

Volumes to Learn: Advancing Therapeutics with Innovative Computed Tomography Image Data Analysis

Michael L. Maitland

Semi-automated methods for calculating tumor volumes from computed tomography images are a new tool for advancing the development of cancer therapeutics. Volumetric measurements, relying on already widely available standard clinical imaging techniques, could shorten the observation intervals needed to identify cohorts of patients sensitive or resistant to treatment. *Clin Cancer Res*; 16(18); 4493–5. ©2010 AACR.

In this issue of *Clinical Cancer Research*, Zhao and colleagues demonstrate that calculating tumor volumes from computed tomography (CT) images could help developers of new cancer therapeutics to learn early which patients have disease that responds to the treatment (1). This report culminates 3 years of publications that suggest that effective cancer therapeutics development could be accelerated if we were to use the CT-imaging data routinely collected in clinical care more efficiently and intelligently.

CT-imaging measurements are the basis for the assessment of new anticancer agents using the Response Evaluation Criteria in Solid Tumors (RECIST) in clinical trials (2). To maximize the interrater reliability of treatment response evaluations, RECIST categorizes the patient as having: complete response, partial response, stable disease, or disease progression. A standardized categorization of disease assessment was important for advancing therapeutics decades ago, when disease response evaluations depended on physical exam and plain film radiography, but in the current era of digital imaging, it has lost its utility.

Few RECIST users are aware of the original study on which some reference cut-points in the categorization are based (3). Moertel and Hanley asked each enrolled oncologist to use "the usual technique and equipment (ruler or caliper) he employed in clinical practice" to measure the sizes of "twelve solid spheres [ranging] from 1.8 to 14.5 cm in diameter ... These [simulated] masses were then arranged ... on a soft mattress and covered with a layer of foam rubber [that] measured 0.5 inches in thickness for the six smaller masses to approximate skin and subcutaneous tissue and 1.5 inches for the six larger masses to approximate

abdominal wall." At the time, this was a thoughtful assessment of the clinical measurement "noise" that interfered with investigators' detection in single arm studies of the signal of disease response to an active treatment. The authors suggested that ineffective agents would have a placebo-like objective response rate of 5 to 10%. They also recommended that the threshold for declaring partial response should be a 50% decrease in bidimensional measurements. The mathematical counterpart of this 50% decline criterion for unidimensional measurements is the 30% employed by RECIST in the current system.

The latest RECIST version acknowledges its limitations but contends that no validated improvement is available for routine use. Reliance on categorical systems has stifled implementation of innovative approaches to exploit digital imaging and better measurement of solid tumors. Shortly after Moertel and Hanley's study of spheres under foam rubber, Lavin proposed quantitative assessment of change in tumor size between randomly assigned groups of patients to assess potential treatments more efficiently than nonrandomized studies using categorical assessments (4). Modeling Lavin's study design for a trial of erlotinib and sorafenib in the second-line treatment of non-small cell lung cancer (NSCLC; ref. 5), Karrison and colleagues suggested the value of the combination could be determined with greater certainty than a single-arm trial without requiring substantially more patients. The design's chief advantage was that the assessment required each patient to remain on her or his assigned treatment arm for only 8 weeks. Patient accrual would be quicker and the total time to completion shorter than trials using alternative end-points such as progression-free survival.

The change in tumor size at an 8-week endpoint for NSCLC treatment was validated in 3,398 NSCLC patients enrolled in four phase III trials (6). The investigators determined the range of tumor growth patterns over time. They then developed a predictive model for each patient and the relationship of this pattern to overall survival. The best predictors for survival in this quantitative model were Eastern Cooperative Oncology Group (ECOG) performance status, baseline tumor size, and percentage

Author's Affiliation: The University of Chicago Medical Center, Chicago, Illinois

Corresponding Author: Michael L. Maitland, The University of Chicago Medical Center, Department of Medicine, Section of Hematology/Oncology, Committee on Clinical Pharmacology and Pharmacogenomics, MC 2115, 5841 South Maryland Avenue, Chicago, IL 60637. Phone: 773-834-8981; Fax: 773-834-0188; E-mail: mmaitlan@medicine.bsd.uchicago.edu.

doi: 10.1158/1078-0432.CCR-10-1548

©2010 American Association for Cancer Research.

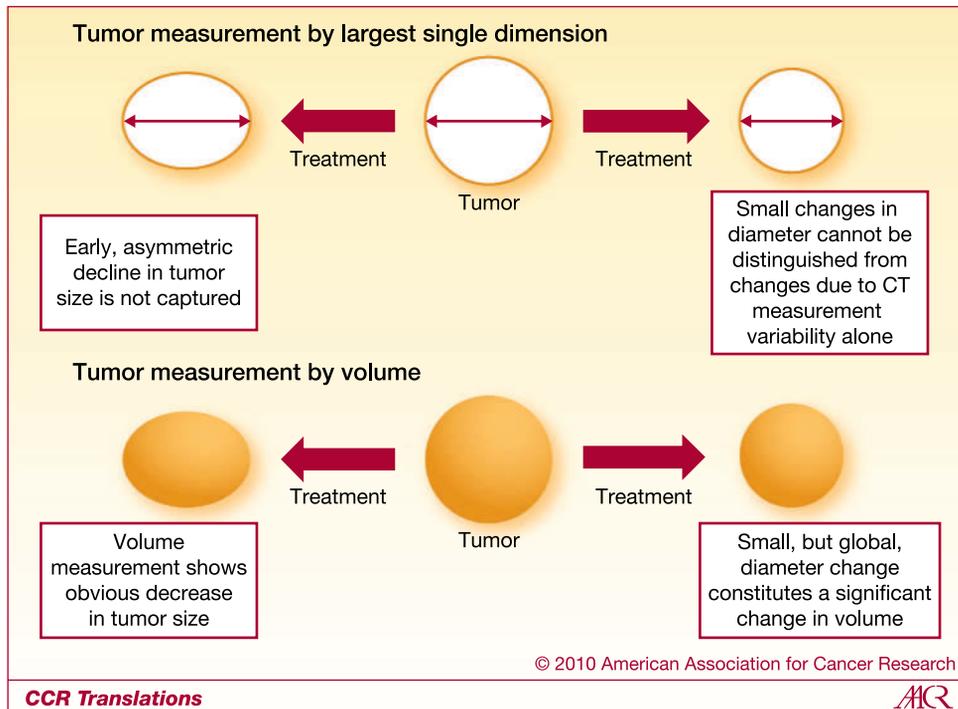


Fig. 1. Tumor volume change as an early response metric. Early response assessment with single-largest dimension measurements is limited by smaller magnitude changes than those seen with conventionally timed response assessment and the variance or "error" in single dimension measurements (top). Volume measurement may allow detection of antitumor effect when it is not apparent with conventional single-largest dimension measurement (bottom).

tumor-size reduction from the baseline measurement to 8 weeks posttreatment. Similarly, in a study of advanced colorectal cancer patients, baseline tumor size and the change in that measurement 7 weeks posttreatment were significant predictors of overall survival (7).

Routine measurement variance makes the 7- to 8-week tumor-size change the earliest for predicting treatment effects on overall survival. Patient position in the scanner and physiologic fluctuations in tumor shape and size over short intervals confound reproducible measurement of treatment effects. As the modern equivalent of the spheres-under-foam study, radiology scientists have carried out "coffee break" studies (8). Patients undergo a CT scan, get off the imager table, and then undergo a second scan. Radiologists measure lesions in both image sets, and the variance is determined. Prior to the current study, coffee break investigations suggested that the ratio of the variance in unidimensional or bidimensional to total measurements was greater than for volume (9). This observation implied that a major advantage of volume measurements would be the routine capacity to detect treatment effects earlier in the course of therapy (Fig. 1). This method also predicted that volume could more readily detect treatment effects in indolent neoplasms in which the changes in tumor size over time are subtle and require prolonged investigations to establish therapeutic benefit. Two independent studies established the therapeutic benefit of sirolimus in indolent angiomyolipomata and tuberous sclerosis with volume measurements (10, 11).

In a clinical trial of 3-weeks gefitinib therapy before curative lobectomy, NSCLC patients had tumor epidermal

growth factor receptor genotyping and pre- and post-treatment CT imaging. Zhao and colleagues used thin slice (1.25-mm sections) CT exams, and, in addition to determining the single longest dimension of each tumor, they calculated tumor volumes with a semi-automated software application. Using genotype as the gold standard to categorize gefitinib sensitivity and resistance, the investigators demonstrated that changes in tumor volume at 3 weeks had better sensitivity and specificity for identifying the tumor genotype than unidimensional measurements. They concluded that early changes in tumor volume measurements could identify patients who will and will not benefit from a novel therapeutic, and that these data could be used in efforts to discover or validate predictive tumor biomarkers.

As a pilot study, applying these results to therapeutics development has limitations. The optimized cut-point for tumor volume changes was specific to this one cohort and had 11% false-positive and 10% false-negative assignments, unacceptable for determining an individual patient's decision to remain on or withdraw from further treatment. As these were early stage NSCLC patients, it could not be determined if this error rate was due to imperfections in tumor genotyping as the "gold standard" or if the assignment accuracy could be improved by observing tumors for a longer timeframe. Although prior studies suggested bidimensional measurements would not be significantly better than unidimensional, this was not addressed in this study. Finally, the sharp contrast in pixel intensity between lung tumors and the adjacent, air-filled parenchyma leads to easier discernment of tumor

boundaries and volume calculations than for tumors in other visceral sites.

The biggest limitation in advancing this research to benefit patients is not technical but operational. After plain films, CT is the most commonly used diagnostic imaging test for patients with solid tumors (12). The popularity of these scans is due to their accessibility, speed, and ease of performance. Although imaging methods that employ new radiotracers or spectroscopic methods offer more detailed functional information, these studies are done orders of magnitude less frequently than CT scans. Similar to the relationship of tumor radius to volume, modest investment

in the infrastructure to share, develop, and implement quantitative data from CT imaging should yield exponential gains in our capacity to develop and deliver personalized cancer therapeutics.

Disclosure of Potential Conflicts of Interest

M. Maitland, consultant, Abbott Laboratories; data-sharing agreements with various companies to develop quantitative models of disease progression in solid tumors.

Received 06/30/2010; accepted 07/01/2010; published OnlineFirst 07/19/2010.

References

1. Zhao B, Oxnard GR, Moskowitz CS, et al. A pilot study of volume measurement as a method of tumor response evaluation to aid biomarker development. *Clin Cancer Res* 2010;16:4647–53.
2. 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.
3. Moertel CG, Hanley JA. The effect of measuring error on the results of therapeutic trials in advanced cancer. *Cancer* 1976;38:388–94.
4. Lavin PT. An alternative model for the evaluation of antitumor activity. *Cancer Clin Trials* 1981;4:451–7.
5. Karrison TG, Maitland ML, Stadler WM, Ratain MJ. Design of phase II cancer trials using a continuous endpoint of change in tumor size: application to a study of sorafenib and erlotinib in non small-cell lung cancer. *J Natl Cancer Inst* 2007;99:1455–61.
6. Wang Y, Sung C, Dartois C, et al. Elucidation of relationship between tumor size and survival in non-small-cell lung cancer patients can aid early decision making in clinical drug development. *Clin Pharmacol Ther* 2009;86:167–74.
7. Claret L, Girard P, Hoff PM, et al. Model-based prediction of phase III overall survival in colorectal cancer on the basis of phase II tumor dynamics. *J Clin Oncol* 2009;27:4103–8.
8. Zhao B, Schwartz LH, Moskowitz CS, Ginsberg MS, Rizvi NA, Kris MG. Lung cancer: computerized quantification of tumor response—initial results. *Radiology* 2006;241:892–8.
9. Mozley PD, Schwartz LH, Bendtsen C, Zhao B, Petrick N, Buckler AJ. Change in lung tumor volume as a biomarker of treatment response: a critical review of the evidence. *Ann Oncol* 2010, Epub 2010 Mar 23.
10. Bissler JJ, McCormack FX, Young LR, et al. Sirolimus for angiomyolipoma in tuberous sclerosis complex or lymphangioleiomyomatosis. *N Engl J Med* 2008;358:140–51.
11. Davies DM, Johnson SR, Tattersfield AE, et al. Sirolimus therapy in tuberous sclerosis or sporadic lymphangioleiomyomatosis. *N Engl J Med* 2008;358:200–3.
12. Dinan MA, Curtis LH, Hammill BG, et al. Changes in the use and costs of diagnostic imaging among Medicare beneficiaries with cancer, 1999–2006. *JAMA* 2010;303:1625–31.