

Oncology & Clinical Trials in the 21st Century



IMAGE SOURCE/GETTY IMAGES

Oliver Bohnsack, MD, and
Annette Schmid

RECIST in the Real World

The when and why of using
the new RECIST 1.1 criteria
without abandoning the old.

In 2000, the Response Evaluation Criteria in Solid Tumors (RECIST 1.0) was first published, setting forth rules defining when cancer patients improve (respond), worsen (progression) or stay the same (stable) during treatment.¹ The criteria emerged from an international collaboration that included the European Organization for Research and Treatment of Cancer, the National Cancer Institute of the United States, and the National Cancer Institute of Canada Clinical Trials Group. Although originally developed for clinical use, RECIST is now employed in the majority of clinical trials evaluating cancer treatments for objective response or progression free survival analysis in solid tumors.

The first formal revision of the criteria (RECIST 1.1) was published in January 2009 by a consortium of academic, government, and industry stakeholders known as the RECIST Working Group to address shortcomings of the original.² The updated version, similar to RECIST 1.0, remains grounded in the anatomical assessment of disease and gauges patient response based on unidimensional measures of tumor lesions in combination with a qualitative assessment of nonmeasurable lesions and new lesions.³ But RECIST 1.1 better reflects the realities of clinical practice, the growing

importance of disease progression as a primary endpoint, and the pivotal role of target lesions in gauging tumor response. Further, RECIST 1.1 opens up new doors for sponsors in terms of patient recruitment and label expansion and could potentially help them achieve some spending and scheduling efficiencies.

This article describes the new RECIST criteria and compares them to the original ones. RECIST 1.1, while meaningful and straightforward, can be difficult to implement outside of leading academic research centers without significant commitment to physician training. The necessity of using RECIST 1.1 in parallel with the original RECIST, potentially for years to come, could create general confusion at investigative sites regarding what structures to measure and how to make the determinations. The pros and cons of using RECIST 1.1 are dependent not only on specifics of individual studies and the indication, but also on sponsors comfort in blazing the regulatory trail with an untested methodology for evaluating efficacy.

Key changes

There are a number of significant changes in RECIST 1.1 compared to 1.0 (see Table 1). One of the most challenging is the measurement of lymph nodes. The RECIST Working Group

Key Differences Between RECIST 1.1 and RECIST 1.0

RECIST 1.1 (2009 criteria)	RECIST 1.0 (2000 criteria)
1. Assessment of lymph nodes	1. No specific recommendations
2. Patients with nonmeasurable disease only eligible for studies when the primary endpoint is tumor progression	2. Measurable disease at baseline required
3. Finding of a new lesion should be unequivocal	3. Not specifically defined
4. Imaging of nontarget lesions not necessary at every protocol-specified time point for declaration of partial response or stable disease	4. Not specifically addressed
5. Patient response to treatment need not be confirmed in randomized trials when the primary endpoint is disease progression	5. Patient's initial response to treatment must be confirmed no sooner than four weeks later
6. FDG-PET is "sometimes reasonable to incorporate" as a complement to CT scanning in assessing disease progression	6. No specific recommendations
7. Maximum of five target lesions total and two per organ	7. Maximum of ten target lesions total and five per organ
8. Lesions too small to measure assigned default value of 5mm	8. No specific recommendations
9. Progressive disease requires > 20% increase in sum of diameters from nadir and an absolute increase of > 5mm	9. Progressive disease requires > 20% increase in sum of the longest diameters from nadir—no absolute increase required

Source: Perceptive Informatics.

Table 1. A comparison of RECIST 1.1 and RECIST 1.0.

precisely defines when nodes are pathologic versus normal and healthy. Under the original guidelines, patients with remaining small but measurable nodes would fail to meet the criteria for “complete response” even if all non-nodal target lesions had disappeared. The new guidelines stipulate 10 mm as the cutoff point in differentiating between a healthy and a pathologic node.

RECIST 1.1 specifies that target nodes be measured in the short axis, perpendicular to the longest diameter. This approach is deemed more reproducible and predictive of malignancy than the customary approach of measuring in the long axis.^{4,5}

Significantly, the new guidelines allow sponsors to enroll patients into oncology studies who have nonmeasurable disease only when the primary endpoint is progression-related. A good example and currently popular research indication for this is gastric cancer. Primary tumors still in place or recurrent in the gastric wall must be assessed qualitatively because the gastric wall moves and changes appearance, making it impossible to reliably and reproducibly measure those lesions.

RECIST 1.1 effectively enlarges the universe of patients that can participate in studies, potentially speeding enrollment, shortening timelines, and increasing product

marketability to new indications. The caveat is the creation of potential subpopulations, including those with and those without measurable disease. Sponsors are left to resolve the puzzle of how to make statistically sound comparisons between qualitative and quantitative assessments of tumor response.

Added language in RECIST 1.1 clarifies that the finding of a new lesion should be unequivocal and not attributable to something other than disease, such as differences in scanning technique or healing of a pre-existing lesion. This is particularly important if lesions show partial or complete response to treatment. For example, necrosis of a liver lesion may be reported on a computed tomography (CT) scan report as a “new” cystic lesion, which it is not.

The new criteria also de-emphasize the importance of nontarget lesions. Target lesions, assigned at baseline, should tell the whole story. Overall response tables offered by RECIST 1.1 (see

Table 2) suggest that what happens to them more often than not determines whether a patient is deemed to have improved or their disease has stabilized, even when the nontargeted lesions are not evaluable.

The elevated status of target lesions could add a degree of freedom to image scheduling. RECIST 1.1 acknowledges that imaging of nontarget lesions is unnecessary at every protocol-specified time point for a declaration of partial response or stable disease. Potentially, this could translate into trials that are less costly as well as more palatable to physicians worried about radiation exposure.

A conservative approach

In the new guidelines, several changes have been made to the criteria as to when confirmatory measurement is required. RECIST 1.0 stipulates that a patient's initial response to treatment must be documented a second time, no sooner than four weeks later. RECIST 1.1 adds that the response need not be confirmed in randomized trials where disease progression is the primary endpoint. However, sponsors might still want to consider independent review if disease progression is not clearly unequivocal to reduce the odds of patients being removed from a trial prematurely or inappropriately. Sponsors may also be

interested in secondary response-related endpoints for which confirmatory scans or regulatory endorsement of the protocol design would be advisable.

Fluorodeoxyglucose-positron emission tomography (FDG-PET) is embraced for the first time by RECIST 1.1, albeit as a functional imaging modality supporting CT scanning rather than independently driving the determination of disease progression. The implication is that FDG-PET is most appropriate for exploratory purposes in early phases on small numbers of patients at a few select sites. Its time has not yet come for use in late-phase studies involving multiple patients and sites due to limited global availability and, more importantly, insufficient standardization. SNM (formerly the Society of Nuclear Medicine) is currently working to reduce variability across nuclear medicine sites worldwide by defining image acquisition and analysis standards.⁶

The selection of five (maximum two per organ) rather than 10 (maximum five per organ) target lesions at baseline is one of the most talked about revisions to RECIST but is noteworthy only because it better mirrors clinical practice. Tumor response to treatment is nearly always assessed on fewer than 10 target lesions. Patients may have no more than five or six lesions available to measure. Data analysis has also shown that study outcomes are unaffected when only five lesions were selected as targets.⁷ Similarly, RECIST 1.1 establishes 5mm as the threshold value that should be assigned to all target lesions or lymph nodes that are “too small to measure.” A 2mm difference to a tiny 4mm lesion is a 50% change, but truly, these 2mm do not accurately reflect a treatment effect, even if those changes could be accurately measured.

Discussing the remaining revisions to RECIST criteria is beyond the scope of this article. These include:

- The measurability of tumors at baseline
- The removal of bone and cystic lesions from the list of nonmeasurable lesions, enlarging the definition of complete response (to encompass lymph nodes)
- Recommendations for measuring lesions that split or coalesce. RECIST 1.1 also specifically addresses how lesion reappearance should impact a prior conclusion about patient response to treatment.

RECIST 1.1. Overall Response Tables

Target Lesion	Nontarget Lesion	New Lesion	Overall Response
CR	CR	No	CR
CR	Non-CR/non-PD	No	PR
CR	NE	No	PR
PR	Non-PD or NE	No	PR
SD	Non-PD or NE	No	SD
Not all evaluated	Non-PD	No	NE
PD	Any	Yes or No	PD
Any	PD	Yes or No	PD
Any	Any	Yes	PD

CR – complete response
PR – partial response

SD – stable disease
PD – progressive disease

NE – non-evaluable

Source: Perceptive Informatics, www.recist.com.

Table 2. A breakdown of how subjects with measurable disease react to treatment of lesions.

Added complexity

Although RECIST 1.1 has several potentially important advantages over the original version, implementing the new criteria could be challenging for both sites and sponsors. Perhaps most concerning is that RECIST 1.1 represents a new layer of complexity and opportunity for error among oncologists and radiologists new to clinical research, since the criteria are not routinely used in daily clinical practice. Only a fraction of physicians are even vaguely familiar with the guidelines, particularly those at nonacademic sites that do not regularly participate in clinical studies.

For seasoned investigative sites, the bigger issue may be simultaneously juggling the old and new criteria, as a mass migration to RECIST 1.1 is at best several years away. There are still multiple trials ongoing with the old RECIST, and sponsors are opting to use the original criteria for some of their upcoming studies—especially if they’re planning trials for an investigational compound that will be compared to a marketed product that was investigated using RECIST 1.0. There are also practical challenges that RECIST 1.1 present for sites, such as differential cataloging and reporting of nodal versus non-nodal target lesions.

Measuring lymph nodes in conformance with the new criteria will be particularly demanding, if only because the methodology is new and entirely untried. While it may be useful for radiologists to have oncology analysis software with automated lesion measurements and volumetric calculations,⁸ the reality is that most sites still do all their measurements manually.

Given the measurement complexities that arise with RECIST 1.1, site training and coaching are crucial.

Making the switch

As with any criterion involving tumor measurements, the key to accurate and reproducible assessment of response to treatment in clinical practice and clinical trials is the involvement of a radiologist experienced in cancer imaging. The assessment of response not only requires precise tumor size measurements but also in-depth understanding of the complications of cancer therapies and a knowledge of the disease-specific patterns of tumor worsening.

Given the measurement complexities that arise with RECIST 1.1, site training and coaching are even more crucial today. Since optimal results from RECIST assessments are dependent on radiologists as well as oncologists, ideally both parties should attend investigator meetings for formal training. Web conferencing, teleconferencing, and prerecorded DVDs are also readily available for site-based RECIST training. The Web site www.recist.com, which is devoted to RECIST and hosted by Perceptive Informatics, is a self-education tool that includes real-life examples of how to apply the new rules.

Sponsors planning studies using central imaging should talk with their imaging vendor before finalizing protocols and, if RECIST 1.1 is under consideration, decide how to accurately train physicians to use the criteria. Some of its features, such as assessment of progression and lymph nodes, may be particularly important for specific trials and indications. On the other hand, for Phase I and II trials, sponsors might want to try out the advantages of RECIST 1.1 and perhaps pursue regulatory endorsement.

Switching to RECIST 1.1 is generally not advised if comparison to historical data or previous trials of the same compound or other indications is planned. The breast cancer treatment Herceptin has been studied extensively utilizing the original RECIST, making an apples-to-apples comparison impossible against any other therapy trying to prove its noninferiority based on RECIST 1.1 assessment guidelines, especially considering the measurement of frequently present lytic bone lesions as target lesions in this patient population, which was not an option previously.

The new criteria are also difficult to implement for treatments aimed at mesothelioma, which grows haphazardly, or prostate cancer, which grows slowly once it metastasizes to the skeletal system. Trials of therapies for gastrointestinal stromal tumors (GIST), in which the tumor may never fully disappear, would also be poor RECIST candidates. So would experimental treatments for hepatocellular carcinoma, due to the relative complexity

of viable tumor tissue versus necrotic tissue available for evaluation. For studies of these types, sponsors may want to consult with their imaging vendor about making indication-specific modifications to RECIST via medical imaging charters for centralized review.

Some study sponsors may simply want to wait to see how regulatory agencies respond to submissions utilizing the new RECIST standards. In general, the U.S. Food & Drug Administration looks for a tight fit between the protocol, statistical analysis plan, and the assessment criteria defined in the medical imaging charter. RECIST, with or without adaptations to it, or any other method or self-developed criteria may be acceptable to the agency as long as it makes scientific sense, is reasonably explained and justified, and supports a sponsor's claim of efficacy.

Consistency is imperative if sponsors hope to get study results approved by regulators, regardless of which set of imaging rules and requirements get adopted. Significant levels of variability in how investigators follow RECIST could effectively erase many of its newfound advantages.

References

1. P. Therasse, S.G. Arbuck, E.A. Eisenhauer et al., "New Guidelines to Evaluate the Response to Treatment in Solid Tumors (RECIST Guidelines)," *Journal of the National Cancer Institute*, 92 (3) 205-216 (2000).
2. E.A. Eisenhauer, P. Therasse, J. Bogaerts et al., "New Response Evaluation Criteria in Solid Tumours: Revised RECIST Guideline (Version 1.1)," *European Journal of Cancer*, 45 (3) 228-247 (2009).
3. J. Verweij, P. Therasse, E. Eisenhauer, "Cancer Clinical Trial Outcomes: Any Progress in Tumour-Size Assessment?" *European Journal of Cancer*, 45 (2) 225-227 (2009).
4. J.E. Husband, L.H. Schwartz, J. Spencer et al., "Evaluation of the Response to Treatment of Solid Tumours—A Consensus Statement of the International Cancer Imaging Society," *British Journal of Cancer*, 90 (12) 2256-2260 (2004).
5. S.H. Kim, S.C. Kim, B.I. Choi, M.C. Han, "Uterine Cervical Carcinoma: Evaluation of Pelvic Lymph Node Metastasis with MR Imaging," *Radiology*, 190 (3) 807-811 (1994).
6. SNM, www.snm.org.
7. J. Bogaerts, R. Ford, D. Sargent et al., "Individual Patient Data Analysis to Assess Modifications to the RECIST Criteria," *European Journal of Cancer* 45 (2) 248-60 (2009).
8. L.H. Schwartz, J. Bogaerts, R. Ford, L. Shankar, P. Therasse, S. Gwyther, E.A. Eisenhauer, "Evaluation of Lymph Nodes with RECIST 1.1," *European Journal of Cancer*, 45 (2) 261-267 (2009).

Oliver Bohnsack, * MD, PhD, MBA, is Senior Medical Director, Head of Oncology, email: oliver.bohnsack@perceptive.com, and **Annette Schmid**, PhD, is Associate Medical Director, Associate Head of Oncology, both for Perceptive Informatics.

*To whom all correspondence should be addressed.