Int J CARS (2013) 8:181–191
DOI 10.1007/s11548-012-0749-7

ORIGINAL ARTICLE

Independent prognostic value of whole-body metabolic tumor burden from FDG-PET in non-small cell lung cancer

Hao Zhang · Kristen Wroblewski · Daniel Appelbaum · Yonglin Pu

Received: 29 January 2012 / Accepted: 7 May 2012 / Published online: 30 May 2012

Abstract *Purpose* To determine whether whole-body metabolic tumor burden, measured as either metabolic tumor volume (MTV_{WB}) or total lesion glycolysis (TLG_{WB}), using FDG-PET/CT is an independent prognostic marker in non-small cell lung cancer (NSCLC).
Methods 328 patients with histologically proven NSCLC were identified for this retrospective analysis. This study was approved by our Institutional Review Board. All patients underwent baseline ^{18}F-FDG-PET/CT scan imaging before therapy. The MTV_{WB}, TLG_{WB}, maximum standardized uptake value (SUV_{maxWB}) and mean standardized uptake value (SUV_{meanWB}) of tumors throughout the whole body were measured from FDG-PET images with semi-automated 3D contouring software.
Results In univariate analysis, there was a statistically significant association of overall survival (OS) with the MTV_{WB} (hazard ratio (HR) = 1.62, $p < 0.001$), TLG_{WB} (HR = 1.47, $p < 0.001$). The patients with a $MTV_{WB} \leq$ median of 65.7 ml and $TLG_{WB} \leq$ median of 205.11 $SUV_{mean} * ml$ had a median OS of 41.1 and 35.4 months compared with 9.5 and 9.7 months for those with a $MTV_{WB} > 65.7$ ml and $TLG_{WB} > 205.11\, SUV_{mean} * ml$, respectively. From a series of multivariate Cox regression models, the MTV_{WB} and TLG_{WB} were significantly better than SUV_{maxWB} and SUV_{meanWB} at prognostication and significantly associated with patients' OS with HRs of 1.50 ($p < 0.001$) and 1.42 ($p < 0.001$), respectively, after adjustment for patient's age, gender and treatment intent as well as the tumor SUV_{maxWB}, histology and stage.
Conclusions MTV_{WB} and TLG_{WB} as metabolic tumor burden measurements in ^{18}F-FDG-PET/CT are independent prognostic markers and are significantly better than SUV_{maxWB} and SUV_{meanWB} at prognostication.

Keywords Non-small cell lung cancer · Tumor burden · F-18 Fluorodeoxyglucose (FDG) · Positron emission tomography (PET)

Introduction

Lung cancer is one of the most common malignant tumors in the world and the leading cause of cancer related death for both men and women [1]. Studies on ^{18}F-FDG-PET/CT have shown its great value in the initial staging of non-small cell lung cancer (NSCLC) and in the later evaluation of the response to treatment [2]. Standardized uptake value (SUV) measurements are commonly used in clinical practice. However, they are affected by many patient-dependent and technical factors, such as a patient's body habitus, blood glucose level, length of FDG uptake period, partial volume effects, definition of region of interest (ROI), image reconstruction method and resolution [3–5]. The metabolic tumor burden measurements including metabolic tumor volume (MTV) [6,7] and total lesion glycolysis (TLG) of tumors [8] have been developed because they incorporate both metabolic activity and tumor volume. The MTV is the tumor volume measured with a segmentation technique on PET scans [6], while TLG can be calculated by multiplying the

H. Zhang
Department of Radiology, First Hospital of Lanzhou University, Lanzhou, Gansu 730000, China

H. Zhang · D. Appelbaum · Y. Pu (✉)
Department of Radiology, University of Chicago, MC 2026, Chicago, IL 60637, USA
e-mail: ypu@radiology.bsd.uchicago.edu

K. Wroblewski
Department of Health Studies, The University of Chicago, Chicago, IL 60637, USA

mean standardized uptake value (SUV_{mean}) by the MTV [8]. The MTV_{WB} has been studied in NSCLC using two patient groups of relatively small size [6,9,10]. Lee et al. [6] found that the baseline MTV_{WB} measured semi-automatically is a statistically significant prognostic index and better than SUV_{max} and SUV_{mean} in the prediction of patient outcome in 19 lung cancer patients. In their recent study, they expanded the cohort to 61 patients with NSCLC and confirmed the significant association of high MTV with decreased overall survival (OS) and progression-free survival in patients who received definitive treatment, but not in the entire cohort of the study who received either definitive or palliative treatment [9]. We have recently found that the baseline metabolic tumor burden as measured with MTV and TLG on FDG-PET is a prognostic measure independent of clinical stage in NSCLC [10,11]. However, in our prior study in non-surgical patients with NSCLC, only the stage of the disease was adjusted for in the multivariate analysis, since it is considered as the most important prognosticator and is widely used by oncologists to estimate prognosis and to choose the most suitable therapy [12]. Other clinical and pathologic factors such as age, gender, performance status and treatment intent of the patients, as well as the histology of the tumor [13], have also been shown to be associated with patients' survival probability [14,15]. Some studies have demonstrated that a high tumor SUV_{max} value is also associated with a poorer survival in NSCLC [16–19]. In the current study we performed survival analysis to include other prognostic factors such as age, gender, and treatment intent of the patients, as well as the histology and SUV measurement of the tumor, in addition to the stage of the disease, to test our hypothesis that metabolic tumor burden as measured with MTV_{WB} and TLG_{WB} is an independent prognostic marker using a large cohort of 328 consecutive patients with NSCLC. In this study we also used the updated Union Internacional Contra la Cancrum (UICC)/American Joint Committee on Cancer (AJCC) staging system for NSCLC (7th edition) [20] in the data analysis since it is now a widely used TNM staging system.

Materials and methods

Patients

This study was approved by our hospital's Institutional Review Board and was compliant with the Health Insurance Portability and Accountability Act. We conducted a retrospective review of the medical records of patients with NSCLC. There were a total of 1,023 cases with NSCLC who were diagnosed and treated in our medical center from January 1, 2004 to December 22, 2007. A total of 328 patients (including surgical and non-surgical patients) with NSCLC treated at the University of Chicago Medical Center (UCMC) were identified for this retrospective study. The inclusion criteria were as follows: 1) all patients underwent a baseline PET/CT scan and had PET-positive tumor(s), 2) they had no known brain metastasis (since our whole-body PET/CT did not cover the whole brain), and 3) they had no history or concurrent diagnosis of another type of cancer. The PET/CT imaging was performed in these patients for initial staging or for the diagnosis of the lung nodule.

These patients were followed semiannually by the Cancer Registry of our medical center. Their survival status was determined through clinical follow-up and the Social Security Death Index. Clinical follow-up and the Illinois State Death Inquiry System were used to determine the cause of death.

PET/CT imaging protocols

The ^{18}F-FDG-PET/CT images were obtained using a PET/CT scanner (Reveal HD, CTI, Knoxville, TN, USA) equipped with high-resolution bismuth germanate detectors and a dual-slice CT scanner in all 328 patients before therapy. The ^{18}F-FDG-PET scans were performed in accordance with National Cancer Institute guidelines [21]. All patients fasted for at least 4 h before intravenous administration of 370–555 MBq of ^{18}F-FDG. The plasma glucose levels of all patients were less than 200 mg/dl before FDG administration. A standard protocol was used for the CT images. Ninety minutes $\pm$ 30 min following injection of the ^{18}F-FDG, a whole-body static PET scan was acquired for about 30 to 35 min, starting at the thighs and proceeding to the skull base. PET scans were obtained with an acquisition time of 3–5 min per cradle position, with a 26.6 % axial overlap at the borders of the field of view to avoid artifacts. PET images were reconstructed using ordered-subsets expectation maximization (OSEM) iterative algorithms with 8 subsets, 2 iterations, and 128 $\times$ 128 pixels. The slice thickness was 2.4 mm, the pixel size within transverse slices was 5.2 mm, with 5 mm full-width at half-maximum (FWHM) three-dimensional (3D) Gaussian smoothing after reconstruction. We used the 3D imaging mode with Fourier rebinning and analytic scatter correction. Monthly concentration calibrations were conducted using a Ge-68 tub phantom or ^{18}F-FDG tub phantom.

Measurements from PET/CT scans

The MTV_{WB}, TLG_{WB}, SUV_{maxWB} and SUV_{meanWB} of whole-body tumors were measured using the commercially available PET Edge tool of the MIMvista software (version 5.1.2, MIMvista Corp, Cleveland, OH, USA) by two board-certified radiologists with PET/CT imaging experience. The software used a gradient-based tumor segmentation method. In comparison with manual and constant threshold methods

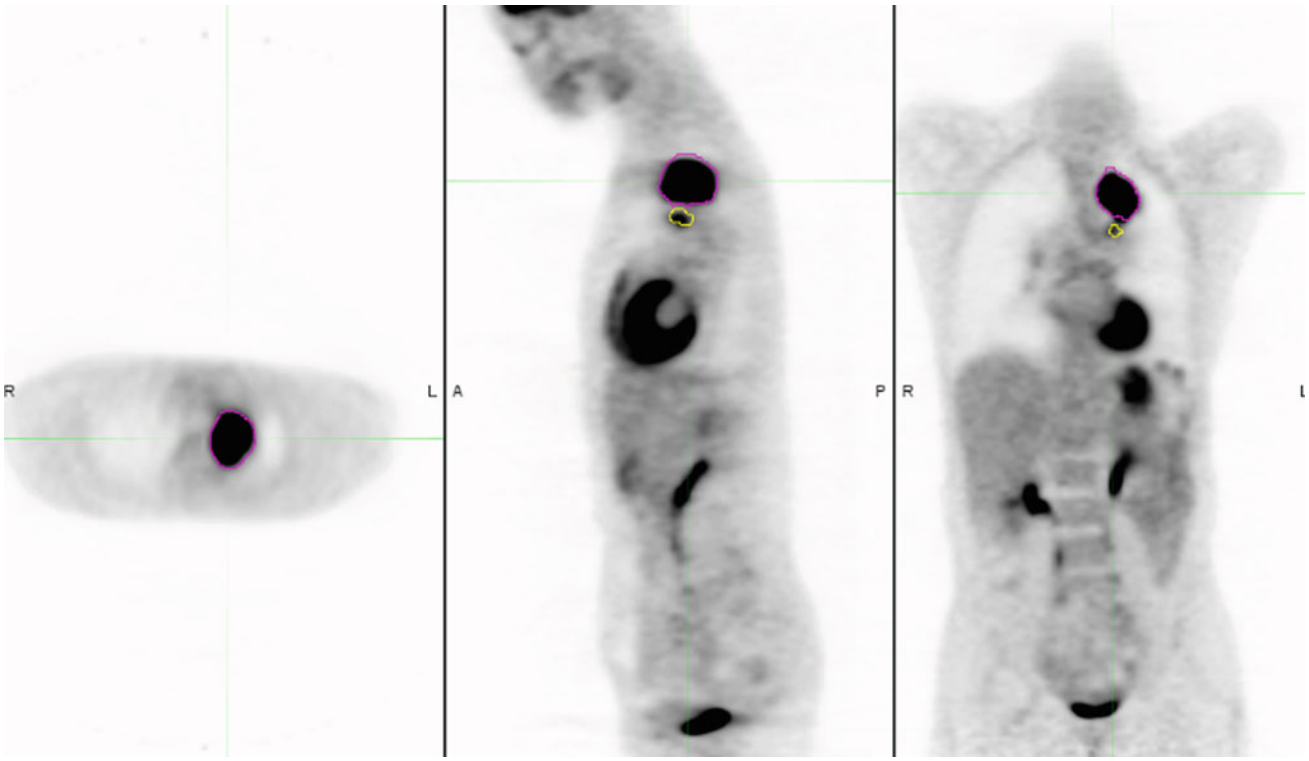

Fig. 1 Axial, sagittal and coronal images from a PET scan of an 56-year-old female with a new diagnosis of non-small cell lung cancer, showing the tumors outlined with the aid of the MIMvista PET Edge tool for the PET measurements

in a phantom study, the gradient-based method was the most accurate and consistent technique for tumor segmentation, having less inter-observer variation and being the most robust for varying imaging conditions [22].

The values of the PET measurements were determined based on reader consensus, which was described in detail in the prior studies (Fig. 1) [10,11]. With the MIMvista PET Edge tool, the radiologists indicated the approximate center of the tumor. The volume of interest (VOI) was drawn automatically after the radiologists had identified the major and minor axes of the tumor in one plane. The software then automatically measured the SUV_{max}, SUV_{mean}, MTV and TLG of a tumor. Adjustment of the estimated tumor surface was sometimes needed to include the entire tumor within the margins of the VOI. This was done visually by the readers using the 2D or 3D ball tool in the MIMvista contouring software. The resultant SUV, MTV and TLG were noted (outputted to an Excel spreadsheet). In this study, the SUV_{max} was defined as maximum activity concentration of FDG in the tumor/(injected FDG dose/body weight) [23]. The SUV_{mean} was defined as the mean concentration of FDG in the tumor/(injected FDG dose/body weight). The whole-body SUV_{max} (SUV_{maxWB}) was the maximum SUV_{max} of all the tumors in the whole body. The MTV_{WB} was the combined MTV of all the tumors in the whole body. The TLG_{WB} was the sum of all the TLG values from all the tumors in the body. The whole-body SUV_{mean} (SUV_{meanWB}) was the mean SUV of all the tumors in the whole body which was calculated by the TLG_{WB} divided by MTV_{WB}.

The (UICC)/(AJCC) staging system for NSCLC (7th edition) was used for staging patients [20]. The clinical stage of the disease was determined by clinical history, physical examination and findings from the contrast infused CT of the chest and abdomen and of the whole-body PET/CT in reference with the original radiology reports on our PACS. Brain MRI was done if clinical symptoms suggested brain metastasis.

Statistical analysis

OS served as the primary endpoint of the study. The OS time was calculated from the date of the initial baseline PET/CT scan to the date the patients died from any cause based on the follow-up and records described above. The patients last known to be alive were censored at the date of last contact.

Nine variables including patient's age, gender, and treatment intent as well as the clinical stage, histology type, SUV_{maxWB}, SUV_{meanWB}, MTV_{WB} and TLG_{WB} of the tumors were used for statistical analysis. Continuous variables including age, SUV_{maxWB}, SUV_{meanWB}, MTV_{WB} and TLG_{WB} were tested for normality and skewness. Categorical variables included patient's gender, and treatment intent

Table 1 Patient and tumor characteristics in 328 patients with NSCLC

Variable	Number	%
Male/female	156/172	47.6/52.4
TNM stage		
Stage IA/IB	47/43	14.3/13.1
Stage IIA/IIB	19/18	5. 8/5.5
Stage IIIA/IIIB	50/40	15.2/12.2
Stage IV	111	33.8
Histology types		
Adenocarcinoma	129	39.3
Squamous cell carcinoma	92	28.1
Large cell carcinoma	21	6.4
Not otherwise specified	78	23.8
Other	8	2.4
No tumor treatment	37	11.3
With tumor treatment	291	88.7
Non-surgery	180	54.9
Radiotherapy alone	28	8.5
Chemotherapy alone	55	16.8
Chemo-radiation	97	29.6
Surgery	111	33.8
Surgery alone	61	18.6
Surgery with radiotherapy	4	1.2
Surgery with chemotherapy	32	9.8
Surgery with chemo-radiation	14	4.3

(see Table 1) as well as the tumor stage (I–IV) and histology type (adenocarcinoma, squamous cell carcinoma, large cell carcinoma, NSCLC not otherwise specified, and other). Natural logarithmic transformations were applied to obtain more normally distributed data for the continuous variables of SUV_{maxWB}, SUV_{meanWB}, MTV_{WB} and TLG_{WB}.

Survival differences between groups were examined and tested using the Kaplan–Meier method and the log-rank test. For descriptive purposes, each PET measurement was dichotomized using the median value as the cutoff point (i.e., $\leq$ median and >median). Univariate and multivariate Cox proportional hazards regression models [24] were used with HRs and their 95 % confidence intervals (CIs) reported. The proportional hazards assumption was tested using Schoenfeld residuals [25]. All variables were included in a series of multivariate Cox proportional hazards regression models. Multicollinearity was assessed using variance inflation factors, and the functional form of the continuous variables was examined using martingale residuals. A C-statistic index (Gönen and Heller's K concordance statistic) was used to assess the predictive performance of the models [26]. To compare C-statistic indices, a z test was calculated based on 500 bootstrap replications. For illustrative purposes, the Kaplan-Meier curves were constructed after creating four roughly equal sized groups using quartiles of each PET measurement. A p value less than 0.05 was considered statistically significant. STATA 12.1 was used for statistical analysis.

Results

Patient and tumor characteristics

The patient and tumor characteristics are summarized in Table 1. There were 156 male and 172 female patients with a median age of 68.3 years (range, 30.0–89.9 years old) in this study. The 328 patients included 129 patients with adenocarcinoma, 92 patients with squamous cell carcinoma, 21 patients with large cell carcinoma, 78 patients with NSCLC of a type that was not further specified, and 8 patients with other types (1 with atypical carcinoid tumor, 1 with mucinous adenocarcinoma, 2 with bronchoalveolar adenocarcinoma, 3 with synchronous adenocarcinoma and squamous cell carcinoma, 1 with synchronous adenocarcinoma and large cell carcinoma).

There were 90 cases with stage I, 37 cases with stage II, 90 with stage III, and 111 with stage IV NSCLC. The mean time from the PET/CT scan to therapy was 11 weeks and 5 weeks for surgical cases and non-surgical cases, respectively. 111 patients underwent surgery. Among those, nine patients had pneumonectomy, 84 patients had lobectomy, 13 patients had resection involving less than entire lobe and 5 had exploratory thoracotomy. Thirty-seven patients had received no cancer-specific treatment, 230 patients had died (70.1 %). Median follow-up time among survivors was 38.4 months (range of 1.5–82.2 months).

Univariate analyses

The results of univariate survival analysis are presented in Table 2. In this study, the survival rates in the entire cohort at 2 years and 5 years were 42.1 and 23.0 %, respectively. Median OS was 16.5 months (95 % CI 14.0–20.4 months).

There was a statistically significant association of better survival with lower stage (Fig. 2), being treated surgically and lower levels of SUV_{max}, SUV_{mean}, MTV_{WB} and TLG_{WB}. In addition, patients with NSCLC of a type that was not further specified had a significantly worse survival than other groups. In the 78 patients with not otherwise specified NSCLC, 17 patients received chemotherapy only, 9 patients received radiation therapy only, 37 patients received both chemotherapy and radiation therapy, 2 received surgical therapy in addition to chemotherapy and/or radiation, and 13 patients had no cancer specific therapy. There was no significant association with age or gender. Patients with an $MTV_{WB} \leq$ median of 65.7 (4.19 on log scale) ml and $TLG_{WB} \leq$ median of 205.11 (5.32 on log scale) $SUV_{mean} *$ ml had

Table 2 Univariate survival analysis in 328 patients with non-small cell lung cancer

Variables	*N*	Survival (months or %)			Log-rank test *p*	Univariate analysis	
		Median	2-year	5-year		HR (95 % CI)	*p*
Gender							
Female	172	19.38	45.60	25.81	0.064	(Reference)	
Male	156	13.93	38.28	19.71		1.28 (0.98, 1.65)	0.065
Age							
≤68.3 (median)	164	15.87	40.72	21.48	0.904	(Reference)	
>68.3	164	18.85	43.65	25.38		0.98 (0.76, 1.27)	0.904
Stage							
I	90	64.69	75.64	52.94	<0.001	(Reference)	
II	37	27.80	54.96	38.17		1.78 (1.04, 3.02)	0.034
III	90	13.64	32.09	10.53		3.46 (2.32, 5.18)	<0.001
IV	111	8.56	18.54	4.41		5.75 (3.90, 8.49)	<0.001
Treatment							
No specific treatment	37	6.62	18.92	6.31	<0.001	(Reference)	
Non-surgery therapy	180	11.93	26.88	5.91		0.71 (0.49, 1.04)	0.077
Surgery	111	65.84	75.34	55.60		0.16 (0.10, 0.26)	<0.001
Histology types							
Adenocarcinoma	129	19.61	47.85	29.52	<0.001	(Reference)	
Squamous cell CA	92	16.52	42.57	23.44		1.19 (0.86, 1.64)	0.301
Large cell CA	21	20.46	49.21	25.84		0.99 (0.55, 1.78)	0.985
Not otherwise specified	78	10.16	27.73	9.83		1.88 (1.36, 2.61)	<0.001
Other	8	44.66	71.43	42.86		0.56 (0.20,1.53)	0.257
$\ln(SUV_{maxWB})$						1.81 (1.49, 2.21)	<0.001
≤2.22(median)	164	26.49	53.12	31.31	<0.001	(Reference)	
>2.22	164	12.13	31.25	14.49		1.75 (1.35,2.27)	<0.001
$\ln(SUV_{meanWB})$						1.79 (1.37,2.33)	<0.001
≤1.27(median)	164	21.57	48.82	28.70	0.013	(Reference)	
>1.27	164	15.18	35.50	16.92		1.39 (1.07, 1.80)	0.013
$\ln(MTV_{WB})$						1.62 (1.48, 1.78)	<0.001
≤4.19(median)	164	41.15	65.67	38.04	<0.001	(Reference)	
>4.19	164	9.51	18.69	8.21		3.14 (2.39, 4.11)	<0.001
$\ln(TLG_{WB})$						1.47 (1.36, 1.59)	<0.001
≤5.32(median)	164	35.44	63.89	36.63	<0.001	(Reference)	
>5.32	164	9.67	20.36	9.61		2.83 (2.17, 3.71)	<0.001

CI Confidence interval, *HR* hazard ratio, *ln* natural log, *N* number of patients. The median of $MTV_{WB} = 65.7$, the median of $SUV_{maxWB} = 9.20$, the median of $SUV_{meanWB} = 3.55$, the median of $TLG_{WB} = 205.11$

a median OS of 41.1 and 35.4 months compared with 9.5 and 9.7 months for those with a $MTV_{WB} > 65.7$ ml and $TLG_{WB} > 205.11$ $SUV_{mean} *$ ml, respectively. See Fig. 3 for Kaplan–Meier curves based on quartiles of SUV_{maxWB}, SUV_{meanWB}, MTV_{WB} and TLG_{WB}. In univariate Cox regression models, $\ln(SUV_{maxWB})$, $\ln(SUV_{meanWB})$, $\ln(MTV_{WB})$ and $\ln(TLG_{WB})$ showed statistically significant prognostic value ($p < 0.001$). The HRs for a one-unit increase of $\ln(SUV_{maxWB})$, $\ln(SUV_{meanWB})$, $\ln(MTV_{WB})$ and $\ln(TLG_{WB})$ were 1.81 (95 % CI 1.49, 2.21), 1.79 (95 % CI 1.37, 2.33), 1.62 (95 % CI 1.48, 1.78), 1.47 (95 % CI 1.36, 1.59), respectively.

Multivariate analyses

The $\ln(SUV_{maxWB})$, $\ln(SUV_{meanWB})$, $\ln(MTV_{WB})$, $\ln(TLG_{WB})$, patient's age, gender, treatment intent, tumor stage, and histology were included in a series of multivariate models to evaluate their joint effect on OS (Table 3). In Model 1, the variables of $\ln(SUV_{maxWB})$, patient's age,

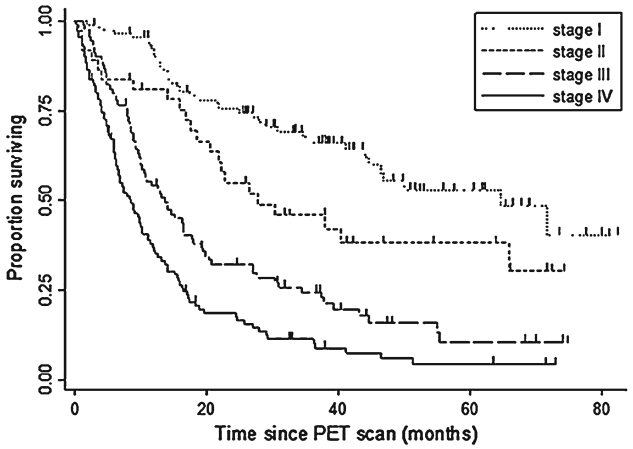

Fig. 2 Kaplan–Meier curves of overall survival after baseline PET/CT grouped according to the clinical stages in 328 patients with stage I–IV non-small cell lung cancer (NSCLC). The *dotted line*, *short dashed line*, *long dashed line* and *solid line* indicate the survival curve of the groups with stages I, II, III and IV NSCLC, respectively

gender, treatment intent, tumor stage and histology were included. In Model 2, the variables of $\ln(SUV_{meanWB})$, patient's age, gender, treatment intent, tumor stage and histology were included. In Model 3, the variables of $\ln(MTV_{WB})$, patient's age, gender, treatment intent, tumor stage and histology were included. In Model 4, the variables of $\ln(TLG_{WB})$, patient's age, gender, treatment intent, tumor stage and histology were included. In Models 1–4, the $\ln(SUV_{maxWB})$, $\ln(SUV_{meanWB})$, $\ln(MTV_{WB})$, or $\ln(TLG_{WB})$ remained as statistically significant prognostic markers for survival with HRs of 1.41 (95 % CI 1.10–1.79, $p = 0.006$), 1.47 (95 % CI 1.04–2.08, $p = 0.028$), 1.49 (95 % CI 1.30–1.70, $p < 0.001$) and 1.36 (95 % CI 1.23–1.51, $p < 0.001$), respectively, after adjusting for the patient's age, gender, treatment intent, tumor stage and histology. The *C*-statistic index (a measure of discriminatory power) for the models using $\ln(SUV_{maxWB})$, $\ln(SUV_{meanWB})$, $\ln(MTV_{WB})$, and $\ln(TLG_{WB})$ separately were 0.73, 0.73, 0.75 and 0.75, respectively. In Model 5,

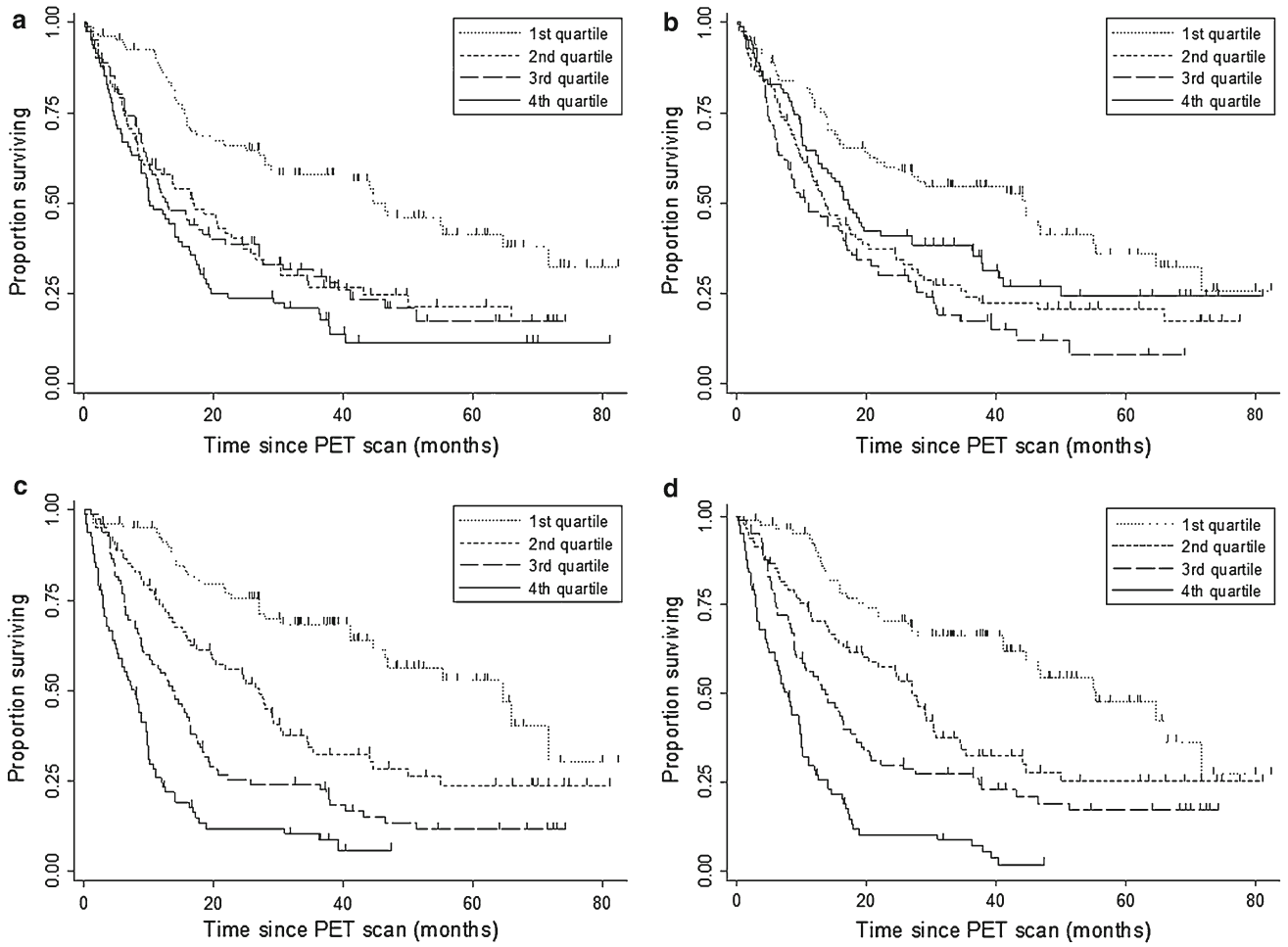

Fig. 3 Kaplan–Meier curves of overall survival after baseline PET/CT grouped according to PET measurements in 328 patients with non-small cell lung cancer (NSCLC). The *dotted line*, *short dashed line*, *long dashed line* and *solid line* indicate the survival curves of the groups with measurements in the 1st, 2nd, 3rd and 4th quartiles. SUV_{maxWB} (**a**), SUV_{meanWB} (**b**), MTV_{WB} (**c**) and TLG_{WB} (**d**)

Table 3 Multivariate Cox regression models for the overall survival analysis of PET measurement of whole-body tumor in NSCLC

Models	Multivariate analysis*		
	Hazard ratio (95 % CI)	p	C-statistic
Model 1			
$\ln(SUV_{maxWB})$	1.41 (1.10, 1.79)	0.006	0.73
Model 2			
$\ln(SUV_{meanWB})$	1.47 (1.04, 2.08)	0.028	0.73
Model 3			
$\ln(MTV_{WB})$	1.49 (1.30, 1.70)	<0.001	0.75**
Model 4			
$\ln(TLG_{WB})$	1.36 (1.23, 1.51)	<0.001	0.75**†
Model 5			
$\ln(SUV_{maxWB})$	0.96 (0.72, 1.27)	0.756	0.75**
$\ln(MTV_{WB})$	1.50 (1.29, 1.75)	<0.001	
Model 6			
$\ln(SUV_{maxWB})$	0.85 (0.62, 1.16)	0.304	0.75**
$\ln(TLG_{WB})$	1.42 (1.24, 1.61)	<0.001	

C-statistic = Gönen and Heller's K concordance statistic. Other abbreviations are as in Table 2

* Models 1–4 are adjusted for patient's age, gender and treatment intent as well as the tumor histology and stage. In Model 5, $\ln(MTV_{WB})$ is adjusted for $\ln(SUV_{maxWB})$ in addition to the factors adjusted for in Model 3. In Model 6, $\ln(TLG_{WB})$ is adjusted for $\ln(SUV_{maxWB})$ in addition to the factors adjusted for in Model 4

** $p < 0.01$ for comparison with C-statistic from Model 1

† $p = 0.877$ for comparison with C-statistic from Model 3

including variables $\ln(SUV_{maxWB})$, $\ln(MTV_{WB})$, patient's age, gender, treatment intent, tumor stage and histology, $\ln(MTV_{WB})$ was still statistically significantly associated with OS with a HR of 1.50 (95 % CI, 1.29–1.75, $p < 0.001$). In Model 6, including variables $\ln(SUV_{maxWB})$, $\ln(TLG_{WB})$, patient's age, gender, treatment intent, tumor stage and histology, $\ln(TLG_{WB})$ was still statistically significantly associated with OS with a HR of 1.42 (95 % CI, 1.24–1.61, $p < 0.001$). In Models 5 and 6, there was no statistically significant association of OS with $\ln(SUV_{maxWB})$. The C-statistic indices for Models 5 and 6 were both 0.75. The C-statistic indices were significantly higher for Models 3–6 than for Model 1 (all $p < 0.01$).

Kaplan–Meier curves of OS after baseline PET/CT grouped according to PET MTV_{WB} measurement within each disease stage group also showed the association of MTV_{WB} with OS independent of stage, in stage II, III or IV (all stage I patients had MTV_{WB} below the median) (Fig. 4). Furthermore, Kaplan–Meier curves of OS after baseline PET/CT grouped according to PET MTV_{WB} measurement within two groups with SUV_{maxWB} equal and below or above the median showed the association of MTV with OS independent of SUV_{maxWB} (Fig. 5). The above analyses was not performed for TLG since there was a statistically significant linear correlation between the $\ln MTV_{WB}$ and $\ln TLG_{WB}$ ($r = 0.96$, $p < 0.001$).

Discussion

Currently, ^{18}F-FDG-PET/CT, as an anatomic and metabolic imaging modality, is considered the best non-invasive tool for diagnosing, staging and treatment monitoring for lung cancer [2]. It is interpreted by nuclear radiologists visually. Semi-quantitative measurements, such as SUV measurement, have been used in clinical practice. The MTV and TLG have been used to measure the whole-body metabolic tumor burden [6,8]. Our previous study [10] suggested that the MTV_{WB} and TLG_{WB} are prognostic markers independent of clinical tumor TNM stage, as defined by the UICC/ AJCC staging system for NSCLC, 6th edition.

Our current study, using a population of 328 patients with NSCLC, demonstrates that metabolic tumor burden measurement as measured with MTV_{WB} and TLG_{WB} from FDG–PET is a prognostic marker, independent of patient's gender, age and treatment intent, as well as the TNM stage (UICC/AJCC staging system for NSCLC, 7th edition), SUV_{maxWB} and histology of the tumor. A possible explanation for the independence of metabolic tumor burden measurement on TNM staging is that the current TNM staging system is based on the resectability of the tumor and a crude surrogate for quantitative volumetric tumor measurement, like dimensional size of primary tumor. In the TNM classification system, the T descriptor provides details regarding primary tumor characteristics among which dimensional size is only loosely correlated with primary tumor volume. The N descriptor in the TNM staging system describes the extent of regional lymph node involvement. However, the term N2 encompasses a spectrum of disease from a micrometastatic deposit in one node to extranodal extension from metastatic deposits in several lymph node stations. Finally, the M descriptor describes the presence or absence of intrathoracic or distant metastases. However, the terms M1a and M1b are used to separate intrathoracic from more advanced extra-thoracic metastasis. In general, the more advanced the disease stage, the less "quantitatively accurate" the TNM classification system turns out to be. There is no contribution of more accurate and more advanced quantitative volumetric tumor measurement, such as MTV_{WB} and TLG_{WB}, even though almost all current imaging modalities used for tumor staging routinely provide 3-dimensional quantitative volumetric imaging. Adding volumetric tumor burden measurement may make NSCLC staging more quantitative and more complete. Tumor histology is related to its biological behavior and may affect the SUV measurement and metabolic tumor burden measurement. However, most tumor types in NSCLC are hypermetabolic with high SUV measurements,

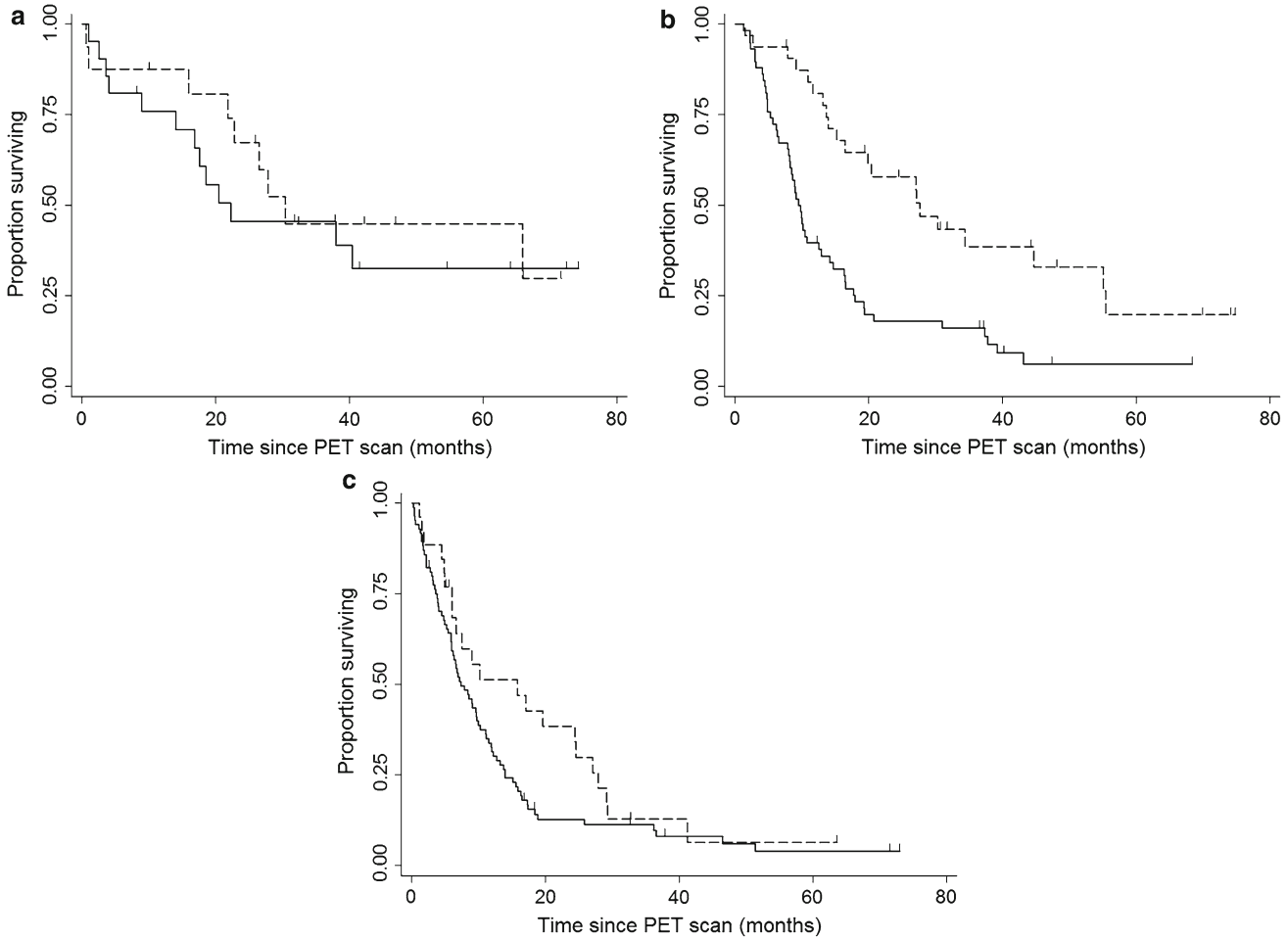

Fig. 4 Kaplan–Meier curves of overall survival after baseline PET/CT grouped according to PET MTV_{WB} measurement within each disease stage group, stage II (**a**), stage III (**b**), stage IV (**c**), showing the association of metabolic tumor volume with overall survival independent of stage. The *dashed lines* indicate the group with MTV_{WB} values equal to or less than the median, and the *solid lines* are the group with values greater than the median of the PET MTV_{WB} measurement

with the exception of adenocarcinoma in situ and minimally invasive adenocarcinoma, formerly called bronchoalveolar carcinoma (BAC). These tumors normally have lower metabolic rates and may not have high FDG activity. The current study included two cases with BAC. The potential clinical implication of the prognostic value of metabolic tumor burden measurement with MTV_{WB} and TLG_{WB} is to divide the patients with NSCLC into subgroups with different prognoses. Since accurate risk stratification in NSCLC is important for determining treatment options and prognosis, incorporating metabolic tumor burden measurement may 1) improve patient treatment selection, 2) aid medical and personal decision making by both clinicians and patients, 3) help clinicians in their routine patient outcome predictions, and 4) assist in the selection of comparable risk patients in clinical trials. On the last point, placing patients predicted to have similar risk factors into the test and control groups using stratified randomization could help improve clinical trials by obtaining more comparable and meaningful results. Incorporating metabolic tumor burden in the clinical staging in NSCLC may potentially impact on selecting patients for adjuvant or even neoadjuvant treatment since a greater metabolic tumor burden is associated with a poor survival outcome. In addition, the metabolic tumor burden could theoretically, be measured on both baseline and follow-up PET/CT studies. It could therefore impact the assessment of tumor response to chemoradiation therapy.

The UICC /AJCC staging system for NSCLC (7th edition) has become popular and has replaced the 6th edition of the staging system in clinical practice [20,27]. Our use of this new staging system assures that our conclusion that MTV_{WB} and TLG_{WB} have independent prognostic value is up-to-date. Based on our current study, the prognostic values of MTV_{WB} and TLG_{WB} are similar and there is a linear correlation between the two parameters. There is no demonstrable advantage of TLG_{WB} over MTV_{WB}. Therefore, we

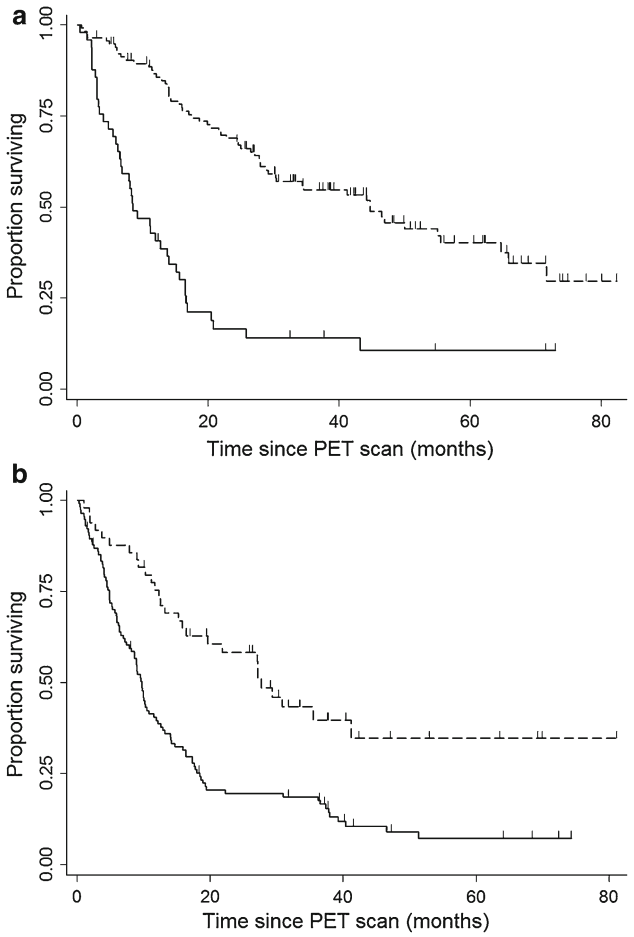

Fig. 5 Kaplan–Meier curves of overall survival after baseline PET/CT grouped according to PET MTV_{WB} measurement within groups with SUV_{maxWB} equal and below (**a**) or above the median (**b**) show the association of metabolic tumor volume with overall survival independent of SUV_{maxWB}. The *dashed lines* indicate the group with MTV values equal to or less than the median and the *solid lines* are the group with values greater than the median of the PET MTV_{WB} measurement

think in the future clinical practice, MTV_{WB} will be sufficient for measuring metabolic tumor burden in NSCLC.

The prognostic value of SUV measurements independent of patient's gender, age and treatment received, as well as the TNM stage and histology, demonstrated in the current study is consistent with prior studies [18,19,28–33]. Some studies demonstrated that SUV measurement was an important factor for predicting survival [28–31], and patients with high SUV_{max} are at increased risk of death following surgery [32,33]. The likelihood of lymph node metastasis increases with an increase in the SUV of a primary lung cancer [34,35]. However, SUV_{maxWB} was no longer statistically significant upon further multivariate Cox regression analysis including MTV_{WB} or TLG_{WB} in our study. This suggests that the SUV measurements do not provide a significant amount of additional prognostic information beyond that provided by MTV_{WB} or TLG_{WB}. Kaplan–Meier curves of all patients based on the quartiles of SUV_{maxWB}, SUV_{meanWB}, MTV_{WB} or TLG_{WB} also indicated the superior prognostic value for MTV_{WB} and TLG_{WB}. Furthermore, the *C*-statistics from the series of multivariate Cox regression analyses demonstrated better discriminatory power for MTV_{WB} and TLG_{WB} than for SUV_{maxWB} and SUV_{meanWB}. Patients with a $MTV_{WB} \leq$ median of 65.7 ml and $TLG_{WB} \leq$ median of 205.11 $SUV_{mean} * ml$ had a median OS of 41.1 and 35.4 months compared with 9.5 and 9.7 months for those with a $MTV_{WB} >$ 65.7 ml and $TLG_{WB} >$ 205.11 $SUV_{mean} * ml$, respectively. In our study, we included gender and age in multivariate Cox regression analysis, despite the fact that they were not significant factors on univariate analysis, because they have been shown to be predictive factors for survival in other studies [36,37].

Based on our previous studies [10], the inter-observer variability of the tumor burden measurement is low. The intraclass correlation coefficients (ICCs) were calculated for the evaluation of inter-observer variability in 77 patients with NSCLC in the previous study [10]. The ICC for $\ln(MTV_{WB})$ was 0.949 with 95% CI of 0.861–0.976. The ICC for $\ln(TLG_{WB})$ was 0.975 with 95% CI of 0.953–0.985. Measuring the MTV_{WB} and TLG_{WB} was not overwhelmingly time-consuming. As reported in our prior study, it took 7 h and 10 min for one observer to complete the measurement of the 193 tumor lesions in the 77 patients with stage I–III NSCLC on whole-body scans, an average of 2.3 min per lesion [10].

There are three limitations to this study. First, although we adjusted for several major factors which have been shown to be associated with patient's survival in NSCLC in the multivariate survival analysis [13], the performance status was not included in our multivariate survival analysis because in all surgical cases included in this study, the performance status was not recorded in their medical records. However, the association of the performance status and OS has not been consistently demonstrated [6,9,13]. Poor performance status may simply be due to high tumor burden [6]. Furthermore, the models generated will need to be validated externally. Second, there are several limitations of ^{18}F-FDG-PET, which can result in false-negative and false-positive results. False-negative PET results may be due to very small tumor size or low tumor grade. False-positive results may be due to active infection or inflammation including granulomatous diseases, such as sarcoidosis and histoplasmosis or tuberculosis, as well as metabolically active brown fat. However, in this study as in our clinical practice, we read the PET/CT (attenuation-corrected, non-attenuation-corrected, MIP and fusion images), as well as the diagnostic CT together, to minimize the false-negative and false-positive results. Finally, since FDG uptake and SUV measurements in lung cancer lesions significantly increase with the duration of the uptake time [38], we expect that the MTV may also increase with the FDG uptake time. However, even with such a limitation, we

have demonstrated a significant association of the metabolic tumor burden with OS of patients with NSCLC.

Conclusions

MTV_{WB} and TLG_{WB}, as metabolic tumor burden measurements in ^{18}F-FDG-PET/CT, are better prognostic markers than SUV_{maxWB} and SUV_{meanWB} and carry information independent of the patient's gender, age, treatment received, TNM stage, SUV_{maxWB} and tumor histology. Our results suggest a complementary role of MTV or TLG to TNM staging in prognostication of NSCLC patients.

Aknowledgments The authors gratefully acknowledge the helpful comments and editorial assistance of Dr. Bill C. Penney. We also gratefully acknowledge the assistance of Drs. Shengri Lao, Ming-Chi Shi and Rony Kampalath, in measuring the whole-body tumor burdens and Ms. Cassie A. Simon for providing us a comprehensive patient's follow-up list.

Conflict of interest None.

References

1. Parkin DM, Bray FI, Devesa SS (2001) Cancer burden in the year 2000. The global picture. Eur J Cancer 37(Suppl 8):S4–S66
2. Maldonado A, Gonzalez-Alenda FJ, Alonso M, Sierra JM (2007) Usefulness of positron emission tomography-computed tomography in respiratory medicine. Arch Bronconeumol 43:562–572
3. Sugawara Y, Zasadny KR, Neuhoff AW, Wahl RL (1999) Reevaluation of the standardized uptake value for FDG: variations with body weight and methods for correction. Radiology 213:521–525
4. Hamberg LM, Hunter GJ, Alpert NM, Choi NC, Babich JW, Fischman AJ (1994) The dose uptake ratio as an index of glucose metabolism: useful parameter or oversimplification?. J Nucl Med 35:1308–1312
5. Weber WA, Schwaiger M, Avril N (2000) Quantitative assessment of tumor metabolism using FDG-PET imaging. Nucl Med Biol 27:683–687
6. Lee P, Weerasuriya DK, Lavori PW, Quon A, Hara W, Maxim PG, Le QT, Wakelee HA, Donington JS, Graves EE, Loo BWJr. (2007) Metabolic tumor burden predicts for disease progression and death in lung cancer. Int J Radiat Oncol Biol Phys 69:328–333
7. Raben D, Chen C, Loo B, Graves E, Kavanagh B (2009) Metabolic tumor volume and oropharynx subsite as prognostic factors in advanced head and neck cancer. Int J Radiat Oncol 75:S412
8. Larson SM, Erdi Y, Akhurst T, Mazumdar M, Macapinlac HA, Finn RD, Casilla C, Fazzari M, Srivastava N, Yeung HW, Humm JL, Guillem J, Downey R, Karpeh M, Cohen AE, Ginsberg R (1999) Tumor treatment response based on visual and quantitative changes in global tumor glycolysis using PET-FDG imaging. The visual response score and the change in total lesion glycolysis. Clin Positron Imaging 2:159–171
9. Lee P, Bazan JG, Lavori PW, Weerasuriya DK, Quon A, Le Q-T, Wakelee HA, Graves EE, Loo BW (2012) Metabolic tumor volume is an independent prognostic factor in patients treated definitively for non-small-cell lung cancer. Clin Lung Cancer 13:52–58
10. Liao S, Penney BC, Wroblewski K, Zhang H, Simon CA, Kampalath R, Shih MC, Shimada N, Chen S, Salgia R, Appelbaum DE, Suzuki K, Chen CT, Pu Y (2012) Prognostic value of metabolic tumor burden on 18F-FDG PET in nonsurgical patients with non-small cell lung cancer. Eur J Nucl Med Mol Imaging 39(1): 27–38
11. Liao S, Penney BC, Zhang H, Suzuki K, Pu Y (2012) Prognostic value of the quantitative metabolic volumetric measurement on 18F-FDG PET/CT in stage IV nonsurgical small-cell lung cancer. Acad Radiol 19(1):69–77
12. Mountain CF (1997) Revisions in the international system for staging lung cancer. Chest 111:1710–1717
13. Grivaux M, Zureik M, Marsal L, Asselain B, Peureux M, Chavaillon JM, Prud'homme A, Carbonnelle M, Goarant E, Maury B, Bedossa A, Blanchon F (2009) Five year survival for lung cancer patients managed in general hospitals. Rev Mal Respir 26:37–44
14. Spira A, Ettinger DS (2004) Multidisciplinary management of lung cancer. N Engl J Med 350:379–392
15. Socinski MA, Morris DE, Masters GA, Lilenbaum R (2003) Chemotherapeutic management of stage IV non-small cell lung cancer. Chest 123:226S–243S
16. Borst GR, Belderbos JSA, Boellaard R, Comans EFI, de Jaeger KD, Lammertsma AA, Lebesque JV (2005) Standardized FDG uptake: a prognostic factor for inoperable non-small cell lung cancer. Eur J Cancer 41:1533–1541
17. Davies A, Tan C, Paschalides C, Barrington S, O'Doherty M, Utley M, Treasure T (2007) FDG-PET maximum standardised uptake value is associated with variation in survival: analysis of 498 lung cancer patients. Lung Cancer 55:75–78
18. van Baardwijk A, Dooms C, van Suylen RJ, Verbeken E, Hochstenbag M, Dehing-Oberije C, Rupa D, Pastorekova S, Stroobants S, Buell U, Lambin P, Vansteenkiste J, De Ruysscher D (2007) The maximum uptake of (18)F-deoxyglucose on positron emission tomography scan correlates with survival, hypoxia inducible factor-1alpha and GLUT-1 in non-small cell lung cancer. Eur J Cancer 43:1392–1398
19. Geus-Oei LF, Oyen WJ (2008) Predictive and prognostic value of FDG-PET. Cancer Imaging 8:70–80
20. UyBico SJ, Wu CC, Suh RD, Le NH, Brown K, Krishnam MS (2010) Lung cancer staging essentials: the new TNM staging system and potential imaging pitfalls. Radiographics 30:1163–1181
21. Shankar LK, Hoffman JM, Bacharach S, Graham MM, Karp J, Lammertsma AA, Larson S, Mankoff DA, Siegel BA, Vanden Abbeele A, Yap J, Sullivan D (2006) Consensus recommendations for the use of 18F-FDG PET as an indicator of therapeutic response in patients in National Cancer Institute trials. J Nucl Med 47:1059–1066
22. Werner-Wasik M, Nelson AD, Choi W, Arai Y, Faulhaber PF, Kang P, Almeida FD, Xiao Y, Ohri N, Brockway KD, Piper JW, Nelson AS (2012) What is the best way to contour lung tumors on PET scans? Multiobserver validation of a gradient-based method using a NSCLC digital PET phantom. Int J Radiat Oncol Biol Phys. 82(3):1164–1171
23. Houseni M, Chamroonrat W, Zhuang J, Gopal R, Alavi A, Zhuang H (2010) Prognostic implication of dual-phase PET in adenocarcinoma of the lung. J Nucl Med 51:535–542
24. Cox DR (1972) Regression models and life-tables (with discussion). J R Stat Soc Series B (Methodol) 34:187–220
25. Grambsch PM, Therneau TM (1994) Proportional hazards tests and diagnostics based on weighted residuals. Biometrika 81: 515–526
26. Gönen M, Heller G (2005) Concordance probability and discriminatory power in proportional hazards regression. Biometrika 92:965–970
27. Goldstraw P, Crowley J, Chansky K, Giroux DJ, Groome PA, Rami-Porta R, Postmus PE, Rusch V, Sobin L (2007) The IASLC Lung Cancer Staging Project: proposals for the revision of the TNM

stage groupings in the forthcoming (seventh) edition of the TNM Classification of malignant tumours. J Thorac Oncol 2:706–714
28. Pelosi E, Billè A, Skanjeti A, Errico L, Arena V, Ardissone F, Borasio P, Mancini M (2011) Prognostic role of the PET parameter maximum standardized uptake value in non small cell lung cancer: analysis in tumour of diameter ≥ and <25 mm. Q J Nucl Med Mol Imaging 55:72–80
29. Oven UBB, Gumus M, Bilici A, Seker M, Ustaalioglu R, Kefeli U, Salepci T, Canpolat N, Mayadagli A, Dogusoy I (2010) Is there a cut-off value for standardized uptake values in positron emission tomography for predicting response to treatment and survival in patients with advanced non-small cell lung cancer? Single center experience. J Balkan Union Oncol 15:529–536
30. Imamura Y, Azuma K, Kurata S, Hattori S, Sasada T, Kinoshita T, Okamoto M, Kawayama T, Kaida H, Ishibashi M, Aizawa H (2011) Prognostic value of SUV_{max} measurements obtained by FDG-PET in patients with non-small cell lung cancer receiving chemotherapy. Lung Cancer 71:49–54
31. Um SW, Kim H, Koh WJ, Suh GY, Chung MP, Kwon OJ, Choi JY, Han J, Lee KS, Kim J (2009) Prognostic value of 18F-FDG uptake on positron emission tomography in patients with pathologic stage I non-small cell lung cancer. J Thorac Oncol 4:1331–1336
32. Nair VS, Barnett PG, Ananth L, Gould MK (2010) PET scan 18F-fluorodeoxyglucose uptake and prognosis in patients with resected clinical stage IA non-small cell lung cancer. Chest 137:1150–1156
33. Hanin FX, Lonneux M, Cornet J, Noirhomme P, Coulon C, Distexhe J, Poncelet AJ (2008) Prognostic value of FDG uptake in early stage non-small cell lung cancer. Eur J Cardiothorac Surg 33: 819–823
34. Nambu A, Kato S, Sato Y, Okuwaki H, Nishikawa K, Saito A, Matsumoto K, Ichikawa T, Araki T (2009) Relationship between maximum standardized uptake value (SUVmax) of lung cancer and lymph node metastasis on FDG-PET. Ann Nucl Med 23:269–275
35. Liu S, Cheng H, Yao S, Wang C, Han G, Li X, Liu C (2010) The clinical application value of PET/CT in adenocarcinoma with bronchioloalveolar carcinoma features. Ann Nucl Med 24:541–547
36. Visbal AL, Williams BA, Nichols FC3rd, Marks RS, Jett JR, Aubry MC, Edell ES, Wampfler JA, Molina JR, Yang P (2004) Gender differences in non-small-cell lung cancer survival: an analysis of 4,618 patients diagnosed between 1997 and 2002. Ann Thorac Surg 78:209–215
37. Agarwal M, Brahmanday G, Chmielewski GW, Welsh RJ, Ravikrishnan KP (2010) Age, tumor size, type of surgery, and gender predict survival in early stage (stage I and II) non-small cell lung cancer after surgical resection. Lung Cancer 68:398–402
38. Demura Y, Tsuchida T, Ishizaki T, Mizuno S, Totani Y, Ameshima S, Miyamori I, Sasaki M, Yonekura Y (2003) 18F-FDG accumulation with PET for differentiation between benign and malignant lesions in the thorax. J Nucl Med 44:540–548