

## Variability of Lung Tumor Measurements on Repeat Computed Tomography Scans Taken Within 15 Minutes

Geoffrey R. Oxnard, Binsheng Zhao, Camelia S. Sima, Michelle S. Ginsberg, Leonard P. James, Robert A. Lefkowitz, Pingzhen Guo, Mark G. Kris, Lawrence H. Schwartz, and Gregory J. Riely

See accompanying editorial on page 3109

All authors: Memorial Sloan-Kettering Cancer Center, Weill Cornell Medical College, New York, NY.

Submitted December 6, 2010; accepted April 29, 2011; published online ahead of print at [www.jco.org](http://www.jco.org) on July 5, 2011.

Supported in part by Grant No. R01-CA125143 (L.H.S.) from the National Cancer Institute, Bethesda, MD.

Authors' disclosures of potential conflicts of interest and author contributions are found at the end of this article.

Corresponding author: Gregory J. Riely, MD, PhD, Memorial Sloan-Kettering Cancer Center, 1275 York Ave, New York, NY 10065; e-mail: [rielyg@mskcc.org](mailto:rielyg@mskcc.org).

© 2011 by American Society of Clinical Oncology

0732-183X/11/2923-3114/\$20.00

DOI: 10.1200/JCO.2010.33.7071

### A B S T R A C T

#### Purpose

We use changes in tumor measurements to assess response and progression, both in routine care and as the primary objective of clinical trials. However, the variability of computed tomography (CT)–based tumor measurement has not been comprehensively evaluated. In this study, we assess the variability of lung tumor measurement using repeat CT scans performed within 15 minutes of each other and discuss the implications of this variability in a clinical context.

#### Patients and Methods

Patients with non–small-cell lung cancer and a target lung lesion  $\geq 1$  cm consented to undergo two CT scans within a period of minutes. Three experienced radiologists measured the diameter of the target lesion on the two scans in a side-by-side fashion, and differences were compared.

#### Results

Fifty-seven percent of changes exceeded 1 mm in magnitude, and 33% of changes exceeded 2 mm. Median increase and decrease in tumor measurements were +4.3% and –4.2%, respectively, and ranged from 23% shrinkage to 31% growth. Measurement changes were within  $\pm 10\%$  for 84% of measurements, whereas 3% met criteria for progression according to Response Evaluation Criteria in Solid Tumors (RECIST;  $\geq 20\%$  increase). Smaller lesions had greater variability of percent measurement change ( $P = .005$ ).

#### Conclusion

Apparent changes in tumor diameter exceeding 1 to 2 mm are common on immediate reimaging. Increases and decreases less than 10% can be a result of the inherent variability of reimaging. Caution should be exercised in interpreting the significance of small changes in lesion size in the care of individual patients and in the interpretation of clinical trial results.

*J Clin Oncol* 29:3114-3119. © 2011 by American Society of Clinical Oncology

### INTRODUCTION

Tumor imaging plays a fundamental role in oncology care and clinical trials, where computed tomography (CT)–based tumor measurement is a primary mechanism for determining response to therapy and time to treatment failure. Improvements in imaging technology mean that detailed tumor measurements are increasingly available for use in clinical decisions. Tumor measurement is commonly performed using a computer interface, which allows radiologists to provide precise measurements. In clinical research, there is increasing interest in specific measurement changes as it has become clear that broad response categories such as Response Evaluation Criteria in Solid Tumors (RECIST) sometimes fail to capture the activity of novel therapies.<sup>1</sup> Phase II trials in particular increas-

ingly present waterfall plots showing individual measurement changes for each patient and, thus, treat response as a continuous rather than categorical variable.<sup>2</sup> Given this increased interest in quantitative tumor measurement, it becomes important to understand what measurement changes are meaningful rather than a result of variability of imaging and measurement.

Multiple factors can contribute to the variability of CT-based tumor measurement. A component of this variability is operator dependant, such as the process of selecting which CT slice to measure and the placement of a linear measurement using a computer interface. Several studies have tried to quantify this variability by having radiologists perform repeat measurements on a single set of CT scans.<sup>3-6</sup> Separately, the process of performing a CT scan can lead to changes in the appearance of a tumor or

surrounding stroma. Lastly, it has been found that the technique used for processing CT imaging data can contribute to measurement variability.<sup>7-9</sup> However, prior studies assessing measurement variability have studied only one step of the imaging process; although these studies may be useful for investigations into reducing measurement variability, they do not have direct applicability to the clinical environment.

In this study, we attempt to quantify the variability of the complete CT measurement process. By imaging patients with lung tumors twice within several minutes and then measuring both images in a similar fashion, we gain an opportunity to quantify the variability inherent in both CT image capture and in the subsequent measurement technique. Prior rescan studies have had limited applicability to clinical oncology because these have looked at small lung nodules of unclear malignant potential.<sup>9-11</sup> Additionally, we previously published a rescan analysis of the clinical study reported here that compared manual measurement to semi-automated measurement; however, radiologist measurements were made independently without allowing comparison of baseline and follow-up images.<sup>12</sup> In the present study, by replicating the clinically standard side-by-side measurement process on two separately obtained scans, we provide the most optimal framework for the interpretation of clinical and research measurement results. We set out to obtain data that would allow clinicians to answer two basic types of imaging questions, demonstrated in the following examples:

- **Clinical:** A patient receiving chemotherapy undergoes CT imaging, and the report reads, “The dominant lung nodule has increased in size, measuring 1.4 cm (previously 1.2 cm)...Impression: Disease progression.” Can you tell the patient with certainty that the tumor has grown?
- **Research:** A journal article describing phase II trial results of a novel treatment reads, “The overall RECIST response rate was 20%. As shown in the waterfall plot, 50% of patients had some tumor shrinkage.” How do you interpret these results?

## PATIENTS AND METHODS

### Patients

Patients with non-small-cell lung cancer (NSCLC) receiving systemic therapy consented to this Institutional Review Board–approved rescan study (ClinicalTrials.gov identifier NCT00579852). Patients were accrued in clinic by their treating oncologist. Patients were eligible for participation if a non-contrast CT scan of the chest was clinically indicated and their most recent CT scan report described a lung lesion  $\geq 1$  cm in diameter; there was no radiologist review before enrollment. Sample size calculation in the study protocol was based on the ability to detect a concordance correlation coefficient between the baseline and follow-up measurements of at least 0.75.

### Imaging

Patients received their initial noncontrast chest CT scan per clinical routine with either a 16-detector or 64-detector scanner during a breath hold. On completion, the patient was instructed to leave the scanner briefly before returning for a second scan obtained in the identical fashion on the same scanner. For each scan, the craniocaudal extent of the scan was separately determined with a scout image of the patient. Parameters for the 16- and 64-detector scanners were as follows: tube voltage, 120 kV (peak) and 120 kV (peak); tube current, 299 to 441 mA and 298 to 351 mA; detector configuration, 16 detectors  $\times$  1.25 mm section gap and 64 detectors  $\times$  0.63 mm section gap; and pitch, 1.375:1 and 0.984:1. The thoracic images were obtained without intravenous contrast material. Images were reconstructed with 1.25-mm

nonoverlapping slice intervals and a sharper convolution kernel and stored in Digital Imaging and Communications in Medicine (DICOM) format. Patient identifiers were removed from the DICOM headers of all CT images analyzed for this study. Through collaboration with the National Cancer Institute, deidentified images have been placed in the public domain as a resource for further investigations and can be accessed at the National Biomedical Imaging Archive.

CT images were viewed at a computer workstation by three experienced radiologists (M.S.G., P.G., and R.A.L.). Radiologists were first asked to perform a measurement of maximum diameter for a selected target lesion on the first scan of each of the 33 patients. Measurements were performed manually using a computer interface. After making this baseline measurement, each radiologist viewed the second scan side-by-side with the first and was asked to measure this follow-up scan in a similar fashion. Radiologists were blinded from knowing how much time had passed between the two scans. The six measurements performed on each patient (by three radiologists on each of two different scans) were averaged to calculate the approximate size of each lesion.

### Biostatistics

The change in size between the two measurements of each tumor by each radiologist was calculated. Because it is unlikely that any real change in tumor burden occurred in the minutes between the two scans, the measured change functions as a gauge of the random variations that can be expected simply by reimaging. Because the two scans were performed at approximately the same time, one of them was randomly selected as the baseline, and the other was designated as the follow-up scan. The change in millimeters and the relative change in measurement (as a percentage) were calculated. The former is the measure of change directly observable, whereas the latter is the metric used commonly in clinical trials.<sup>13</sup> This random assignment step was repeated 1,000 times, resulting in 1,000 distributions of the change (in millimeters) and of the relative change. We report the mean and standard deviation of measurement change by averaging these statistics over the 1,000 distributions.

To assess what range of measurements could be expected as a result of variability, the 95% limits of agreement for change in millimeters were calculated as the mean change  $\pm 2$  standard deviations. The 95% limits of agreement can inform clinical practice because changes that fall within these limits can be considered as potentially arising as a result of measurement variability, rather than true change in tumor size. The relationship between this measurement error and lesion size was examined by fitting two separate generalized linear models with a normal probability distribution, which accounted for the intraclass correlation resulting from the fact that each scan is measured by three different radiologists. In each model, the positive measurement change (in millimeters and percent change, respectively) was modeled as a linear function of the lesion size.

## RESULTS

Between January and September of 2007, 33 patients with NSCLC consented to participation in the study. Two patients were excluded from analysis because, after undergoing measurement per protocol, the mean target lesion size was determined to be less than 1 cm. Another patient did not follow study protocol and was excluded because more than 1 day elapsed between CT scans. The characteristics of the 30 remaining patients and their 30 target lesions are listed in Table 1. The mean lesion size was 3.7 cm (range, 1.0 to 8.0 cm). The median time interval between the two scans was 8 minutes (range, 5 to 14 minutes). Twenty-seven patients (90%) were imaged with a 16-detector scanner, and three patients (10%) were imaged with a 64-detector scanner. Repeat measurements were made of each of the 30 lesions by three radiologists, totaling 90 paired measurements.

### Absolute Measurement Change in Millimeters

The distribution of the 90 measurement changes is shown in Figure 1A. The standard deviation of measurement change was 2.4

Demographic or Clinical Characteristic	No. of Patients (N = 30)	%
Median age, years	64	
<b>Histology</b>		
Adenocarcinoma	26	87
Squamous	2	7
NSCLC, NOS	2	7
<b>Stage</b>		
III	2	7
IV	27	90
Recurrent	1	3
<b>Lesion diameter, cm</b>		
1-2	6	20
2-3	7	23
3-5	11	37
5-7	4	13
> 7	2	7

Abbreviations: NOS, not otherwise specified; NSCLC, non-small-cell lung cancer.

mm, indicating that 95% of measurement changes fell between  $-4.8$  mm and  $+4.8$  mm. Fifty-seven percent of measurement changes had a magnitude greater than 1 mm, and 33% of changes had a magnitude greater than 2 mm, with half of these appearing as positive changes and half appearing as negative changes (Table 2).

### Relative Measurement Change

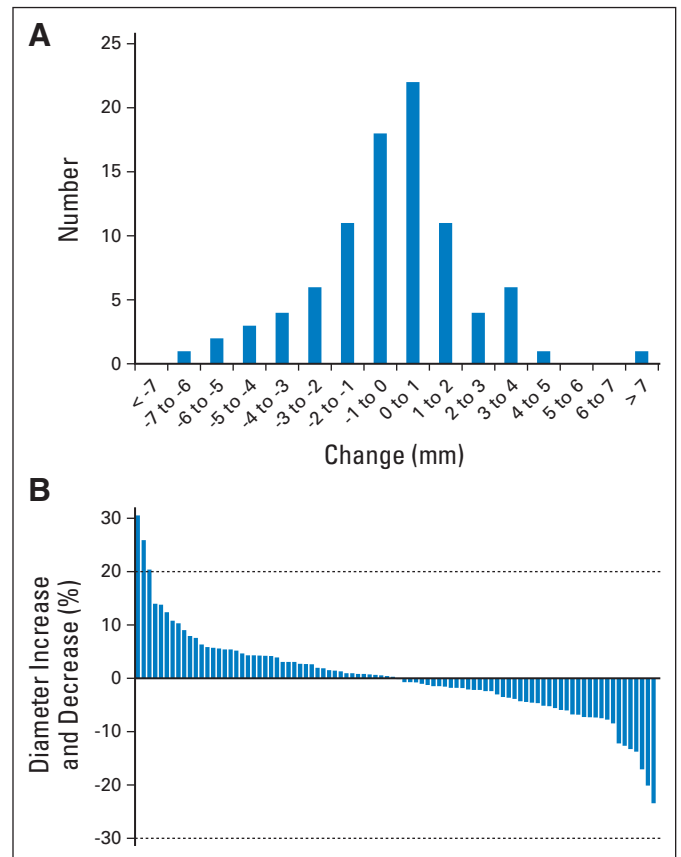
A waterfall plot of the 90 relative changes in tumor measurement is shown in Figure 1B, ranging from 23% shrinkage to 31% growth. The median increase was  $+4.3\%$ , and the median decrease was  $-4.2\%$ . Three percent of changes met the RECIST threshold for progressive disease (20% increase), whereas none met the RECIST threshold for partial response (30% decrease). Eighty-four percent of the tumor measurement changes were between  $-10\%$  and  $+10\%$ .

### Variability and Lesion Size

Table 3 lists the standard deviation of measurement change calculated from tumors of different sizes. Larger tumors tended to have larger magnitude measurement changes in millimeters ( $P = .06$ ; Fig 2). In contrast, relative change (percent increase or percent decrease) was found to be significantly larger for smaller tumors ( $P = .005$ ; Fig 2). For tumors smaller than 3 cm, 6% of the changes met RECIST criteria for disease progression, as opposed to 1% of the changes for tumors larger than 3 cm. The range of potential measurement changes as a result of variability are listed in Table 3 for three tumors of different sizes, using 95% limits of agreement calculated from the standard deviation. Although standard deviation of absolute measurement increases somewhat with increased tumor size, variability of percent change decreases.

## DISCUSSION

This study represents the first time that the variability of CT-based tumor measurement has been fully quantified with the use of repeat CT imaging and conventional side-by-side measurement.



**Fig 1.** (A) Distribution of measurement changes found on repeat computed tomography scans performed within 15 minutes of each other, in millimeters; there was a greater than 1-mm magnitude of change in the majority of lesions (57%). (B) A plot of relative change of longest dimension of indicator lesion (waterfall plot) is also shown; 3% of changes met Response Evaluation Criteria in Solid Tumors for progression, but none met criteria for partial response (dashed lines).

The variability we found is directly applicable to measurements of lung tumors and is likely to be observed in tumor measurements in other organs. We believe our data have important relevance to the clinical care of patients with parenchymal metastases and to the interpretation of clinical trial results that rely heavily on radiographic response end points.

Our data demonstrate that CT measurement of lung lesions has variability of a clinically meaningful magnitude, with standard deviation of 2.4 mm of change between two CT scans of a tumor performed within a short period of time. This measurement variation is multifactorial in nature, partly because of differences in the appearance of the

Magnitude of Change (positive or negative, mm)	Frequency (%)
$\geq 1.0$	57
$\geq 2.0$	33
$\geq 3.0$	20
$\geq 4.0$	10
$\geq 5.0$	4

**Table 3.** Differences in Measurement Variability Depending on Lesion Size, As Calculated From Repeat CT Scans Performed Within 15 Minutes of Each Other

Size of Tumor (cm)	Standard Deviation (mm)	Example Tumor		
		Size (cm)	Range As a Result of Variability (cm)*	% Change As a Result of Variability
1-3	2.0	2	1.6-2.4	± 20
3-5	2.3	4	3.5-4.5	± 12
5-8	3.3	7	6.3-7.7	± 9

NOTE. For a lesion that in fact measures 4 cm, for example, the variability of CT imaging can lead to measurements ranging from 3.5 to 4.5 cm.

Abbreviation: CT, computed tomography.

\*As calculated from the 95% limits of agreement.

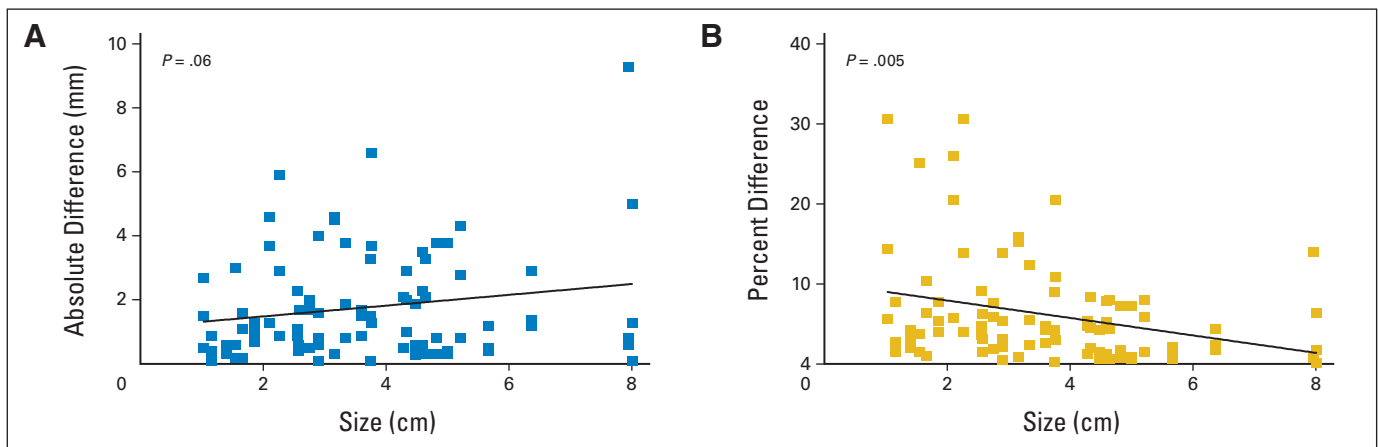
tumor or surrounding tissue (Appendix Fig A1, online only) and partly because of operator-dependent differences in the placement of a measurement. In total, the multistep process of repeat imaging and measurement of a tumor will lead to differences of  $\geq 4$  mm approximately 10% of the time (Table 2). Or put in other terms, for a lesion that in fact measures 4 cm, the variability of CT imaging can lead to measurements ranging from 3.5 to 4.5 cm (Table 3).

These data can assist clinical oncologists as they determine when a patient has developed disease progression. Although disease progression in clinical trials is defined by RECIST as an increase in summed tumor diameter of  $\geq 20\%$ ,<sup>13</sup> the criteria state that, “it is not intended that these RECIST guidelines play a role in [clinical] decision making, except if determined appropriate by the treating oncologist.” In clinical practice, any clear evidence of tumor growth could be judged a treatment failure, supporting the discontinuation of a therapy and possible change to another. Some oncologists may interpret a diameter increase of 1 or 2 mm as evidence of clear tumor growth; however, we found that reimaging led to measurement differences exceeding 2 mm 33% of the time, with half of these (17%) showing a more than 2 mm increase in tumor diameter. Therefore, we believe the variability inherent in CT imaging requires that clinicians consider other factors, such as changes in size of other lesions or patient toxicity, when using CT measurements to identify tumor progression.

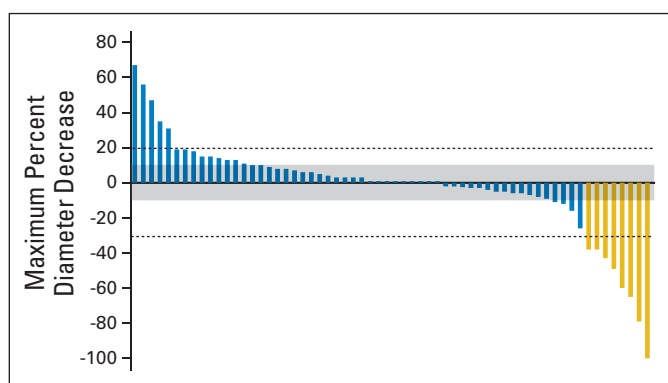
Our findings indicate that the inherent variability of conventional unidimensional CT measurement can at times lead to the appearance of RECIST progression ( $\geq 20\%$  diameter increase), con-

sidering that 3% of measurement changes calculated from the repeat CTs met this criterion. This was more common with lesions measuring between 1 and 3 cm; in 6% of these lesions, the measurement change from reimaging resulted in an appearance of a  $\geq 20\%$  increase in diameter. One strategy for avoiding cases of variability being misclassified as progression was adopted by RECIST 1.1, which now dictates that a  $\geq 20\%$  increase only qualifies as disease progression if there is “an absolute increase of at least 5 mm” in summed diameter measurements,<sup>13</sup> and our data support this concept of a minimal change requirement.

This work also may have important implications for the interpretation of tumor response, particularly when measurement change is considered as a continuous variable, a technique increasingly considered as a way of better expressing a therapy’s antitumor activity.<sup>1-2,14</sup> One analysis often used is the waterfall plot, which displays the magnitude of each patient’s best response as a percent measurement change (Fig 1B). To gauge its prevalence in the literature, we searched PubMed for phase I and II trials treating the major CT-measured carcinomas (lung, colorectal, pancreas, and renal) that were published in the *Journal of Clinical Oncology* in 2009. The search found 41 articles; nine articles focusing on radiotherapy or toxicity of therapy were excluded. Of the remaining 32 clinical trials, 15 (47%) included data addressing response as a continuous variable; 14 showed waterfall plots, and nine quantified the number of patients with a reduction in measurements. Although these statistics are used frequently, they can only be meaningfully interpreted (and meaningfully reported) if the



**Fig 2.** Correlation between measurement change and lesion size, in millimeters (left) and as a percentage (right). There is a trend toward greater measurement variability in larger tumors using absolute measurement differences ( $P = .06$ ), whereas there is significantly more measurement variability in smaller tumors using percent measurement difference ( $P = .005$ ). Most of those increases exceeding 20% were from tumors smaller than 30 mm.



**Fig 3.** Relative change of longest dimension of indicator lesions, using data from a recently completed phase II study of targeted therapy in non–small-cell lung cancer.<sup>15</sup> The gray zone demarcates changes of less than 10%, highlighting the fact that changes in this range may be a result of variability alone.

expected variability of CT imaging and measurement is known. In the present study, we found a median decrease of 4.2% with the reimaging of a tumor after a short interval of time, indicating that half of measurement decreases as a result of variability alone will exceed 4.2%. As shown in the waterfall plot of our data (Fig 1B), 84% of percent measurement differences fell between  $-10\%$  and  $+10\%$ . Tumor size changes of small magnitudes are commonly reported in clinical trial results, yet the implication of these changes remains unclear considering that these differences may be solely a result of imaging variability.

As an example, in Figure 3, we present a waterfall plot of data from a recently completed phase II trial of targeted therapy in NSCLC (published separately).<sup>15</sup> No standard statistical method exists for comparing waterfall plots, but many authors describe the proportion of patients with tumor shrinkage, including tumors with any evidence of diameter decrease. However, our data would suggest that many measurement decreases, particularly those less than 10%, may be indistinguishable from variability-related changes. Rather than calculating a tumor shrinkage rate, we would recommend that tumors with diameter changes between 0% and 10% decrease be considered relatively unchanged; for this reason, we have added a gray zone to Figure 3 to minimize the significance of changes with a magnitude less than 10%. The diameter decreases between 10% and 30% are less likely to be a result of variability and could potentially represent true antitumor effect, although the clinical implications of such a minor response would need to be investigated further. It is worth remembering that the historical roots of the RECIST response criteria date back to a variability study performed in 1976<sup>16,17</sup>; perhaps the improved accuracy of modern measurement could in part be a basis for reconsidering what qualifies as a tumor response. Interestingly, several groups have found that a 10% decrease in tumor diameter may be correlated with better outcomes in some cancers.<sup>18-21</sup>

Because this study measured only a single lesion for each patient, we are not able to quantify how summed measurement of multiple lesions might affect variability, although we can estimate this effect using the standard deviation of measurement variability from earlier. Although variance increases proportionally with a number ( $n$ ) of independent summed measurements with the same standard deviation ( $\sigma$ ), standard deviation (the square root of variance) increases at a square root proportion ( $\sqrt{n} \times \sigma$ ). This means the relative magnitude of the standard deviation can decrease with summed measurement,

particularly if one assumes no correlation between the measurement error for two tumors on an individual CT scan. To illustrate this, we can consider a patient with multiple 15-mm lung tumors. Measurement of a single 15-mm tumor has a standard deviation of approximately 2.0 mm (Table 3), meaning that 95% of tumor measurements can be expected to lie within  $15 \pm 4.0$  mm; as a percentage of tumor size, this is equal to a 95% limits of agreement of  $-27\%$  and  $+27\%$ . Yet when four 15-mm tumors are measured, the standard deviation increases only by a factor of two, to equal 4.0 mm; therefore, 95% of tumor measurements will lie within  $60 \pm 8.0$  mm, equal to a 95% limits of agreement of  $-13\%$  and  $+13\%$  (Appendix Table A1, online only). This demonstrates how one can increase the relative accuracy of summed measurements by measuring a greater number of similarly sized lesions.

In conclusion, this rescans study of lung lesions in patients with advanced NSCLC found a clinically important magnitude of measurement variability inherent in repeat CT imaging. This variability is greatest in the measurement of small tumors and has important implications for accurate determination of disease progression. Apparent changes in tumor diameter exceeding 1 to 2 mm are common on reimaging and alone may not be indicative of progression. Relative changes less than 10% may be indistinguishable from changes caused by variability alone and are unproven as a marker of efficacy in clinical trials.

#### AUTHORS' DISCLOSURES OF POTENTIAL CONFLICTS OF INTEREST

Although all authors completed the disclosure declaration, the following author(s) indicated a financial or other interest that is relevant to the subject matter under consideration in this article. Certain relationships marked with a "U" are those for which no compensation was received; those relationships marked with a "C" were compensated. For a detailed description of the disclosure categories, or for more information about ASCO's conflict of interest policy, please refer to the Author Disclosure Declaration and the Disclosures of Potential Conflicts of Interest section in Information for Contributors.

**Employment or Leadership Position:** None **Consultant or Advisory Role:** Lawrence H. Schwartz, Novartis (C), GlaxoSmithKline (C) **Stock Ownership:** None **Honoraria:** None **Research Funding:** Lawrence H. Schwartz, AstraZeneca **Expert Testimony:** None **Other Remuneration:** None

#### AUTHOR CONTRIBUTIONS

**Conception and design:** Binsheng Zhao, Mark G. Kris, Lawrence H. Schwartz, Gregory J. Riely

**Financial support:** Mark G. Kris, Lawrence H. Schwartz

**Administrative support:** Binsheng Zhao, Mark G. Kris, Lawrence H. Schwartz

**Provision of study materials or patients:** Michelle S. Ginsberg, Robert A. Lefkowitz, Pingzhen Guo, Mark G. Kris, Gregory J. Riely

**Collection and assembly of data:** Geoffrey R. Oxnard, Binsheng Zhao, Michelle S. Ginsberg, Robert A. Lefkowitz, Pingzhen Guo, Lawrence H. Schwartz, Gregory J. Riely

**Data analysis and interpretation:** Geoffrey R. Oxnard, Binsheng Zhao, Camelia S. Sima, Leonard P. James, Lawrence H. Schwartz, Gregory J. Riely

**Manuscript writing:** All authors

**Final approval of manuscript:** All authors

## REFERENCES

1. Karrison TG, Maitland ML, Stadler WM, et al: 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* 99:1455-1461, 2007
2. LoRusso PM, Anderson AB, Boerner SA, et al: Making the investigational oncology pipeline more efficient and effective: Are we headed in the right direction? *Clin Cancer Res* 16:5956-5962, 2010
3. Hopper KD, Kasales CJ, Van Slyke MA, et al: Analysis of interobserver and intraobserver variability in CT tumor measurements. *AJR Am J Roentgenol* 167:851-854, 1996
4. Thiesse P, Ollivier L, Di Stefano-Louineau D, et al: Response rate accuracy in oncology trials: Reasons for interobserver variability: Groupe Francais d'Immunotherapie of the Federation Nationale des Centres de Lutte Contre le Cancer. *J Clin Oncol* 15:3507-3514, 1997
5. 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* 21:2574-2582, 2003
6. Armato SG 3rd, Oxnard GR, MacMahon H, et al: Measurement of mesothelioma on thoracic CT scans: A comparison of manual and computer-assisted techniques. *Med Phys* 31:1105-1115, 2004
7. Zhao B, Schwartz LH, Moskowitz CS, et al: Pulmonary metastases: Effect of CT section thickness on measurement—Initial experience. *Radiology* 234:934-939, 2005
8. Schwartz LH, Ginsberg MS, DeCorato D, et al: Evaluation of tumor measurements in oncology: Use of film-based and electronic techniques. *J Clin Oncol* 18:2179-2184, 2000
9. Wang Y, de Bock GH, van Klaveren RJ, et al: Volumetric measurement of pulmonary nodules at low-dose chest CT: Effect of reconstruction setting on measurement variability. *Eur Radiol* 20:1180-1187, 2010
10. Goodman LR, Gulsun M, Washington L, et al: Inherent variability of CT lung nodule measurements in vivo using semiautomated volumetric measurements. *AJR Am J Roentgenol* 186:989-994, 2006
11. Wormanns D, Kohl G, Klotz E, et al: Volumetric measurements of pulmonary nodules at multi-row detector CT: In vivo reproducibility. *Eur Radiol* 14:86-92, 2004
12. Zhao B, James LP, Moskowitz CS, et al: Evaluating variability in tumor measurements from same-day repeat CT scans of patients with non-small cell lung cancer. *Radiology* 252:263-272, 2009
13. Eisenhauer EA, Therasse P, Bogaerts J, et al: New Response Evaluation Criteria in Solid Tumours: Revised RECIST guideline (version 1.1). *Eur J Cancer* 45:228-247, 2009
14. Cohen EE, Sharma M, Janisch LA, et al: A phase I study of sirolimus (rapamycin) and bevacizumab in patients with advanced malignancies. *J Clin Oncol* 28:226s, 2010 (suppl; abstr 2591)
15. Price KA, Azzoli CG, Krug LM, et al: Phase II trial of gefitinib and everolimus in advanced non-small cell lung cancer. *J Thorac Oncol* 5:1623-1629, 2010
16. Maitland ML: Volumes to learn: Advancing therapeutics with innovative computed tomography image data analysis. *Clin Cancer Res* 16:4493-4495, 2010
17. Miller AB, Hoogstraten B, Staquet M, et al: Reporting results of cancer treatment. *Cancer* 47:207-214, 1981
18. Le Cesne A, Van Glabbeke M, Verweij J, et al: Absence of progression as assessed by response evaluation criteria in solid tumors predicts survival in advanced GI stromal tumors treated with imatinib mesylate: The intergroup EORTC-ISG-AGITG phase III trial. *J Clin Oncol* 27:3969-3974, 2009
19. Piessevaux H, Buyse M, De Roock W, et al: Radiological tumor size decrease at week 6 is a potent predictor of outcome in chemorefractory metastatic colorectal cancer treated with cetuximab (BOND trial). *Ann Oncol* 20:1375-1382, 2009
20. 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* 59:856-862, 2011
21. 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* 16:4647-4653, 2010