

iRECIST: guidelines for response criteria for use in trials testing immunotherapeutics



Lesley Seymour, Jan Bogaerts, Andrea Perrone, Robert Ford, Lawrence H Schwartz, Sumithra Mandrekar, Nancy U Lin, Saskia Litière, Janet Dancey, Alice Chen, F Stephen Hodi, Patrick Therasse, Otto S Hoekstra, Lalitha K Shankar, Jedd D Wolchok, Marcus Ballinger, Caroline Caramella, Elisabeth G E de Vries, on behalf of the RECIST working group

Tumours respond differently to immunotherapies compared with chemotherapeutic drugs, raising questions about the assessment of changes in tumour burden—a mainstay of evaluation of cancer therapeutics that provides key information about objective response and disease progression. A consensus guideline—iRECIST—was developed by the RECIST working group for the use of modified Response Evaluation Criteria in Solid Tumours (RECIST version 1.1) in cancer immunotherapy trials, to ensure consistent design and data collection, facilitate the ongoing collection of trial data, and ultimate validation of the guideline. This guideline describes a standard approach to solid tumour measurements and definitions for objective change in tumour size for use in trials in which an immunotherapy is used. Additionally, it defines the minimum datapoints required from future trials and those currently in development to facilitate the compilation of a data warehouse to use to later validate iRECIST. An unprecedented number of trials have been done, initiated, or are planned to test new immune modulators for cancer therapy using a variety of modified response criteria. This guideline will allow consistent conduct, interpretation, and analysis of trials of immunotherapies.

Introduction

Changes in tumour burden (termed response) are often used as surrogates of survival or quality of life;¹ consequently, validated and consistent criteria for defining response to treatment are crucial. In 2000, the Response Evaluation Criteria in Solid Tumours (RECIST) working group simplified the 1981 WHO response criteria² after validation in a large data warehouse.³ In 2009, RECIST was refined to RECIST version 1.1.⁴ The RECIST working group ensures that RECIST undergoes continuous testing, validation, and updates.^{5–7}

Immune modulators are one of the most important classes of new anticancer therapeutics.^{8–10} Cytotoxic T-lymphocyte antigen-4 (CTLA-4), programmed death-1 (PD-1), and programmed death ligand-1 (PD-L1) pathways are the most intensively studied,^{11–17} and drugs that are active in these pathways have, since 2011, received marketing authorisation (for some drugs the authorisation is conditional, pending the completion of other studies) for melanoma, lung, bladder, renal, and head and neck cancer.^{18–23} The novel mechanism of action of these drugs, with immune and T-cell activation, is postulated to lead to unusual patterns of response that resemble tumour flare but are more pronounced and more frequent than previously described responses. In early trials of immune-based therapeutics in melanoma, investigators described unique response patterns, termed pseudoprogression. Some patients whose disease met the criteria for disease progression based on traditional response criteria such as RECIST (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.^{24–28} In 2009, modified response criteria based on WHO criteria (which include the collection of bidimensional measurements of target

lesions) were proposed—the immune-related response criteria (irRC).²⁹ The major modification involved the inclusion of the measurements of new target lesions (each must be at least 5×5 mm in size; with a maximum of ten visceral lesions in total, up to five new lesions per organ, and five new cutaneous lesions) into disease assessments. In 2013, researchers published revised irRC using unidimensional measurements based on the original RECIST.³⁰ Subsequent recommendations, some published in abstract form, seem to incorporate RECIST 1.1 recommendations.^{31–33} These recommendations are often referred to as irRECIST, but have not always been consistently applied, leading to concerns about the comparability of data and results across trials, difficulty with pooling databases, and poor clarity regarding whether new lesions were measured, and if so, how many were captured, and whether measures were incorporated into tumour burden. Recent trials (since 2010) have generally used RECIST-based immune criteria to assess responses to immunotherapies.

Because of the need to standardise 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 drugs 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.^{31,32} Additionally, substantial variability in which criteria were used was seen across clinical trials within pharmaceutical companies and cooperative groups, leading to serious concerns about interpretation of pooled datasets. Finally, most trials that used immune-modified criteria used

Lancet Oncol 2017; 18: e143–52

This online publication has been corrected. The corrected version first appeared at thelancet.com/oncology on April 30, 2019

Canadian Cancer Trials Group, Queen's University, Kingston, ON, Canada

(Prof L Seymour MD, Prof J Dancey MD); EORTC Headquarters, Brussels, Belgium (J Bogaerts PhD, S Litière PhD); Translational Medicine, Merck & Co, Kenilworth, NJ, USA (A Perrone MD); Clinical Trials Imaging Consulting, LLC, Belle Mead, NJ, USA (R Ford MD); Department of Radiology, Columbia University Medical Center, New York, NY, USA (Prof L H Schwartz MD); New York Presbyterian Hospital, New York, NY, USA (Prof S Mandrekar PhD); Division of Biomedical Statistics and Informatics, Mayo Clinic, Rochester, MN, USA (Prof S Mandrekar PhD); Department of Medical Oncology, Dana-Farber Cancer Institute, Harvard Medical School, Boston, MA, USA (N U Lin MD, Prof F S Hodi MD); Early Clinical Trials Development Program, Division of Cancer Treatment and Diagnosis, National Cancer Institute, Bethesda, MD, USA (A Chen MD); Institut de Recherche International Servier, Paris, France (P Therasse MD); Department of Radiology and Nuclear Medicine, VU University Medical Center, Amsterdam, Netherlands (Prof O S Hoekstra MD); Diagnostic Imaging Branch, National Cancer Institute, Bethesda, MD, USA (L K Shankar MD); Department of Medicine, Memorial Sloan Kettering Cancer Center, New York, NY, USA (J D Wolchok MD); Weill Cornell Medical and Graduate Colleges, New York, NY, USA (M Ballinger PhD); Ludwig

Institute for Cancer Research, New York, NY, USA (M Ballinger); Genentech Inc, San Francisco, CA, USA (M Ballinger); Department of Radiology, Gustav Roussy Cancer Campus, Villejuif, France (C Caramella MD); and Department of Medical Oncology, University Medical Center Groningen, Groningen, Netherlands (Prof E G E de Vries MD)

Correspondence to: Prof Lesley Seymour, Canadian Cancer Trials Group, Queen's University, Kingston, ON K7L 3N6, Canada lseymour@ctg.queensu.ca

For more on the RECIST working group see <http://www.eortc.org/recist/>

See Online for appendix

independent imaging review by a commercial entity for those criteria, rather than investigator assessments. We think that response criteria should be applicable across all cancer clinical trials, including those done in the academic sector, where costly independent review is not feasible.

On the basis of these observations, the RECIST working group decided to develop a guideline for 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 health-care team.

Terminology

iRECIST is based on RECIST 1.1. Responses assigned using iRECIST have a prefix of “i” (ie, immune)—eg, “immune” complete response (iCR) or partial response (iPR), and unconfirmed progressive disease (iUPD) or confirmed progressive disease (iCPD) to differentiate them from responses assigned using RECIST 1.1. Similar nomenclature is used for stable disease (iSD). New lesions are assessed and subcategorised into those that

qualify as target lesions (new lesion, target) or non-target lesions (new lesion, non-target).

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 (figure 1) and to review existing approaches to assess response in immune modulator trials, and also to identify points of consensus and items that needed further discussion. Members of the subcommittee included clinical, statistical, and imaging experts in methodology and immunotherapy, representatives from the pharmaceutical companies developing immunotherapeutics, and key regulatory authorities (appendix p 1). On June 2, 2016, a formal meeting was held in Chicago (IL, USA), with invited presentations from regulatory authorities, pharmaceutical companies with immune modulator drugs in development, and academic groups, followed by a structured discussion. Before the meeting, the 52 invited participants were polled to enable the identification of questions that needed to be addressed, as well as the response criteria routinely used by participants. Ten respondents provided responses before the meeting (including some pooled responses) and all eight presenters identified additional areas of interest in their presentations. After review and discussion during the meeting, the group identified a list of important questions to be addressed by iRECIST (panel 1). Notably, all participants confirmed that RECIST 1.1 was used for primary endpoints, with immune-modified response criteria being used in an exploratory manner, with very few exceptions; in one instance, immune-modified criteria were used as a coprimary endpoint. The most commonly used immune-modified criteria were variations of iRECIST. There was more variability in independent imaging review and the period of time during which response data were collected after RECIST 1.1 progression or cessation of protocol therapy. Further calls and meetings were held to develop and plan the full validation of iRECIST (figure 1).

iRECIST

The continued use of RECIST 1.1 is recommended to define whether tumour lesions, including lymph nodes, are measurable or non-measurable, as well as for the management of bone lesions, cystic lesions, and lesions with previous local treatment (eg, radiotherapy; table 1). Similarly, no changes have been made to the recommendations regarding the method of measurement, although clinical examination and chest radiograph are rarely used, with the availability of more modern imaging techniques (eg, CT scans and MRI). The principles used to establish objective tumour response are largely unchanged from RECIST 1.1, but the major change for iRECIST is the concept of resetting the bar

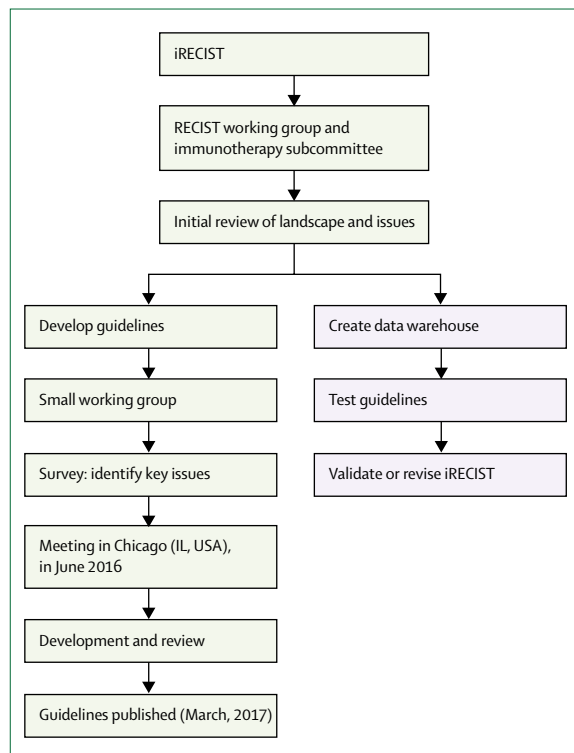


Figure 1: Process for developing and validating iRECIST consensus guidelines Blue shaded boxes represent steps still in progress. RECIST=Response Evaluation Criteria in Solid Tumours.

if RECIST 1.1 progression is followed at the next assessment by tumour shrinkage.

iRECIST defines iUPD on the basis of RECIST 1.1 principles; however, iUPD requires confirmation, which is done on the basis of observing either a further increase in size (or in the number of new lesions) in the lesion category in which progression was first identified in (ie, target or non-target disease), or progression (defined by RECIST 1.1) in lesion categories that had not previously met RECIST 1.1 progression criteria. However, if progression is not confirmed, but instead tumour shrinkage occurs (compared with baseline), which meets the criteria of iCR, iPR, or iSD, then the bar is reset so that iUPD needs to occur again (compared with nadir values) and then be confirmed (by further growth) at the next assessment for iCPD to be assigned. If no change in tumour size or extent from iUPD occurs, then the timepoint response would again be iUPD. This approach allows atypical responses, such as delayed responses that occur after pseudoprogression, to be identified, further understood, and better characterised (tables 1–3, figure 2, appendix pp 2–4). Sample case record forms and protocol sections are included in the appendix pp 5–19. In the next few paragraphs, we only briefly summarise sections of RECIST 1.1 that are unchanged; readers should refer to RECIST 1.1 for full descriptions.⁴

Assessment of target, non-target, and new lesions

Most RECIST 1.1 recommendations are unchanged for timepoint response, including the management of lymph nodes, lesions that become too small to measure, lesions that split or coalesce, and the definition of complete response, partial response, stable disease, and

progressive disease. Each timepoint response is based on the assessment 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, as long as iCPD was not confirmed. iUPD is defined by RECIST 1.1 criteria for progressive disease; iUPD can be assigned multiple times as long as iCPD is not confirmed at the next assessment. Progression is confirmed in the target lesion category if the next imaging assessment after iUPD (4–8 weeks later) confirms a further increase in sum of measures of target disease from iUPD, with an increase of at least 5 mm. However, the criteria for iCPD (after iUPD) are not considered to have been met if complete response, partial response, or stable disease criteria (compared with baseline and as defined by RECIST 1.1) are met at the next assessment after iUPD. The status is reset (unlike RECIST 1.1, in which any progression precludes later complete response, partial response, or stable disease). iCR, iPR, or iSD should then be assigned; and if no change is detected, then the timepoint response is iUPD.

The assessment of non-target lesions at each timepoint follows similar principles. iUPD (but not iCPD) can have been documented before iCR or when the criteria for neither CR nor PD have been met (referred to as non-iCR/non-iUPD) and can be assigned several times, as long as iCPD was not confirmed. iUPD is defined by RECIST 1.1 criteria; however, iUPD can be assigned multiple times as long as iCPD is not confirmed at the next assessment. Progressive disease in the non-target lesion category is confirmed if subsequent imaging, done 4–8 weeks after iUPD, shows a further increase from iUPD. The criteria for iCPD are not judged to have

	RECIST 1.1	iRECIST
Definitions of measurable and non-measurable disease; numbers and site of target disease	Measurable lesions are ≥ 10 mm in diameter (≥ 15 mm for nodal lesions); maximum of five lesions (two per organ); all other disease is considered non-target (must be ≥ 10 mm in short axis for nodal disease)	No change from RECIST 1.1; however, new lesions are assessed as per RECIST 1.1 but are recorded separately on the case report form (but not included in the sum of lesions for target lesions identified at baseline)
Complete response, partial response, or stable disease	Cannot have met criteria for progression before complete response, partial response, or stable disease	Can have had iUPD (one or more instances), but not iCPD, before iCR, iPR, or iSD
Confirmation of complete response or partial response	Only required for non-randomised trials	As per RECIST 1.1
Confirmation of stable disease	Not required	As per RECIST 1.1
New lesions	Result in progression; recorded but not measured	Results in iUPD but iCPD is only assigned on the basis of this category if at next assessment additional new lesions appear or an increase in size of new lesions is seen (≥ 5 mm for sum of new lesion target or any increase in new lesion non-target); the appearance of new lesions when none have previously been recorded, can also confirm iCPD
Independent blinded review and central collection of scans	Recommended in some circumstances—eg, in some trials with progression-based endpoints planned for marketing approval	Collection of scans (but not independent review) recommended for all trials
Confirmation of progression	Not required (unless equivocal)	Required
Consideration of clinical status	Not included in assessment	Clinical stability is considered when deciding whether treatment is continued after iUPD

"i" indicates immune responses assigned using iRECIST. RECIST=Response Evaluation Criteria in Solid Tumours. iUPD=unconfirmed progression. iCPD=confirmed progression. iCR=complete response. iPR=partial response. iSD=stable disease.

Table 1: Comparison of RECIST 1.1 and iRECIST

been met if RECIST 1.1 criteria for complete response or non-iCR/non-iUPD are met after a previous iUPD. The status is reset (unlike RECIST 1.1) and iCR, or

non-iCR/non-iUPD is assigned; if no change is detected, the timepoint response 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 assessed and discounted if not malignant. These principles of RECIST 1.1 remain useful and clearly identify the management of new lesions that are considered to be potentially artefactual: “If a new lesion is equivocal, for example because of its small size, continued therapy and follow-up assessment will clarify whether 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”.⁴

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, treatment should be continued. New lesions should be assessed and categorised as measurable or non-measurable using RECIST 1.1 principles. Five lesions (no more than two per organ) should be measured and recorded as a new lesion target, but should not be included in the sum of measures of the original target lesions identified at baseline (appendix p 17). Other

Panel 1: Key questions identified by the RECIST working group

- How to define the date of progression in scenarios in which initial progression by RECIST 1.1 is followed by response and later progression
- How to define best overall response when initial progression is established with RECIST 1.1
- How to manage response and progression in trials comparing standard non-immunotherapy drugs against immunotherapeutics
- Whether or not progression should be confirmed with a second scan; and if so, which timepoint denotes the date of progression?
- New lesions: when to measure, how many to measure, and whether all should be measured at each subsequent assessment
- Optimal timing of frequency of response assessment
- How to manage therapeutic interventions such as surgery or radiotherapy after response

	Timepoint response with no previous iUPD in any category	Timepoint response with previous iUPD in any category*
Target lesions: iCR; non-target lesions: iCR; new lesions: no	iCR	iCR
Target lesions: iCR; non-target lesions: non-iCR/non-iUPD; new lesions: no	iPR	iPR
Target lesions: iPR; non-target lesions: non-iCR/non-iUPD; new lesions: no	iPR	iPR
Target lesions: iSD; non-target lesions: non-iCR/non-iUPD; new lesions: no	iSD	iSD
Target lesions: iUPD with no change, or with a decrease from last timepoint; non-target lesions: iUPD with no change, or decrease from last timepoint; new lesions: yes	Not applicable	New lesions confirm iCPD if new lesions were previously identified and they have increased in size (≥5 mm in sum of measures for new lesion target or any increase for new lesion non-target) or number; if no change is seen in new lesions (size or number) from last timepoint, assignment remains iUPD
Target lesions: iSD, iPR, iCR; non-target lesions: iUPD; new lesions: no	iUPD	Remains iUPD unless iCPD is confirmed on the basis of a further increase in the size of non-target disease (does not need to meet RECIST 1.1 criteria for unequivocal progression)
Target lesions: iUPD; non-target lesions: non-iCR/non-iUPD, or iCR; new lesions: no	iUPD	Remains iUPD unless iCPD is confirmed on the basis of a further increase in sum of measures ≥5 mm; otherwise, assignment remains iUPD
Target lesions: iUPD; non-target lesions: iUPD; new lesions: no	iUPD	Remains iUPD unless iCPD is confirmed based on a further increase in previously identified target lesion iUPD in sum of measures ≥5 mm or non-target lesion iUPD (previous assessment need not have shown unequivocal progression)
Target lesions: iUPD; non-target lesions: iUPD; new lesions: yes	iUPD	Remains iUPD unless iCPD is confirmed on the basis of a further increase in previously identified target lesion iUPD sum of measures ≥5 mm, previously identified non-target lesion iUPD (does not need to be unequivocal), or an increase in the size or number of new lesions previously identified
Target lesions: non-iUPD or progression; non-target lesions: non-iUPD or progression; new lesions: yes	iUPD	Remains iUPD unless iCPD is confirmed on the basis of an increase in the size or number of new lesions previously identified

Target lesions, non-target lesions, and new lesions defined according to RECIST 1.1 principles; if no pseudoprogression occurs, RECIST 1.1 and iRECIST categories for complete response, partial response, and stable disease would be the same. *Previously identified in assessment immediately before this timepoint. “i” indicates immune responses assigned using iRECIST. iCR=complete response. iPR=partial response. iSD=stable disease. iUPD=unconfirmed progression. non-iCR/non-iUPD=criteria for neither CR nor PD have been met. iCPD=confirmed progression. RECIST=Response Evaluation Criteria in Solid Tumours.

Table 2: Assignment of timepoint response using iRECIST

	Timepoint response 1	Timepoint response 2	Timepoint response 3	Timepoint response 4	Timepoint response 5	iBOR
Example 1	iCR	iCR, iUPD, or NE	iCR, iUPD, or NE	iUPD	iCPD	iCR
Example 2	iUPD	iPR, iSD, or NE	iCR	iCR, iUPD, or NE	iCR, iPR, iSD, iUPD, iCPD, or NE	iCR
Example 3	iUPD	iPR	iPR, iSD, iUPD, or NE	iPR, iSD, iUPD, NE, or iCPD	iPR, iSD, iUPD, NE, or iCPD	iPR
Example 4	iUPD	iSD or NE	iPR	iPR, iSD, iUPD, or NE	iPR, iSD, iUPD, iCPD, or NE	iPR
Example 5	iUPD	iSD	iSD, iUPD, or NE	iSD, iUPD, iCPD, or NE	iSD, iUPD, iCPD, or NE	iSD
Example 6	iUPD	iCPD	Any	Any	Any	iCPD
Example 7	iUPD	iUPD (no iCPD)	iCPD	Any	Any	iCPD
Example 8	iUPD	NE	NE	NE	NE	iUPD

Eight examples are presented for patients with target disease at baseline, but many more scenarios exist following the same principles. Table assumes a randomised study in which confirmation of complete response or partial response is not required. For patients with non-target disease only at baseline, only iCR or non-complete response or non-progression of disease can be assigned at each timepoint (not shown in the table for ease of presentation). "i" indicates immune responses assigned using iRECIST. iBOR=best overall response. iCR=complete response. iPR=partial response. NE=not evaluable. iUPD=unconfirmed progression. iCPD=confirmed progression. iSD=stable disease. RECIST=Response Evaluation Criteria in Solid Tumours.

Table 3: Scenarios of assignments of best overall response using iRECIST

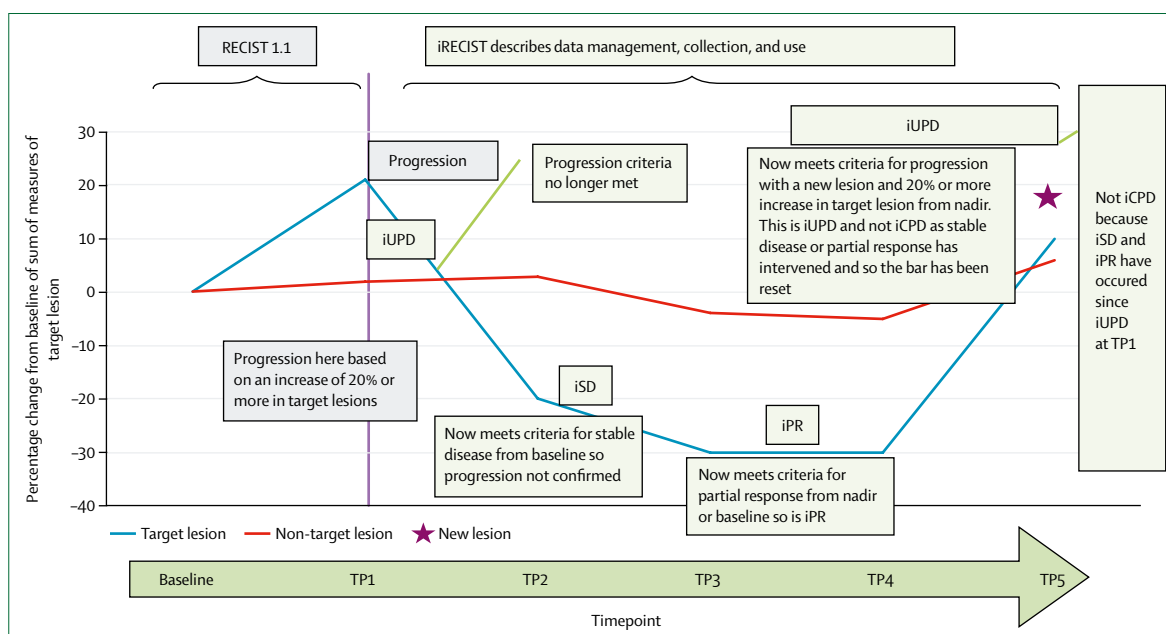


Figure 2: RECIST 1.1 and iRECIST: an example of assessment

Prefix "i" indicates immune responses assigned using iRECIST; others without "i" are confirmed by RECIST 1.1. RECIST=Response Evaluation Criteria in Solid Tumours. iCR=complete response. iCPD=complete progression. iPR=partial response. iSD=stable disease. iUPD=unconfirmed progression. TP=timepoint.

measurable and non-measurable lesions are recorded as new lesion non-target. Trialists might choose to measure and record more than five new lesions for research purposes, but this method is not believed to be practical for general use. New lesions do not need to meet the criteria for new lesion target to result in iUPD (or iCPD); new lesion non-target can also drive iUPD or iCPD. Progressive disease is confirmed (iCPD) in the new lesion category if the next imaging assessment, done at 4–8 weeks after iUPD, confirms additional new lesions or a further increase in new lesion size from iUPD (sum of measures increase in new lesion target ≥ 5 mm, any increase for new lesion non-target).

Notably, if iUPD criteria were met on the basis of progression in the target or non-target disease, or the

appearance of new lesions, then RECIST 1.1-assigned progression in another lesion category in the confirmatory scan also confirms iCPD.

Continued treatment after iUPD

The existing literature describes pseudoprogression as an increase in the size of lesions, or the visualisation of new lesions, followed by a response, which might be durable. Although well described, differentiating transient pseudoprogression from true progression, potentially requiring a change in therapy, can be challenging. Although early discontinuation of an effective drug is not desirable, continued long-term treatment with a non-effective drug past true progression might delay the initiation of potentially effective salvage therapy.

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 timepoint response could be:
 - iUPD: no change noted in any category of lesion
 - iSD, iPR, or iCR. Here, iUPD (followed by iCPD) should occur again
 - iCPD, if the category in which iUPD was met at the last timepoint response shows a further increase in tumour burden as evidenced (as applicable) by a ≥ 5 mm increase in sum of measures of target or new target lesions, further increase in non-target or new non-target 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 progression Prefix "i" indicates immune responses assigned using iRECIST. RECIST=Response Evaluation Criteria in Solid Tumours. iCR=complete response. iCPD=complete progression. iPR=partial response. iSD=stable disease. iUPD=unconfirmed progression.

We recommend that clinical trials in which treatment beyond initial RECIST 1.1-defined progression (ie, iUPD) is permitted should only allow patients who are clinically stable to continue on treatment until the next assessment (≥ 4 weeks later); this next imaging assessment should be no longer than 8 weeks later, to ensure that patients remain fit for salvage therapies. A longer timeframe before the next assessment might be reasonable if pseudoprogression is well described in the tumour type (eg, melanoma treated with a CTLA4 inhibitor), especially if no effective salvage therapies are available (eg, *BRAF* wild-type melanoma) but should be justified in the trial protocol. All decisions regarding continuation or discontinuation of therapy should be made by 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.

An assignment of clinical stability requires that no worsening of performance status has occurred, that no clinically relevant increases in disease-related symptoms such as pain or dyspnoea occur that are thought to be associated with disease progression (these symptoms are generally understood to mean a requirement for increased palliative intervention), and that no requirement for intensified management of disease-related symptoms exists, including increased analgesia, radiotherapy, or other palliative care.

The imaging findings and the recommendation to continue with treatment despite iUPD should be discussed with the patient before a decision is made about whether or not to continue therapy. Patients who have iUPD and are not clinically stable should be designated as not clinically stable in the case report form.

This designation will allow the best overall response to be calculated and the date of iUPD to be used in estimates of progression-free survival.

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

Timepoint and best overall response

Although the principles of the assignment of the timepoint response and best overall response closely follow RECIST 1.1, and reflect assessment of target and non-target lesions as well as the presence of new lesions, the possibility of pseudoprogression adds complexity (tables 1–3, panel 2, appendix pp 2–4). The timepoint response is calculated using the response assigned for each category of lesion (as for RECIST 1.1), but takes into account the last timepoint response.

The algorithm for patients with no previous iUPD is identical to RECIST 1.1. For patients with iUPD at the last timepoint response, the next timepoint response is dependent on the status of all lesions, including target, non-target, new lesion target, and new lesion non-target; on whether any increase in size has occurred (either a further increase in size or a sufficient increase to assign a new iUPD if the criteria were not previously met); or the appearance of additional new lesions.

For iRECIST, the best overall response (iBOR) is the best timepoint 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 best overall response of iSD, iPR, or iCR (tables 1–3, appendix pp 2–4), meaning that iPR or iSD can be assigned (timepoint response or iBOR) even if new lesions have not regressed, or if unequivocal progression (non-target 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 approach is also recommended for iRECIST. The duration of iCR and iPR is from the timepoint when the criteria for iCR or iPR are first met, whereas the duration of iSD is still calculated from baseline.

The protocol should establish how missing response assessments will be handled. Assessments that are not done or are not evaluable should be disregarded. For example, an iUPD followed by an assessment that was not done or not evaluable, and then another unconfirmed progressive disease, would be indicative of iCPD. Protocols should clearly specify whether assessments done after protocol therapy is discontinued can be considered in identification of iBOR; it might be

reasonable to include assessments done several 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 should also specify how any new therapy introduced before progression (eg, radiotherapy or surgery) will affect iBOR designation. Other RECIST 1.1 recommendations, including the management of missing assessments, remain unchanged, including requiring that the statistical analysis plan should indicate how missing data or assessments will be addressed in the determination of response and progression.

Frequency of tumour reassessment

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

Tumour reassessment can be done earlier than originally planned (but only between 4 and 8 weeks after iUPD) to confirm iUPD (or, in non-randomised trials, to confirm iCR or iPR ≥ 4 weeks after the scan showing complete or partial response). If progression is not confirmed, reassessment should continue as originally planned (ie, if scans were to be done at 8, 16, and 24 weeks, and a scan was done at 12 weeks to confirm response, then the next scans should be done at 16 weeks and 24 weeks, as planned). If patients continue on treatment per protocol after iCPD, assessments should continue to be done, at the same planned schedule, until protocol treatment is discontinued.

Ideally, all imaging done after protocol treatment has been discontinued should continue to be recorded on the case report form until subsequent therapies are initiated, as the protocol and informed consent document permit. 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 at which progression criteria are met (ie, the date of iUPD) provided that iCPD is confirmed at the next assessment (appendix pp 2–4 and 19). 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 judged to be clinically stable, or no further response assessments are done (because of patient refusal, protocol non-compliance, or patient death); the next timepoint responses are all iUPD, and iCPD never occurs; or the patient dies from their cancer. The case report form collects the reason why confirmatory response assessment was not done at any timepoint, such as not clinically stable, centre error, patient refusal, or patient death.

For protocols that permit crossover, or if 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 that iCPD be used especially for scenarios with immunotherapy in both treatment groups and when pseudoprogression is anticipated.

Adjuvant trials of immune modulators given after curative surgery for melanoma or lung cancer are ongoing (NCT 02437279, 02388906, 02595944, 02504372, and 02273375) but have yet to report their results. Suspected new lesions in the curative setting should always be investigated thoroughly and preferably have a biopsy taken before the designation of relapse is assigned. If taking a biopsy sample is not technically feasible, then it would seem 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 centralised blinded review of imaging studies is not planned) is recommended for all studies using an imaging-based endpoint (ie, response or progression-free survival) if feasible. Although the iRECIST guideline requires the recording of the measurements of up to five new lesions, it might eventually be necessary to record additional lesions to obtain a more precise estimate of progression. Central collection of images will allow further assessment by an independent radiologist if necessary. If real-time central review is planned, the protocol should clearly explain how treatment decisions will be made.

We recommend that phase 3 clinical trials continue to incorporate both RECIST 1.1 and iRECIST (table 1) and that RECIST 1.1 should continue to be used to define the primary efficacy outcomes (progression-free survival, disease progression, and best overall response). Exploratory analyses using the iPD date (ie, the first date of iUPD that is subsequently confirmed) can be defined in the statistical analysis plan. Early-phase trials can 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 information is especially important for trials that compare an immune modulator treatment with a non-immune modulator treatment.

Discussion: next steps and validation

Immunotherapeutics are a major advance in the treatment of an escalating number of cancers. The increasing testing and use of these drugs in multiple clinical settings, including adjuvant, first, second, and subsequent lines of therapy will require the use of progression-based endpoints. RECIST 1.1 might not always adequately capture the unique patterns of response that have been well described in clinical trials of these drugs in a low proportion of patients, typically reported as 10% or less, mainly in melanoma studies.^{32–34} The true frequency in trials of other malignancies (including non-small-cell lung cancer) is unclear because most trials have reported RECIST 1.1-based response rates,³⁵ but might be less common based on anecdotal reports. Similarly, whether this pattern is unique to drugs active in the CTLA4–PD-1–PD-L1 pathway is currently unknown. Trials testing immunotherapeutics in combination with standard therapies, especially when they are compared with these standard therapies alone, further confound the assessment of progression-based endpoints.

RECIST 1.1 already addresses the management of equivocal progression, including suspected new lesions, which might 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 that new lesions might have developed, or the concept of the isodense lesion 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 intention was never to use those recommendations to manage pseudoprogression described with immune modulators.

Although modified response criteria have been used, a formal guideline is clearly needed, with robust plans for prospective testing and consistent data collection and validation. Trials have not always been consistent in the definition of the response criteria to be used, have used trial-specific modifications of response criteria in which new lesion measurements can or cannot be included in the assessment of response, and response assessments after progression defined by RECIST 1.1 are not always done. Those data are crucial to understand the dynamics of tumour response to immunotherapeutics, including whether immunotherapeutics with different mechanisms of action have varying effects.

Although some progress has been made in understanding tumour dynamics with immunotherapeutics, progress in this area has undoubtedly been limited by reluctance toward data sharing across trials, companies, and immunotherapeutics. Publications have been based on trials done by individual pharmaceutical companies or commercial organisations. In the development of this guideline, virtually all major pharmaceutical companies developing immunotherapeutics participated and have shared their experiences, protocols, response criteria,

and, most importantly, their data. The iRECIST team also included members of the European Medicines Agency and the US Food and Drug Administration.

Although this guideline is consensus based, it is not yet validated because the data warehouse is still being created with initial trial data already in place. The guideline includes all available knowledge on response dynamics, allowing appropriate management of true pseudoprogression, but importantly, it also safeguards patients: although pseudoprogression is now well described, it still only occurs in fewer than one in ten patients. Treatment past radiographic progression might be appropriate only in a small number of patients, and the continuation of treatment past true progression could reduce subsequent effective therapies if the patient is no longer fit enough to tolerate any further treatment.

iRECIST requires the confirmation of progression to rule out or confirm pseudoprogression. Although 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 pseudoprogression is common, patients might be exposed to a higher risk (of continuing ineffective therapy or increasing exposure to radiotherapy) or cost (for the potentially ineffective therapy or the costs of imaging). We recommend that these criteria are used for clinical trial protocols rather than to guide clinical practice. Treatment beyond RECIST 1.1-based progression should be considered only in carefully selected scenarios in which the patient is stable (or improving) symptomatically and if there is just a short period remaining before reassessment.

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 timepoint response and best overall response have several benefits. The association between the site of the new lesion and progression-free survival and the value of adding new lesion measurements to the sum of measures can be explored. Continuing to record RECIST 1.1 allows comparison with reported immunotherapy trials that have used RECIST 1.1, as well as chemotherapy trials, while also allowing treatment past progression and collecting data that will allow further testing and validation of iRECIST. Differences in trial outcomes using RECIST 1.1 versus iRECIST could occur, and the interpretation will be informative. Our proposed plan will enable identification of such situations, and hopefully clarification of the underlying mechanisms. Additionally, in the future, quantification of the differences in outcome estimation between RECIST 1.1 and iRECIST will be possible, enabling better informed decisions for future changes to RECIST guidelines.

This strategy will also be useful for trials comparing immunotherapy-based with non-immunotherapy-based therapeutics. RECIST 1.1 and iRECIST should yield almost identical results for non-immunotherapy treatments,

Search strategy and selection criteria

This paper describes a consensus guideline, rather than a formal literature review. However, a database search was done using PubMed in August, 2016, with the following search terms: “immune response criteria” (limited to cancer, clinical trials, and publications in English language; 234 citations), “iRC” (23 citations), and “pseudoprogression” (limited to cancer, clinical trials, and publications in English language; 39 citations).

based on the RECIST warehouses; whereas an immune modulator warehouse and associated sensitivity analysis of endpoints will enable the quantification of potential added benefit for the immunotherapy component. Although comparison of iRECIST in such situations incorporates an element of bias by construction, confirmation and validation of the guideline by overall survival results might gain additional importance.

Our 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 regarded as exploratory in such trials, although earlier phase trials might consider using primarily iRECIST.

The creation of a data warehouse is underway and updates are available from EORTC where the warehouse is held. Meanwhile 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 improve treatments for 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. AP owns stocks in Merck & Co. 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, Sun Advanced Pharma Research Company, Median Therapeutics, Optimer Biotechnology Inc, Vascular Biogenics, Ignyta, Immunocellular CBT Pharmaceuticals, Tracon, DNatrix, CytRx, Kolltan, and Samsung Bioepsis, outside the submitted work. LHS reports grants and consulting fees from Novartis, grants from Astellas, Eli Lilly, Merck, Pfizer, consulting fees from GSK, outside the submitted work. JD reports grants from Merck, AstraZeneca, Pfizer, Ottawa Hospital Research Institute, Novartis, outside the submitted work. FSH reports grants from Bristol-Myers Squibb, and personal fees from Merck, Novartis, Genentech, and EMD Serono, outside the submitted work. Additionally, FSH has a patent MICA Related Disorders with royalties paid, and a patent Tumor Antigens and Uses Thereof issued. JDW reports grants from Bristol Myers Squibb, Merck, and Genentech during the conduct of the study; consulting fees from Bristol Myers Squibb and Merck, and is on the scientific advisory board for Medimmune, outside the submitted work. MB reports personal fees from Genentech outside the submitted work, and owns stocks in Roche. CC reports personal fees from Roche, BMS, and AstraZeneca, outside

the submitted work. EGEdV reports consulting fees from Synthon, Medivation, and Merck; and grants from Novartis, Amgen, Roche/Genentech, Servier, Chugai, Synthon, AstraZeneca, and Radius Health, outside the submitted work. JB, SM, NUL, SL, AC, PT, OSH, and LKS declare no competing interests.

Acknowledgments

AC and LKS are employees of the National Institutes of Health/National Cancer Institute. This publication was supported by the Canadian Cancer Society Research Institute (grant #021039), the EORTC Cancer Research Fund, and the National Cancer Institute (grant number 5U10-CA11488-45). The contents of this paper were presented in part at the EORTC–NCI–AACR 2016 Meeting (Munich, Germany; Nov 29–Dec 2, 2016). We gratefully acknowledge the thoughtful participation of the following in this initiative: Patricia Keegan (US Food and Drug Administration, Silver Spring, MD, USA); Francesco Pignatti (European Medicine Agency, London, UK); Wendy Hayes (Bristol-Myers Squibb, Princeton, NJ, USA); Eric Rubin (Merck & Co, Kenilworth, NJ, USA). We also received written comments from Darragh Halpenny, Jean-Yves Blay, Florian Lordick, Silke Gillissen, 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, and Glen Laird. We thank Anouk Funke for her assistance with this manuscript.

References

- 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**: 1992–94.
- Miller AB, Hoogstraten B, Staquet M, Winkler A. Reporting results of cancer treatment. *Cancer* 1981; **47**: 207–14.
- 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.
- 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.
- 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**: 260–66.
- 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.
- Schwartz LH, Litière S, de Vries E, et al. RECIST 1.1—update and clarification: from the RECIST committee. *Eur J Cancer* 2016; **62**: 132–37.
- 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.
- Pardoll, Drew M. The blockade of immune checkpoints in cancer immunotherapy. *Nat Rev Cancer* 2012; **12**: 252–64.
- Sidaway P. Bladder cancer: atezolizumab effective against advanced-stage disease. *Nat Rev Urol* 2016; **13**: 238.
- Holt GE, Podack ER, Raez LE. Immunotherapy as a strategy for the treatment of non-small cell lung cancer. *Therapy* 2011; **8**: 43–54.
- 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.
- 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.
- Postow MA, Chesney J, Pavlick AC, et al. Nivolumab and ipilimumab versus ipilimumab in untreated melanoma. *N Engl J Med* 2015; **372**: 2006–17.
- 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.

For the EORTC RECIST data warehouse see www.eortc.org/RECIST

- 16 Lynch TJ, Bondarenko I, Luft A, et al. Ipilimumab in combination with paclitaxel and carboplatin as first-line treatment in stage IIB/IV non-small cell lung cancer: results from a randomized, double-blind, multi-center phase II study. *J Clin Oncol* 2012; **30**: 2046–54.
- 17 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.
- 18 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.
- 19 Topalian SL, Hodi FS, Brahmer JR, et al. Safety, activity, and immune correlates of anti-PD-1 antibody in cancer. *N Engl J Med* 2012; **366**: 2443–54.
- 20 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.
- 21 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.
- 22 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.
- 23 Motzer RJ, Escudier B, McDermott DF, et al. Nivolumab versus everolimus in advanced renal-cell carcinoma. *N Engl J Med* 2015; **373**: 1803–13.
- 24 Fife BT, Bluestone JA. Control of peripheral T-cell tolerance and autoimmunity via CTLA-4 and PD-1 pathways. *Immunol Rev* 2008; **224**: 166–82.
- 25 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.
- 26 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.
- 27 Tarhini AA. Tremelimumab: a review of development to date in solid tumors. *Immunotherapy* 2013; **5**: 215–29.
- 28 Hoos A, Parmiani G, Hege K, et al. A clinical development paradigm for cancer vaccines and related biologics. *J Immunother* 2007; **30**: 1–15.
- 29 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.
- 30 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.
- 31 Bohnsack O, Ludajic K, Hoos A. Adaptation of the immune-related response criteria: irRECIST. *Ann Oncol* 2014; **25** (suppl 4): iv361–iv372.
- 32 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**: 1510–17.
- 33 Chiou VL, Burotto M. Pseudoprogression and immune-related response in solid tumors. *J Clin Oncol* 2015; **33**: 3541–43.
- 34 Kurra V, Sullivan RJ, Gainor JF, et al. Pseudoprogression in cancer immunotherapy: rates, time course and patient outcomes. *Proc Am Soc Clin Oncol* 2016; **34**: abstr 6580.
- 35 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.