

METABOLIC TUMOR BURDEN PREDICTS FOR DISEASE PROGRESSION AND DEATH IN LUNG CANCER

PERCY LEE, M.D.,* DILANI K. WEERASURIYA, B.S.,* PHILIP W. LAVORI, PH.D.,† ANDREW QUON, M.D.,‡
WENDY HARA, M.D.,* PETER G. MAXIM, PH.D.,* QUYNH-THU LE, M.D.,* HEATHER A. WAKELEE, M.D.,§
JESSICA S. DONINGTON, M.D.,|| EDWARD E. GRAVES, PH.D.,* AND BILLY W. LOO, JR., M.D., PH.D.*

*Department of Radiation Oncology and †Department of Health Research and Policy, Stanford University, Stanford, CA; ‡Department of Radiology, Division of Nuclear Medicine, Stanford University, Stanford, CA; §Department of Medicine, Division of Oncology, Stanford University, Stanford, CA; and ||Department of Cardiothoracic Surgery, Stanford University, Stanford, CA

Purpose: In lung cancer, stage is an important prognostic factor for disease progression and survival. However, stage may be simply a surrogate for underlying tumor burden. Our purpose was to assess the prognostic value of tumor burden measured by ¹⁸F-fluorodeoxyglucose–positron emission tomography (FDG-PET) imaging.

Patients and Methods: We identified 19 patients with lung cancer who had staging PET-CT scans before any therapy, and adequate follow-up (complete to time of progression for 18, and death for 15 of 19). Metabolically active tumor regions were segmented on pretreatment PET scans semi-automatically using custom software. We determined the relationship between times to progression (TTP) and death (OS) and two PET parameters: total metabolic tumor volume (MTV), and standardized uptake value (SUV).

Results: The estimated median TTP and OS for the cohort were 9.3 months and 14.8 months. On multivariate Cox proportional hazards regression analysis, an increase in MTV of 25 ml (difference between the 75th and 25th percentiles) was associated with increased hazard of progression and of death (5.4-fold and 7.6-fold), statistically significant ($p = 0.0014$ and $p = 0.001$) after controlling for stage, treatment intent (definitive or palliative), age, Karnofsky performance status, and weight loss. We did not find a significant relationship between SUV and TTP or OS.

Conclusions: In this study, high tumor burden assessed by PET MTV is an independent poor prognostic feature in lung cancer, promising for stratifying patients in randomized trials and ultimately for selecting risk-adapted therapies. These results will need to be validated in larger cohorts with longer follow-up, and evaluated prospectively. © 2007 Elsevier Inc.

Lung cancer, Positron emission tomography (PET), Metabolic tumor volume (MTV), Prognostic factors, Automatic image analysis.

INTRODUCTION

Currently, lung cancer stage is the single most prognostic factor in predicting the outcomes of patients with lung cancer (1–3). Other potentially important prognostic factors for lung cancer include weight loss, performance status, age, and gender (4). However, these prognostic factors may simply be a surrogate for or correlate with the underlying tumor burden, which may be a more direct predictor of disease progression and survival. Until recently, it has been difficult to quantify tumor burden directly and systematically. Crude surrogates such as primary tumor size have been demonstrated to correlate inversely with the duration of survival in resected lung cancer (5). Accordingly, it has been demon-

strated that gross tumor volume (GTV) determined by manual contouring on computed tomography (CT) images as part of three-dimensional (3D) conformal radiation treatment planning predicts for overall and cause-specific survival as well as local tumor control (6).

Positron emission tomography (PET) imaging using the tracer ¹⁸F-fluorodeoxyglucose (FDG) has revolutionized the staging of lung cancer and has become the standard of care for this purpose. Fused FDG-PET and CT imaging is beginning to have a considerable impact on radiation therapy treatment planning, altering the tumor volume delineation in more than 50% of patients when compared with CT-based treatment planning alone (7, 8). Recently, the degree of tumor uptake of FDG on PET as assessed by the standardized uptake

Note—An online CME test for this article can be taken at <http://astro.astro.org> under Continuing Education.

Reprint requests to: Billy W. Loo, Jr., M.D., Ph.D., Department of Radiation Oncology, Stanford University, Stanford Cancer Center, 875 Blake Wilbur Drive, Stanford, CA 94305-5847. Tel: (650) 736-7143; Fax: (650) 725-8231; E-mail: BWLoo@Stanford.edu

Presented at the 47th Annual Meeting of the American Society for Therapeutic Radiology and Oncology (ASTRO), October 16–20, 2005, Denver, CO.

Conflict of interest: none.

Received Jan 26, 2007, and in revised form April 14, 2007. Accepted for publication April 19, 2007.

value (SUV) was shown to be an independent prognostic factor in non-small cell lung cancer (NSCLC) (9–12). However, total body tumor burden reflected by the volume of tumor tissue demonstrating increased FDG uptake on PET, or metabolic tumor volume (MTV), is a novel potential prognostic factor in lung cancer that has not yet been investigated.

Thus, the objective of this study was to test the hypothesis that tumor burden as characterized by MTV is an independent prognostic factor that can predict disease progression in lung cancer.

PATIENTS AND METHODS

Patients

We conducted a retrospective review of the medical records of patients who underwent FDG-PET-CT scanning at Stanford Hospital and Clinics. We conducted this study under the review and approval of the Stanford institutional review board. Between January 2003 and November 2004, 1565 FDG-PET scans were performed on the PET-CT scanner. Of the patients scanned, we identified 80 who were evaluated for NSCLC and small cell lung cancer (SCLC). This group includes patients who only had follow-up scans for restaging but not initial staging scans, and patients who only had imaging at our institution but no further treatment or follow-up. Of the 80 patients, 25 had initial staging scans, and of these we identified 19 patients who were treated and had subsequent clinical follow-up, thus forming the cohort of this study.

PET imaging protocol

All scans were performed on a GE Discovery LS PET-CT scanner (GE Medical Systems, Milwaukee, WI). Each patient fasted for at least 8 h before imaging. After ensuring that blood glucose was <180 mg/dl, patients were injected with 12 to 18 mCi of FDG. After a tracer uptake time of 45 to 60 min, patients underwent PET/CT imaging. Initially, frontal and lateral x-ray projection images were acquired to act as localizers. Using these images, a whole body scan volume was defined and CT data were collected in helical acquisition mode. Using the same scan locations, generally spanning about seven bed positions, PET data were acquired in two dimensional (2D) mode, for 3 to 5 min of acquisition time per bed position. The PET data were then reconstructed with an ordered set expectation maximization (OSEM) algorithm, using the CT images for attenuation correction. At the conclusion of the examination, all reconstructed image data were transferred to a Xeleris workstation (GE Medical Systems, Milwaukee, WI) for clinical evaluation and also to a local server for tumor volume data analysis (described below). The complete whole body PET-CT examination requires approximately 90 min, including patient setup, tracer uptake, and CT and PET image acquisition.

Measurement of tumor volume

Computer-aided metabolic tumor volume measurement was performed using RT_Image, a software application developed at our institution to analyze functional imaging data for radiation therapy applications, using the Interactive Data Language (IDL; Research Systems, Inc., Boulder, CO) (13). The FDG-PET data were read into the program in DICOM format. Intensity values were automatically converted to SUVs. The images were viewed as maximum intensity projection (MIP) images to allow rapid visual identification of the hypermetabolic lesions by a radiation oncologist experienced in PET-CT based treatment planning (PL), using the diagnostic

nuclear medicine reports as a reference. The user then selected each lesion interactively by clicking on its projection using a graphical user interface. This is the only step in the segmentation process requiring user interaction.

Each tumor thus identified by the user was then segmented automatically in three dimensions by the software using the following procedure. First, the voxel of maximum intensity along the selected projection line is used as the starting point for a region growing procedure. The algorithm then finds the voxel of local maximum intensity within a specified radius (default value of 1 cm) of the starting voxel. The region growing algorithm then defines the segmented volume as all voxels connected to the local maximum intensity voxel that have an intensity greater than a specified fraction of the maximum intensity. The threshold intensity value used in this study was 50% of the local maximum intensity, which has been identified as a reasonable choice in phantom studies (7).

Once all of the hypermetabolic tumor foci are segmented, the software calculates the metabolic tumor volume (MTV), defined as the total volume of all tumors in the body in milliliters. Of note, the necrotic center of tumors, when present, is included as part of the MTV. Maximum and average SUV within the MTV are also calculated automatically. Figure 1 shows examples of patients with small and large MTV.

Statistical analysis

These data were analyzed using the free software environment R (version 2.2.0) with Harrell's "Design" package (14). Actuarial curves were estimated using the Kaplan-Meier method. Time to progression was calculated as the interval from the date of initial PET-CT scan to the date of the first finding based on imaging indicating local or distant disease progression that led to additional confirmatory testing (*e.g.*, biopsy or imaging). The Cox proportional hazards (CPH) model was used to evaluate prognostic variables in our study for univariate and multivariate prediction of freedom from progression (FFP); tests were based on the likelihood-ratio (LR) statistic. Prognostic factors analyzed included PET MTV, maximum and average SUV, and stage, treatment intent (definitive or palliative), age, Karnofsky performance status (KPS), and weight loss (<5% or ≥5% of baseline). We analyzed MTV, SUV, age, and KPS as continuous variables, whereas we analyzed stage, treatment intent, and weight loss as categorical variables in the CPH model. The proportional hazards assumption was tested with the "cox.zph" method (15) and the linearity assumption was tested by fitting cubic splines (neither assumption was rejected).

RESULTS

Patient characteristics

Medical records of these 19 patients were reviewed for patient age, gender, tumor histology, date of initial PET-CT scan, American Joint Committee on Cancer (AJCC) stage (TNM), treatment type, treatment intent, date of treatment, radiation dose and target volume, date of local recurrence, date of distant progression, date of last follow-up, KPS, and weight loss. These characteristics are summarized in Table 1.

Follow-up to the time of progression was complete for 18 of 19 patients. Only 1 of the 19 patients had no evidence of disease at the last follow-up (censoring time for progression: 6.6 months). There were no treatment-related or other non-cancer deaths as the first event. Nine of the 19 patients had their first site of progression locally, 8 progressed first at

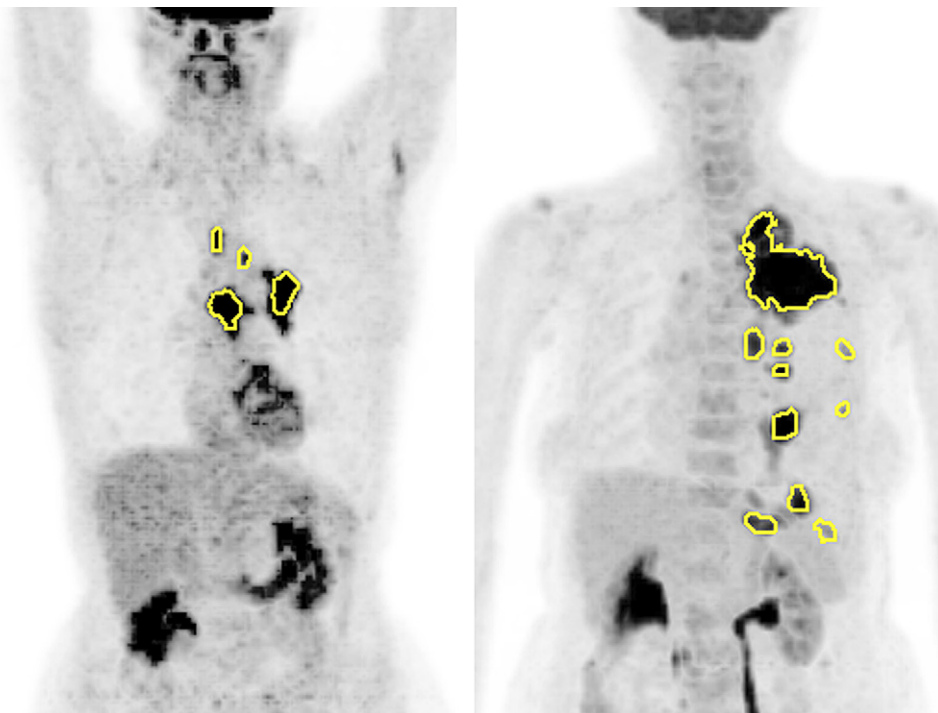


Fig. 1. Small and large metabolic tumor volume. Maximum intensity projection views of ^{18}F -fluorodeoxyglucose-positron emission tomography scans for two patients, with overlay of segmented metabolic tumor volumes (MTV). (Left:) Patient with Stage IIIA non-small-cell lung cancer (NSCLC) (T1N2M0) with small MTV (MTV: 14.7 ml, Progression free survival: 22.2 months). (Right:) A patient with Stage IIIB (T4N2M0 with malignant pleural nodules) NSCLC with large MTV (MTV: 106 ml, Progression free survival: 2.0 months).

Table 1. Patient characteristics

Parameter	No. of patients
Gender	
Male	7
Female	12
Histology	
Non-small-cell lung cancer	18
AJCC Stage	
Small-cell lung cancer	1
I	3
II	1
III	10
IV	5
Treatment	
Radiation therapy (20–66 Gy)	13
Surgery	7
Chemotherapy	13
Treatment intent	
Definitive	12
Palliative	7 (1 Stg II, 1 Stg III, 5 Stg IV)
KPS	
70	4
80	5
90	8
100	2
Weight loss ($\geq 5\%$ of baseline)	
Yes	4
No	15
Age	Median, 72 y; range, 49–86 y
Time to start of treatment (from PET-CT)	Median, 2.9 wk; 90th percentile, 6.9 wk; range, 0–14.8 wk

Abbreviations: AJCC = American Joint Committee on Cancer; KPS = Karnofsky performance status; PET-CT = positron emission tomography-computed tomography; Stg = Stage.

a distant site, and 1 progressed both locally and distantly. The estimated (Kaplan-Meier) median time to progression (TTP) and the 6 month FFP for the cohort were 9.3 months (range, 0.7–25 months) and 74%, respectively. Figure 2 shows the progression-free survival (PFS) and overall survival (OS) curves of the whole cohort, with 95% confidence bands. Follow-up for mortality was complete for all 19 patients, with 15 dead and 4 alive at the time of this analysis. The estimated (Kaplan-Meier) median OS and the 1-year OS for the cohort were 14.8 months (range, 1.2–38 months) and 68%, respectively. The median MTV for the cohort was 27 ml (range, 0.8–106 ml). The median values for maximum and average SUV were 14.8 (range, 6.5–30.5) and 4.8 (range, 0.8–12.6), respectively.

Prognostic value

On univariate analysis for FFP, MTV (LR = 9.8, $df = 1$, $p = 0.002$) had a significant effect on the hazard of progression. An increase in MTV of 25 ml (the difference between the 75th and 25th percentiles) was associated with a 2.8-fold increase in hazard of progression. Neither maximum nor average SUV, which were strongly correlated with each other, had a significant effect (LR = 2.57, $df = 1$, $p = 0.11$; and LR = 1.43, $df = 1$, $p = 0.23$, respectively). Patients with Stage IV disease had a 4.8-fold higher hazard than patients with Stage I to III (LR = 5.71, $df = 1$, $p = 0.017$); note that this result is based on a data-dependent pooling of the Stage I to Stage III patients, who had indistinguishable PFS. Similarly, patients treated with palliative intent had a 6.8-fold higher

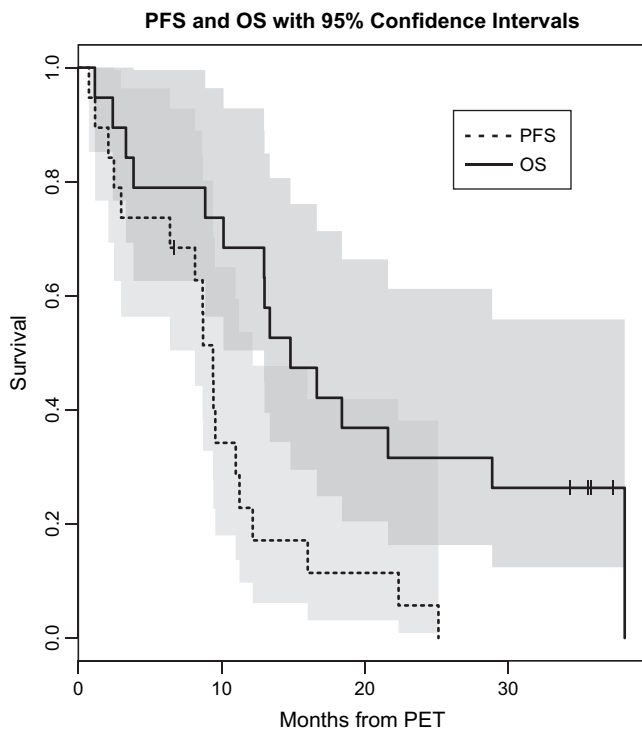


Fig. 2. Progression-free and overall survival. Progression-free survival (PFS, dashed line) and overall survival (OS, solid line) for all 19 patients and the 95% confidence intervals (gray shading). PET = positron emission tomography.

hazard than patients treated definitively (LR = 9.06, $df = 1$, $p = 0.0026$). Of note, all patients with Stage IV were treated with palliative intent, whereas only one patient with Stage II and one with Stage III were treated with palliative intent.

The effect of MTV remained significant (3.4-fold hazard increase across the quartiles, LR = 8.2, $df = 1$, $p = 0.004$) after controlling for stage and treatment intent, and even after controlling for all 5 clinical factors (stage, treatment intent, age, KPS, and weight loss, with a 5.4-fold hazard increase, LR = 10.15, $df = 1$, $p = 0.0014$). Figure 3 shows the relationship between PFS time (log scale) and MTV, with a superimposed locally weighted regression line (“lowess”). The censored time is indicated in the plot; because the unobserved value must lie above the plotted point it is evident that the relationship between MTV and PFS would only be strengthened by observing it exactly. The flexible regression line in Fig. 3 suggests that the relationship may be flatter in the lower range of MTV values. This can also be seen in Fig. 4, showing the PFS by tertiles of MTV, where it is evident that the lower and middle thirds do not differ appreciably.

The prognostic value of MTV for OS was similar. The hazard increase across the quartiles was 2.9 (LR = 9.94, $df = 1$, $p = 0.0016$, univariate analysis). Figure 5 shows OS by tertiles of MTV. Only stage and treatment intent were also predictive, and on multivariate analysis, the effect of MTV was maintained, correcting for stage and treatment intent (hazard increase 3.3 across the quartiles, LR = 6.98, $df = 1$, $p = 0.0083$) or for all five clinical variables (hazard increase 7.6 across the quartiles, LR = 10.3, $df = 1$, $p = 0.001$).

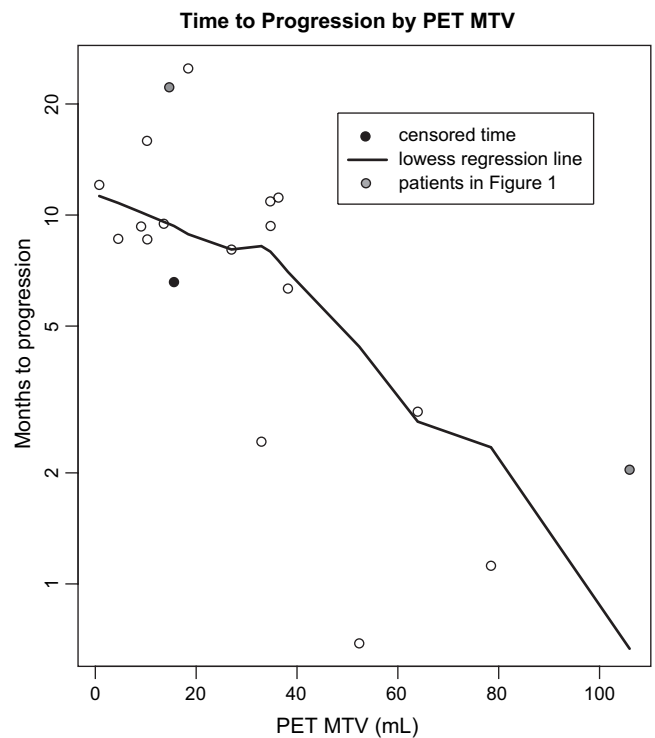


Fig. 3. Time to progression (TTP) (log scale) by metabolic tumor volume (MTV). The censored time is indicated in the plot (solid circle); since its unobserved value must lie above the plotted point it is evident that the inverse relationship between MTV and TTP would only be strengthened by observing it exactly. Also indicated are the data points corresponding to the patients in Fig. 1 (shaded circles).

We note that the number of events (progression or death) is small relative to the number of predictors considered in the multivariate models. Therefore we interpret the large hazard ratios and small p values for MTV in these analyses as indicating only that the prognostic effect of MTV is not accounted for by the other variables. We do not suggest that they are reliable adjusted estimates in and of themselves; stable estimates of the multivariate adjusted effects of MTV will require datasets with at least five times the number of events.

DISCUSSION

In this pilot study, we have demonstrated that an FDG-PET based measure of tumor burden, MTV, is highly prognostic for disease progression and death in lung cancer, independent of other established prognostic factors, namely stage, treatment intent, age, weight loss, and performance status. This is consistent with the hypothesis that some of these prognostic factors, especially stage, may simply be a surrogate for the underlying and more prognostically significant tumor burden. Indeed, a recent report demonstrated that although T-stage and stage grouping are highly correlated with tumor volume, when tumors were categorized into four groups based on tumor volume, there was only 55% concordance with T-stage, and 67% concordance with stage grouping, indicating a large range of tumor volumes within each stage and

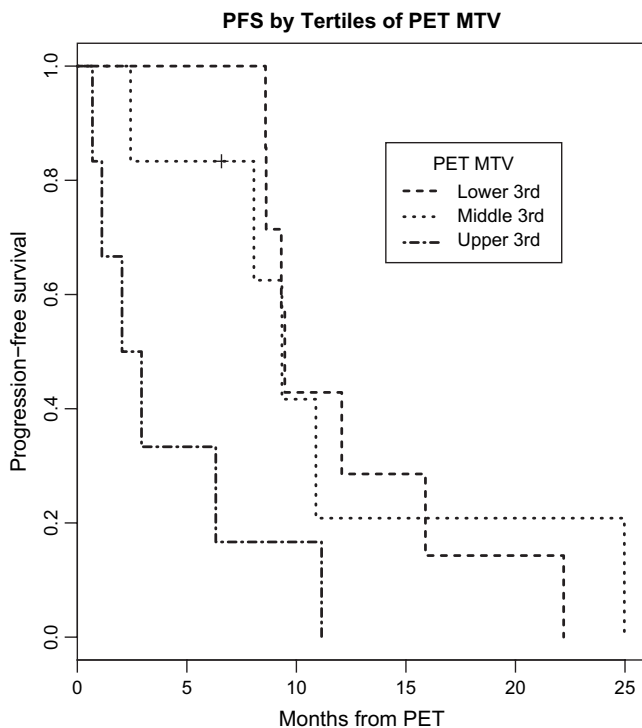


Fig. 4. Progression-free survival by tertiles of metabolic tumor volume (MTV). PET = positron emission tomography.

supporting that tumor volume may be an independent prognostic factor aside from stage (16).

Our results are consistent with those of Bradley *et al.*, who reported the prognostic value of gross tumor volume, determined on CT scans acquired as part of 3-D conformal radiation treatment planning, for overall and cause-specific survival as well as local control in patients with NSCLC (6). Our study differs from theirs in that our tumor volumes are derived from FDG-PET scans.

In contrast to some other studies, we failed to detect a significant prognostic effect of PET SUV parameters (maximum and average SUV). This may be because of the small sample size analyzed. For example, in a large study of 498 patients reported by Davies, *et al.*, the survival difference at 12 months between patients in the top and bottom quintiles of maximum SUV was only 18% (12), an effect that would have been too small for us to detect in this study given our limited number of patients and somewhat narrower range of maximum SUV values.

A key advantage of using FDG-PET in this context is the high tumor to background intensity ratio. This greatly facilitates computer-aided measurements such as the one used in this study. Because the user interaction is limited to visually identifying the hypermetabolic tumor foci, leaving the other-

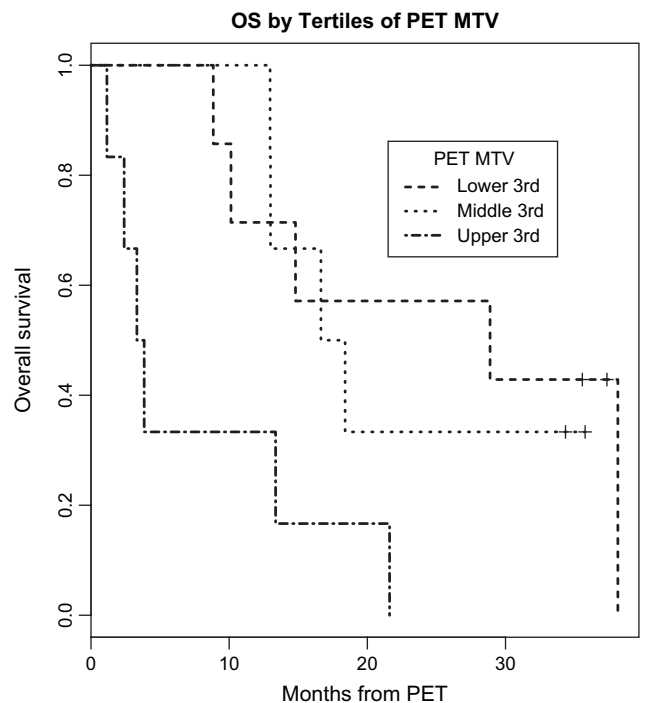


Fig. 5. Overall survival by tertiles of metabolic tumor volume (MTV). PET = positron emission tomography.

wise tedious segmentation steps to the automatic algorithm, tumor burden measurements can be made rapidly in a high throughput fashion and consistently with minimal interobserver variability. This overcomes the main limitations of current CT-based measurements with respect to broad applicability. Of note, the choice of the threshold intensity value used for PET segmentation can affect the absolute value of the volume measurements slightly, but not the consistency of the measurements as long as the same threshold is used. In addition, PET scans tend to be whole body studies, allowing a comprehensive assessment of tumor burden, with the exception of the brain.

The main limitations of our study are the relatively low number of patients in our cohort, the heterogeneity of the patients and treatments, and the retrospective design. However, despite these limitations, we were able to obtain highly significant results demonstrating a correlation between high MTV and disease progression. We plan to validate our findings in a larger patient cohort, as well as prospectively. A larger cohort will also better define the relationship in the lower range of MTV values. We anticipate that this method will be useful for stratifying patients in prospective clinical trials into high and low risk sub-populations within stage groups, and ultimately for selecting patients for risk-adapted therapies.

REFERENCES

1. Mountain CF. Revisions in the International System for Staging Lung Cancer. *Chest* 1997;111:1710-1717.
2. Adebajo SA, Bowser AN, Moritz DM, *et al.* Impact of revised stage classification of lung cancer on survival: A military experience. *Chest* 1999;115:1507-1513.
3. van Rens MT, de la Riviere AB, Elbers HR, *et al.* Prognostic assessment of 2,361 patients who underwent pulmonary resection for non-small cell lung cancer, stage I, II, and IIIA. *Chest* 2000; 117:374-379.

4. Brundage MD, Davies D, Mackillop WJ. Prognostic factors in non-small cell lung cancer: A decade of progress. *Chest* 2002;122:1037–1057.
5. Mery CM, Pappas AN, Burt BM, *et al.* Diameter of non-small cell lung cancer correlates with long-term survival: Implications for T stage. *Chest* 2005;128:3255–3260.
6. Bradley JD, Ieumwananonthachai N, Purdy JA, *et al.* Gross tumor volume, critical prognostic factor in patients treated with three-dimensional conformal radiation therapy for non-small-cell lung carcinoma. *Int J Radiat Oncol Biol Phys* 2002;52:49–57.
7. Ciernik IF, Dizendorf E, Baumert BG, *et al.* Radiation treatment planning with an integrated positron emission and computer tomography (PET/CT): A feasibility study. *Int J Radiat Oncol Biol Phys* 2003;57:853–863.
8. Bradley JD, Thorstad WL, Mutic S, *et al.* Impact of FDG-PET on radiation therapy volume delineation in non-small-cell lung cancer. *Int J Radiat Oncol Biol Phys* 2004;59:78–86.
9. Cerfolio RJ, Bryant AS, Ohja B, *et al.* The maximum standardized uptake values on positron emission tomography of a non-small cell lung cancer predict stage, recurrence, and survival. *J Thorac Cardiovasc Surg* 2005;130:151–159.
10. Sasaki R, Komaki R, Macapinlac H, *et al.* [18F]fluorodeoxyglucose uptake by positron emission tomography predicts outcome of non-small-cell lung cancer. *J Clin Oncol* 2005;23:1136–1143.
11. Eschmann SM, Friedel G, Paulsen F, *et al.* Is standardised (18)F-FDG uptake value an outcome predictor in patients with stage III non-small cell lung cancer? *Eur J Nucl Med Mol Imaging* 2006;33:263–269.
12. Davies A, Tan C, Paschalides C, *et al.* FDG-PET maximum standardised uptake value is associated with variation in survival: Analysis of 498 lung cancer patients. *Lung Cancer* 2007;55:75–78.
13. Graves EE, Quon A, Loo BW Jr. RT Image: An open-source tool for investigating PET in radiation oncology. *Technol Cancer Res Treat* 2007;6:111–122.
14. R Development Core Team. R: A language and environment for statistical computing. Vienna: R Foundation for Statistical Computing; 2006.
15. Grambsch PM, Therneau TM. Proportional hazards tests and diagnostics based on weighted residuals. *Biometrika* 1994;81:515–526.
16. Ball D, Burmeister B, Graham P, *et al.* Stage is not a reliable indicator of tumor volume in non-small cell lung cancer: A preliminary analysis of the Trans-Tasman Radiation Oncology Group 99-05 Database. *J Thorac Oncol* 2006;1:667–672.