seymour, MD[1], Jan Bogaerts, PhD[2], Andrea Perrone, MD[3], Robert Ford, MD[4], Prof Lawrence H Schwartz, MD[5], Prof Sumithra Mandrekar, PhD[6], Nancy U Lin, MD[7], Saskia Litière, PhD[2], Prof Janet Dancey, MD[1], Alice Chen, MD[8], Prof F. Stephen Hodi, MD[9], Patrick Therasse, MD[10], Prof Otto S Hoekstra, MD[11], Lalitha K Shankar, MD[12], Jedd D Wolchok, MD[13], Marcus Ballinger, PhD[14], Caroline Caramella, MD[15], Prof Elisabeth GE de Vries, MD[16] on behalf of the RECIST Working Group.

[1]Canadian Cancer Trials Group, Queen's University, Kingston, Ontario, Canada
[2]EORTC Headquarters, Brussels, Belgium
[3]Translational Medicine, Merck & Co, Kenilworth, New Jersey, USA
[4]Clinical Trials Imaging Consulting, LLC, New Jersey, USA
[5]Department of Radiology, Columbia University Medical Center, New York, New York, USA, New York Presbyterian Hospital New York, New York, USA
[6]Division of Biomedical Statistics and Informatics, Mayo Clinic, Rochester, Minnesota, USA
[7]Department of Medical Oncology, Dana-Farber Cancer Institute, Harvard Medical School, Boston, Massachusetts, USA
[8]Early Clinical Trials Development Program, Division of Cancer Treatment and Diagnosis, National Cancer Institute, Bethesda, Maryland, USA
[9]Department of Medical Oncology, Dana-Farber Cancer Institute, Harvard Medical School, Boston, Massachusetts, USA
[10]Institut de Recherche International Servier, Paris, France
[11]Department of Radiology and Nuclear Medicine, VU University Medical Center, Amsterdam, The Netherlands
[12]Diagnostic Imaging Branch, National Cancer Institute, Bethesda, Maryland, USA
[13]Department of Medicine, Memorial Sloan Kettering Cancer Center, New York, New York, USA; Weill Cornell Medical and Graduate Colleges, New York, New York, USA; Ludwig Institute for Cancer Research, New York, New York, USA
[14]Genentech Inc., South San Francisco, California, USA
[15]Department of Radiology, Gustav Roussy Cancer Campus, Villejuif, France
[16]Department of Medical Oncology, University Medical Center Groningen, Groningen, The Netherlands

**Correspondence to:**
Prof. Lesley Seymour,
Canadian Cancer Trials Group,
Queen's University,
10 Stuart Street, Kingston,
Ontario K7L 3N6, Canada
lseymour@ctg.queensu.ca
Tel number: 00-1-613 5336430


WORD COUNT: 4783

## Abstract

Tumours respond differently to immunotherapeutics compared to chemotherapeutic drugs, raising questions about the assessment of changes in tumour burden, a mainstay of evaluation of cancer therapeutics to inform objective response and disease progression. A consensus guideline, iRECIST, was developed for the use of modified Response Evaluation Criteria in Solid Tumours (RECIST, V1.1) in cancer immunotherapy trials, to ensure consistent design and data collection and facilitate the ongoing collection of trial data and ultimate validation.

The RECIST Working Group held conference calls and meetings to discuss plans for warehouse creation to validate iRECIST. Key questions were identified, and issues and concerns with response evaluation in immunotherapeutic trials were defined. At a formal kick-off meeting, attendees were asked to provide details of their current approach to evaluating response in immunotherapeutics trials. Thereafter a consensus guideline was drafted and monthly meetings served to develop, review and agree on the guideline and plans for validation.

The guideline describes a standard approach to solid tumour measurements and definitions for objective change in tumour size for use in trials where an immunotherapeutic is used. In addition, it defines the minimum data to be collected for future trials and those currently in development, to facilitate the compilation of a data warehouse to be used to later validate iRECIST.

An unprecedented number of trials have been conducted, initiated or are planned testing new immune modulators for cancer therapy using a variety of modified response criteria. This guideline, developed by a multidisciplinary group including academic, commercial and regulatory members, will allow consistent conduct, interpretation and analysis of trials of immunotherapies. RECIST 1.1 should continue to be used as the primary criteria for response based endpoints for randomised studies planned for licensing applications; iRECIST should be considered exploratory in such trials, although earlier phase trials may consider using primarily iRECIST.

**Funding:**

This publication was supported by the Canadian Cancer Society Research Institute (grant #021039), the EORTC Cancer Research Fund and the NCI grant number 5U10-CA11488-45.

WORD COUNT: 300/300

## Introduction

Changes in tumour burden are frequently used as surrogates of survival / quality of life.[1] Validated and consistent criteria are critical. In 2000 the RECIST (Response Evaluation Criteria in Solid Tumours) Working Group (RWG) simplified the 1981 World Health Response Criteria (WHO)[2] after validation in a large data warehouse.[3] In 2009 RECIST was refined (RECIST 1.1).[4] The RWG ensures RECIST undergoes continuous testing, validation and updates.[5–8]

Immune modulators (IMs) are one of the most significant classes of new anticancer therapeutics.[9–11] Cytotoxic T-lymphocyte antigen-4 (CTLA-4), programmed death-1 (PD-1) and programmed death ligand-1 (PD-L1) pathways are the most actively studied[12–18] and agents active on those pathways have recently received marketing authorisation (in some cases conditional pending the completion of other studies) for melanoma, lung, bladder, renal and head and neck cancers.[19–24] The novel mechanism of action of these agents, with immune and T cell activation, is postulated to lead to unusual patterns of response, which appeared more pronounced and more frequent than had been described before (such as tumour 'flare'). In early trials of immune based therapeutics in melanoma, investigators described unique response patterns, termed 'pseudoprogression' (PSPD). Some patients whose disease met the criteria for disease progression based on traditional response criteria such as RECIST (with an increase in the sum of measures of target lesions, unequivocal increase in non-target disease or the appearance of new lesions) were noted to have late, but deep and durable, responses.[25–29] Modified response criteria were proposed, based on WHO criteria, (which collect bi-dimensional measurements of target lesions) - the so-called immune-related response criteria (irRC).[30] The major modification involved the inclusion of the measurements of new target lesions (≥ 5 × 5 mm up to 5 new lesions per organ: 5 new cutaneous lesions and 10 visceral lesions) into disease assessments. More recently, researchers published revised irRC using unidimensional measurements, based on the original RECIST.[31] Subsequent recommendations, some

published in abstract form, appear to incorporate RECIST 1.1 recommendations.[32–34] These are often referred to as irRECIST, but have not always been consistently applied leading to concerns regarding comparability of data and results across trials, difficulty with pooling databases, and a lack of clarity whether new lesions were measured, if so, how many were captured, and whether measures were incorporated into tumour burden. Current and recent trials have generally used RECIST based immune criteria.

Because of the need to standardize and validate response criteria, the RECIST Working Group prospectively planned to create a warehouse of data from trials of immunotherapeutics to test and validate RECIST 1.1 and suggest modifications if required. During the planning and initial collection of the immunotherapeutic warehouse, it was apparent that most trials testing these agents have typically used RECIST 1.1 to define the primary and secondary efficacy based endpoints, and reserved irRC or their modified definition of RECIST for exploratory endpoints.[32,33] In addition, as noted above, and described in figure 1, there was considerable variability across clinical trials, including within pharmaceutical companies as well as cooperative groups, as to which criteria were used leading to serious concerns about interpretation of pooled datasets. Finally, virtually all trials used immune-modified criteria utilised independent imaging review by a commercial entity for those criteria, rather than investigator assessments. We believe that response criteria should be applicable across all cancer clinical trials, including those conducted in the academic sector where costly independent review is not feasible.

Based on those observations it was decided to develop a Guideline describing the use of a modified RECIST to ensure consistent design and data collection that would facilitate the ongoing collection of clinical trial data and ultimate validation, if indicated, of a modified RECIST 1.1 for immune based therapeutics (termed iRECIST). These guidelines are not intended to define or guide clinical practice or

treatment decisions, but rather to provide a consistent framework for the management of data collected in clinical trials of immune based therapies. Treatment decisions rest with the patient and their healthcare team.

## Terminology

iRECIST is based on RECIST 1.1. Responses assigned using iRECIST have a prefix of "i" (i.e. immune complete or partial response (iCR; iPR)) and progression as unconfirmed and confirmed (iUPD, iCPD) in order to differentiate them from responses assigned using RECIST 1.1. Similar nomenclature is used for stable disease (iSD). New lesions are evaluated and subcategorised into those that qualify as target lesions (New Lesion Target (NLT), while all others are referred to as non-target (NL non-target; NLNT). Response assessments are referred to as time-points, while best overall response is unchanged.

## Development of the Guideline

The RECIST Working Group formed a subcommittee and held a series of conference calls and face-to-face meetings in 2015 and 2016 to discuss plans for the development and validation of iRECIST and to review current approaches to evaluating response in IM trials, to identify points of consensus and items that required further discussion. Members of the subcommittee included clinical, statistical and imaging experts in methodology and immunotherapy, as well as representatives from the pharmaceutical companies developing immunotherapeutics and key regulatory authorities (supplementary materials: table S1). In June 2016 a formal meeting was held in Chicago, with invited presentations from regulatory authorities, pharmaceutical companies with IM agents in development and academic groups, followed by a structured discussion. Prior to the meeting the 52 invited participants were polled in order to initiate the identification of questions to be addressed as well as the response criteria in use. Ten respondents provided responses prior to the meeting while all 8 presenters identified additional areas of interest. After

review and discussion during the meeting, a list of critical questions to be addressed by iRECIST were identified. Those key questions are shown in panel 1. Interestingly, all participants confirmed that RECIST 1.1 was used for primary endpoints, with immune modified response criteria being used in an exploratory fashion, with very few exceptions; in one instance immune modified criteria were used as a co-primary endpoint. The most commonly used immune modified criteria were variations of irRECIST. There was more variability in independent imaging review and the period of time that response data were collected after RECIST 1.1 progression and or cessation of protocol therapy. Further calls and meetings were held to develop and plan the later full validation of iRECIST (figure 1).

***Search strategy and selection criteria***

This paper describes a consensus guideline, rather than a literature review. However, a database search was conducted using PubMed with the following search terms: immune response criteria (limited to cancer, clinical trials and English; 234 citations), irRC (23 citations) and pseudoprogression (limited to cancer, clinical trials and English; 39 citations).

## iRECIST

The continued use of RECIST 1.1 is recommended to define whether tumor lesions, including lymph nodes, are measurable or non-measurable, as well as the management of bone lesions, cystic lesions and lesions with prior local treatment (such as radiation) (table 1). Similarly, there are no changes to the recommendations regarding the method of measurement, although it is recognized that clinical examination and chest X-ray are rarely used with available modern imaging techniques. The principles used to determine objective tumor response are largely unchanged from RECIST 1.1, while a major change of iRECIST is the concept of 'resetting the bar' if RECIST 1.1 progression is followed at the next assessment by tumor shrinkage.

iRECIST defines iUPD based on RECIST 1.1 principles; however iUPD requires confirmation; confirmation is based on observing either further increase in size (or in the number of new lesions) in the lesion category (i.e. target, non-target disease) where progression was first identified, or progression (defined by RECIST 1.1) in lesion categories that had not previously met RECIST 1.1 progression criteria. If, however, progression is not confirmed as described above, but instead tumour shrinkage (compared to baseline) meeting the criteria of iCR, iPR or iSD, then the bar is reset so that iUPD must occur again (compared to nadir values) and then be confirmed (by further growth) at the next assessment for iCPD to be assigned. If there is no change in tumour size/extent from iUPD, then the time-point response (TPR) would again be iUPD. This approach allows atypical responses, such as delayed responses that occur after PSPD, to be identified, further understood and better characterized (tables 1-3 and S2 and figure 2). Sample case record forms and protocol sections are included in Appendix 2.

Below, sections of RECIST 1.1 that are unchanged are only briefly summarized; readers should refer to RECIST 1.1 for full descriptions.

***Evaluation of target, non-target and new lesions***

Most RECIST 1.1 recommendations are unchanged for TPR including the management of lymph nodes, lesions that become too small to measure, lesions that split or coalesce and the definition of CR, PR, SD and PD. Each TPR is based on the evaluation of target lesions, non-target lesions and new lesions.

For target lesions, iCR, iPR and iSD can all be assigned after iUPD has been documented, providing that iCPD was not confirmed. iUPD is defined by RECIST 1.1 criteria for PD; iUPD may be assigned multiple times as long as iCPD is not confirmed at the next assessment. PD is confirmed in the target

lesion category if the next imaging assessment, (4 weeks but no more than 8 weeks later) after iUPD confirms further increase in sum of measures (SOM) of target disease from iUPD, with an increase of at least 5 mm. See below for conditions under which iCPD would also be declared based on new lesions and or non-target (NT) lesions.). However, the criteria for iCPD (after iUPD) are not considered to have been met if CR, PR or SD criteria (compared to baseline and as defined by RECIST 1.1) are met at the next assessment after iUPD. The status is 'reset' (unlike RECIST 1.1 where any PD precludes later CR, PR or SD). iCR, iPR or iSD should then be assigned. If no change is detected, then the TPR is iUPD.

The evaluation of non-target lesions at each TP follows similar principles. iUPD (but not iCPD) may have been documented prior to iCR or Non-iCR/non-iUPD and may be assigned multiple times providing iCPD was not confirmed. iUPD is defined by RECIST 1.1 criteria; however, iUPD may be assigned multiple times as long as iCPD is not confirmed at the next assessment. PD in the NT lesion category is confirmed if subsequent imaging, conducted at least 4 weeks (but no more than 8 weeks) after iUPD shows further increase from iUPD. (See above and below for conditions under which iCPD would also be declared for target and new lesions). The criteria for iCPD are not considered to have been met if RECIST 1.1 defined CR or non-iCR/non-iUPD criteria are met after a prior iUPD. The status is 'reset' (unlike RECIST 1.1) and iCR, or non-iCR/non-iUPD is assigned; if no change is detected the TPR is iUPD.

RECIST 1.1 defines the appearance of new malignant lesions as denoting true disease progression, providing that other lesions (artefacts or benign intercurrent disease) are appropriately evaluated and discounted if not malignant. These principles of RECIST 1.1 remain useful and clearly identify the management of new lesions which are considered to be potentially artefactual: "If a new lesion is equivocal, for example because of its small size, continued therapy and follow-up evaluation will clarify if it represents truly new disease. If repeat scans confirm there is definitely a new lesion, then progression

should be declared using the date of the initial scan".

However, many aspects of new lesion assessment are unique to iRECIST. If a new lesion is identified (thus meeting the criteria for iUPD), and the patient is clinically stable, therapy should be continued (see below). New lesions should be assessed and categorized as measurable or non-measurable using RECIST 1.1 principles. Five lesions, no more than two per organ, should be measured and recorded as New Lesions-Target (NLT), but should NOT be included in SOM of the original target lesions identified at baseline (Supplementary materials: Appendix 1 and 2). Other measurable and non-measurable lesions are recorded as New Lesion-Non-Target (NLNT). Trialists may choose to measure and record more than 5 new lesions for research purposes, but this is not felt to be practical for general usage. New lesions do not need to meet the criteria for NLT in order to result in iUPD (or iCPD); NLNT can also drive iUPD or iCPD. PD is confirmed (iCPD) in the New Lesion category if the next imaging assessment, conducted at least 4 weeks (but more than 8 weeks) after iUPD, confirms additional new lesions or further increase in new lesion size from iUPD (SOM increase in NLT $\geq$ 5 mm, any increase for NLNT).

Note that if iUPD criteria were met based on progression in one of: T or NT disease or the appearance of new lesions, then RECIST 1.1 defined progression in another lesion category in the confirmatory scan also confirms iCPD.

***Continued treatment after iUPD***

As noted earlier, the literature describes 'PSPD', with an increase in the size of lesions, or the visualization of new lesions, followed by a deep response, including complete response, that may be durable. Although well described, differentiating transient PSPD from true progression requiring a change in therapy can be challenging. Although early discontinuation of an effective agent is not desirable,

continued long term treatment with a non-effective agent past true progression may delay the initiation of potentially effective salvage therapy.

Because of this dilemma, we recommend that clinical trials where treatment past initial RECIST 1.1 defined progression (i.e iUPD) is permitted only allow patients who are clinically stable to continue on treatment until the next assessment (at least 4 weeks later); the next imaging assessment should be no longer than 8 weeks later to ensure patients remain fit for salvage therapies. A longer time frame before the next assessment may be reasonable if PSPD is well described in the tumour type (for example melanoma treated with a CTLA4 inhibitor) especially if no effective salvage therapies are available (for example, BRAF wild type melanoma) but must be justified in the protocol. All decisions regarding continuing or discontinuing therapy rest with the patient and their health care provider; iRECIST describes what data are to be collected, submitted and analysed in clinical trials of immune based therapies.

Clinical stability requires no worsening of performance status, no clinically relevant increase in disease related symptoms such as pain or dyspnoea felt related to disease progression (generally understood to mean a requirement for increased palliative intervention as below) and no requirement for intensified management of disease related symptoms including increased analgesia, radiation or other palliative care.

The imaging findings and the recommendation to continue with treatment despite iUPD must be discussed with the patient prior to a decision to continue treatment being taken. Patients who have iUPD and are not clinically stable should be designated as ‘not clinically stable’ in the CRF. This will allow the BOR to be calculated and the date of iUPD to be used in estimations of progression free survival.

If the 'confirmatory' scan done confirms iCPD, but the investigator/patient believes that continued treatment is appropriate, imaging should continue, and data collected to allow further elucidation of tumour growth dynamics with IMs. For the same reason, and if feasible, it is recommended that even patients who discontinue therapy for iCPD should continue to have disease assessments performed until the start of other systemic or local therapies.

***Time-point and best overall response***

Although the principles of the assignment of the TPR and best overall response (BOR) closely follow RECIST 1.1, and reflect assessment of target and NT lesions as well as the presence of new lesions, the possibility of PSPD adds complexity (tables 1-3 and S2 and panel 2). The TPR is calculated using the response assigned for each category of lesion described above (as for RECIST 1.1), but takes into account the last TPR.

The algorithm for patients with no prior iUPD is identical to RECIST 1.1. For patients with iUPD at the last TPR, the next TPR is dependent on: the status of all lesions, including T, NT, NLT and NLNT; on whether any increase in size has occurred (either further increase in size OR sufficient to assign a new iUPD where the criteria were not previously met); or the appearance of additional new lesions.

For iRECIST, the iBOR is the best TPR response recorded from the start of the study treatment until the end of treatment taking into account any requirement for confirmation. iUPD will not override a subsequent BOR of iSD, iPR or iCR (tables 1-3 and S2). This also means that iPR or iSD can be assigned (TPR or iBOR) even if new lesions have not regressed, or if unequivocal progression (NT lesions) remains unchanged, providing that the criteria for iCPD are not met.

Confirmation of response is not required when using RECIST 1.1, except in non-randomised trials, and this is also recommended for iRECIST. The duration of iCR and iPR is from the time the criteria for iCR or iPR are first met while the duration of iSD is still calculated from baseline.

The protocol should define how missing response assessments will be handled. It is recommended that assessments that are not done (ND), or are not evaluable (NE) should be disregarded. In that scenario an iUPD followed by ND or NE and then another unconfirmed PD would be indicative of iCPD. Protocols should also be clear whether assessments done after protocol therapy is discontinued can be considered in determination of iBOR; it may be reasonable to include assessments done some weeks or months after protocol treatment has been discontinued if late responses are anticipated (such as with a CTLA4 inhibitor) and patients have not received other systemic or local therapies. Protocols must also specify how any new therapy (such as radiation or surgery), introduced before progression will affect iBOR designation. Other RECIST 1.1 recommendations, including the management of missing assessments remain unchanged, including requiring that the analysis plan must address how missing data/assessments will be addressed in the determination of response and progression.

***Frequency of tumour re-evaluation***

In general, follow up response assessment every 6 to 12 weeks is recommended, depending on the frequency of treatment visits - as recommended for RECIST 1.1. The protocol should specify which anatomic locations are evaluated at baseline and follow up and whether bone scans should be repeated each response assessment or only to confirm iPR, iCR, or as clinically indicated. For all trials, especially comparative ones, response assessments should be performed on a calendar schedule and not be impacted by delays in therapy or the requirement for earlier confirmatory scans.

Tumour re-evaluation may be performed earlier than originally planned (but not less than 4 weeks or longer than 8 weeks after iUPD) to confirm iUPD (or to confirm iCR or iPR 4 or more weeks later in non-randomised trials). If progression is not confirmed, re-evaluation should continue as originally planned (i.e. if scans were to be done at 8, 16, 24 weeks, and a scan was done at 12 weeks to confirm response, then the next scans should be performed at 16 weeks, as planned). If patients continue on protocol treatment after iCPD, assessments should continue to be performed, in the same planned schedule, until protocol treatment is discontinued.

Ideally, all imaging performed after protocol treatment has been discontinued should continue to be recorded on the CRF until subsequent therapies are initiated, and the protocol and informed consent document written to facilitate this. These data will allow further refinement of iRECIST.

***Statistical and protocol considerations***

The event date to be used for calculation of progression free survival (iPFS) should be the first date that progression criteria are met (i.e. the date of iUPD) providing that iCPD is confirmed at the next assessment (table S2). If iUPD occurs, but is disregarded because of later iSD, iPR or iCR, that iUPD date should not be used as the progression event date.

If progression is not confirmed and there is no subsequent iSD, iPR or iCR then the iUPD date should still be used in the following scenarios: if the patient stops protocol treatment because they were not considered to be clinically stable, or no further response assessments are done (patient refusal or protocol non-compliance or patient death); the next TPRs are all iUPD, and iCPD never occurs; or the patient dies of cancer. The CRF collects the reason why confirmatory response assessment was not performed at any TP, such as ‘not clinically stable’, ‘centre error’ or ‘patient refusal or death’.

For protocols which permit crossover, or where intermittent schedules are being tested, the protocol should clearly specify whether iUPD or iCPD would be used for a treatment decision leading to crossover and how data subsequent to crossover will be managed and analysed. In general, we suggest iCPD be used especially for scenarios with immunotherapy in both arms and where PSPD is anticipated.

Adjuvant trials of IMs are ongoing but have yet to report. Suspected new lesions in the curative setting should always be investigated thoroughly and preferably biopsied prior to the designation of relapse being assigned. If biopsy is not technically feasible, then it would appear to be reasonable to follow the principles of iRECIST with a follow-up scan to confirm relapse in patients who are clinically stable.

The collection of anonymised imaging (even if centralized blinded review of imaging studies is not planned) is recommended for all studies using an imaging based endpoint (i.e. response or PFS) if feasible. Although the iRECIST guideline requires the recording of the measurements of up to 5 new lesions, it may be eventually necessary to record additional lesions to have a more precise estimation of progression. Central collection of images will allow further evaluation by an independent radiologist if required. If real time central review is planned, the protocol must clearly define how treatment decisions will be made.

It is recommended that phase III clinical trials continue to incorporate both RECIST 1.1 and iRECIST (see table 1 for comparisons) and that RECIST 1.1 should continue to be used to define the primary efficacy outcomes (PFS, PD, BOR). Exploratory analyses using the iPD date (i.e. the first date of iUPD which is subsequently confirmed) can be defined in the statistical analysis plan. Earlier trials may consider using iRECIST as the primary criteria. The protocol should carefully explain which will be the

primary criteria used to assess response, and which would be exploratory. This is particularly important for trials which compare an IM arm to a non-IM arm.

## Discussion: next steps and validation

Immunotherapeutics are a major advance in the treatment of an increasing number of cancers. The increasing testing and use of these agents in multiple clinical settings, including adjuvant, first, second and subsequent lines of therapy will require the use of progression based endpoints. RECIST 1.1 may not always adequately capture the unique patterns of response that have been well described in clinical trials of these agents in a low percentage of patients, typically reported as 10% or less, mainly reported in melanoma studies.[33-35] The true frequency in trials of other malignancies (including non-small cell lung cancer) is unclear as the majority of trials have reported RECIST 1.1 based response rates,[36] but may be less common based on anecdotal reports. Similarly, it is unknown whether this pattern is unique to CTLA4 / PD-1/PD-L1 pathway active drugs or will be reported with emerging immunotherapeutics. Trials testing immunotherapeutics in combination with standard therapies, especially when they are compared to standard therapies, further confound the assessment of progression based endpoints.

RECIST 1.1 already addresses the management of 'equivocal' progression including suspected new lesions, which may explain, at least in part, the continued use of RECIST 1.1 to define response based primary endpoints. RECIST 1.1 deals with mainly technical differences in scans that give the appearance there may be new lesions or the concept of the isodense lesions at baseline that becomes more visible after the start of therapy since it becomes internally more necrotic as opposed to a true new lesion. However, the intent was never to use those recommendations to manage 'PSPD' described with IMs.

Although modified response criteria have been used, it is clear that a need exists for a formal guideline, with robust plans for prospective testing, and consistent data collection and validation. Recent and current trials have not always been consistent in the definition of the response criteria to be used, have used trial specific modifications of response criteria where new lesion measurements may or may not be included in

the assessment of response, and response assessments after progression defined by RECIST 1.1 are not always collected. Those data are critical to understand the dynamics of tumour response to immunotherapeutics, including whether immunotherapeutics with different mechanisms of action have different effects.

Although some progress has been made in understanding tumour dynamics with immunotherapeutics, progress in the field has undoubtedly been limited by limited data sharing across trials, companies and immunotherapeutics. Publications have been based on trials conducted by individual pharmaceutical companies or commercial organizations. In the development of this guideline, virtually all major pharmaceutical companies developing immunotherapeutics participated and have shared their experiences, protocols, response criteria, and most critically, their data. The iRECIST team also included members of the European Medicines Agency and the Food and Drug Administration.

Although this guideline is consensus based, it is not yet validated as the data warehouse is presently being created with initial trial data already in place. It includes all current knowledge on response dynamics, allowing appropriate management of true 'PSPD' but importantly also safeguards patients. Although PSPD is now well described, it remains unusual, occurring in less than 1 in 10 patients. Treatment past radiographic progression may be appropriate only in a small number of patients, and the continuation of treatment past true progression may limit subsequent effective therapies in the patient is no longer fit enough to tolerate.

iRECIST requires the confirmation of progression in order to rule out – or confirm – PSPD. While this recommendation is in keeping with that of RECIST 1.1 to continue treatment and repeat imaging in the case of a mixed response or equivocal findings, if PSPD is common, patients may be exposed to a higher

risk (from continuing ineffective therapy or increased exposure to radiation) or cost (for the potentially ineffective therapy or the costs of imaging). We recommend these criteria are used for clinical trial protocols rather than to guide clinical practice. Treatment past RECIST 1.1 based progression should only be considered in carefully selected scenarios, when the patient is stable (or improving) symptomatically and where there is a short period before reassessment. We believe this is a reasonable balance.

Although at first glance the recommendation to collect measurements of new lesions as defined in this guideline seems onerous, the collection of these measurements and the recording of both RECIST 1.1 and iRECIST for TPR and BOR has several advantages. The relationship between the site of the new lesion and PFS as well as the value of adding new lesion measurements to the SOM can be explored. Continuing to record RECIST 1.1 allows comparison to reported immunotherapy trials (which used RECIST 1.1), as well as chemotherapy trials in the same setting, while at the same time allowing treatment past progression as well as collecting data that will allow further testing and validation of iRECIST. Differences in trial outcomes using RECIST 1.1 versus iRECIST may occur, and the interpretation will be informative. Our proposed plan will enable identification of such situations, and hopefully clarification of underlying mechanisms. Also, in the future it will be possible to quantify the differences in outcome estimation between RECIST 1.1 and iRECIST enabling better informed decisions for future RECIST changes.

This strategy will also be useful for trials comparing immunotherapy to non-immunotherapy based therapeutics. RECIST 1.1 and iRECIST should yield almost identical results for non-immunotherapy treatments, based on the current RECIST warehouses; while an IM warehouse and associated sensitivity analysis of endpoints will permit the quantification of potential added benefit for the immunotherapy component. While comparison of iRECIST in such situations incorporates an element of bias by

construction, confirmation by overall survival results may gain additional importance.

Our current recommendation for the design of randomised studies planned for licensing applications is to continue to use RECIST 1.1 as the primary criteria for response based endpoints. iRECIST should be considered exploratory in such trials, although earlier phase trials may consider using primarily iRECIST.

The creation of a data warehouse is underway, while the implementation of this guideline, and the continued sharing of anonymised, patient level data will allow the formal validation of iRECIST, ensuring that response based guidelines remain robust and enable the rapid and robust future development of new cancer therapeutics to better treat our patients.

## Contributors

All authors contributed to the literature search and writing of the report.

## Declaration of interests

LS reports grants from AstraZeneca and Merck, outside the submitted work; JB has nothing to disclose; AP reports other from Merck & Co, outside the submitted work; RF reports personal fees from Amgen, Abbvie, Aptiv, Aragon, BMS, Bioclinica, Celldex, Celsion, Clovis, Covance, Biomedical systems, ACR Image Metrix, Exelixis, Genentech, Janssen, Kyowa, Loxo, ICON Medical Imaging, Eisai, EMD Serono, Imaging Endpoints, Mlrati, Celgene, Merck, Novartis, Novocure, Roche, Pfizer, Quintile, Tokai, ONO, Red Hill, Radiant Sage, Orbimed Advisors, Cinven, Virtualscopics, others not listed, outside the submitted work; LHS reports grants and other from Novartis, grants from Astellas, Eli Lilly, Merck, Pfizer, other from GSK, outside the submitted work; SM has nothing to disclose; NUL has nothing to disclose; SL has nothing to disclose; JD reports grants from Merck, AstraZeneca, Pfizer, Ottawa Hospital Research Institute, Novartis, outside the submitted work; AC has nothing to disclose; FSH reports grants from Bristol-Myers Squibb, personal fees from Merck, Novartis, Genentech and EMD Serono, outside the submitted work. In addition, FSH has a patent MICA Related Disorders with royalties paid, and a patent Tumor Antigens and Uses Thereof issued; PT has nothing to disclose; OSH has nothing to disclose; LKS has nothing to disclose; JDW reports grants from Bristol Myers Squibb, Merck, Genentech, during the conduct of the study; other (consultant) from Bristol Myers Squibb, Merck and other (SAB) from Medimmune, outside the submitted work; MB reports personal fees from Genentech, other from Roche, outside the submitted work; CC reports personal fees from Roche, BMS, AstraZeneca, outside the submitted work; EGEdV reports other from Synthon, Medivation and Merck, grants from Novartis, Amgen, Roche/Genentech, Servier, Chugai, Synthon, AstraZeneca, Radius Health, outside the submitted work.

**Acknowledgments**

We gratefully acknowledge the thoughtful participation of the following in this initiative:

- Patricia Keegan, MD, member of the Food and Drug Administration
- Francesco Pignatti, European Medicine Agency
- Wendy Hayes, Bristol-Myers Squibb, Princeton, New Jersey, USA
- Eric Rubin, Merck & Co, Kenilworth, New Jersey, USA

We also received written comments from Darragh Halpenny, Jean-Yves Blay, Florian Lordick, Silke Gillessen, Hirokazu Watanabe, Jose Pablo Maroto Rey, Pietro Quaglino, Howard Kaufman, Denis Lacombe, Corneel Coens, Catherine Fortpied, Jessica Menis, Francisco Vera-Badillo, Jean Powers, Michail Ignatiadis, Eric Gauthier, Michael O'Neal, Caroline Malhaire, Laure Fournier, Glen Laird.

We thank Anouk Funke for her assistance with this manuscript.

Drs Chen and Shankar are employees of the National Institutes of Health/National Cancer Institute.

This paper has not been submitted or published elsewhere but has been presented in part at the EORTC-NCI-AACR 2016 Meeting (Munich, November 29 – December 2, 2016).

## References


1 Kim C, Prasad V. Cancer drugs approved on the basis of a surrogate end point and subsequent overall survival: An analysis of 5 years of US Food and Drug Administration approvals. *JAMA Intern Med* 2015; **175**: 1992–4.

2 Miller AB, Hoogstraten B, Staquet M, Winkler A. Reporting results of cancer treatment. *Cancer* 1981; **47:** 207–14.

3 Therasse P, Arbuck SG, Eisenhauer EA, et al. 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; **92**: 205–16.

4 Eisenhauer EA, Therasse P, Bogaerts J, et al. New response evaluation criteria in solid tumours: revised RECIST guideline (version 1.1). *Eur J Cancer* 2009; **45**: 228–47.

5 Liu Y, Litière S, de Vries EG, et al. The role of response evaluation criteria in solid tumor in anticancer treatment evaluation: results of a survey in the oncology community. *Eur J Cancer* 2014; **50**: 260–6.

6 Schwartz LH, Seymour L, Litière S, et al. RECIST 1.1 - Standardisation and disease-specific adaptations: Perspectives from the RECIST Working Group. *Eur J Cancer* 2016; **62**: 138–45.

7 Schwartz LH, Litière S, de Vries EG, et al. RECIST 1.1 – IUPDate and clarification: From the RECIST committee. *Eur J Cancer* 2016; **62**: 132–7.

8 http://www.eortc.org/recist/

9 Brahmer J, Reckamp KL, Baas P, Crinò L, et al. Nivolumab versus docetaxel in advanced squamous-cell non–small-cell lung cancer. *N Engl J Med* 2015; **373**: 123–35.

10 Pardoll, Drew M. The blockade of immune checkpoints in cancer immunotherapy. *Nat Rev Cancer* 2012:**12**:252–264.

11 Hahn NM. Immune therapy in genitourinary cancers: the beginning of the end or the end of the beginning? Presented at the American Society of Clinical Oncology (ASCO) Annual meeting 2015.

12 Holt GE, Podack ER, Raez LE. Immunotherapy as a strategy for the treatment of non-small cell lung cancer. *Therapy* 2011; **8**: 43–54.

13 Sharma P, Wagner K, Wolchok JD, Allison JP. Novel cancer immunotherapy agents with survival benefit: recent successes and next steps. *Nat Rev Cancer* 2011; **11**: 805–12.

14 Antonia S, Goldberg SB, Balmanoukian A, et al. Safety and antitumor activity of durvalumab plus tremelimumab in non-small cell lung cancer: a multicentre, phase 1b study. *Lancet Oncol* 2016; **3**: 299–308.

15 Postow MA, Chesney J, Pavlick AC, et al. Nivolumab and ipilimumab versus ipilimumab in untreated melanoma. *N Engl J Med* 2015; **372**: 2006–17.

16 Reck M, Bondarenko I, Luft A, et al. Ipilimumab in combination with paclitaxel and carboplatin as first-line therapy in extensive-disease-small-cell lung cancer: results from a randomized, double-blind, multicenter phase 2 trial. *Ann Oncol* 2013; **24**: 75–83.

17 Lynch TJ, Bondarenko I, Luft A, et al. Ipilimumab in combination with paclitaxel and carboplatin as first-line treatment in stage IIIB/IV non-small cell lung cancer: results from a randomized, double-blind, multi-center phase II study. *J Clin Oncol* 2012; **30**: 2046–54.

18 Larkin J, Chiarion-Sileni V, Gonzalez R, et al. Combined nivolumab and ipilimumab or monotherapy in untreated melanoma. *N Engl J Med* 2015; **373**: 23–34.

19 Hodi FS, O'Day SJ, McDermott DF, et al. Improved survival with ipilimumab in patients with metastatic melanoma. *N Engl J Med* 2010; **363**: 711–23.

20 Topalian SL, Hodi FS, Brahmer JR, et al. Safety, activity, and immune correlates of anti-PD-1 antibody in cancer. *N Eng J Med* 2012; **366**: 2443–54.

21 Borghaei H, Paz-Ares L, Horn L, et al. Nivolumab versus docetaxel in advanced nonsquamous non-small-cell lung cancer. *N Engl J Med* 2015; **373**: 1627–39.

22 Garon EB, Rizvi N, Hui R, et al. Pembrolizumab for the treatment of non-small cell lung cancer. *N Engl J Med* 2015; **372**: 2018–28.

23 Rosenberg JE, Hoffman-Censits J, Powles T, et al. Atezolizumab in patients with locally advanced and metastatic urothelial carcinoma who have progressed following treatment with platinum-based chemotherapy: a single-arm, multicentre, phase 2 trial. *Lancet* 2016; **387**: 1909–20.

24 Motzer RJ, Escudier B, McDermott DF, et al. Nivolumab versus everolimus in advanced renal-cell

carcinoma. *N Engl J Med* 2015; **373**: 1803–13.
25 Fife BT, Bluestone JA. Control of peripheral T-cell tolerance and autoimmunity via CTLA-4 and PD-1 pathways. *Immunol Rev* 2008; **224**: 166–182.
26 Fife BT, Pauken KE, Eagar TN, et al. Interactions between PD-1 and PD-L1 promote tolerance by blocking the TCR-induced stop signal. *Nat Immunol* 2009; **10**: 1185–92.
27 Iwai Y, Ishida M, Tanaka Y, et al. Involvement of PD-L1 on tumor cells in the escape from host immune system and tumor immunotherapy by PD-L1 blockade. *Proc Natl Acad Sci USA* 2002; **99**: 12293–97.
28 Tarhini AA. Tremelimumab: a review of development to date in solid tumors. *Immunotherapy* 2013; **5**: 215–29.
29 Hoos A, Parmiani G, Hege K, et al. A clinical development paradigm for cancer vaccines and related biologics. *J Immunother* 2007; **30**: 1–15.
30 Wolchok JD, Hoos A, O'Day S, et al. Guidelines for the evaluation of immune therapy activity in solid tumors: immune-related response criteria. *Clin Cancer Res* 2009;**15**:7412–20.
31 Nishino M, Giobbie-Hurder A, Gargano M, Suda M, Ramaiya NH, Hodi FS. Developing a common language for tumor response to immunotherapy: Immune-Related Response Criteria using unidimensional measurements. *Clin Cancer Res*. 2013; **19**: 3936–43.
32 Bohnsack O, Ludajic K, Hoos A. Adaptation of the immune-related response criteria: irRECIST. *Ann Oncol* 2014; **25 (suppl 4)**: iv361–iv372.
33 Hodi FS, Hwu WJ, Kefford R, et al. Evaluation of Immune-Related Response Criteria and RECIST v1.1 in patients with advanced melanoma treated with pembrolizumab. *J Clin Oncol* 2016; **34**: 1510–7.
34 Chiou VL, Burotto M. Pseudoprogression and Immune-Related Response in Solid Tumors. *J Clin Oncol* 2015; **33**: 3541–3543.
35 Kurra V, Sullivan RJ, Gainor JF et al, Pseudoprogression in cancer immunotherapy: Rates, time course and patient outcomes. *J Clin Onco*l **34**, 2016 (suppl; abstr 6580).
36 Reck M, Rodríguez-Abreu D, Robinson AG et al. Pembrolizumab versus Chemotherapy for PD-L1–Positive Non–Small-Cell Lung Cancer *N Engl J Med* 2016; **375**:1823-33.

***Panel* 1: Key questions identified**

Defining the date of progression in scenarios where initial progression by RECIST 1.1 is followed by response and later progression.

Defining best overall response when initial progression is seen with RECIST 1.1.

Managing response and progression in trials comparing standard agents to immunotherapeutics.

Managing response and progression in trials combining standard agents with immunotherapeutics.

Whether progression must be confirmed with a second scan; if so which is the date of progression?

New lesions: when to measure, how many to measure, and whether all must be measured each subsequent assessment.

Optimal timing of frequency of response assessment.

Management of therapeutics interventions such as surgery or radiation after response.

***Panel* 2: Key principles to be considered**

If the criteria for iUPD have never been met, principles follow RECIST 1.1.

However if the criteria for iUPD have been met, the next TPR could be:

- iUPD – no change noted in any category of lesion.
- iSD, iPR or iCR. Here iUPD (followed by iCPD) must occur again.
- iCPD, if the category in which iUPD was met at the last TPR shows further increase in tumour burden as evidenced (as applicable) by ≥ 5 mm increase in SOM of target or NLT lesions, further increase in NT or NLNT lesions, or an increase in the number of new lesions.
- iCPD of a category which did NOT meet criteria for iUPD now meets the criteria for RECIST 1.1 PD.

| | RECIST 1.1 | iRECIST |
|---|---|---|
| Definitions of measurable and non-measurable disease; numbers and site of target disease | Measurable lesions are 10 mm or more in long diameter (15 mm for nodal lesions); maximum of 5 lesions (2 per organ); all other disease considered not-target (must be 10 mm of longer in short axis for nodal disease) | No change; however, NEW lesions are evaluated as per RECIST 1.1 but are recorded separately on the CRF (but not included in the sum of lesions for target lesions identified at baseline) |
| CR, PR or SD | Cannot have met criteria for PD prior to CR, PR or SD | May have had iUPD (1 or more instances), but not iCPD, prior to iCR, iPR or iSD |
| Confirmation of CR, PR | Only required for non-randomised trials | As per RECIST 1.1 |
| Confirmation of SD | Not required | As per RECIST 1.1 |
| New lesions | Results in PD. Recorded but not measured | Results in iUPD but iCPD is only assigned based on this category if at next assessment<br>• Additional NL appear or<br>• Increase in size of NLs (≥5 mm for sum of NLT or any increase in NLNT)<br><br>NLs, where none have previously been recorded can also confirm iCPD |
| Independent blinded review and central collection of scans | Recommended in some circumstances | Collection of scans (but not independent review ) recommended for all trials |
| Confirmation of PD | Not required (unless equivocal) | Required |
| Consideration of clinical status | Not included in assessment | Clinical stability (see definition) is considered in whether treatment is continued after iUPD |

***Table* 1: RECIST 1.1 and iRECIST**

NT = non-target, T = target; NL = new lesions; NLT = new lesion target; NLNT = new lesion non target; iUPD = unconfirmed immune PD; iCPD = confirmed immune PD; SOM= sum of measures. iCR – immune complete response; iPR – immune partial response; iSD – immune stable disease

| **Time-point response** | | | | |
|---|---|---|---|---|
| Target Lesions* | Non-Target Lesions* | New Lesions* | Time Point Response | |
| | | | No prior iUPD** | Prior iUPD**; *** |
| iCR | iCR | No | iCR | iCR |
| iCR | Non-iCR/Non-iUPD | No | iPR | iPR |
| iPR | Non-iCR/Non-iUPD | No | iPR | iPR |
| iSD | Non-iCR/Non-iUPD | No | iSD | iSD |
| iUPD with no change OR decrease from last TP | iUPD with no change OR decrease from last TP | Yes | NA | NLs confirms iCPD if NLs were previously identified and increase in size (≥5 mm in SOM for NLT or any increase for NLNT) or number; If no change in NLs (size or number) from last TP, remains iUPD |
| iSD, iPR, iCR | iUPD | No | iUPD | Remains iUPD unless iCPD confirmed based in further increase in size of NT disease (need not meet RECIST 1.1 criteria for unequivocal PD) |
| iUPD | Non-iCR/Non-iUPD; iCR | No | iUPD | Remains iUPD unless iCPD confirmed based on:<br>○ further increase in SOM of at least 5 mm, otherwise remains iUPD |
| iUPD | iUPD | No | iUPD | Remains iUPD unless iCPD confirmed based on further increase in:<br>○ previously identified T lesion iUPD in SOM ≥5 mm and / or<br>○ NT lesion iUPD (prior assessment - need not be unequivocal PD) |
| iUPD | iUPD | Yes | iUPD | Remains iUPD unless iCPD confirmed based on further increase in<br>○ previously identified T lesion iUPD SOM ≥5 mm and / or<br>○ previously identified NT lesion iUPD (need not be unequivocal) and /or<br>○ size or number of new lesions previously identified |
| Non-iUPD/PD | Non-iUPD/PD | Yes | iUPD | Remains iUPD unless iCPD confirmed based on<br>○ increase in size or number of new lesions previously identified |
| * Using RECIST 1.1 principles. If no PSPD occurs, RECIST 1.1 and iRECIST categories for CR, PR and SD would be the same. ** in any lesion category. *** previously identified in assessment immediately prior to this TP. | | | | |

***Table 2*: Assigning time-point response for iRECIST**

iCR – immune complete response; iPR – immune partial response; iSD – immune stable disease; iUPD – immune unconfirmed progression; iCPD – immune confirmed progression; NL – new lesion; NLT – new lesion target; NLNT – new lesion non target; T – target; TP – time point; NA = not applicable; NE = not evaluable / evaluated

| **Best overall response** | | | | | |
|---|---|---|---|---|---|
| TPR1 | TPR2 | TPR3 | TPR4 | TPR5 | iBOR |
| iCR | iCR, iPR, iUPD, NE | iCR, iPR, iUPD, NE | iUPD | iCPD | iCR |
| iUPD | iPR, iSD, NE | iCR | iCR, iUPD, NE | iCR, iPR, iSD, iUPD, iCPD, NE | iCR |
| iUPD | iPR | iPR, iSD, iUPD, NE | iPR, iSD, iUPD, NE, iCPD | iPR, iSD, iUPD, NE, iCPD | iPR |
| iUPD | iSD, NE | iPR | iPR, iSD, iUPD, NE | iPR, iSD, iUPD, iCPD, NE | iPR |
| iUPD | iSD | iSD, iUPD, NE | iSD, iUPD, iCPD, NE | iSD, iUPD, iCPD, NE | iSD |
| iUPD | iCPD | Anything | Anything | Anything | iCPD |
| iUPD | iUPD (no iCPD) | iCPD | Anything | Anything | iCPD |
| iUPD | NE | NE | NE | NE | iUPD |

- Examples only – many more scenarios exist but follow the same principles
- Table assumes a randomised study where confirmation of CR or PR is not required
- For patients with non-target disease only at baseline, only iCR or non-CR/non-PD can be assigned at each TPR but is not shown in the table for ease of presentation

***Table 3*: Assigning best overall response for iRECIST**

iCR – immune complete response; iPR – immune partial response; iSD – immune stable disease; iUPD – immune unconfirmed progression; iCPD – immune confirmed progression; NL – new lesion; NLT – new lesion target; NLNT – new lesion non target; T – target; TP – time point; NA = not applicable; NE = not evaluable / evaluated


***Figure* 1: Process for developing and validating guidelines. Shaded boxes are in progress.**

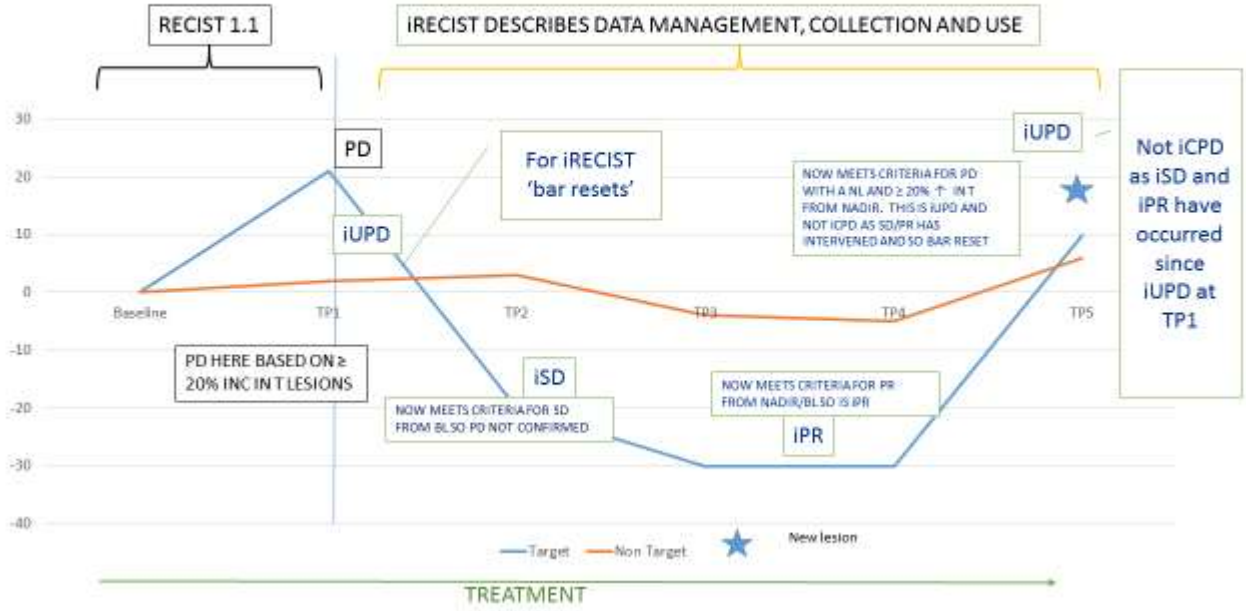


***Figure* 2: RECIST 1.1 and iRECIST: an example of evaluation.**