Are We Ready for the 10% Solution?

HELEN X. CHEN,^a LARRY V. RUBINSTEIN,^b LALITHA K. SHANKAR,^c JEFFREY S. ABRAMS^a

^aCancer Therapy Evaluation Program, ^bBiometric Research Branch, and ^cCancer Imaging Program, Division of Cancer Treatment and Diagnosis, National Cancer Institute, Bethesda, Maryland, USA

Disclosures of potential conflicts of interest may be found at the end of this article.

The Response Evaluation Criteria in Solid Tumors (RECIST) criteria, which categorize quantitative tumor size changes into complete response, partial response (PR), stable disease (SD), or progressive disease (PD) [1], provide standardized, objective measurements of tumor response to therapy and have been used extensively over the last decade for early to late drug development and regulatory approvals, as well as for treatment decisions in individual patient care. The RECIST criteria were generated based on data from cytotoxic chemotherapy trials; their limitations for the assessment of molecularly targeted agents have been increasingly recognized with the advance of new therapies. In particular, the cutoff of 30% change in the sum of longest diameters (Δ SLD) as the criterion of response has been criticized for not adequately capturing potentially effective therapies. As exemplified by antiangiogenic therapies, such as sorafenib in renal cell carcinoma (RCC) and hepatocellular carcinoma, drugs with very low rates of RECIST-defined responses (10% or less) can still succeed and confer significant clinical benefit. For that reason, many clinical trials also use SD as an additional indicator of therapeutic effect. However, inclusion of SD and progression-free survival (PFS) in the efficacy readout often requires randomized trials to distinguish the drug effect from the natural course of the tumor. For patients receiving therapy, the implication of SD is often uncertain, because the criterion encompasses a wide range of tumor size changes, from 29% reduction to 19% increase. Exploration and validation of optimal criteria of tumor burden changes or functional imaging parameters as markers of drug effect and/or clinical benefit have the promise to improve the efficiency of both development of new therapeutics and therapeutic management of individual patients.

A study published in this issue of *The Oncologist* [2] represents one of the many retrospective analyses of the relation between tumor size changes and clinical outcomes in patients treated with vascular endothelial growth factor (VEGF) pathway-targeting agents, focusing on a specific patient care question: what degree of tumor size change early in the course of therapy may predict the clinical outcome in the patient and therefore provide guidance for decisions on further treatment? The analysis was based on 66 patients treated with 1 of the 6 different VEGF-pathway inhibitors, and thresholds of -30% (as in RECIST) or -10% Δ SLDs were tested for their ability to classify patients with good or poor outcomes. This analysis concluded that $\geq 10\%$ reduction in SLD (responders) at the first scan was associated with

significantly better outcomes, compared with that of non-responders (those who did not achieve 10% SLD reduction). Time to treatment failure was 8.4 months versus 4.1 months, and overall survival was 35 months versus 15 months, both with p values $< .01$. In contrast, the RECIST threshold of -30% Δ SLDs at the first scan failed to predict patient outcome (TTF of 6.9 months versus 5.5 months). It further suggested that -10% Δ SLDs at first scan could be used for treatment decisions as to whether the anti-VEGF therapies should be continued, although the negative predictive value of “nonresponse,” by either the -10% or -30% cutoffs, was not discussed.

As already recognized by the authors, there are multiple limitations of this series, including limited sample size and small numbers in each marker subgroup, heterogeneity of the VEGF-targeting therapies ranging from tyrosine-kinase inhibitors (TKIs) to monoclonal antibodies, and inconsistency in the timing of first scans (20–170 days from start of therapy) on which the cutoff optimization was based. Despite the limitations, this study joined a large body of independent retrospective studies that collectively demonstrated a significant correlation between tumor size changes at the first scan and the clinical outcome [3–5]. One of the largest series was reported by Thiam et al. [5] based on 334 patients with advanced RCC treated on the sunitinib arm in a phase III trial for sunitinib versus interferon- α . It tested a series of Δ SLD thresholds at the first scan at 6 weeks (-45% , -30% , -20% , -10% , 0% , $+10\%$) for their correlations with PFS and found that Δ SLDs of -10% provided the optimal cutoff that distinguished the PFS outcomes (median PFS of 5.6 months versus 11 months). The -10% cutoff was also examined in a number of other retrospective studies in independent patient cohorts and was consistently found to be significantly associated with outcomes ([5–7] and this study).

What is the clinical utility of this finding? Although $\geq 10\%$ shrinkage is clearly associated with significantly better outcome, the practical concern for a given patient is the likelihood of benefiting from therapy if that threshold is not reached. As an inherent limitation of post-treatment biomarkers, it is not possible to distinguish between predictive and prognostic markers because all marker subgroups would have received the same therapy. PFS of 5 months was often selected as an arbitrary landmark to estimate presence or lack of therapeutic benefit from VEGF-pathway inhibitors in first-line metastatic RCC. In several studies, including Thiam et al. [5] and the current study, median PFS values in nonresponders by the

Correspondence: Jeffrey S. Abrams, M.D., National Cancer Institute, Cancer Therapy Evaluation Program, 9609 Medical Center Drive, Rockville, Maryland 20850, USA. Telephone: 240-276-6515; E-Mail: jeff.abrams@nih.gov Received March 27, 2014; accepted for publication April 8, 2014; first published online in *The Oncologist Express* on April 22, 2014. ©AlphaMed Press 1083-7159/2014/\$20.00/0 <http://dx.doi.org/10.1634/theoncologist.2014-0126>

–10% criterion were 5–6 months, indicating that although these patients did not do well on average, 50% of the patients were progression-free for longer than 5–6 months. Furthermore, in tumors in which the target pathways remain relevant through progression, continuation of therapy may be beneficial whichever PD criteria are used. Given that the treatment benefit in an individual patient cannot be ruled out with high certainty, the value of this early post-treatment marker would depend on the availability of better, less toxic treatment options.

Is this new criterion of response ($\geq 10\%$ SLD reduction) better than RECIST in guiding treatment decisions? The answer depends on how RECIST is used. It is clear that objective response ($\geq 30\%$ SLD reduction) is not a good predictor of outcomes. However, treatment decisions in general practice are not based on response, but rather on progression. Patients not achieving PR, but in SD, by RECIST definition (Δ SLDs in the range of -29% to $+19\%$) would continue therapy until PD. To demonstrate the potential advantage of the -10% cutoff over the current practice, outcome differences using the cutoffs of RECIST PR/SD versus -10% should be compared. However, such comparisons are not always feasible because the numbers of patients with PD on first scans are usually very small, at least for anti-VEGF therapies in RCC. On the other hand, the use of the SD category as a basis for continuing therapy is conceivably problematic, because the tumor size could have increased by up to 19%. Indeed, lack of tumor shrinkage (0% SLD reduction) or $+10\%$ SLD at the first scan was associated with an extremely poor outcome in patients treated with sunitinib, with a median PFS of 1.5 months [5]. Further studies with sufficient power would be required to confirm this finding.

What is the general implication of the findings on the use of tumor imaging in drug development? This study, among many others, reinforced the notion that clinically viable targeted agents may have low response rates as defined by RECIST, but tumor shrinkage with smaller magnitudes is common. These studies further suggest that modification of the response criteria to capture and categorize minor responses may provide a more sensitive readout of the therapeutic effect

and potentially improve the accuracy of early drug evaluation. On the other hand, finding the optimal cutoff can be challenging, because the threshold may differ for the same agents in different indications or for different agents in the same indication. For example, although studies on VEGFR TKIs in metastatic RCC identified 10% as the optimal cutoff, similar tests for everolimus in second-line metastatic RCC failed to identify any cutoffs (from -45% to $+20\%$) that separate better and poorer outcomes [8]. Technical limitations, such as measurement variability, may also limit the choice of thresholds that can be adopted. Efforts are under way by the RECIST committee to expand the database to include large randomized clinical trials of targeted agents, as well as studies using molecular imaging with fluorodeoxyglucose-positron emission tomography (FDG-PET), with the plan to evaluate and update the guidelines and criteria to meet the needs of drug development and patient care in the era of novel therapies.

In summary, the current study and others focusing on VEGF-pathway inhibitors in advanced RCC highlight the observation that tumor shrinkage below the threshold of RECIST can be associated with significant clinical benefit, and that flexibility in the size change categorization should be considered in future modification of the response criteria. However, whether -10% or other cutoffs would be optimal for treatment decisions in individual patients is uncertain and would require further, sufficiently powered, prospective studies. It should be noted that criteria for one agent or one disease setting may not be generalizable to others. As newer therapeutic agents become available in various clinical settings, it may be necessary to further modify existing tumor size-based response criteria, verify promising molecular and functional imaging methods, and investigate new technologies. Finally, before optimal criteria are defined for dichotomous characterization of tumor responses, it would be desirable to collect tumor size measurements as a continuous variable and evaluate the reporting of antitumor activity based on several cutoffs, rather than just one, for responses or progression.

DISCLOSURES

The authors indicated no financial relationships.

REFERENCES

1. 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; 92:205–216.
2. Krajewski KM, Franchetti Y, Nishino M et al. 10% tumor diameter shrinkage on the first follow-up computed tomography predicts clinical outcome in patients with advanced renal cell carcinoma treated with angiogenesis inhibitors: A follow-up validation study. *The Oncologist* 2014;19:507–514.
3. van der Veldt AA, Haanen JB, van den Eertwegh AJ et al. Targeted therapy for renal cell cancer: Current perspectives. *Discov Med* 2010;10: 394–405.
4. Krajewski KM, Guo M, Van den Abbeele AD et al. Comparison of four early posttherapy imaging changes (EPTIC; RECIST 1.0, tumor shrinkage, computed tomography tumor density, Choi criteria) in assessing outcome to vascular endothelial growth factor-targeted therapy in patients with advanced renal cell carcinoma. *Eur Urol* 2011; 59:856–862.
5. Thiam R, Fournier LS, Trinquart L et al. Optimizing the size variation threshold for the CT evaluation of response in metastatic renal cell carcinoma treated with sunitinib. *Ann Oncol* 2010; 21:936–941.
6. Abel EJ, Culp SH, Tannir NM et al. Early primary tumor size reduction is an independent predictor of improved overall survival in metastatic renal cell carcinoma patients treated with sunitinib. *Eur Urol* 2011;60:1273–1279.
7. Smith AD, Lieber ML, Shah SN. Assessing tumor response and detecting recurrence in metastatic renal cell carcinoma on targeted therapy: Importance of size and attenuation on contrast-enhanced CT. *AJR Am J Roentgenol* 2010;194:157–165.
8. Oudard S, Thiam R, Fournier LS et al. Optimisation of the tumour response threshold in patients treated with everolimus for metastatic renal cell carcinoma: Analysis of response and progression-free survival in the RECORD-1 study. *Eur J Cancer* 2012;48:1512–1518.

EDITOR'S NOTE: See the related article, “10% Tumor Diameter Shrinkage on the First Follow-Up Computed Tomography Predicts Clinical Outcome in Patients With Advanced Renal Cell Carcinoma Treated With Angiogenesis Inhibitors: A Follow-Up Validation Study,” by Katherine M. Krajewski et al., on page 507 of this issue.