



Published in final edited form as:

Eur J Radiol. 2016 March ; 85(3): 524–533. doi:10.1016/j.ejrad.2015.12.019.

Comparison of CT volumetric measurement with RECIST response in patients with lung cancer

SA Hayes^{1,*}, MC Pietanza^{2,a}, D O'Driscoll^{3,b}, J Zheng^{4,c}, CS Moskowitz^{5,d}, MG Kris^{6,e}, and MS Ginsberg^{7,f}

SA Hayes: hayess@mskcc.org; MC Pietanza: pietanzm@mskcc.org; D O'Driscoll: deemeagle@gmail.com; J Zheng: zhengj@mskcc.org; CS Moskowitz: moskowc1@mskcc.org; MG Kris: krism@mskcc.org; MS Ginsberg: ginsberm@mskcc.org

¹Department of Radiology, Memorial Sloan Kettering Cancer Center, 1275 York Avenue, New York, NY 10065, U.S ²Department of Medicine, Memorial Sloan Kettering Cancer Center, 300 East 66th Street Breast & Imaging Center, NY 10065, U.S ³Department of Radiology, Mater Misericordiae University Hospital, Eccles Street, Dublin 7, Ireland ⁴Department of Epidemiology and Biostatistics, Memorial Sloan Kettering Cancer Center, 307 East 63rd Street, New York, NY 10065, U.S ⁵Department of Epidemiology and Biostatistics, Memorial Sloan Kettering Cancer Center, 307 East 63rd Street, New York, NY 10065, U.S ⁶Department of Medicine, Memorial Sloan Kettering Cancer Center, 300 East 66th Street Breast & Imaging Center, NY 10065, U.S ⁷Department of Radiology, Memorial Sloan Kettering Cancer Center, 1275 York Avenue, New York, NY 10065, U.S

Abstract

Purpose—To examine the correlations between uni-dimensional RECIST and volumetric measurements in patients with lung adenocarcinoma and to assess their association with overall survival (OS) and progression-free survival (PFS).

Materials and Methods—In this study of patients receiving chemotherapy for lung cancer in the setting of a clinical trial, response was prospectively evaluated using RECIST 1.0. Retrospectively, volumetric measurements were recorded and response was assessed by two different volumetric methods at each followup CT scan using a semi-automated segmentation algorithm. We subsequently evaluated the correlation between the uni-dimensional RECIST measurements and the volumetric measurements and performed landmark analyses for OS and

*Corresponding author. Tel.: +1 646-888-4521; Fax: +1 212-794-4010.

^aTel.: +1 646-888-4203

^bTel.: +21 1 803 2000

^cTel.: +1 646-735-8113

^dTel.: +1 646-735-8117

^eTel.: +1 646-888-4205

^fTel.: +1 212-639-7292

Conflict of interest

None

Publisher's Disclaimer: This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final citable form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

PFS at the completion of the first and second follow-ups. Kaplan-Meier curves together with log-rank tests were used to evaluate the association between the different response criteria and patient outcome.

Results—Forty-two patients had CT scans at baseline, after the first follow up scan and second followup scan, and then every 8 weeks. The uni-dimensional RECIST measurements and volumetric measurements were strongly correlated, with a Spearman correlation coefficient (ρ) of 0.853 at baseline, $\rho = 0.861$ at the first followup, $\rho = 0.843$ at the 2nd followup, and $\rho = 0.887$ overall between-subject. On first follow-up CT, partial responders and non responders as assessed by an “ellipsoid” volumetric criteria showed a significant difference in OS ($p=0.008$, 1-year OS of 70% for partial responders and 46% for non responders). There was no difference between the groups when assessed by RECIST criteria on first follow-up CT ($p=0.841$, 1-year OS rate of 64% for partial responders and 64% for non responders).

Conclusion—Volumetric response on first follow-up CT may better predict OS than RECIST response.

Clinical Relevance Statement—Assessment of tumor size and response is of utmost importance in clinical trials. Volumetric measurements may help to better predict OS than uni-dimensional RECIST criteria.

Keywords

Volumetric; RECIST; computed tomography; response; lung cancer

Introduction

Accurate radiological measurement of treatment response is imperative in both clinical practice and clinical trials, and ultimately, also may help predict overall outcome.

The Response Evaluation Criteria In Solid Tumors (RECIST 1.0) was first introduced in 2000 [1] and established uni-dimensional measurement of target lesions on cross-sectional imaging as the standard method for evaluating treatment response. These guidelines were updated in 2009 when RECIST 1.1 was published, with the aim of improving ease of use and providing more accurate assessment of tumor response[2]. World Health Organization (WHO) guidelines published in 1985 also used linear measurements, but recommended two perpendicular measurements[3]. These guidelines were practical to use, particularly in the past when many radiologists used printed film images for reporting, and were widely adopted in oncologic research. However, both systems have limitations in the assessment of treatment response. RECIST guidelines exclude the use of lesions smaller than 1.0 cm. Linear measurements have been shown to be subject to significant intra- and inter-observer variation[4, 5]. For irregularly-shaped non-small cell lung cancers (NSCLC), inter-observer variation resulting in response misclassification has been found to be as high as 45% for uni-dimensional measurements and 53% for bi-dimensional measurements[4].

More recent advancements in multi-detector row computed tomography has allowed for the volumetric measurement of tumors [6, 7]. Studies have indicated that CT measurements are accurate for determining volume [8], possibly with better repeatability and reproducibility

than for linear measurement [9, 10]. Dinkel et al found that using computer-assisted size assessment of lung tumors reduced interobserver variability to about half to one third compared with traditional manual measurements[11]. Published literature has suggested that volumetric measurements can better predict response and outcome than linear measurements [12]. CT volumetry, as well as other advanced methods such as CT perfusion, dynamic contrast-enhanced and diffusion weighted MR and metabolic tumor volume in PET CT, have all shown potential for the assessment of tumor response, but standardization and validation of these newer techniques is needed before they can be widely adopted [13].

Here, we assess volumetric measurements on CT to examine the correlations between uni-dimensional RECIST and volumetric measurements in patients with lung adenocarcinoma receiving first-line chemotherapy in the setting of a clinical trial and to assess their association with overall survival (OS) and progression-free survival (PFS).

Materials and Methods

Patients included in this study were all successfully accrued to a single-arm, open-label, phase II single-institution study evaluating chemotherapy for stage IV lung cancer (ClinicalTrials.gov ID: NCT00807573) that had been reviewed and approved by the institutional review board. Written informed consent was provided by all patients. All patients had pathologically confirmed lung adenocarcinomas with stage IV disease at diagnosis or evidence of metastatic recurrence after definitive local therapy. Inclusion in the study also required Karnofsky performance status of $\geq 70\%$, and measurable disease per RECIST 1.0. Adequate organ and marrow function were necessary. Patients were excluded if they had received systemic therapy for advanced lung cancers or radiation therapy to greater than 25% of the bone marrow within 30 days of starting treatment. While prior neoadjuvant or adjuvant chemotherapy was permitted if it did not contain paclitaxel, pemetrexed or bevacizumab, at least 6 months had to have elapsed from last administration. Additional exclusion criteria included squamous cell carcinoma, small cell carcinoma, hemoptysis; symptomatic brain metastases with evidence of hemorrhage; history of abdominal fistula, gastrointestinal perforation or intra-abdominal abscess; and myocardial infarction or stroke within 6 months.

Forty-four patients were treated with a chemotherapeutic regimen including paclitaxel, pemetrexed, and bevacizumab. Each patient had a thin section CT scan of the chest and other relevant sites of disease at baseline, following cycles 1 and 2 of chemotherapy, and every 8 weeks thereafter. Cycles of therapy consisted of 28 days each. Response to chemotherapy was assessed based on uni-dimensional RECIST 1.0 measurements. Retrospectively, we obtained institutional review board approval to assess tumor volumes on each CT scan and correlated these to the previously measured RECIST. Target lesions for measurement were selected based on RECIST 1.0 guidelines [1], as the study was written and opened to accrual in 2008, prior to the publication of RECIST 1.1 guidelines [2].

Multidetector CT was performed using LightSpeed 16 CT scanners (GE Medical Systems, Milwaukee, Wis). Chest, abdomen and pelvis were scanned from the supraclavicular regions to symphysis pubis, using a single breath hold for the chest. Scan parameters are listed in

Table 1. The images were obtained with intravenous and oral contrast, unless the patient had a contraindication to iodinated contrast. CT images were reconstructed at 1.25mm slice thickness. These thin-section images were directly downloaded from the CT workstation onto a research server, where de-identified DICOM (Digital Imaging and Communications in Medicine) images were stored. The images were then transferred onto an Ultra 10 workstation (Sun Microsystems, Santa Clara, California) for segmentation.

The target lesions in the lung were measured using a novel semi-automated segmentation algorithm, as described previously [7], which had been adapted in house from a program used for the segmentation of pulmonary nodules which had been developed by Zhao et al [14]. Other segmentation algorithms also developed by these authors were used to assess target lesions in other organs, such as lymph nodes and liver [15, 16]. The initial automated segmentation of the target lesions was performed by a technologist. All segmentation results, included the longest diameters and segmented target lesions, were visually inspected for errors by a board-certified cardiothoracic radiologist (M.S.G.) with >15 years of experience in CT interpretation. An example of the segmentation, showing the volumetric outline and the axial measurements recorded is shown in Figure 1. The RECIST uni-dimensional measurements, as well as other patient data, were recorded prospectively by the same radiologist in a novel computer software system designed to enable real-time collection and review of clinical data during trials, as previously described by Pietanza et al[17]. Volumetric measurements of the target lesions were performed using the segmentation algorithms. The sum of the volumes of the target lesions was recorded for each CT scan, similar to a RECIST read. RECIST reads were performed separately to the volumetric segmentation and the radiologist was blinded to the RECIST results at the time of volumetric segmentation.

Response assessment

Treatment response was assessed on follow-up CT scans using uni-dimensional RECIST and two different methods of three-dimensional volumetric response assessment, as a consensus for volumetric response criteria is still lacking. For the first volumetric method, volumetric spherical, we used volumetric response cut-offs based on simple mathematical extrapolation of RECIST to spherical volumes, as initially described by Therasse et al. in the 2000 RECIST guidelines[1]. Here, follow-up CT scans were categorized as complete response (CR, disappearance of the lesions), partial response (PR, 30% decrease in diameter, correlating geometrically to 65% decrease in volume), progressive disease (PD) (20% increase in diameter, correlating geometrically to a 73% increase in volume), or otherwise stable disease (SD). For the second method, volumetric ellipsoid, we utilized an alternative criteria proposed by Schiavon et al in 2012[18], who found that extrapolating RECIST to an ellipsoid, rather than spherical, volume better correlated with survival. The ellipsoid volumetric criteria, when calculated, were the same numerically as RECIST cut-off: partial response (30% decrease in volume), stable disease, or disease progression (20% increase in volume). The three comparison methods are outlined in Table 2.

For all methods, patients who achieved partial response were designated as partial responders and patients with stable disease or progressive disease were labeled as non-

responders. No patient in our study had a complete response. Also, for each technique at each follow-up, we calculated percent change relative to baseline, or to the smallest measurement to date if the method indicated progressive disease (as per RECIST guidelines).

The response assessments occurred during the fourth week of cycle one and cycle two, and during the fourth week of every two cycles thereafter. The range for obtaining the scans was approximately 7 days. However, all imaging assessments were made according to cycles administered and not time on study. Therefore, if a cycle was held and delayed, the scans were delayed as well. In order to estimate the survival in each response category, we performed landmark analysis and choose as landmark times the point when all patients had completed each follow scan.

Statistical analysis

The original clinical trial (ClinicalTrials.gov ID: NCT00807573) was designed to evaluate efficacy of the chemotherapeutic regimen. In this study, we included all evaluable patients from the clinical trial and examined the correlation between RECIST and volumetric measurements as well as associations between response assessments and overall survival.

Scatter plots and between-subject Spearman correlation coefficients proposed by Bland and Altman [19] were used to evaluate the correlation between the uni-dimensional RECIST measurements and volumetric measurements. Overall survival (OS) was defined as the time interval between the starting date of the first treatment and date of death. Patients still alive were censored at the last followup. Progression free survival (PFS) was defined from the starting date of the first treatment to the date of progression, or death when no progression was observed. Patients alive and progression free were censored at the last followup.

The association of response assessments with OS and PFS was evaluated using landmark analyses when all patients had completed the first follow-up scan (by 34 days after the start of treatment) and the second follow-up scan (by 72 days after the start of treatment). OS and PFS were estimated using the Kaplan-Meier method. The log-rank test was used to compare outcomes between partial responders (CR+PR) and non-responders (SD+PD). We examined association between OS/PFS and patient characteristics and clinical factors including sex, performance status (90 vs 70/80), smoking status (current smoker vs others), and mutation status (any vs none). Multivariate OS/PFS analyses were then performed for volumetric response adjusting for univariately significant clinical variables. Cox proportional hazards regression was used for continuous measurements, comparing continuous RECIST and continuous volumetric measurements (measurement percentage changes in continuous values). We tested for differences in the ability of the different response assessments to predict survival by fitting marginal proportional hazards regression models separately in partial responders and non-responders and using a robust score statistic. [20]

Results

Forty-four patients with Stage IV adenocarcinoma of the lung were enrolled in the study between January 2009 and September 2011. Of these, two patients did not undergo post-

treatment imaging; one was removed from the study as a result of toxicity and the other withdrew consent, leaving a total of 42 patients who had subsequent imaging. Follow-up for all 42 patients was completed by 50.4 months, at which time 13 of the 42 patients remained alive with disease and 29 had died. Table 3 summarizes the baseline patient characteristics.

As none of the patients experienced progression of disease or death within 34 days, all 42 patients were included in the landmark analyses at this time point. However, four patients did not have a second follow-up scan while on study, as two had hypersensitivity reactions to one of the chemotherapy drugs in the regimen, one developed clinical disease progression and one died of respiratory failure, likely due to disease. Therefore, thirty-eight patients had undergone the second follow-up CT scan by day 72, and the second landmark analysis was then performed at the 3 month time point. A total of 181 surveillance CT scans were obtained on 42 patients (range, 1–12 scans, median 4 CT scans). The total number of target lesions analyzed at baseline and subsequent follow-up CT was 127, with a median of 3 target lesions per patient and a range of 1–8 target lesions. The target lesion distribution is outlined in Table 4.

Overall study outcomes

The overall response rate (complete and partial) for the described therapeutic regimen was 52% (23/44; 95% CI 37 to 68%) based on RECIST 1.0. Response could not be determined in the 2 patients who did not have a follow-up imaging study. The median progression free and overall survivals for the entire cohort (n=42) were 8 months (95% CI 6 to 12) and 17 months (95% CI 11 to 33), respectively.

Correlation between volumetric measurement and RECIST measurement

The volumetric measurements and the uni-dimensional (RECIST) measurements (sum of longest diameters) were strongly correlated with a Spearman correlation coefficient (ρ) of 0.853 at baseline ($p < 0.001$), $\rho = 0.861$ at the first followup ($p < 0.001$), and $\rho = 0.843$ at the 2nd followup ($p < 0.001$).

As no patient had progressive disease at the first or second follow-up CT scan using either of the volumetric response assessment methods (volumetric ellipsoid or volumetric spherical), percent change was calculated relative to baseline measurements for both volumetric techniques and was identical for both approaches at these specific evaluations. Consequently, in Figure 2 we show results for a single volumetric percent change at first and second follow-up. The correlation between the volumetric percent change and the RECIST percent change was strong (1st followup $\rho = 0.890$, $p < 0.001$ and 2nd followup $\rho = 0.884$, $p < 0.001$).

Response assessment

The number of patients in each response category was calculated for each of the different methods (RECIST, volumetric ellipsoid and volumetric spherical) at first and second follow-up CT scans. At the first follow-up scan, only 14 patients (33.3%) were found to have a partial response when assessed by RECIST and volumetric spherical criteria; all other patients were classified as stable disease. However when assessed using the volumetric

ellipsoid criteria, more patients were classified as partial response rather than stable disease (partial response = 31 patients, 73.8%). Similarly at the second follow-up more patients were also found to have a partial response when assessed by the volumetric ellipsoid criteria (32 patients, 84.2%) as compared to the RECIST criteria (20 patients, 52.6%) and volumetric spherical criteria (19 patients, 50%). No patient was classified as having progressive disease by any of the methods on either the first or second follow-up scan.

Survival outcomes

With a median followup of 27.9 months (range, 17.5 – 50.4 months) on survivors, the median OS was 17.3 months (95% CI: 10.8 – 31.7 months) in 42 patients.

Table 5 outlines the survival analysis, analyzed using continuous volumetric measurements and RECIST group as the predictor at first follow-up (34 days) and at second follow-up (3 months). The results of the Cox proportional hazards regression model are also shown. Survival curves were plotted comparing RECIST and volumetric ellipsoid criteria (Figures 3 and 4). Using the completion of first followup scan as the landmark time, there was a significant difference in OS of partial responders and non-responders as assessed by volumetric ellipsoid criteria ($p=0.008$) (Figure 3b) with a 1-year OS rate of 71.0% for partial responders and 45.5% for non-responders. However, no significant difference in OS was observed between the groups when assessed by RECIST ($p=0.841$) (Figure 3a), with a 1-year OS rate of 64.3% for partial responders and 64.3% for non-responders. When the volumetric spherical criteria were applied, there was a suggestion of a difference in OS between partial responders and non-responders at 1-year (67.9% vs 57.1%, respectively), but the difference was not statistically significant over time ($p=0.269$).

On univariate analysis, the only clinical variable significantly associated with OS was performance status (PS) (hazard ratio 0.37, 95% CI: 0.17–0.80, $p=0.012$), and smoking status was the only variable associated with PFS (hazard ratio 4.13, 95% CI: 1.62–10.51, $p=0.003$). The other variables were not significantly associated with survival on the univariate analysis and were not included in the multivariate analysis. In multivariate analysis adjusting for PS, partial responders by volumetric ellipsoid criteria had a trend of better OS than non-responders at first follow-up (34 days) (HR=0.46, 95% CI: 0.20–1.05, $p=0.065$), Table 6.

We compared the ability of the two different volumetric response assessments to predict OS at first follow-up. Partial responders based on volumetric ellipsoid criteria had significantly better OS compared to partial responder patients based on the volumetric spherical criteria ($p=0.008$, HR=0.5, 95% CI: 0.3–1.2). Further, patients who were non-responders based on volumetric ellipsoid criteria had a significantly increased risk of dying relative to patients who were non-responders based on the volumetric spherical criteria ($p=0.001$, HR=2.3, 95% CI: 1.4–3.7). In comparing volumetric ellipsoid criteria to RECIST, there was no significant difference in OS for patients identified as partial responders by either method ($p=0.170$, HR=0.7, 95% CI: 0.5–1.2), while a significantly increased mortality risk was observed for non-responders based on volumetric ellipsoid criteria relative to non-responders from RECIST ($p=0.007$, HR=2.0, 95% CI: 1.2–3.2).

In contrast, using response status defined at the second follow-up scan, we found little differences in OS between any of the methods for either partial responders or non-responders (Table 5). A non-significant numerical difference was seen using the volumetric ellipsoid criteria ($p=0.094$), though the numbers were small; the 1-year OS rate was 68.8% for partial responders and 50.0% for non responders. However, there was no significant difference in OS between patients identified as partial responders or non-responders when using RECIST ($p=0.457$, 1-year OS rate 65.0% for partial responders and 66.7% for non responders) or volumetric spherical criteria ($p=0.442$, 1-year OS rate 63.2% for partial responders and 68.4% for non responders). A graphical comparison of OS evaluated at second follow-up using RECIST and the volumetric ellipsoid criteria is shown in Figure 4.

We also analyzed the ability of RECIST and the two volumetric response criteria to predict progression-free survival at first and second follow-up. The results of the PFS analysis are presented in Table 7 along with the results for continuous measurements; however no significant difference were seen between the partial-responder and non-responder groups, except a trend of difference at the first followup. In multivariate analysis adjusting for smoking status, the non-responder group as classified by volumetric ellipsoid was associated with a better PFS than the partial responder group (HR=0.48, 95%CI: 0.23–1.00, $p=0.049$).

Discussion

Our study showed that on the first two follow-up scans, volumetric ellipsoid response criteria categorized more patients as partial responders than RECIST or volumetric spherical criteria and categorized fewer patients as non-responders. Larger differences in overall survival between the partial and non-responders were seen when classified by volumetric ellipsoid method than by the other methods, particularly at first follow-up. These results suggest that early radiological response assessed using volumetric ellipsoid criteria, particularly as measured at the first follow-up, is significantly associated with overall survival; in contrast to early response assessed using RECIST where we found no evidence of an association. This finding is of clinical importance, as it suggests the potential use of volumetric response criteria to identify, at an earlier stage of treatment, those responders who will have a long term benefit and non-responders who might benefit from changing therapy. As more therapeutic agents and targeted therapies become available for lung cancer, it will become increasingly important to optimize radiological methods to identify the non responders at the earliest opportunity, to optimize their personalized cancer care.

We found that using extrapolated volumetric cut-off criteria based on a hypothesized ellipsoid shape was a better prognostic indicator of overall survival than using criteria based on a spherical volume, similar to the findings of Schiavon et al. who also investigated this technique[18]. This makes clinical sense, as few tumors are perfectly spherical in shape and an irregular volume such as an ellipsoid is likely to be closer to the true shape of the lesion. This technique also most closely approximates RECIST in implementation with PR measured as decrease of 30% in volume and PD as 20% increase. We suggest that the volumetric measurements more accurately represent the tumor burden and have an advantage over conventional RECIST tumor measurements in detecting partial response, which was seen as early as 34 days in our patient group. Our findings are supported by a

previous study evaluating the same lung segmentation algorithm used in this paper, which determined that the volumetric segmentation identified a larger number of patients with absolute volume changes, compared to axial uni-dimensional or bi-dimensional measurements [7]. Our results also correlate with the findings of a study by Force et al [12] who showed that volumetric analysis detected changes earlier than RECIST. They found that PR was detected earlier and more often with volumetric analysis than with RECIST in a cohort of patients with thymic and lung cancers. Several previous studies have investigated the correlation between RECIST response and survival in NSCLC [21, 22], however these studies did not specifically evaluate the ability of early response to predict survival as we did; which we believe has the potential to be of clinical benefit to patients. We did not use our results to propose new volumetric cut-offs. Given the lack of consensus in the published literature for volumetric response criteria, our aim in this study was to attempt to validate already proposed methods rather than propose additional new criteria.

The reproducibility of volumetric measurements in lung tumors has been widely published previously [10, 23]. Specifically, the reproducibility of volume measurements using the same lung segmentation algorithm we used has previously shown to be very high; volume differences outside the range of -12.1 to 13.4 % could be considered a true change in tumor volume [24]. This cutoff range is much narrower than the cut off values used in either of the two volumetric response assessment methods we evaluated in this study (volumetric spherical -65% , $+73\%$ and volumetric ellipsoid -30% , $+20\%$). The hepatic segmentation algorithm we used has also previously been shown to have high accuracy and reproducibility [15].

Several previous reports have evaluated the association between tumor volume and survival in patients with NSCLC, however most of those previous studies were in patients who received radiotherapy as well as chemotherapy [25–27], unlike our study. Disease burden, as measured by primary tumor and metastatic nodal volume, has been found to be associated with overall survival in patients with Stage III NSCLC [25] and in 270 patients with inoperable Stages I–IIIB NSCLC [26]. A small study of 13 patients with resectable locally advanced NSCLC found that both pre- and post-treatment gross tumor volume predicted PFS [27], using automatic deformable image registration software for volumetric measurements. Nishino et al evaluated a group of 56 NSCLC patients with EGFR mutations treated with first-line tyrosine kinase inhibitors [28] and similar to our results found that an early decrease in tumor volume was associated with overall survival.

The clinical trial on which this study was based used RECIST 1.0 guidelines rather than RECIST 1.1, as the study opened for accrual in 2008 prior to the publication of the updated RECIST guidelines [2]. While this is a potential limitation of our study, several studies comparing RECIST 1.0 to RECIST 1.1 in NSCLC patients have shown very strong or near perfect agreement in response assessment between the two systems [29, 30], although some difference in time to progression were identified. The major factors influencing response assessment in these studies were the inclusion of new lesions identified on PET CT and differences in criteria for lymph node assessment in RECIST 1.1 [31]. A study by Nishino et al in 2010 [30] comparing RECIST 1.0 and RECIST 1.1 suggested that RECIST 1.1 was more reproducible than RECIST 1.0, however both were concordant in terms of response

assessment. The improved reproducibility of RECIST 1.1 over RECIST 1.0 in that study was felt to be due to the lower number of target lesions in RECIST 1.1 compared to 1.0. As the median number of target lesions in our patients was 3 (less than the 5 target lesion maximum in RECIST 1.1), we feel that the reproducibility of RECIST reads in our study was unlikely to be significantly affected by using the older system. However, we accept that some studies suggest an advantage for RECIST 1.1, including a study by Sun et al [32] which suggested that RECIST 1.1 may reflect tumor burden more accurately than RECIST 1.0 in patients with NSCLC being treated with tyrosine kinase inhibitors.

The patients in this study were treated with bevacizumab, a monoclonal antibody directed against vascular endothelial growth factor (VEGF). Bevacizumab is now routinely used in patients with advanced non-squamous NSCLC. Several studies have shown that lung cancers treated with this drug often develop cavitation [33–35] which could potentially impact response assessment, although Marom et al [34] suggested that this tumoral cavitation was of minimal clinical significance, with no significant difference in overall or progression-free survival between cavitating and non-cavitating groups. We feel that any potential effect cavitation could have on response assessment would apply equally to RECIST measurements and volumetric measurements and thus would not affect the overall comparisons between the two, as both of the measurement methods currently evaluate the outer border of the tumor. The authors of one of these studies [33] proposed an alternative method of response assessment which excluded the cavitory component; however it is unknown whether this would aid in predicting outcome.

We advise caution in the application of our results to patients treated with newer targeted therapies for NSCLC. These drugs, for example erlotinib in EGFR-mutated NSCLC, can cause distinct immune-related patterns of response that differ from those of conventional chemotherapy, including initial pseudo-progression and delayed clinical responses. Both the RECIST and volumetric response assessment criteria used in this paper could misclassify some patients on immunotherapeutic agents as progressive disease. In an attempt to address this issue a new response criteria was proposed by Wolchok et al in 2009, the Immune-Related Response Criteria [36], however this has not yet been established as standard of care.

Measurement variability has been shown to be a potential issue in volumetry [37, 38]. Segmentation errors in particular have the potential to cause large inaccuracies in lesion volume measurement and have been shown to account for a large degree of inter-scan variability of nodule volumetry[37]. As the segmentation of each lesion in our study was visually inspected and verified by a radiologist, we feel the potential for segmentation errors was minimized. The lack of an independent gold standard for lesion volume measurements was a limitation of this study, as none of these target lesions were resected and correlated with pathological tumor size. Given that patients with stage IV NSCLC are not generally considered for surgical resection, this potential limitation will apply to any study performed on this patient group.

We used the same target lesions for both the RECIST and the volumetric response assessment, as the use of different target lesions has been found to be a major source of

variability in RECIST 1.1 response assessment [39]. A recent study published by Oubel et al also found that consensual target lesion selection reduced inter-observer variability in volumetric response assessments [40].

One of the benefits of using RECIST guidelines in assessing tumor response is its ease of use, without the need for additional software to measure tumor size. Volumetric software is not yet widely available, can be costly and has not been adopted into daily clinical practice in most radiology departments. There is also a lack of standardization and external validation across many of the volumetric software systems being used in current research. Additional software costs to departments may hamper widespread adoption of volumetric measurements. The software used in our study was developed in house and is not commercially available, therefore the cost per study was difficult to quantify, however a technologist was needed to implement the segmentation algorithm on each target lesion, so added labor costs are a consideration. The segmentation in our study was semi-automated and required the radiologist to assess and adjust the segmentation on each target lesion, which was more time consuming than a RECIST response assessment read would have been. For these reasons, although we feel the volumetric technique has promise as a potential future tool in response assessment, we accept that further advances and validation of volumetric software will be needed before this can be applied to clinical practice.

Imaging modalities other than CT have also been studied extensively in response assessment in recent years. While these modalities offer future potential, response criteria cutoffs and standardization of techniques remain an issue for both PET CT and MRI. The role of PET CT in measuring response has been widely studied in lung cancer. While there is good evidence that decreased 18F-FDG uptake in lung tumors is associated with increased survival, as summarized by Hicks in 2009 [41], consensus is lacking among researchers regarding standardization of measures of response. Lack of reproducibility in PET CT methodology remains a real concern [41, 42]. Similarly, MRI response assessment, particularly with dynamic contrast-enhanced MRI and diffusion-weighted imaging, has shown promising results in correlating tumor response and survival [43–48], but is also hampered by issues of standardization across image acquisition, post-processing and analysis and also by the range of measurement variability. While standardization of different volumetric software will be an issue, reproducibility of individual volumetric measurement algorithms has been shown to be high and we feel that overall the obstacles facing volumetric response assessment are less than those weighted against molecular imaging or MRI.

Overall, the main limitation of our study was the small sample size, which may have reduced our ability to detect significant differences in outcomes, particularly in the multivariate analysis. However, we believe that the differences in survival curves between the volumetric ellipsoid criteria and RECIST, particularly at the first follow-up, indicate a definite advantage for volumetric measurement rather than linear measurements. Larger studies will be needed to validate these findings in the future and to assess for more subtle difference between the techniques that our study may have lacked the power to reveal.

Conclusion

In conclusion, our findings show that volumetric ellipsoid measurement criteria may detect partial responders and non responders to treatment at an earlier stage than RECIST and have the potential to improve clinical care in cancer patients in the future. Volumetric measurements tools are continuing to advance and are becoming more widely available in clinical radiology settings. Further validation of the advantages of volumetric measurements over linear measurements to assess radiological response to treatment will be necessary before this can be adopted as a gold standard in clinical trials.

Acknowledgments

We acknowledge the support of the MSKCC Biostatistics Core (P30 CA008748).

References

1. Therasse P, Arbuck SG, Eisenhauer EA, Wanders J, Kaplan RS, Rubinstein L, 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; 92:205–16. [PubMed: 10655437]
2. Eisenhauer EA, Therasse P, Bogaerts J, Schwartz LH, Sargent D, Ford R, et al. New response evaluation criteria in solid tumours: revised RECIST guideline (version 1.1). *European journal of cancer (Oxford, England : 1990).* 2009; 45:228–47.
3. Miller AB, Hoogstraten B, Staquet M, Winkler A. Reporting results of cancer treatment. *Cancer.* 1981; 47:207–14. [PubMed: 7459811]
4. Erasmus JJ, Gladish GW, Broemeling L, Sabloff BS, Truong MT, Herbst RS, et al. Interobserver and intraobserver variability in measurement of non-small-cell carcinoma lung lesions: implications for assessment of tumor response. *Journal of clinical oncology : official journal of the American Society of Clinical Oncology.* 2003; 21:2574–82. [PubMed: 12829678]
5. Zhao B, Tan Y, Bell DJ, Marley SE, Guo P, Mann H, et al. Exploring intra- and inter-reader variability in uni-dimensional, bi-dimensional, and volumetric measurements of solid tumors on CT scans reconstructed at different slice intervals. *European journal of radiology.* 2013; 82:959–68. [PubMed: 23489982]
6. Gavrielides MA, Kinnard LM, Myers KJ, Petrick N. Noncalcified lung nodules: volumetric assessment with thoracic CT. *Radiology.* 2009; 251:26–37. [PubMed: 19332844]
7. 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. [PubMed: 17114630]
8. Yankelevitz DF, Reeves AP, Kostis WJ, Zhao B, Henschke CI. Small pulmonary nodules: volumetrically determined growth rates based on CT evaluation. *Radiology.* 2000; 217:251–6. [PubMed: 11012453]
9. Goodman LR, Gulsun M, Washington L, Nagy PG, Piacsek KL. Inherent variability of CT lung nodule measurements in vivo using semiautomated volumetric measurements. *AJR American journal of roentgenology.* 2006; 186:989–94. [PubMed: 16554568]
10. Mozley PD, Bendtsen C, Zhao B, Schwartz LH, Thorn M, Rong Y, et al. Measurement of tumor volumes improves RECIST-based response assessments in advanced lung cancer. *Translational oncology.* 2012; 5:19–25. [PubMed: 22348172]
11. Dinkel J, Khalilzadeh O, Hintze C, Fabel M, Puderbach M, Eichinger M, et al. Inter-observer reproducibility of semi-automatic tumor diameter measurement and volumetric analysis in patients with lung cancer. *Lung cancer.* 2013; 82:76–82. [PubMed: 23932487]
12. Force J, Rajan A, Dombi E, Steinberg SM, Giaccone G. Assessment of objective responses using volumetric evaluation in advanced thymic malignancies and metastatic non-small cell lung cancer.

Journal of thoracic oncology : official publication of the International Association for the Study of Lung Cancer. 2011; 6:1267–73.

13. Nishino M, Hatabu H, Johnson BE, McLoud TC. State of the art: Response assessment in lung cancer in the era of genomic medicine. *Radiology*. 2014; 271:6–27. [PubMed: 24661292]
14. Zhao B, Reeves AP, Yankelevitz DF, Henschke CI. Three-dimensional multicriterion automatic segmentation of pulmonary nodules of helical computed tomography images. *Optical Engineering*. 1999; 38:1340–7.
15. Zhao B, Schwartz LH, Jiang L, Colville J, Moskowitz C, Wang L, et al. Shape-constraint region growing for delineation of hepatic metastases on contrast-enhanced computed tomograph scans. *Investigative radiology*. 2006; 41:753–62. [PubMed: 16971799]
16. Yan J, Zhao B, Wang L, Zelenetz A, Schwartz LH. Marker-controlled watershed for lymphoma segmentation in sequential CT images. *Medical physics*. 2006; 33:2452–60. [PubMed: 16898448]
17. Pietanza MC, Basch EM, Lash A, Schwartz LH, Ginsberg MS, Zhao B, et al. Harnessing technology to improve clinical trials: study of real-time informatics to collect data, toxicities, image response assessments, and patient-reported outcomes in a phase II clinical trial. *Journal of clinical oncology : official journal of the American Society of Clinical Oncology*. 2013; 31:2004–9. [PubMed: 23630218]
18. Schiavon G, Ruggiero A, Schoffski P, van der Holt B, Bekers DJ, Eechoute K, et al. Tumor volume as an alternative response measurement for imatinib treated GIST patients. *PLoS one*. 2012; 7:e48372. [PubMed: 23133631]
19. Bland JM, Altman DG. Calculating correlation coefficients with repeated observations: Part 2--Correlation between subjects. *BMJ (Clinical research ed)*. 1995; 310:633.
20. Moskowitz CS, Pepe MS. Quantifying and comparing the predictive accuracy of continuous prognostic factors for binary outcomes. *Biostatistics (Oxford, England)*. 2004; 5:113–27.
21. Jain RK, Lee JJ, Ng C, Hong D, Gong J, Naing A, et al. Change in tumor size by RECIST correlates linearly with overall survival in phase I oncology studies. *Journal of clinical oncology : official journal of the American Society of Clinical Oncology*. 2012; 30:2684–90. [PubMed: 22689801]
22. Takeda M, Okamoto I, Nakagawa K. Survival outcome assessed according to tumor response and shrinkage pattern in patients with EGFR mutation-positive non-small-cell lung cancer treated with gefitinib or erlotinib. *Journal of thoracic oncology : official publication of the International Association for the Study of Lung Cancer*. 2014; 9:200–4.
23. Nishino M, Guo M, Jackman DM, DiPiro PJ, Yap JT, Ho TK, et al. CT tumor volume measurement in advanced non-small-cell lung cancer: Performance characteristics of an emerging clinical tool. *Academic radiology*. 2011; 18:54–62. [PubMed: 21036632]
24. Zhao B, James LP, Moskowitz CS, Guo P, Ginsberg MS, Lefkowitz RA, et al. Evaluating variability in tumor measurements from same-day repeat CT scans of patients with non-small cell lung cancer. *Radiology*. 2009; 252:263–72. [PubMed: 19561260]
25. Alexander BM, Othus M, Caglar HB, Allen AM. Tumor volume is a prognostic factor in non-small-cell lung cancer treated with chemoradiotherapy. *International journal of radiation oncology, biology, physics*. 2011; 79:1381–7.
26. Dehing-Oberije C, De Ruyscher D, van der Weide H, Hochstenbag M, Bootsma G, Geraedts W, et al. Tumor volume combined with number of positive lymph node stations is a more important prognostic factor than TNM stage for survival of non-small-cell lung cancer patients treated with (chemo)radiotherapy. *International journal of radiation oncology, biology, physics*. 2008; 70:1039–44.
27. Kozak MM, Murphy JD, Schipper ML, Donington JS, Zhou L, Whyte RI, et al. Tumor volume as a potential imaging-based risk-stratification factor in trimodality therapy for locally advanced non-small cell lung cancer. *Journal of thoracic oncology : official publication of the International Association for the Study of Lung Cancer*. 2011; 6:920–6.
28. Nishino M, Dahlberg SE, Cardarella S, Jackman DM, Rabin MS, Hatabu H, et al. Tumor volume decrease at 8 weeks is associated with longer survival in EGFR-mutant advanced non-small-cell lung cancer patients treated with EGFR TKI. *Journal of thoracic oncology : official publication of the International Association for the Study of Lung Cancer*. 2013; 8:1059–68.

29. Nishino M, Cardarella S, Jackman DM, Ramaiya NH, Rabin MS, Hatabu H, et al. RECIST 1.1 in NSCLC patients with EGFR mutations treated with EGFR tyrosine kinase inhibitors: comparison with RECIST 1.0. *AJR American journal of roentgenology*. 2013; 201:W64–71. [PubMed: 23789698]
30. Nishino M, Jackman DM, Hatabu H, Yeap BY, Cioffredi LA, Yap JT, et al. New Response Evaluation Criteria in Solid Tumors (RECIST) guidelines for advanced non-small cell lung cancer: comparison with original RECIST and impact on assessment of tumor response to targeted therapy. *AJR American journal of roentgenology*. 2010; 195:W221–8. [PubMed: 20729419]
31. Choi HC, Kim JH, Kim HS, Jung SG, Hwang SM, Ju SB, et al. Comparison of the RECIST 1.0 and RECIST 1.1 in Non-Small Cell Lung Cancer Treated with Cytotoxic Chemotherapy. *Journal of Cancer*. 2015; 6:652–7. [PubMed: 26078796]
32. Sun JM, Ahn MJ, Park MJ, Yi JH, Kim TS, Chung MJ, et al. Accuracy of RECIST 1.1 for non-small cell lung cancer treated with EGFR tyrosine kinase inhibitors. *Lung cancer*. 2010; 69:105–9. [PubMed: 19959251]
33. Crabb SJ, Patsios D, Sauerbrei E, Ellis PM, Arnold A, Goss G, et al. Tumor cavitation: impact on objective response evaluation in trials of angiogenesis inhibitors in non-small-cell lung cancer. *Journal of clinical oncology : official journal of the American Society of Clinical Oncology*. 2009; 27:404–10. [PubMed: 19047292]
34. Marom EM, Martinez CH, Truong MT, Lei X, Sabloff BS, Munden RF, et al. Tumor cavitation during therapy with antiangiogenesis agents in patients with lung cancer. *Journal of thoracic oncology : official publication of the International Association for the Study of Lung Cancer*. 2008; 3:351–7.
35. Nishino M, Cryer SK, Okajima Y, Sholl LM, Hatabu H, Rabin MS, et al. Tumoral cavitation in patients with non-small-cell lung cancer treated with antiangiogenic therapy using bevacizumab. *Cancer imaging : the official publication of the International Cancer Imaging Society*. 2012; 12:225–35. [PubMed: 22743083]
36. Wolchok JD, Hoos A, O'Day S, Weber JS, Hamid O, Lebbe C, et al. Guidelines for the evaluation of immune therapy activity in solid tumors: immune-related response criteria. *Clinical cancer research : an official journal of the American Association for Cancer Research*. 2009; 15:7412–20. [PubMed: 19934295]
37. Gietema HA, Schaefer-Prokop CM, Mali WP, Groenewegen G, Prokop M. Pulmonary nodules: Interscan variability of semiautomated volume measurements with multisection CT-- influence of inspiration level, nodule size, and segmentation performance. *Radiology*. 2007; 245:888–94. [PubMed: 17923508]
38. Wormanns D, Kohl G, Klotz E, Marheine A, Beyer F, Heindel W, et al. Volumetric measurements of pulmonary nodules at multi-row detector CT: in vivo reproducibility. *European radiology*. 2004; 14:86–92. [PubMed: 14615902]
39. Keil S, Barabasch A, Dirrichs T, Bruners P, Hansen NL, Bieling HB, et al. Target lesion selection: an important factor causing variability of response classification in the Response Evaluation Criteria for Solid Tumors 1.1. *Investigative radiology*. 2014; 49:509–17. [PubMed: 24651664]
40. Oubel E, Bonnard E, Sueoka-Aragane N, Kobayashi N, Charbonnier C, Yamamichi J, et al. Volume-based response evaluation with consensual lesion selection: a pilot study by using cloud solutions and comparison to RECIST 1.1. *Academic radiology*. 2015; 22:217–25. [PubMed: 25488429]
41. Hicks RJ. Role of 18F-FDG PET in assessment of response in non-small cell lung cancer. *Journal of nuclear medicine : official publication, Society of Nuclear Medicine*. 2009; 50(Suppl 1):31s–42s.
42. Graham MM, Badawi RD, Wahl RL. Variations in PET/CT methodology for oncologic imaging at U.S. academic medical centers: an imaging response assessment team survey. *Journal of nuclear medicine : official publication, Society of Nuclear Medicine*. 2011; 52:311–7.
43. de Langen AJ, van den Boogaart V, Lubberink M, Backes WH, Marcus JT, van Tinteren H, et al. Monitoring response to antiangiogenic therapy in non-small cell lung cancer using imaging markers derived from PET and dynamic contrast-enhanced MRI. *Journal of nuclear medicine : official publication, Society of Nuclear Medicine*. 2011; 52:48–55.

44. Dingemans AM, de Langen AJ, van den Boogaart V, Marcus JT, Backes WH, Scholtens HT, et al. First-line erlotinib and bevacizumab in patients with locally advanced and/or metastatic non-small-cell lung cancer: a phase II study including molecular imaging. *Annals of oncology : official journal of the European Society for Medical Oncology / ESMO*. 2011; 22:559–66. [PubMed: 20702788]
45. Fujimoto K, Abe T, Muller NL, Terasaki H, Kato S, Sadohara J, et al. Small peripheral pulmonary carcinomas evaluated with dynamic MR imaging: correlation with tumor vascularity and prognosis. *Radiology*. 2003; 227:786–93. [PubMed: 12714678]
46. Kelly RJ, Rajan A, Force J, Lopez-Chavez A, Keen C, Cao L, et al. Evaluation of KRAS mutations, angiogenic biomarkers, and DCE-MRI in patients with advanced non-small-cell lung cancer receiving sorafenib. *Clinical cancer research : an official journal of the American Association for Cancer Research*. 2011; 17:1190–9. [PubMed: 21224376]
47. Ohno Y, Nogami M, Higashino T, Takenaka D, Matsumoto S, Hatabu H, et al. Prognostic value of dynamic MR imaging for non-small-cell lung cancer patients after chemoradiotherapy. *Journal of magnetic resonance imaging : JMRI*. 2005; 21:775–83. [PubMed: 15906341]
48. Yabuuchi H, Hatakenaka M, Takayama K, Matsuo Y, Sunami S, Kamitani T, et al. Non-small cell lung cancer: detection of early response to chemotherapy by using contrast-enhanced dynamic and diffusion-weighted MR imaging. *Radiology*. 2011; 261:598–604. [PubMed: 21852569]

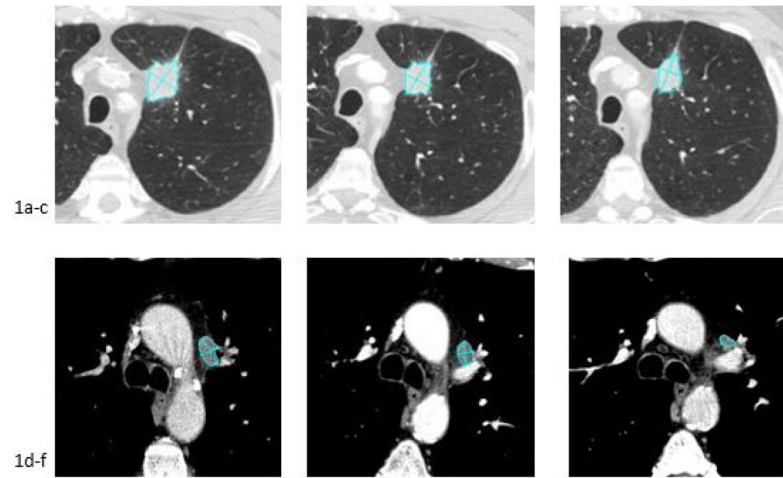


Fig. 1. 66 year old male with metastatic lung cancer. CT images demonstrate the volumetric segmentation and axial diameters recorded for a left upper lobe tumor at baseline (1a), first follow-up (1b) and second follow-up (1c) and a metastatic lymph node at baseline (1d), first follow-up (1e) and second follow-up (1f).

Fig. 2a

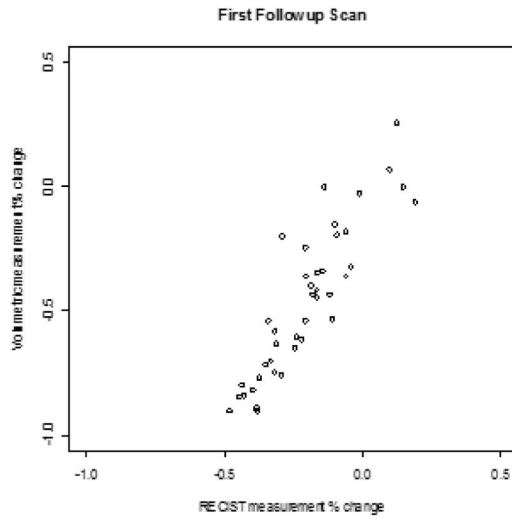


Fig. 2b

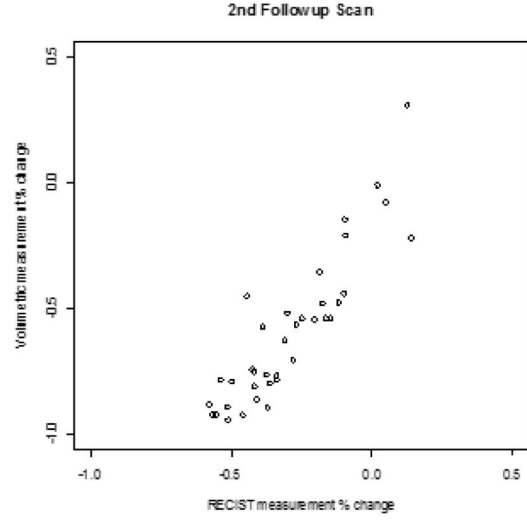


Fig. 2. Scatter plots showing the correlation between volumetric percentage change and RECIST measurement percentage change at first (Fig. 2a) and second (Fig. 2b) follow-ups

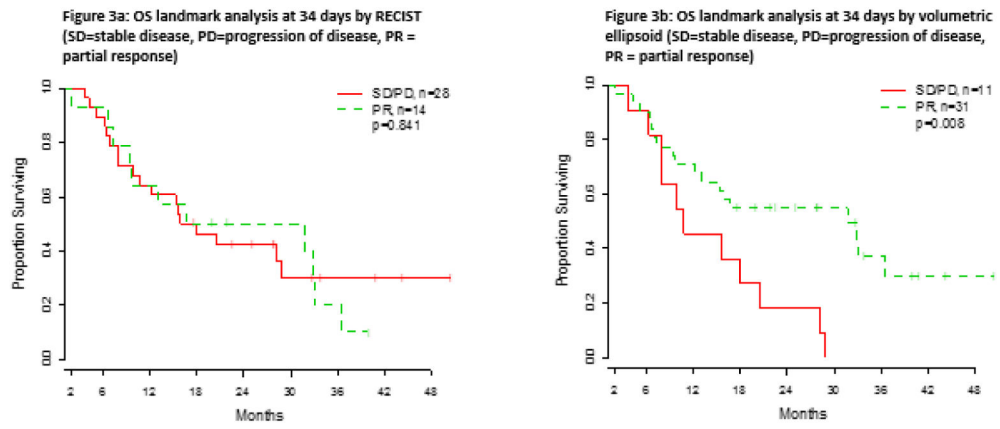


Fig 3. Comparison of overall survival evaluated at first follow-up (34 days) using a) RECIST and b) volumetric ellipsoid criteria

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

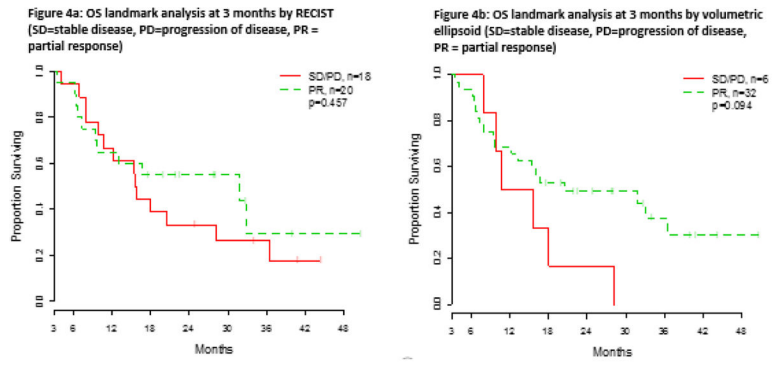


Fig. 4. Comparison of overall survival evaluated at second follow-up (3 months) using a) RECIST and b) volumetric ellipsoid criteria

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

Table 1

CT Scanning Parameters

Parameter	Value
Detector row configuration	16 × 1.25
Pitch/Table Speed	1.375/27.50 mm
Collimation	2.5 × 2.5 mm
Reconstruction algorithm	1.25 mm slice thickness, lung and soft tissue windows

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

Table 2

Comparison of treatment response assessment criteria

Criteria	Partial Response (PR)	Progressive Disease (PD)	Stable Disease (SD)
RECIST (2D)	Decrease by 30%	Increase by 20%	Neither PR nor PD criteria met
Volumetric spherical (3D)	Decrease by 65%	Increase by 73%	Neither PR nor PD criteria met
Volumetric ellipsoid (3D)	Decrease by 30%	Increase by 20%	Neither PR nor PD criteria met

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

Table 3

Baseline Patient characteristics

Patient Characteristics (N = 42)	
Median Age, years (range)	60 (31–77)
Sex	
Male	22 (52%)
Female	20 (48%)
Performance Status	
70	5 (12%)
80	18 (43%)
90	19 (45%)
Race/Ethnicity	
White	39 (93%)
Asian	2 (5%)
Other	1 (2%)
Smoker	
Current	7 (17%)
Former	28 (67%)
Mutation Status	
KRAS mutations	16 (36%)
ALK rearranged	3 (7%)
BRAF mutation	2 (5%)
EGFR exon 20 Insertion	1 (2%)
HER2/PIK3CA	1 (2%)
None	15 (34%)
Not tested	6 (14%)

Table 4

Target lesion distribution

Target Lesion Distribution	
Lung	57 (45%)
Lymph node	48 (38%)
Adrenal	11 (9%)
Liver	8 (6%)
Soft tissue	2 (2%)
Peritoneum	1 (1%)

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

Table 5

Landmark overall survival (OS) analysis at first and second follow-ups

	First follow-up (N = 42)		Second follow-up (N = 38)	
	1-year OS (95%CI)	p value	1-year OS (95%CI)	p value
RECIST				
Partial responder	64.3% (34.3%, 83.3%)	0.841	65.0% (40.3%, 81.5%)	0.457
Non-responder	64.3% (43.8%, 78.9%)		66.7% (40.4%, 83.4%)	
Volumetric ellipsoid				
Partial responder	71.0% (51.6%, 83.7%)	0.008	68.8% (49.7%, 81.8%)	0.094
Non-responder	45.5% (16.7%, 70.7%)		50.0% (11.1%, 80.4%)	
Volumetric spherical				
Partial responder	67.9% (47.3%, 81.8%)	0.269	63.2% (37.9%, 80.4%)	0.442
Non-responder	57.1% (28.4%, 78.0%)		68.4% (42.8%, 84.4%)	
	Hazard Ratio (95%CI)	p value	Hazard Ratio (95%CI)	p value
Continuous RECIST	1.27 (0.12, 13.66)	0.843	1.52 (0.17, 13.56)	0.707
Continuous Volumetric	2.18 (0.57, 8.33)	0.256	2.66 (0.69, 10.30)	0.156

Volumetric ellipsoid = volumetric measurement groups based on ellipsoid volumetric criteria.

Volumetric spherical = volumetric measurement groups based on spherical volumetric criteria.

Non-responder = Stable disease or progression of disease

P value = measured between partial responders and non-responders in each response criteria category

Continuous Volumetric = volumetric measurement percent change in continuous value.

Table 6

Landmark multivariate analysis at 34 days

	OS		PFS		
PS	Hazard Ratio (95%CI)	p value	Smoking Status	Hazard Ratio (95%CI)	p value
70/80	1	0.05	Former/never smoker	1	0.003
90	0.44 (0.20, 1.00)		Current Smoker	4.32 (1.69, 11.03)	
Volumetric Ellipsoid					
Partial responder	1	0.065	Partial responder	1	0.049
Non-responder	0.46 (0.20, 1.05)		Non-responder	0.48 (0.23, 1.00)	

Table 7

Landmark progression-free survival (PFS) analysis at first and second follow-ups

RECIST	First follow-up (N = 42)		Second follow-up (N = 38)	
	1-year PFS (95%CI)	p value	1-year PFS (95%CI)	p value
Partial responder	42.9% (17.7%, 66.0%)	0.651	40.0% (19.3%, 60.0%)	0.609
Non-responder	35.7% (18.9%, 53.0%)		35.3% (14.4%, 57.0%)	
Volumetric ellipsoid				
Partial responder	41.9% (24.7%, 58.3%)	0.058	38.7% (22.0%, 55.1%)	0.159
Non-responder	27.3% (6.5%, 53.9%)		33.3% (4.6%, 67.6%)	
Volumetric spherical				
Partial responder	39.3% (21.7%, 56.5%)	0.544	38.9% (17.5%, 60.0%)	0.893
Non-responder	35.7% (13.0%, 59.4%)		36.8% (16.5%, 57.5%)	
	Hazard Ratio (95%CI)	p value	Hazard Ratio (95%CI)	p value
Continuous RECIST	1.29 (0.15, 10.90)	0.817	1.15 (0.15, 8.86)	0.897
Continuous Volumetric	1.99 (0.58, 6.84)	0.278	2.44 (0.56, 10.67)	0.232

Volumetric ellipsoid = volumetric measurement groups based on ellipsoid volumetric criteria.

Volumetric spherical = volumetric measurement groups based on spherical volumetric criteria.

Non-responder = Stable disease or progression of disease

P value = measured between partial responders and non-responders in each response criteria category

Continuous Volumetric = volumetric measurement percent change in continuous value.