

Response Criteria in Oncologic Imaging: Review of Traditional and New Criteria¹

Temel Tirkes, MD • Margaret A. Hollar, DO • Mark Tann, MD • Marc D. Kohli, MD • Fatih Akisik, MD • Kumaresan Sandrasegaran, MD

ONLINE-ONLY SA-CME

See www.rsna.org/education/search/RG

LEARNING OBJECTIVES

After completing this journal-based SA-CME activity, participants will be able to:

- Discuss imaging for the assessment of tumor response in oncology patients.
- List the criteria that would be most accurate for different tumor types.
- Describe tumor response to treatment on the basis of specific criteria.

TEACHING POINTS

See last page

There has been a proliferation and divergence of imaging-based tumor-specific response criteria over the past 3 decades whose purpose is to achieve objective assessment of treatment response in oncologic clinical trials. The World Health Organization (WHO) criteria, published in 1981, were the first response criteria and made use of bidimensional measurements of tumors. The Response Evaluation Criteria in Solid Tumors (RECIST) were created in 2000 and revised in 2009. The RECIST criteria made use of unidimensional measurements and addressed several pitfalls and limitations of the original WHO criteria. Both the WHO and RECIST criteria were developed during the era of cytotoxic chemotherapeutic agents and are still widely used. However, treatment strategies changed over the past decade, and the limitations of using tumor size alone in patients undergoing targeted therapy (including arbitrarily determined cutoff values to categorize tumor response and progression, lack of information about changes in tumor attenuation, inability to help distinguish viable tumor from nonviable components, and inconsistency of size measurements) necessitated revision of these criteria. More recent criteria that are used for targeted therapies include the Choi response criteria for gastrointestinal stromal tumor, modified RECIST criteria for hepatocellular carcinoma, and Immune-related Response Criteria for melanoma. The Cheson criteria and Positron Emission Tomography Response Criteria in Solid Tumors make use of positron emission tomography to provide functional information and thereby help determine tumor viability. As newer therapeutic agents and approaches become available, it may be necessary to further modify existing anatomy-based response-assessment methodologies, verify promising functional imaging methods in large prospective trials, and investigate new quantitative imaging technologies.

©RSNA, 2013 • radiographics.rsna.org

Abbreviations: FDG = 2-[fluorine 18]fluoro-2-deoxy-D-glucose, GIST = gastrointestinal stromal tumor, HCC = hepatocellular carcinoma, IrRC = Immune-related Response Criteria, mRECIST = modified RECIST, PERCIST = Positron Emission Tomography Response Criteria in Solid Tumors, RECIST = Response Evaluation Criteria in Solid Tumors, SLD = sum of the longest diameters, SPD = sum of the products of diameters, SUV = standardized uptake value, WHO = World Health Organization

RadioGraphics 2013; 33:1323-1341 • Published online 10.1148/rg.335125214 • Content Codes: **CT** **GN** **NM** **OI**

¹From the Department of Radiology and Clinical Sciences, Indiana University School of Medicine, 550 N University Blvd, UH0663, Indianapolis, IN 46202. Presented as an education exhibit at the 2011 RSNA Annual Meeting. Received December 10, 2012; revision requested January 3, 2013, and received February 22; accepted March 6. For this journal-based SA-CME activity, the author F.A. has disclosed a financial relationship (see p 1340); the other authors, editor, and reviewers have no financial relationships to disclose. **Address correspondence to T.T.** (e-mail: atirkes@iupui.edu).

Introduction

Monitoring tumor response to treatment is an integral and increasingly important function of oncologic imaging. With oncology patients now receiving more complex therapies, there is a growing need to develop imaging methods to act as surrogate end points to replace the more traditional end points of morbidity or mortality. Distinguishing as early as possible between patients who are responding to a particular treatment and those who are not can maximize the effectiveness of patient care. Currently, imaging assessment of treatment response is more germane to the drug development process through clinical trials than to routine clinical use. **The ability to marry imaging findings with new clinical end points has become important in cancer therapy trials conducted to assess a new generation of targeted molecules for cancer treatment. This paves the way for much more rapid drug evaluation and, potentially, clinical decision making.**

Teaching Point

The first criteria to be proposed for the standardization of methodologies for assessing treatment response were the World Health Organization (WHO) criteria and Response Evaluation Criteria in Solid Tumors (RECIST). Both sets of criteria were developed to assess response to cytotoxic chemotherapeutic agents and to monitor only changes in tumor size during the course of treatment. The use of tumor size alone has certain pitfalls and limitations that have been observed in various clinical trials, especially those in which targeted therapies are used for specific tumors (eg, gastrointestinal stromal tumor [GIST] or hepatocellular carcinoma

[HCC]). **Over the years, the WHO and RECIST criteria have been modified by combining changes in size and the morphologic and metabolic features of specific tumors to overcome the limitations of the traditional criteria.**

Teaching Point

In this article, we discuss the use of a variety of traditional and new criteria for the evaluation of tumor response at oncologic imaging.

WHO Criteria

In 1981, the WHO published the first tumor response criteria as a standard for assessing treatment response (1). The WHO criteria introduced the concept of assessing tumor burden on the basis of the sum of the products of diameters (SPD) (ie, longest overall tumor diameter and longest diameter perpendicular to the longest overall diameter) and determining response to therapy by evaluating the changes from baseline during treatment. These changes were categorized into four groups: (a) complete response (tumor not detected for at least 4 weeks); (b) partial response ($\geq 50\%$ reduction in the SPD from baseline [confirmed at 4 weeks]); (c) progressive disease ($\geq 25\%$ increase in tumor size in one or more lesions); and (d) stable disease (neither partial response, complete response, nor progressive disease) (Table 1) (2). A common criticism of the WHO criteria is that, because the SPD is used, tumors could easily be taken to represent progressive disease on the basis of minor changes in tumor size or even measurement error. For example, an increase of only 12% in each dimension would result in a 25% increase in tumor size. In addition, the original WHO criteria were not explicit as to how many lesions should be measured, how small a lesion could be measured, or how progression should be defined. During the decades that followed the introduction of the original

Table 1
Comparison of WHO, RECIST 1.1, Choi, mRECIST, and PERCIST Tumor Response Criteria

Response	WHO*	RECIST 1.1	Choi†	mRECIST‡	PERCIST§
Complete response	No lesions detected for at least 4 weeks	Disappearance of all target lesions or lymph nodes <10 mm in the short axis	Disappearance of all target lesions	Disappearance of arterial phase enhancement in all target lesions	Disappearance of all metabolically active tumors
Partial response	≥50% decrease in SPD (confirmed at 4 weeks)	>30% decrease in sum of longest diameters (SLD) of target lesions	≥10% decrease in tumor size or ≥15% decrease in tumor attenuation at computed tomography (CT); no new lesions	>30% decrease in SLD of “viable” target lesion (arterial phase enhancement)	>30% (0.8-unit) decline in SUL peak between the most intense lesion before treatment and the most intense lesion after treatment
Progressive disease	≥25% increase in SPD in one or more lesions; new lesions	>20% increase in SLD of target lesions with an absolute increase of ≥5 mm; new lesions	≥10% increase in SLD of lesions; does not meet the criteria for partial response by virtue of tumor attenuation, new intratumoral nodules, or an increase in the size of the existing intratumoral nodules	>20% increase in SLD of “viable” target lesion (arterial phase enhancement)	>30% (0.8-unit) increase in SUL peak or confirmed new lesions
Stable disease	None of the above	None of the above	None of the above	None of the above	None of the above

Note.—SUL = lean body mass–normalized standardized uptake value (SUV).

*Measurements are calculated as the SPD.

†Used for GIST.

‡Modified RECIST (used for HCC).

§Positron Emission Tomography Response Criteria in Solid Tumors, used with 2-[fluorine 18]fluoro-2-deoxy-D-glucose (FDG) positron emission tomography (PET). The four response categories are complete metabolic response, partial metabolic response, progressive metabolic disease, and stable metabolic disease.

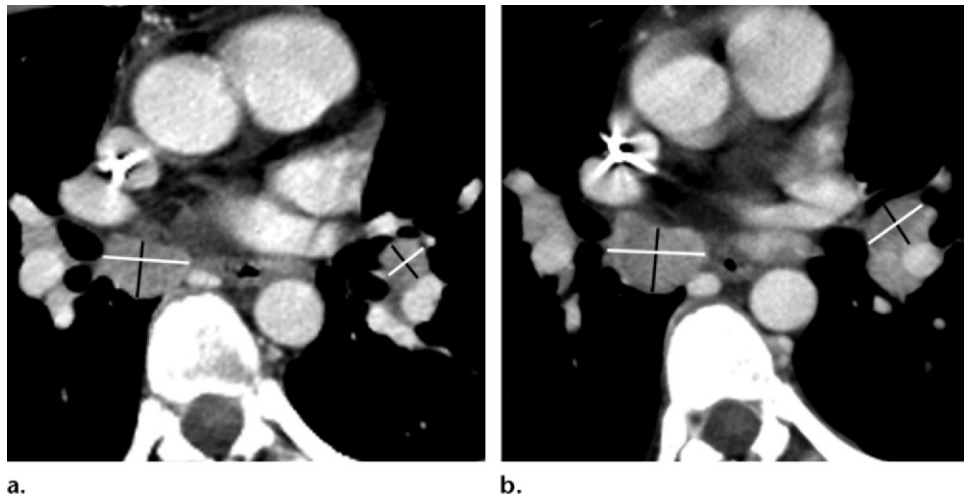


Figure 1. Comparison of treatment response according to WHO, RECIST 1.0, and RECIST 1.1 criteria in a 39-year-old woman with breast cancer. White lines = longest diameter, black lines = longest perpendicular diameter. **(a)** Axial CT image shows two metastatic lymph nodes. The WHO criteria make use of the SPD; RECIST 1.0 uses the SLD of all target lesions; and RECIST 1.1 uses the shortest diameters of the lymph nodes and the longest diameter of the target lesion. In this baseline study, SPD = 455, RECIST 1.0 SLD = 35 mm, and RECIST 1.1 SLD = 28 mm. **(b)** Follow-up CT image shows an increase of a few millimeters in the size of the lymph nodes. SPD increased to 569 (25% change), RECIST 1.0 SLD increased to 39 mm (11% change), and RECIST 1.1 SLD increased to 33 mm (18% change). On the basis of these measurements, treatment response would be categorized as progressive disease by the WHO criteria and as stable disease by both RECIST 1.0 and 1.1. Because the size of the lymph nodes increased in short diameter more than in long diameter, there is a 7% difference in SLD between RECIST 1.0 and 1.1. As seen in this example, the treatment response category can vary depending on which criteria are used. Because the WHO criteria make use of the product of the diameters, they have been criticized as yielding results that are overly sensitive to small changes in tumor size or possible measurement errors. RECIST 1.1 added the requirement that target lymph nodes be at least 15 mm in short-axis diameter.

WHO criteria, cooperative groups and pharmaceutical companies often modified these criteria to accommodate new treatments or to address areas that were unclear in the original document.

RECIST Criteria

In 2000, the WHO, the National Cancer Institute, and the European Organization for Research and Treatment of Cancer proposed the new RECIST criteria (3). The original RECIST criteria (RECIST version 1.0) were largely based on a retrospective statistical evaluation of measurements obtained in eight pharmaceutical-sponsored clinical trials involving 569 patients (4). The RE-

CIST criteria have been adopted by academic institutions, regulatory authorities, and the pharmaceutical industry, in which the primary end points are objective response or progression. Key features of the original RECIST criteria include definitions of minimum size of measurable lesions, use of a measurement in only one dimension (ie, longest diameter), and details on how to use the new imaging technologies such as spiral CT. Progressive disease was defined as the appearance of new lesions or a greater than 20% increase in the smallest SLD (versus an increase of 25% or more according to the WHO criteria) (Fig 1). In addition, suspicious findings must be unequivocal for a diagnosis of progressive disease.

Table 2
Summary of Key Changes for WHO, RECIST 1.0, and RECIST 1.1 Criteria

Criterion	WHO	RECIST 1.0	RECIST 1.1
Definition of “measurable” lesions	Should be measurable in two dimensions, no minimum lesion size	Minimum size = 10 mm at spiral CT, 20 mm at conventional CT	Minimum size = 10 mm at CT
Method of measurement	SPD	Longest diameter	Longest diameter (except in lymph nodes)
Lymph nodes	Unspecified	Unspecified	Short axis: target lesions ≥ 15 mm, nontarget lesions = 10–15 mm, nonpathologic lesions < 10 mm
Definition of progressive disease	$\geq 25\%$ increase in SPD	20% increase in SLD or new lesions, unequivocal progression considered to indicate progressive disease	$> 20\%$ increase in SLD; ≥ 5 -mm increase in size; new lesions; detailed description of unequivocal progression
Number of lesions measured	N/A	10 lesions (≤ 5 in any one organ)	Five lesions (≤ 2 in any one organ)
New lesions	N/A	N/A	Provides guidance as to when a lesion is considered new (ie, representative of progressive disease)
Guidance for imaging studies	N/A	CT, MRI, chest radiography	CT, MRI, FDG PET

Note.—MRI = MR imaging, N/A = not applicable.

RECIST Version 1.1

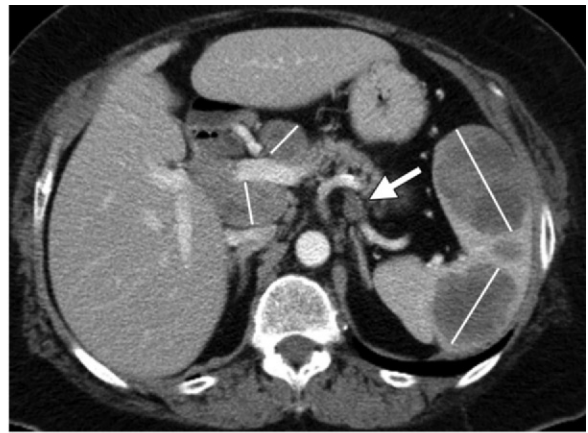
A number of questions and issues have arisen since the introduction of the original RECIST criteria, including the assessment of lymph nodes and the use of newer imaging technologies such as multidetector CT and magnetic resonance (MR) imaging. The RECIST Working Group revised the original criteria in 2009 to address these issues (5). RECIST version 1.1 was established based on the analysis of a significantly larger database of over 6500 patients (6).

Methods of Measurement

Some of the modifications and additions that accompanied the new criteria are listed in Table 2. All target lesions must be measured in their longest dimension, except for the lymph nodes, whose *shortest* diameter is used to define pathologic enlargement. To be considered measurable,

target lesions must be at least 10 mm in longest diameter and lymph nodes must be at least 15 mm in the short axis (Fig 2a). Lesions less than 10 mm in longest diameter and lymph nodes less than 15 mm in the short axis were not considered to be target lesions. Most measurements are made in the axial plane, but some tumors (eg, paraspinal lesions) may be measured in the coronal or sagittal plane if the CT reconstructions in these planes are isotropic or the images are MR images (Fig 2b).

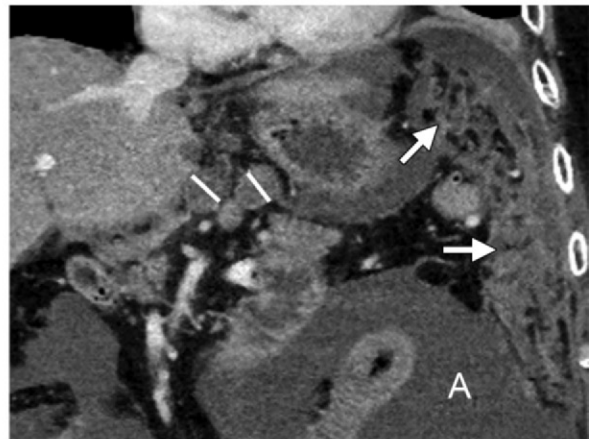
CT or MR imaging was recommended for assessment of the lytic or mixed lytic-blastic skeletal lesions to measure the soft-tissue component, as long as this component meets the criteria described earlier. Blastic bone lesions were considered “nonmeasurable.” Solid lesions rather than



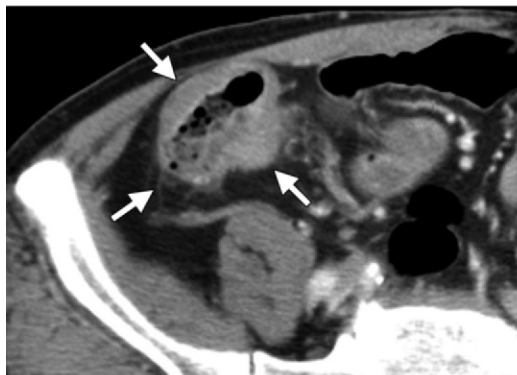
a.



b.



c.



d.



e.

Figure 2. Target lesion selection and measurements with RECIST 1.1. **(a)** Number and size of target lesions. Axial contrast material–enhanced image in a 57-year-old woman with a history of metastatic melanoma shows three metastatic lesions in the spleen, as well as three metastatic lymph nodes adjacent to the portal vein and posterior to the splenic artery. According to RECIST 1.1, up to two of the largest, most well-defined lesions per organ are measured in the longest dimension (long lines in the spleen). The two largest nodes in the porta hepatis qualify as measurable lesions because their shortest axes (short lines) exceed 15 mm. The third lymph node, located posterior to the pancreas (arrow), measures 9 mm in the short axis and therefore cannot be used as a target lesion according to RECIST 1.1. **(b)** Measurement plane. Sagittal T2-weighted MR image in a 56-year-old woman shows an intramedullary mass of the cervical cord extending from C2 to C6. The longest diameter of the mass (line) was measured in the sagittal rather than axial plane. RECIST 1.1 allows sagittal or coronal MR images to be used for target lesion measurements. It is also acceptable to measure these lesions in the coronal or sagittal plane at CT if reconstructed images are created using an isotropic image matrix. **(c)** Nonmeasurable versus measurable lesions. Coronal contrast-enhanced image in a 66-year-old patient with metastatic sigmoid adenocarcinoma shows evidence of peritoneal carcinomatosis (arrows) and lymph nodes (lines) in the gastrohepatic ligament. Peritoneal disease has indistinct borders and can often be obscured by a significant amount of ascites (*A*). It would be difficult to measure lesion size accurately on follow-up images; therefore, the lesions should be considered nonmeasurable. The lymph nodes in the gastrohepatic ligament may be used as the target lesions as long as they meet the appropriate selection criteria. **(d)** Axial CT image in a 67-year-old patient with infiltrative adenocarcinoma of the cecum shows a mass circumferentially involving the bowel wall (arrows). This primary tumor should not be used as a target lesion, since it arises from a hollow organ and its longest dimension cannot be defined on axial images, making it nonmeasurable. **(e)** Coalescing lesions. Axial contrast-enhanced image in a 43-year-old woman with a history of ovarian carcinoma shows four low-attenuation metastatic liver lesions. The two largest lesions were selected as target lesions in a prior study; however, they have progressed and currently abut one another. According to RECIST 1.1 guidelines, these two lesions should still be measured separately and in the longest dimension (lines), as long as there is a clear plane of separation (arrow) between them. This separation allows measurement of the lesions in their longest axes. If two lesions have completely coalesced such that they are no longer separable, the longest diameter of the merged lesion should be measured.

cystic metastases were recommended as target lesions. Target lesions, located in a previously irradiated area, are not considered measurable unless there has been a change in lesion size.

Assessment of Disease Progression

At baseline, a maximum of five lesions (up to two lesions in any one organ) are identified as target lesions. If the largest lesion does not lend itself to reproducible measurement, the next largest lesion that can be measured reproducibly should be selected (Fig 2c, 2d). The SLD is calculated (long axis for nonnodal lesions, short axis for nodal lesions) for all target lesions and reported as the baseline SLD. This baseline value is used as a reference for assessing objective tumor response at future time points. All other lesions (or sites of disease), including pathologic lymph nodes, are identified as nontarget lesions, and their presence should also be recorded at baseline.

Evaluation of Target and Nontarget Lesions

Target lesions, including lymph nodes that become “too small to measure,” should still be measured and their presence recorded at each subsequent evaluation. To qualify for characterization as complete response, each lymph node must be less than 10 mm in the short axis. If the nonnodal lesions “fragment,” the longest diameters of the fragmented portions should be added together to calculate the longest diameter of the target lesion. Similarly, as lesions coalesce, a plane may be maintained between them to aid in determining the longest diameter of each lesion (Fig 2e). If two lesions have completely coalesced such that they are no longer separable, the vector of the longest diameter should be the maximum longest diameter.

Table 3
Assessment of Treatment Response with RECIST 1.1 at Each Time Point during Treatment

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

Note.—CR = complete response, NE = not amenable to evaluation, PD = progressive disease, PR = partial response, SD = stable disease.

The appearance of new lesions denotes disease progression; therefore, it is important to comment on these lesions. However, the finding of a new lesion should be unequivocal—that is, not attributable to differences in scanning technique, change in imaging modality, or findings thought to represent something other than tumor. Equivocal new lesions (eg, lesions that are too small to measure) should be reassessed at follow-up examinations to determine whether they truly represent new disease. It is sometimes reasonable to incorporate FDG PET to assess possible new disease.

If the nontarget lesions demonstrate a change at follow-up, characterization as unequivocal progression requires that there be substantial worsening so that even in the presence of stable disease or partial response, the treating physician would think it necessary to change therapy.

Imaging Considerations

RECIST 1.1 recommended maintaining standard image acquisition parameters to allow optimal comparison between studies. CT should be performed with a section thickness of 5 mm or less. CT of the chest, abdomen, and pelvis should be performed contiguously throughout the entire anatomic region of interest. For detection of possible new lesions, follow-up studies should cover all areas in which metastatic spread of the primary tumor in question is known to occur. Specific attention should be given to a consistent dose and rate of administration of intravenous contrast material. Most solid tumors may be scanned with a single-

phase sequence after contrast material administration. Multiphase CT scans are necessary to improve lesion conspicuity for some hypervascular tumors (eg, HCC or neuroendocrine tumors).

MR imaging offers superior soft-tissue contrast and spatial and temporal resolution compared with CT, but it is also more costly and less readily available. There are many image acquisition variables involved in MR imaging, which may impact lesion conspicuity and measurement. Therefore, the same image acquisition protocol should be used on similar MR imaging hardware for baseline and follow-up examinations.

Ultrasonography (US) should not be used in clinical trials to measure tumor regression or progression because US evaluation is subjective and operator dependent.

Chest radiographic measurement of lesions surrounded by pulmonary parenchyma is acceptable but not preferred, since it represents a summation of densities. Chest CT would be the preferred method secondary to decreased sensitivity of lesion detection at chest radiography.

Time Point Response

The revised RECIST guidelines are useful for the assessment of stable disease, tumor progression, or time to progression in clinical trials. It is assumed that a new response assessment based on new imaging findings should be conducted at each follow-up examination (Table 3). New overall response status is assigned according to the status of the target, nontarget, and new lesions. For example, a patient may be accorded partial response status at the first follow-up examination and stable disease status at the second follow-

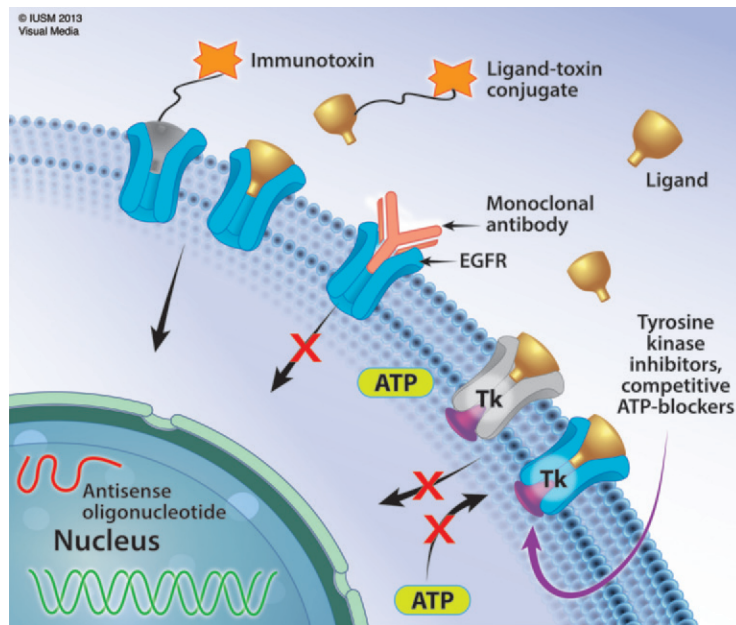


Figure 3. Mechanism of action of targeted cancer therapies. Targeted cancer therapies interfere with the proliferation and spread of cancer cells in different ways. Many of these therapies focus on proteins that are involved in cell signaling pathways, which form a complex communication system that governs basic cellular functions and activities such as cell division, cell movement, cell responses to specific external stimuli, and even cell death. The U.S. Food and Drug Administration has approved almost 40 targeted therapy drugs, and many others are currently under clinical trials. Among the drugs that inhibit tyrosine kinase (*Tk*) enzyme are imatinib and sunitinib. Examples of monoclonal antibodies include trastuzumab and pertuzumab, which bind to HER-2 protein on human epidermal growth factor receptor (*EGFR*) 2. Another monoclonal antibody is bevacizumab, which binds to vascular endothelial growth factor receptor. *ATP* = adenosine triphosphate.

up examination during the course of treatment. However, if a new lesion appears at a subsequent study, the status would change to progressive disease. Best overall response is defined as the best response across all time points.

Limitations of RECIST Criteria

Although the RECIST criteria have been used extensively since their introduction, concerns about using change in tumor size as the only criterion have not been fully addressed, even in RECIST 1.1. Studies of the reliability of measurements have found that tumor size measurements made at CT are often inconsistent (7,8). In one study, the difference between measurements made by two readers was significant enough to result in misclassification rates of 29.75% for progressive disease and 13.75% for partial response (7). Even repeated measurements made by the same observer were associated with significant variability, with potential misclassification rates of

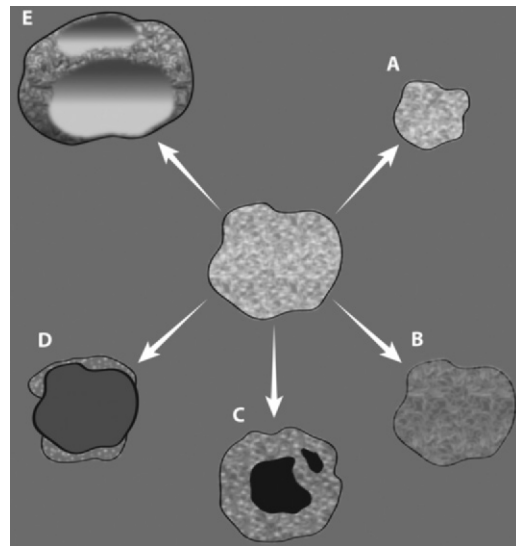
9.5% for progressive disease and 3% for partial response. Evaluation of tumor response on the basis of RECIST criteria may also be limited by problems in defining the margins of ill-defined or irregular lesions (eg, bone marrow disease).

Tumor Response Criteria in Targeted Cancer Therapies

Targeted cancer therapies make use of drugs that block the growth and spread of cancer by interfering with specific molecules involved in tumor growth and progression. These agents have significantly changed the treatment of cancer over the past 10 years. The U.S. Food and Drug Administration has approved many targeted cancer therapies for the treatment of specific types of cancer. The mechanisms of action of targeted therapies differ from those of traditional cytotoxic chemotherapy (Fig 3).

Teaching Point

Figure 4. Different treatment responses to targeted therapies. Treatment response observed with targeted therapies can be different from that observed with conventional chemotherapies and can include decrease in lesion size (*A*), decrease in lesion vascularity with or without a significant change in size (*B*), stability or decrease in size with cavitations (lung tumors) (*C*), cystic changes with or without a decrease in size (*D*), and intratumoral hemorrhage with or without a change in size (*E*).



Some agents can induce apoptosis; however, some agents stop progression. Because of differences in the mechanism of action, tumors treated with targeted therapies do not necessarily demonstrate the same radiographic findings as tumors treated with standard cytotoxic therapies (Fig 4) (9). Therefore, traditional anatomic size-based criteria can lead to the miscategorization of treatment response for tumors like GIST, HCC, or melanoma when treated with targeted therapies.

Choi Response Criteria

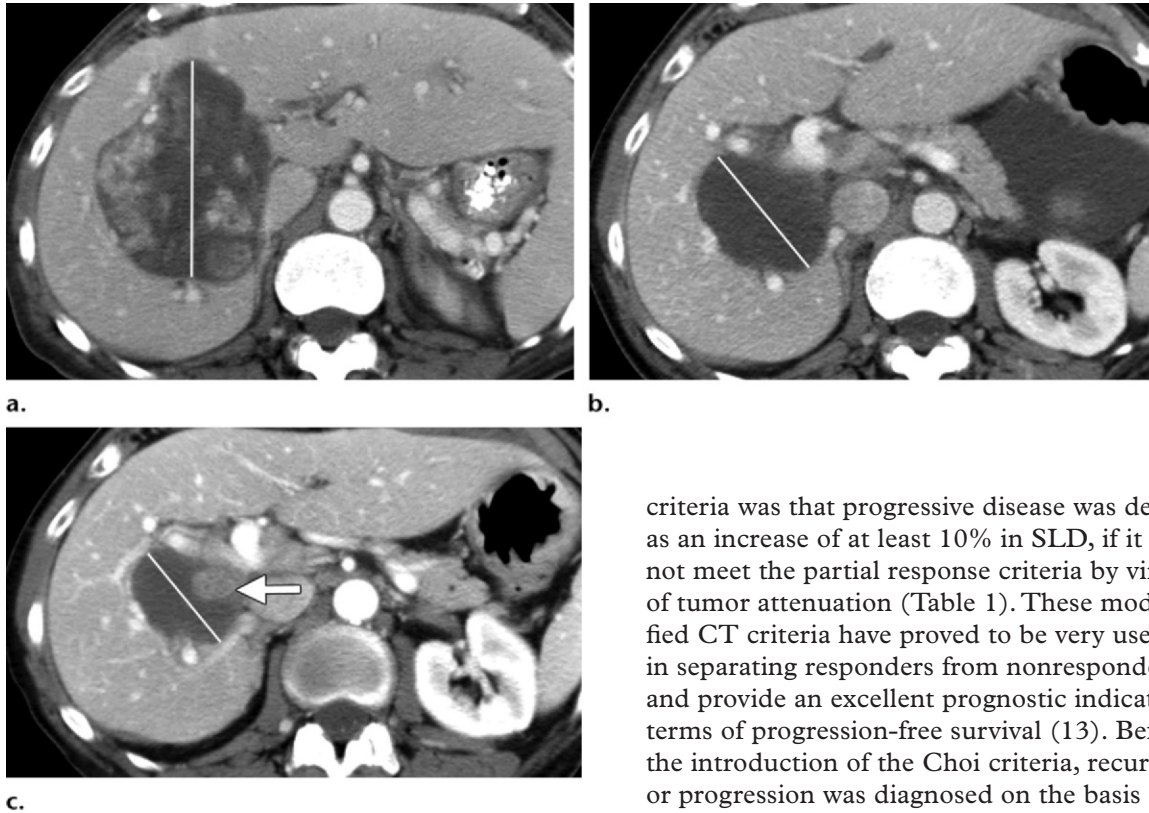
The therapeutic options for advanced GISTs were limited until the introduction of imatinib, a competitive inhibitor of tyrosine kinase receptor that has demonstrated remarkable efficacy. During the course of treatment with imatinib, tumor size usually decreases; however, changes in tumor dimension do not necessarily reflect tumor response (Fig 5) (10,11). In some cases, size can actually increase secondary to internal hemorrhage, necrosis, or myxoid degeneration (10). Decrease in tumor size is usually minimal dur-

ing the early stages of posttreatment, whereas dramatic changes in internal characteristics (eg, tumor attenuation, nodularity, and number of vessels) will occur. **The Choi response criteria for GIST proposed that tumor attenuation could provide an additional measure of response to imatinib therapy.** The response can be seen very early during treatment (10).

PET has been found to be highly sensitive in detecting early response and to be useful in predicting long-term response to imatinib in patients with metastatic GIST (12). There is good correlation between the responses based on overall tumor burden, CT attenuation, and maximum SUV (SUV_{max}) at FDG PET (10). However, the availability of PET is still limited, and, in up to 21% of patients, pretreatment glucose uptake is not sufficient to be detected with FDG PET (10). In an attempt to achieve better response evaluation with CT, partial response status was redefined as a decrease in SUV at FDG PET (<70% from baseline or $SUV_{max} < 2.5$) (12). Among patients in whom treatment response was seen at FDG PET, 97% had a decrease in tumor size of at least 10% or a decrease in tumor attenuation of at least

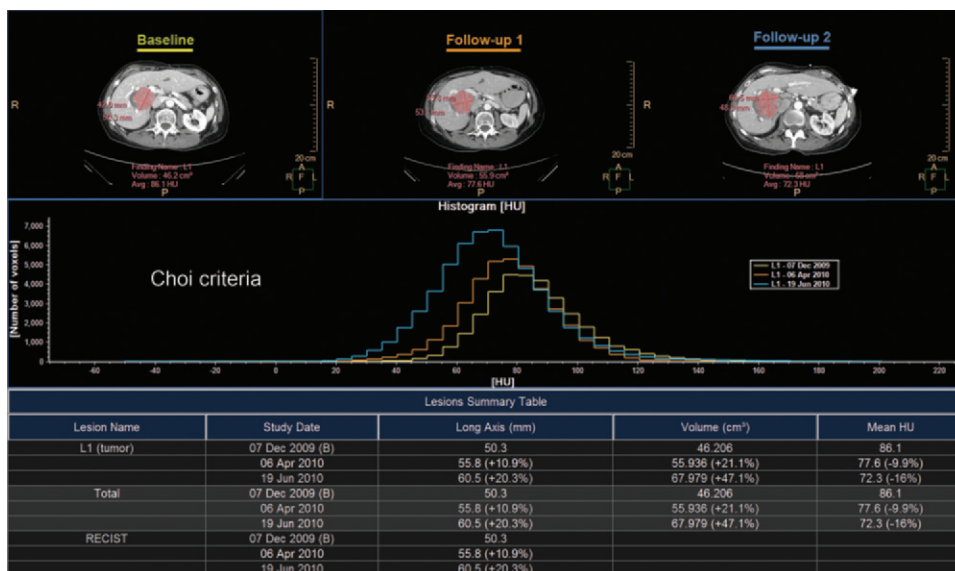
Teaching Point

Figure 5. Limitations of size criteria in assessing the activity of newer cancer therapies that stabilize disease. **(a)** On an axial contrast-enhanced baseline CT image in a 60-year-old woman with a history of metastatic GIST, the largest metastasis is in the right hepatic lobe and demonstrates nodular internal enhancement. It was selected as the target lesion and measured in the longest axis (line). **(b)** On an axial contrast-enhanced CT image after targeted therapy, the longest axis of the tumor (line) is 38% shorter than on the baseline image. According to both the RECIST and Choi criteria, this change represents a partial response to therapy. In addition, the average attenuation of the tumor decreased 61%. **(c)** Follow-up CT image shows a further decrease in the size of the tumor (49% smaller than on the baseline image). However, there is a new enhancing nodule in the tumor (arrow), which should be considered to represent recurrent disease. On the basis of RECIST, this would still be categorized as partial response, since these criteria do not take changes in intratumoral morphology into account. On the other hand, with the Choi criteria, tumor attenuation as well as tumor size are measured, and a new increase in tumor attenuation (secondary to the enhancing nodule) would be correctly categorized as progressive disease.

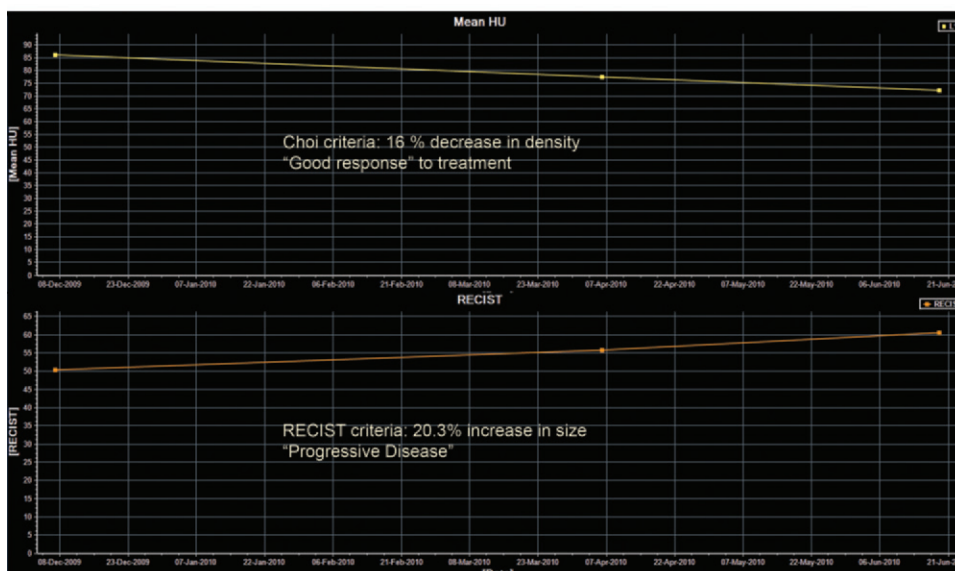


15% at CT after 8 weeks of imatinib treatment (11). On the basis of these results, the new criteria used a combination of tumor attenuation ($\geq 15\%$ decrease) and modified tumor size ($\geq 10\%$ decrease) to assess partial response. Another important difference from the RECIST

criteria was that progressive disease was defined as an increase of at least 10% in SLD, if it does not meet the partial response criteria by virtue of tumor attenuation (Table 1). These modified CT criteria have proved to be very useful in separating responders from nonresponders and provide an excellent prognostic indicator in terms of progression-free survival (13). Before the introduction of the Choi criteria, recurrence or progression was diagnosed on the basis of an increase in tumor size and identification of new lesions at either local or distant sites. In GISTs, an increase in tumor size is still important; however, recurrence may occur within the treated hypoattenuating tumor without a change in tumor size (Fig 6) (14).



a.



b.

Figure 6. Choi versus RECIST response criteria in evaluation of a 60-year-old patient with metastatic GIST who underwent targeted therapy with imatinib. **(a)** Screen shot from the multimodality tumor-tracking software Intellispace Portal (Philips Healthcare, Best, the Netherlands) shows longitudinal comparisons of tumor attenuation and size from three CT examinations (one baseline scan, two follow-up scans) performed during the course of treatment. For each scan, the tumor is semiautomatically segmented (red lines) to measure its attenuation and longest diameter. Hounsfield units (HU) are displayed in histograms to assess tumor attenuation using the Choi criteria. Each histogram depicts the number of voxels along the vertical axis and tumor attenuation along the horizontal axis. Note the increase in size (increased number of voxels) and decrease in average attenuation (histogram mean value shifting to the left). (Courtesy of Jeffrey H. Yanof, PhD.) **(b)** Line charts show average tumor attenuation according to the Choi criteria (top) and size according to RECIST (bottom) over time. There was an increase in size from the baseline study as demonstrated on the RECIST graph (50.3 mm at baseline versus 60.5 mm on the most recent study [20.3% increase]). According to RECIST, an increase in size of over 20% would be consistent with progressive disease. During the course of treatment, the size of the GIST may increase secondary to internal hemorrhage, necrosis, or myxoid degeneration. The mean Hounsfield units (HU) chart demonstrates decreasing tumor attenuation from baseline to follow-up (86.1 vs 72.3 HU [16% decrease]). According to the Choi criteria, a decrease in tumor attenuation of 15% or more is considered a partial response. In this case, use of RECIST would lead to underestimation of tumor response.

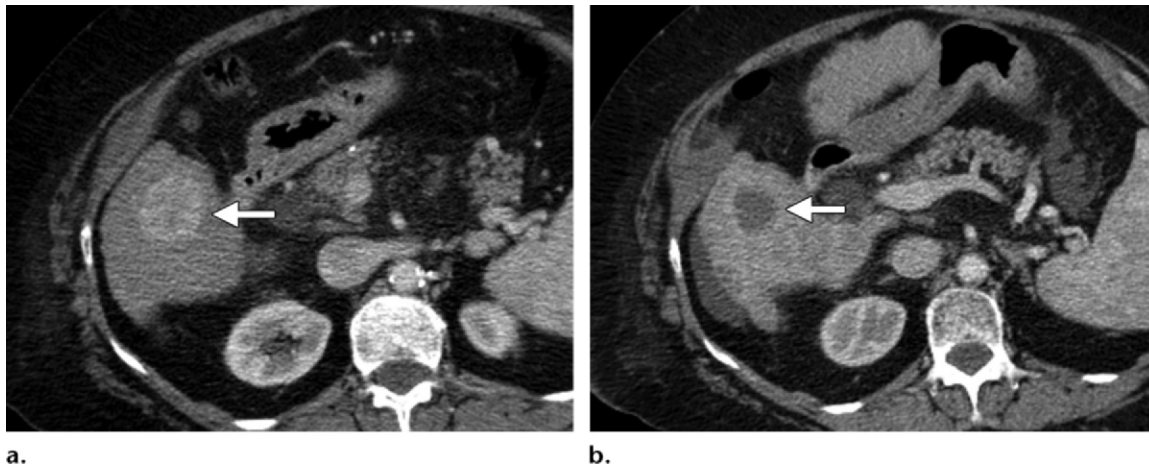


Figure 7. Use of mRECIST versus RECIST in a 76-year-old woman with nonalcoholic steatohepatitis–related cirrhosis complicated by HCC. **(a)** Axial contrast-enhanced arterial phase CT image shows a 3.2-cm hypervascular tumor (arrow) in the right hepatic lobe. The patient underwent yttrium-90 glass microsphere transarterial therapy (TheraSphere; Nordion, Ottawa, Ontario, Canada). **(b)** Arterial phase CT image 6 weeks after intervention shows loss of hypervascularity in the tumor (arrow), compatible with complete response according to mRECIST. On the basis of mRECIST, viable tumor is defined as contrast material uptake during dynamic arterial phase CT or MR imaging. On the other hand, the decrease in tumor size from 3.2 cm to 2.5 cm (22% change) is compatible with stable disease according to RECIST. During assessment of the HCC, it is important to perform all CT scans consistently during the hepatic arterial phase. Modified RECIST required optimization of image acquisition protocols and consistent use of the same protocol throughout all follow-up examinations.

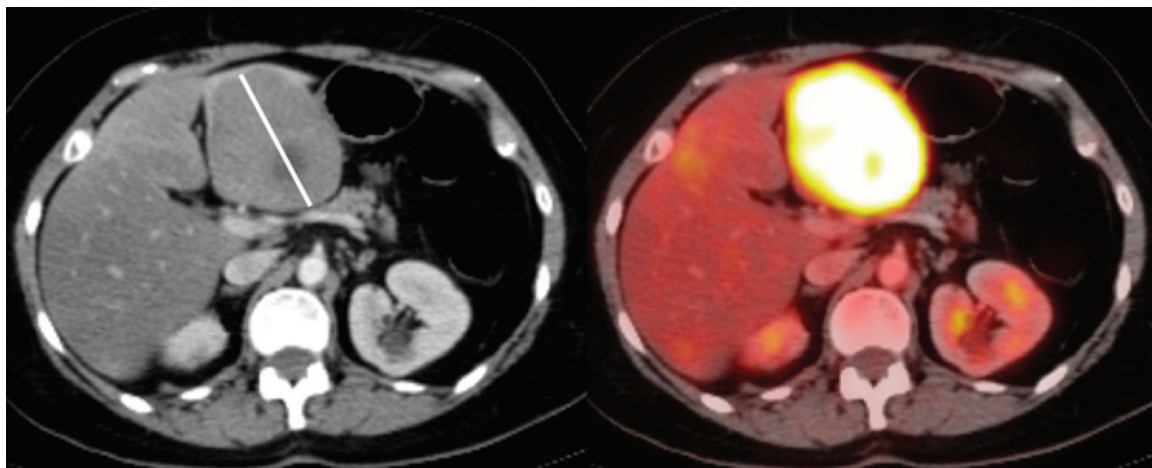
Attempts have been made to use the Choi response criteria in the assessment of other solid tumors. A recent study found that the Choi criteria may be helpful in assessing early metastatic renal cell carcinoma treated with sunitinib, but the use of these criteria did not change patient management (15). A pilot study showed that the Choi criteria were superior to the RECIST criteria in assessing the response of soft-tissue sarcoma to chemotherapy and radiation therapy (16). Nevertheless, more studies are needed for further evaluation.

Modified RECIST

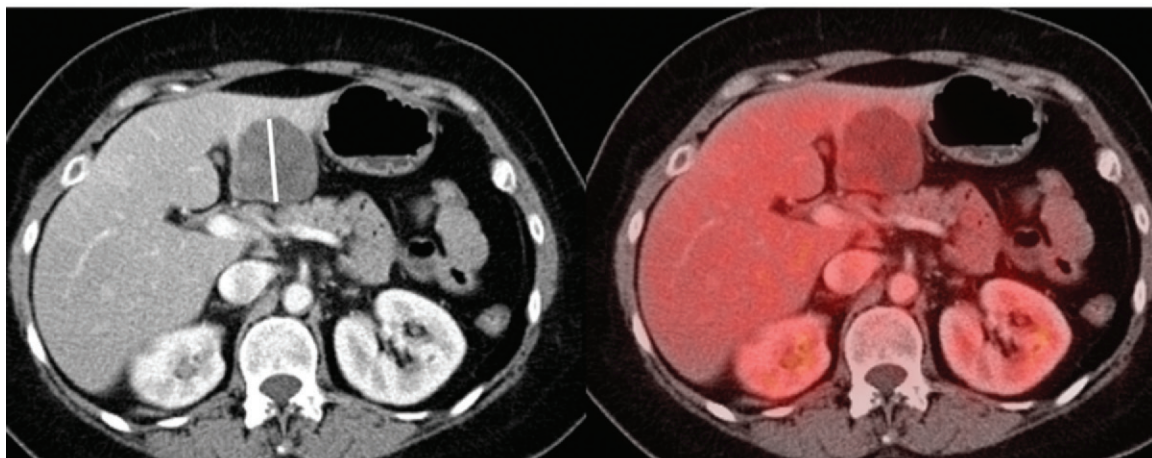
Assessments based solely on changes in tumor size can also be misleading when applied to HCC being treated with targeted therapies (eg, with sorafenib) or interventional therapies (17–19). In 2000, a panel of experts on HCC convened by the European Association for the Study of the Liver proposed that estimation of viable tumor with contrast-enhanced imaging should be the optimal method for assessing treatment response (20). The new criteria, referred to as mRECIST, were subsequently endorsed by the American Association for the Study of Liver Diseases (21). Viable tumor was defined as uptake of contrast agent during arterial phase dynamic CT or MR imaging (Table 1). On the basis of this assumption, the disappearance of arterial phase enhance-

ment in all target lesions was considered to represent a complete response (Fig 7).

The new guidelines emphasized the optimization of image acquisition protocols and consistency in the use of the same protocol throughout follow-up (18). Patients can be followed up with either contrast-enhanced spiral CT or contrast-enhanced MR imaging. The liver must be imaged using a dual-phase protocol with either modality. Delayed equilibrium phase imaging may be useful, but it is not mandatory and should be performed only if it is part of clinical practice. The viable tumor should be measured during the arterial phase. To be selected as a target lesion, the lesion should be classified as a measurable lesion according to RECIST criteria, suitable for repeat measurement and showing enhancement during the arterial phase. Infiltrative-type HCC should be considered a nontarget lesion if the mass is not well defined and does not appear to be amenable to accurate measurement. Because of variability in internal necrosis, the longest diameter of the viable tumor can be located in a plane different from that in which the baseline diameter was measured. Malignant portal vein thrombosis should be considered a nonmeasurable lesion, since the tumor may be obscured by the presence of a bland thrombus during the course of



a.



b.

Figure 8. Assessment of tumor response according to PERCIST versus RECIST in a 69-year-old woman with metastatic GIST who was treated with imatinib. **(a)** Axial pretreatment contrast-enhanced CT image (left) and fused PET/CT image (right) show an FDG-avid tumor in the left lobe that measured 7.2 cm (line) on the baseline image. **(b)** On contrast-enhanced CT (left) and PET/CT (right) images 2 months after treatment, the tumor has decreased to 4.3 cm (line on CT image) and shows no increased metabolic activity. According to PERCIST, the disappearance of all metabolically active tumors is considered to represent complete metabolic response. RECIST does not take tumor metabolic activity into account, resulting in categorization of this change as partial response on the basis of size change alone, leading to underestimation of treatment response. The tumor did not recur for 2 years after treatment and continued to decrease in size.

treatment. The presence of a new lesion is considered to represent disease progression. A new lesion must have a maximum diameter of over 1 cm and show the typical vascular pattern of HCC at dynamic imaging (ie, hypervascularity in the arterial phase with washout in the portal venous or late venous phase). Otherwise, new lesions should be considered equivocal and monitored for interval growth at subsequent scans.

PERCIST Criteria

Although a range of factors have been associated with FDG uptake, there appears to be a rather

strong relationship between FDG uptake and number of cancer cells in a substantial number of studies (22,23). **Because many newer cancer therapies may be more cytostatic than cytotoxic, good tumor response may be associated predominantly with a decrease in metabolism, without a major reduction in tumor size** (Fig 8). Therefore, metabolic response as a leading indicator of tumor response may be even more predictive of outcome than morphologic criteria. It is in this context that the PERCIST criteria were proposed in 2009 (24) to refine and validate quantitative approaches to monitoring PET tumor response.

There are two basic approaches for assessing metabolic changes brought about by treatment:

Teaching Point

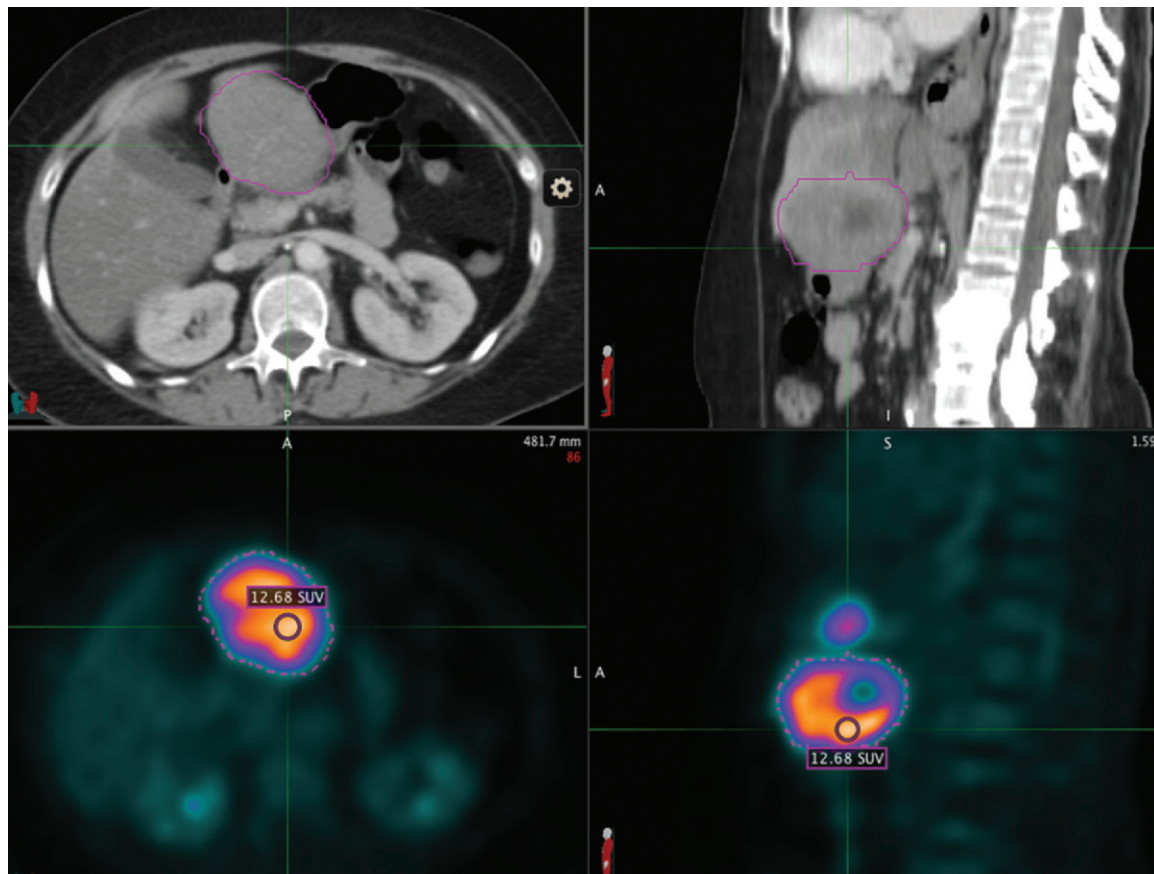


Figure 9. Use of tumor tracking–software to measure SUV. Axial (top left) and sagittal (top right) contrast-enhanced CT images and corresponding PET images (bottom) demonstrate SUV measurement using tumor-tracking software (MIM Software, Cleveland, Ohio). The border of the tumor has been semiautomatically outlined in purple. A 1-cm³ region of interest (circle on PET images) has been positioned in the area of highest recorded metabolic activity within the selected volume, providing SUV peak. Use of standard protocol for selection of the region of interest is a key aspect of reproducible SUV measurement.

qualitative and quantitative. The quantitative method was preferred by PERCIST secondary to insufficient data on the reproducibility of the reporting of qualitative treatment response among readers. Standardized quantitative assessment of metabolic tumor response with PET necessitates a consistent and reliable measurement of tumor activity. This requires identical patient preparation and adequate scan quality that is similar between the baseline and follow-up studies, which should be performed on the same scanner with comparable injected doses of FDG and uptake times. PERCIST recommended using SUL (lean body mass–normalized SUV [SUV_{lbm}]) owing to its reduced dependence on patient weight compared with standard body weight–normalized SUV (SUV_{bw}). SUL peak, which should be measured using a 1-cm³ (or 1.2-cm-diameter) fixed-dimension region of interest centered over the area of highest uptake in the tumor, was the preferred method mainly because of its widespread use. By strictly

defining the dimensions and position of the volume of interest, measurement variation can be eliminated, and averaging of multiple voxels reduces susceptibility to noise (Fig 9). PERCIST also recommends comparing the variability in SUL between studies performed in the same patient by using liver activity as the standard of reference. In addition, by stipulating that variation in SUL should be less than 20% (0.3 SUL mean units), the influence of nonpathologic variability in PET quantification across multiple time points can be reduced. An interval of at least 10 days between the last chemotherapy session and the next FDG PET study is advised. Longer and more variable time intervals after external-beam radiation therapy (8–12 weeks) have been recommended.

In PERCIST, response to therapy is evaluated as a continuous variable and expressed as a percentage change in SUL peak for the most

active lesion at each time point between the pre- and posttreatment PET/CT studies (Table 1). A complete metabolic response is defined as visual disappearance of all metabolically active tumors. A partial metabolic response is defined as a 0.8-unit (>30%) decline in SUL peak between the most intense lesion before treatment and the most intense lesion after treatment, which may not be the same lesion. A 0.8-unit (>30%) increase in SUL peak or the appearance of a new lesion is classified as progressive metabolic disease. There is no definitive recommendation regarding how many lesions should be measured. Although PERCIST has specific criteria for response based on a single target lesion, the collection of additional data on five lesions was recommended so as to develop a database suitable for future studies.

It is crucial to adhere to a standardized PET/CT scanning protocol that is characterized by consistency in injected dose, postinjection delay, reconstruction parameters, and SUV normalization technique, among other variables.

There are difficulties with imaging standardization across PET centers and tumor types. Combined with uncertainty concerning the timing of assessment relative to treatment, the use of quantitative measurements of FDG uptake for evaluating response remains difficult. However, there is a growing body of evidence suggesting that FDG PET is becoming established as a clinical technique for assessing tumor response, especially in FDG-avid lymphoma subtypes.

Cheson Response Criteria for Malignant Lymphomas

Lymphoma is another disease in which there has been development of specific tumor assessment criteria. Again, this is due to disease-specific complexity, wherein the basic assessment of interval change in size alone may not accurately reflect disease status (25). Masses often do not regress completely after curative treatment because of residual fibrosis and necrotic debris, and this stability in size does not necessarily represent viable tumor.

First created in 1997 and later revised in 2007 (26), the Cheson response criteria allow analysis

of both the size and metabolic activity of tumors during the course of treatment. The 2007 revision was necessary to incorporate advances in treatment and image-based evaluation. The major changes between the 1997 and 2007 versions were (a) discontinuation of gallium scintigraphy in favor of PET (reflecting the widespread preference of the latter modality for evaluation of tumor response), and (b) inclusion of evaluation with flow cytometry and immunohistochemistry (Table 4).

For uniformity in tumor measurements, the use of the SPD obtained in up to six dominant nodes or masses is recommended. These target lesions need to show a decrease in activity of at least 50% to attain partial response status. A size of 1.5 cm is used as a cutoff point for the inclusion of new target lesions (in lymph nodes, 1 cm in the short axis) to grant relapsed disease or progressive disease status. The presence of a post-treatment residual mass that has not disappeared or shown an interval decrease in size is considered to represent complete response, as long as the mass has become PET negative.

Immune-related Response Criteria

For cytotoxic agents, WHO and RECIST guidelines assumed that an early increase in tumor size or the appearance of new lesions signaled progressive disease, resulting in discontinuation of treatment. However, in studies with immunotherapeutic agents (eg, ipilimumab), clinical experience showed that complete response, partial response, or stable disease status could still be achieved after an increase in overall tumor burden. Therefore, conventional response criteria may not allow adequate assessment of the activity of immunotherapeutic agents. Patients whose performance status is stable and whose laboratory values have not significantly deteriorated should be considered for repeat confirmation imaging before true progressive disease status is declared and the immunotherapeutic agent is withdrawn (27).

In 2004 and 2005, a series of international workshops hosted by the Cancer Vaccine Consortium in collaboration with the International Society of Biologic Therapy of Cancer proposed additions to the WHO criteria that would allow the evaluation of unique response patterns of

Table 4
Definitions of Treatment Response According to Cheson Criteria

Response	Nodal Masses	Spleen and Liver	Bone Marrow
Complete response	All previously enlarged FDG-avid or PET-positive lymph nodes regressed to normal size (≤ 1.5 cm in greatest diameter)	Regressed in size and not palpable at physical examination, disappearance of nodules	Clearance of infiltrate at repeat biopsy; if findings at morphologic analysis are indeterminate, immunohistochemical findings should be negative
Partial response	$\geq 50\%$ decrease in SPD of up to six largest dominant masses, no increase in size of other nodes; FDG avid or PET positive before therapy, one or more nodes PET positive at previously involved site, or variably FDG avid or PET negative with regression at CT	$\geq 50\%$ decrease in SPD of nodules, no increase in size of liver or spleen	Irrelevant if findings are positive before therapy, cell type should be specified
Stable disease	FDG avid or PET positive before therapy, PET positive at prior sites of disease and no new sites at CT or PET, or variably FDG avid or PET negative with no change in size of previous lesions at CT
Relapse or progressive disease	Appearance of one or more new lesions > 1.5 cm in any axis, $\geq 50\%$ increase in SPD of more than one node, or $\geq 50\%$ increase in the longest diameter of a previously identified node > 1 cm in the short axis; lesions PET positive if FDG-avid lymphoma or PET positive before therapy	$> 50\%$ increase from nadir in SPD of any previous lesions	New or recurrent involvement

immunotherapeutic agents (28). Subsequently, the Immune-related Response Criteria (IrRC) were developed during clinical trials in patients with advanced melanoma who were receiving ipilimumab, a human monoclonal antibody that blocks cytotoxic T lymphocyte antigen-4 (CTLA-4) (27). The core novelty of the IrRC is the incorporation of measurable new lesions into a new concept of “total tumor burden” and comparison of this variable with baseline measurements. With the IrRC criteria, both the index and measurable new lesions are taken into account (in contrast to conventional WHO criteria, which do not require the measurement of new lesions and do not include new lesion measurements in the characterization of evolving tumor burden). At baseline tumor assess-

ment, the SPD of all index lesions (up to five lesions in a single organ; maximum, 10 visceral and five cutaneous index lesions) is calculated. At each subsequent time point, the SPD of the index lesions and any possible new measurable lesions are added together to calculate the total tumor burden. With this new concept, the tumor response categories have been modified from those of the WHO criteria (Table 5). The main difference between the WHO criteria and the IrRC criteria is that the former always classify new measurable lesions as progressive disease. According to the IrRC criteria, these lesions are not always viewed as progressive disease and can result in discontinuation of treatment.

Table 5
WHO versus IrRC Criteria

Response	WHO	IrRC
Complete response	Disappearance of all lesions at two consecutive observations ≥ 4 weeks apart	Disappearance of all lesions (including nonindex lesions) at two consecutive observations ≥ 4 weeks apart, no new nonmeasurable disease
Partial response	50% decrease in SPD of all lesions (confirmed at 4 weeks)	50% decrease in tumor burden (confirmed at 4 weeks)
Progressive disease	25% increase in SPD or new lesions (measurable or nonmeasurable)	25% increase in tumor burden (confirmed at 4 weeks), new measurable lesions included within tumor burden
Stable disease	None of the above	None of the above

Although potentially representing another improvement over conventional criteria for immunotherapeutic agents, the IrRC criteria have their own challenges; therefore, further prospective evaluation is warranted, particularly regarding the association with overall survival (27).

Future Trends

Radiology will continue to adapt the new tumor response concepts observed with the current and future targeted therapy agents. With the advent of molecular medicine in the era of individualized medicine, the ultimate goal of research in oncology is to tailor treatments to both the specific type of cancer and the patient. Tumor response criteria should be chosen based on treatment and type of tumor. Validation of functional biomarkers, including but not limited to FDG PET, is essential to ensure that imaging continues to keep up with the new treatment concepts in oncology.

Time will tell whether these tumor response criteria are incorporated into daily radiology practice; however, as the number of criteria increases, the resultant growing complexity makes such incorporation less likely. Many multimodal-

ity tumor-tracking software packages are commercially available, but at present they are not integrated with current clinical image-viewing workstations and are sold as separate third-party software solutions. If radiologists are to use these criteria in daily practice apart from the sponsored clinical trials, simplified and integrated software and hardware solutions will be required.

Acknowledgment.—The authors thank Jeffrey H. Yanof, PhD, for providing the screen shot from the tumor-tracking software.

Disclosures of Conflicts of Interest.—**F.A.:** *Related financial activities:* none. *Other financial activities:* consultant for Repligen and Siemens.

References

1. World Health Organization. WHO handbook for reporting results of cancer treatment. Geneva, Switzerland: World Health Organization, 1979.
2. Miller AB, Hoogstraten B, Staquet M, Winkler A. Reporting results of cancer treatment. *Cancer* 1981; 47(1):207–214.
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(3):205–216.

4. James K, Eisenhauer E, Christian M, et al. Measuring response in solid tumors: unidimensional versus bidimensional measurement. *J Natl Cancer Inst* 1999;91(6):523–528.
5. 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(2):228–247.
6. Bogaerts J, Ford R, Sargent D, et al. Individual patient data analysis to assess modifications to the RECIST criteria. *Eur J Cancer* 2009;45(2):248–260.
7. Erasmus JJ, Gladish GW, Broemeling L, et al. Interobserver and intraobserver variability in measurement of non-small-cell carcinoma lung lesions: implications for assessment of tumor response. *J Clin Oncol* 2003;21(13):2574–2582.
8. Zhao B, Schwartz LH, Moskowitz CS, Ginsberg MS, Rizvi NA, Kris MG. Lung cancer: computerized quantification of tumor response—initial results. *Radiology* 2006;241(3):892–898.
9. Zhao B, Schwartz LH, Larson SM. Imaging surrogates of tumor response to therapy: anatomic and functional biomarkers. *J Nucl Med* 2009;50(2):239–249.
10. Choi H, Charnsangavej C, de Castro Faria S, et al. CT evaluation of the response of gastrointestinal stromal tumors after imatinib mesylate treatment: a quantitative analysis correlated with FDG PET findings. *AJR Am J Roentgenol* 2004;183(6):1619–1628.
11. Choi H, Charnsangavej C, Faria SC, et al. Correlation of computed tomography and positron emission tomography in patients with metastatic gastrointestinal stromal tumor treated at a single institution with imatinib mesylate: proposal of new computed tomography response criteria. *J Clin Oncol* 2007;25(13):1753–1759.
12. Van den Abbeele AD, Badawi RD. Use of positron emission tomography in oncology and its potential role to assess response to imatinib mesylate therapy in gastrointestinal stromal tumors (GISTs). *Eur J Cancer* 2002;38(suppl 5):S60–S65.
13. Choi H. Response evaluation of gastrointestinal stromal tumors. *Oncologist* 2008;13(suppl 2):4–7.
14. Hong X, Choi H, Loyer EM, Benjamin RS, Trent JC, Charnsangavej C. Gastrointestinal stromal tumor: role of CT in diagnosis and in response evaluation and surveillance after treatment with imatinib. *RadioGraphics* 2006;26(2):481–495.
15. van der Veldt AA, Meijerink MR, van den Eertwegh AJ, Haanen JB, Boven E. Choi response criteria for early prediction of clinical outcome in patients with metastatic renal cell cancer treated with sunitinib. *Br J Cancer* 2010;102(5):803–809.
16. Stacchiotti S, Collini P, Messina A, et al. High-grade soft-tissue sarcomas: tumor response assessment—pilot study to assess the correlation between radiologic and pathologic response by using RECIST and Choi criteria. *Radiology* 2009;251(2):447–456.
17. Vossen JA, Buijs M, Kamel IR. Assessment of tumor response on MR imaging after locoregional therapy. *Tech Vasc Interv Radiol* 2006;9(3):125–132.
18. Llovet JM, Di Bisceglie AM, Bruix J, et al. Design and endpoints of clinical trials in hepatocellular carcinoma. *J Natl Cancer Inst* 2008;100(10):698–711.
19. Forner A, Ayuso C, Varela M, et al. Evaluation of tumor response after locoregional therapies in hepatocellular carcinoma: are response evaluation criteria in solid tumors reliable? *Cancer* 2009;115(3):616–623.
20. Bruix J, Sherman M, Llovet JM, et al. Clinical management of hepatocellular carcinoma: conclusions of the Barcelona 2000 EASL conference. European Association for the Study of the Liver. *J Hepatol* 2001;35(3):421–430.
21. Bruix J, Sherman M; Practice Guidelines Committee, American Association for the Study of Liver Diseases. Management of hepatocellular carcinoma. *Hepatology* 2005;42(5):1208–1236.
22. Bos R, van Der Hoeven JJ, van Der Wall E, et al. Biologic correlates of (18)fluorodeoxyglucose uptake in human breast cancer measured by positron emission tomography. *J Clin Oncol* 2002;20(2):379–387.
23. Brücher BL, Weber W, Bauer M, et al. Neoadjuvant therapy of esophageal squamous cell carcinoma: response evaluation by positron emission tomography. *Ann Surg* 2001;233(3):300–309.
24. Wahl RL, Jacene H, Kasamon Y, Lodge MA. From RECIST to PERCIST: evolving considerations for PET response criteria in solid tumors. *J Nucl Med* 2009;50(suppl 1):122S–150S.
25. Jochelson M, Mauch P, Balikian J, Rosenthal D, Canellos G. The significance of the residual mediastinal mass in treated Hodgkin's disease. *J Clin Oncol* 1985;3(5):637–640.
26. Cheson BD, Pfistner B, Juweid ME, et al. Revised response criteria for malignant lymphoma. *J Clin Oncol* 2007;25(5):579–586.
27. 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(23):7412–7420.
28. Hoos A, Parmiani G, Hege K, et al. A clinical development paradigm for cancer vaccines and related biologics. *J Immunother* 2007;30(1):1–15.

Response Criteria in Oncologic Imaging: Review of Traditional and New Criteria

Temel Tirkes, MD • Margaret A. Hollar, DO • Mark Tann, MD • Marc D. Kohli, MD • Fatih Akisik, MD • Kumaresan Sandrasegaran, MD

RadioGraphics 2013; 33:1323–1341 • Published online 10.1148/rg.335125214 • Content Codes: CT GN NM OI

Page 1324

The ability to marry imaging findings with new clinical end points has become important in cancer therapy trials conducted to assess a new generation of targeted molecules for cancer treatment. This paves the way for much more rapid drug evaluation and, potentially, clinical decision making.

Page 1324

Over the years, the WHO and RECIST criteria have been modified by combining changes in size and the morphologic and metabolic features of specific tumors to overcome the limitations of the traditional criteria.

Page 1331

Although the RECIST criteria have been used extensively since their introduction, concerns about using change in tumor size as the only criterion have not been fully addressed, even in RECIST 1.1.

Page 1332

The Choi response criteria for GIST proposed that tumor attenuation could provide an additional measure of response to imatinib therapy.

Page 1336

Because many newer cancer therapies may be more cytostatic than cytotoxic, good tumor response may be associated predominantly with a decrease in metabolism, without a major reduction in tumor size.