Eur J Nucl Med Mol Imaging (2015) 42:241–251
DOI 10.1007/s00259-014-2903-7

ORIGINAL ARTICLE

Prognostic value of volumetric parameters of ^{18}F-FDG PET in non-small-cell lung cancer: a meta-analysis

Hyung-Jun Im · Kyoungjune Pak · Gi Jeong Cheon · Keon Wook Kang · Seong-Jang Kim · In-Joo Kim · June-Key Chung · E. Edmund Kim · Dong Soo Lee

Received: 6 May 2014 / Accepted: 25 August 2014 / Published online: 6 September 2014

Abstract
Purpose We conducted a comprehensive systematic review of the literature on volumetric parameters from ^{18}F-FDG PET and a meta-analysis of the prognostic value of metabolic tumour volume (MTV) and total lesion glycolysis (TLG) in patients with lung cancer.
Methods A systematic search of MEDLINE and EMBASE was performed using the keywords "positron emission tomography (PET)", "lung cancer", and "volume". Inclusion criteria were: ^{18}F-FDG PET used as an initial imaging tool; studies limited to non-small-cell lung cancer (NSCLC); volume measurement of lung cancer; patients who had not undergone surgery, chemotherapy, or radiotherapy before the PET scan; and studies that reported survival data. Event-free survival and overall survival were evaluated as outcomes. The impact of MTV and TLG on survival was measured in terms of the hazard ratio (HR) effect size. Data from each study were analysed using Review Manager 5.2.
Results Thirteen eligible studies including 1,581 patients were analysed. Patients with high MTV showed a worse prognosis with an HR of 2.71 (95 % CI 1.82 – 4.02, $p<0.00001$) for adverse events and an HR of 2.31 (95 % CI 1.54 – 3.47, $p<0.00001$) for death. Patients with high TLG also showed a worse prognosis with an HR of 2.35 (95 % CI 1.91 – 2.89, $p<0.00001$) for adverse events and an HR of 2.43 (95 % CI 1.89 – 3.11, $p<0.00001$) for death. The prognostic value of MTV and TLG remained significant in a subgroup analysis according to TNM stage as well as the methods for defining cut-off values and tumour delineation.
Conclusion Volumetric parameters from ^{18}F-FDG PET are significant prognostic factors for outcome in patients with NSCLC. Patients with a high MTV or TLG are at higher risk of adverse events and death. MTV and TLG were significant prognostic factors in patients with TNM stage I/II and stage III/IV NSCLC.

Keywords PET · Volume · Lung cancer · Prognosis

Hyung-Jun Im and Kyoungjune Pak contributed equally to this work.

Electronic supplementary material The online version of this article (doi:10.1007/s00259-014-2903-7) contains supplementary material, which is available to authorized users.

H.-J. Im · K. Pak · G. J. Cheon (✉) · K. W. Kang · J.-K. Chung · D. S. Lee
Department of Nuclear Medicine, Seoul National University Hospital, 101 Daehak-roJongno-gu Seoul 110-744, South Korea
e-mail: larrycheon@gmail.com

H.-J. Im · E. E. Kim · D. S. Lee
Department of Molecular Medicine and Biopharmaceutical Sciences, Graduate School of Convergence Science and Technology, and College of Medicine or College of Pharmacy, Seoul, South Korea

K. Pak · S.-J. Kim · I.-J. Kim
Department of Nuclear Medicine and Biomedical Research Institute, Pusan National University Hospital, Busan, South Korea

G. J. Cheon · K. W. Kang · J.-K. Chung · D. S. Lee
Cancer Research Institute, Seoul National University Hospital, Seoul, South Korea

E. E. Kim
Department of Radiological Science, University of California at Irvine, California, CA, USA

Introduction

Lung cancer is the first leading cause of cancer death among both men and women, and is expected to account for 26 % of all female cancer deaths and 28 % of all male cancer deaths in 2013 [1]. Non-small-cell lung cancer (NSCLC) accounts for 80 % of lung cancers [2]. The standard care for the treatment of early NSCLC is surgical resection and/or radiation therapy

according to the patient's eligibility for surgery [3]. For advanced NSCLC, chemotherapy or chemoradiotherapy is the principal treatment modality [4]. Despite standard treatment, overall survival (OS) in NSCLC is very poor even in low-stage disease (50 % in stage IA) and becomes progressively worse with increasing TNM stage (2 % in stage IV) [5, 6]. However, there is no established prognostic marker except TNM stage and performance status [7].

PET/CT using ^{18}F-FDG has become a valuable tool in the differential diagnosis of a solitary pulmonary nodule and a standard modality for staging and monitoring treatment response in lung cancer [8, 9]. To quantify a lesion's metabolism, maximum standardized uptake value (SUVmax) is widely used in clinical practice. It provides a semiquantitative measure of the normalized concentration of radioactivity in a lesion [10]. In a meta-analysis the European Lung Cancer Working Party for the International Association for the Study of Lung Cancer Staging Project demonstrated the prognostic value of the SUV of the primary tumour in NSCLC [11, 12]. However, SUVmax is not recommended for risk stratification in the 7th edition of the American Joint Committee on Cancer cancer staging manual [13], and is also not considered to be a prognostic biomarker in the National Comprehensive Cancer Network guidelines (version 3, 2014) [14]. The reasons for this are that SUVmax is a single voxel value that may not represent total tumour metabolism [15] and it is not certain whether SUVmax is a reliable independent prognostic marker or whether it provides additional risk stratification over T staging [13, 16].

Instead of SUVmax, metabolic tumour volume (MTV) and total lesion glycolysis (TLG), which are volumetric indices derived from ^{18}F-FDG PET, have been proposed for risk stratification of lung cancer patients [17]. On ^{18}F-FDG PET images, tumour can be delineated by a specific threshold SUV or with other methods such as the gradient or the fuzzy C-mean method, with MTV referred to a volume of the delineated tumour [3, 18, 19]. TLG is calculated by multiplying MTV by the mean SUV of all voxels in the MTV, and represents both the degree of ^{18}F-FDG uptake and the size of the tumour, in other words the whole metabolic and volumetric burden of the tumour [10, 20–22]. Growing interest in volumetric indices has led to the development of commercially available tools that enable the rapid and simple measurement of the indices for tumour analysis [20]. In fact, MTV and TLG are considered to be more reliable markers reflecting tumour burden and aggressiveness and are thus better candidates as prognostic markers in a variety of types of malignancy including lung cancer [15, 23, 24]. However, there are also several conflicting results regarding the prognostic value of volumetric parameters in NSCLC [3, 25]. Therefore, we designed a meta-analysis to assess the prognostic value of MTV and TLG in patients with NSCLC.

Materials and methods

Data search and study selection

We performed a systematic search of MEDLINE (inception to November 2013) and EMBASE (inception to November 2013) for English language publications using the keywords "positron emission tomography", "lung", and "volume." All searches were limited to human studies. Inclusion criteria were: ^{18}F-FDG PET used as an initial imaging tool; studies limited to NSCLC; volume measurement of lung cancer; patients who had not undergone surgery, chemotherapy, or radiotherapy before the ^{18}F-FDG PET scan; and articles that reported survival data. Reviews, abstracts, and editorial material were excluded. Two authors conducted the searches and screening independently. Any discrepancies were resolved by consensus.

Data extraction and quality assessment

Data were extracted from the publications independently by two reviewers (K. Pak and H.J. Im) and the following information was recorded: first author, year of publication, country, study design, number of patients, TNM staging, treatment, and endpoints. Three reviewers scored each publication according to a quality scale, which was based on that used in previous studies [11, 26]. This quality scale was grouped into four categories: scientific design, generalizability, analysis of results, and PET reports (Supplementary Table 1). A value of 0, 1 or 2 was attributed to each item. Each category had a maximum score of 10 points.

Statistical analysis

We followed the same methodology as used in our previous study [27]. The primary outcome was event-free survival (EFS). Disease-free survival, recurrence-free survival and progression-free survival were obtained as primary outcomes and newly defined as EFS, which was measured from the date of initiation of therapy to the date of recurrence or metastasis [28]. The secondary endpoint was OS, defined as the time from initiation of therapy until death from any cause. The relationships between MTV and TLG and survival were measured in terms of the hazard ratio (HR) effect size. Survival data were extracted using the following methodology as suggested by Parmar et al. [29]. We extracted a univariate HR estimate and 95 % confidence intervals (CIs) directly from each study, if provided by the authors. Otherwise, p values of the log-rank test, 95 % CI, number of events and number at risk were extracted to estimate the HR indirectly. Survival rates on Kaplan-Meier curves were read using Engauge Digitizer version 3.0 (http://digitizer.sourceforge.net) to reconstruct the HR estimate and its variance, assuming that

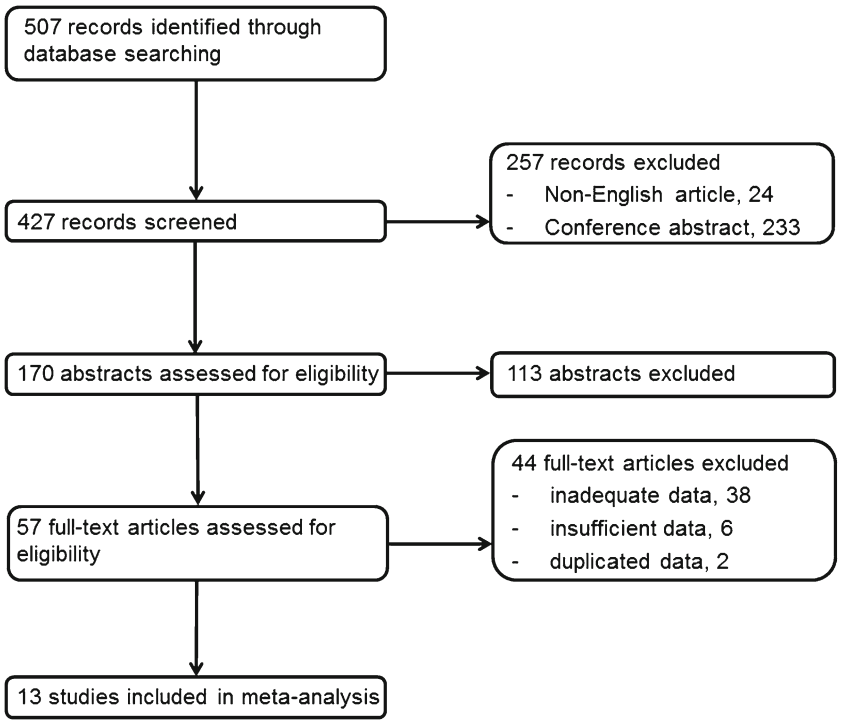

Fig. 1 Flowchart for the identification of eligible studies

patients were censored at a constant rate during follow-up. An HR greater than 1 implied worse survival in patients with a high MTV or TLG, whereas an HR less than 1 implied a survival benefit in patients with a high MTV or TLG. Heterogeneity between studies was assessed in term so χ^2 test and I^2 statistics, as described by Higgins et al. [30]. Funnel plots were used to assess publication bias graphically [31]. Survival data were also extracted in relation to SUVmax from the same studies included in this meta-analysis as mentioned above. *P* values less than 0.05 were considered statistically significant. Data from each study were analysed using Review Manager (RevMan, version 5.2; The Nordic Cochrane Centre, The Cochrane Collaboration, 2012, Copenhagen).

Results

Study characteristics

The electronic search identified 507 articles. After excluding 24 articles in languages other than English, 233 conference abstracts, and 113 studies that did not meet the inclusion criteria based on title and abstract, and reviewing the full text of 57 articles, 13 studies including 1,581 patients were eligible for inclusion in this study (Fig. 1). All 13 studies were of a retrospective design. We excluded whole-body MTV or TLG data from this meta-analysis. Either MTV [32–34] or TLG [2] was measured in four studies, and both [3, 24, 25, 35–40] were measured in nine studies. The volume of interest (VOI) was defined as the primary lung cancer lesion. Four threshold methods were adapted to segment the VOI in each study. A fixed SUV of 2.5 [2, 24, 25, 33, 34, 38, 40] or 7 [32] was used in eight studies. The gradient segmentation method was applied in two studies [3, 36], and 50 % of SUVmax was used in two studies [35, 39]. In one study, a threshold was determined using mediastinal background average SUV plus 2 standard deviations [37]. In each study, patients were divided into two groups (high and low volume) based on cut-off values. To determine cut-off values receiver operating characteristics in six studies [25, 33, 34, 38–40], median values in four studies [3, 32, 35, 36], maximally selected rank statistics in two studies [24, 37], and maximizing the profile partial likelihood [2] in one study were applied. The cut-off values of MTV ranged between 0.3 and 68.3 cm^3 and those of TLG ranged from 9.6 to 525. Visual inspection of the funnel plot suggested no evidence of publication bias. Study characteristics are summarized in Tables 1 and 2.

Primary outcome: EFS

EFS was analysed based on eight studies investigating the prognostic value of MTV. The combined HR for adverse events was 2.71 (95 % CI 1.82 – 4.02, $p<0.00001$). There was significant heterogeneity ($\chi^2=15.82$, $p=0.03$; $I^2=56$ %). Eight studies investigating the prognostic value of TLG were included in the second analysis of EFS. Using a fixed-effect model, the pooled HR was 2.43 (95 % CI 1.95 – 3.02, $p<0.00001$; $I^2=0$ %), indicating that tumours with a high TLG are associated with progression and recurrence. Forest

Table 1 Studies included in the meta-analysis

Reference	Year of publication	Country	Study design	Quality score (%)	No. of patients	TNM stage	Treatment	Endpoints	Volumetric parameters	Tumour delineation (threshold)	Determination of cut-off value	Cut-off values	
												MTV (cm^3)	TLG
[35]	2013	Italy	Retrospective	68.4	99	I	Op	DFS	MTV/TLG	SUVmax (50 %)	Median	2.9	9.6
[3]	2013	USA	Retrospective	68.4	50	I	SBRT	RFS/OS	MTV/TLG	Gradient method	Median	4.3	14.9
[24]	2014	Korea	Retrospective	73.7	194	IIIA	Op/op+CTx or RTx	OS	MTV/TLG	SUV (2.5)	Maximally selected rank statistics	36	226
[36]	2013	USA	Retrospective	57.9	39	I-IV	–	OS	MTV/TLG	Gradient method	Median	9.7	74
[37]	2013	Korea	Retrospective	78.9	529	I-II	Op/op+CTx or RTx	DFS/OS	MTV/TLG	Backgound[a] + 2SD	Maximally selected rank statistics	16	70
[38]	2012	Korea	Retrospective	76.3	57	IV	CTx	PFS	MTV/TLG	SUV (2.5)	ROC	68.3	194.2
[2]	2012	Japan	Retrospective	68.4	81	III/IV	CTx	PFS/OS	TLG	SUV (2.5)	Maximizing the profile partial likelihood	–	33
[32]	2012	USA	Retrospective	73.7	54	I	SABR	PTC/PFS/OS/CSS	MTV	SUV (7)	Median	0.3	–
[33]	2012	Taiwan	Retrospective	65.8	60	I	Op/Op+CTx	DFS	MTV	SUV (2.5)	ROC	9.8	–
[39]	2012	Taiwan	Retrospective	76.3	105	I-IV	Op/CTx/RTx/CCRT	PFS/OS	MTV/TLG	SUVmax (50 %)	ROC	60	525
[34]	2011	China	Retrospective	78.9	120	III-IV	RTx/CCRT	OS	MTV	SUV (2.5)	ROC	34	–
[40]	2013	Korea	Retrospective	76.3	102	I-II	Op	DFS	MTV/TLG	SUV (2.5)	ROC	10.8	39.7
[25]	2012	Korea	Retrospective	81.6	91	I-III	Op	RFS/OS	MTV/TLG	SUV (2.5)	ROC	9.6/11.6 (RFS/OS)	18.8/13.8 (RFS/OS)

Op operation, *CTx* chemotherapy, *RTx* radiotherapy, *CCRT* concurrent chemoradiotherapy, *SBRT* stereotactic body radiotherapy, *SABR* stereotactic ablative radiotherapy, *DFS* disease-free survival, *OS* overall survival, *PTC* primary tumour control, *CSS* cancer-specific survival, *RFS* recurrence-free survival

[a] Mediastinal background mean SUV

Table 2 PET protocols of included studies

Reference	SUV normalization	Blood sugar (mg/dL)	Fasting time (h)	Uptake time (min)	Scan time (min/bed)	SUV formula	MTV, TLG method	Reconstruction method	Attenuation correction	Dose (MBq)
[35]	ND	<180	6	60	2.5	ND	D	OSEM	CT	370
[3]	ND	<200	4 – 6	60	3 – 5	ND	D	ND	ND	370 – 444
[24]	ND	<150	6	45	5	ND	D	OSEM	CT	370
[36]	ND	100±16	ND	84±32	2 – 4	ND	D	ND	CT	455±96
[37]	Body weight	<150	6	45	4	ND	D	OSEM	CT	370
[38]	ND	<150	6	60	2.5	ND	D	OSEM	CT	7.4/kg
[2]	Body weight	<150	4	60	2.5	D	D	RAMLA	Transmission scan	266
[32]	ND	80 – 160	4 – 8	45–60	3 – 5	ND	D	ND	CT	444 – 666
[33]	Lean body mass	<120	4	45	2	D	D	ND	CT	370
[39]	ND	<150	6	60	ND	ND	D	OSEM	CT	370
[34]	Body weight	<200	6	45	ND	D	D	OSEM	Transmission scan	370
[40]	ND	<150	6	60	3	ND	D	OSEM	CT	8.1/kg
[25]	ND	<120	6	60	3	ND	D	Iterative method	ND	ND

ND not described, *D* described, *OSEM* ordered subsets and expectation maximization, *RAMLA* row-action maximum likelihood algorithm

plots of the HR in studies investigating the prognostic value of MTV and TLG are presented in Figs. 2 and 3.

Subgroup analyses were performed in relation to tumour delineation method, cut-off value, and TNM stage. According to three variables, eligible studies were divided into two subgroups. Two studies [36, 39] included patients with stage I to IV, and thus were excluded from the subgroup meta-analysis of TNM stage. Each subgroup analysis showed significant HR for events (Table 3).

Secondary outcome: OS

The survival analysis was based on seven studies investigating the prognostic value of MTV. The combined HR was 2.31 (95 % CI 1.54 – 3.47, $p<0.0001$; $\chi^2=18.97$, $p=0.004$; $I^2=68$ %; Fig. 4). Six studies investigating the prognostic value of TLG were included in the analysis of OS. The pooled HR for death was 2.49 (95 % CI 1.94 – 3.18, $p<0.00001$; Fig. 5). There was no evidence of significant heterogeneity ($I^2=28$ %. $\chi^2=6.99$, $p=0.22$). Subgroup meta-analyses in relation to cut-off value, tumour delineation method, and TNM stage were performed. Each subgroup analysis showed a significant HR for death (Table 3).

Combined SUVmax data

Survival data were extracted from studies investigating the value SUVmax in predicting EFS (seven studies) and OS (four studies). The HR for adverse events was 2.12 (95 % CI 1.30 – 3.47, $p=0.003$) with significant heterogeneity ($\chi^2=16.59$, $p=0.01$; $I^2=64$ %). The pooled HR for death was 1.2 (95 % CI 1.05 – 1.38, $p=0.008$) with significant heterogeneity (I^2 of 72 %, $\chi^2=10.89$, $p=0.01$; Table 4).

Discussion

In the present meta-analysis, the prognostic value of volumetric indices from ^{18}F-FDG PET in NSCLC patients was evaluated by analysing the HR for EFS and OS in patients with high MTV and/or TLG compared to those with low MTV and/or TLG. The pooled results showed that patients with high

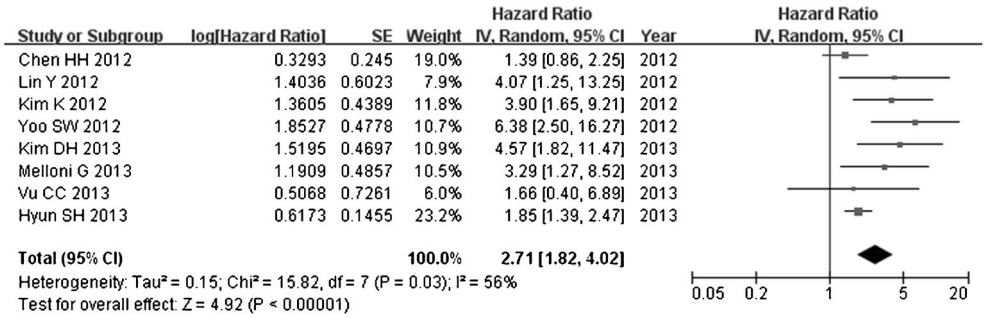

Fig. 2 Forest plots of hazard ratios for events in studies investigating the prognostic value of MTV. Hazard ratios for events in individual studies together with the pooled result are shown (*error bars* 95 % CI, *SE* standard error)

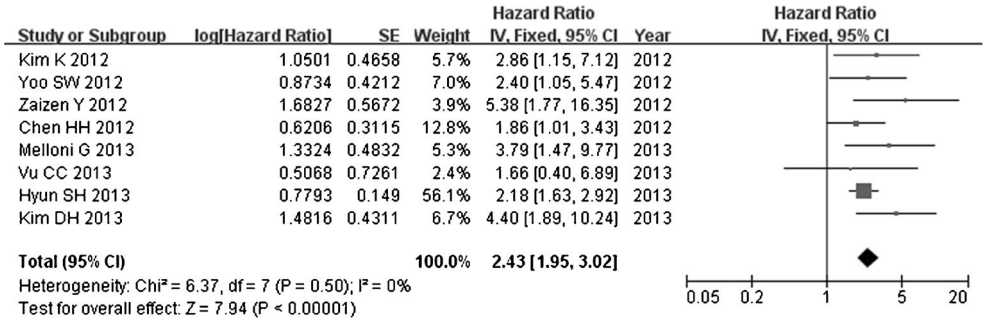

Fig. 3 Forest plots of hazard ratio for events in studies investigating the prognostic value of TLG. Hazard ratios for events in individual studies together with the pooled result are shown (*error bars* 95 % CI, *SE* standard error)

Table 3 Subgroup analyses of volumetric parameters of ^{18}F-FDG PET

Endpoint	Volumetric parameter	Factor	No. of studies	HR	95 % CI of HR	Heterogeneity, I^2 (%)	Model used
Event-free survival	MTV	Cut-off values					
		ROC	5	3.39	1.74 – 6.63	69	Random effects
		Others	3	1.93	1.47 – 2.52	0	Fixed effect
		Tumour delineation					
		SUV 2.5	4	4.66	2.89 – 7.51	0	Fixed effect
		Others	4	1.78	1.41 – 2.26	0	Fixed effect
		TNM stage					
		I/II	5	2.13	1.66 – 2.74	29	Fixed effect
		III/IV	1	6.38	2.50 – 16.27	27	Fixed effect
	TLG	Cut-off values					
		ROC	4	2.70	1.76 – 4.15	0	Fixed effect
		Others	4	2.28	1.79 – 2.91	0	Fixed effect
		Tumour delineation					
		SUV 2.5	4	2.75	1.91 – 3.98	0	Fixed effect
		Others	4	2.19	1.71 – 2.81	0	Fixed effect
		TNM stage					
		I/II	4	2.41	1.86 – 3.12	15	Fixed effect
		III/IV	2	3.19	1.65 – 6.20	24	Fixed effect
Overall survival	MTV	Cut-off values					
		Median	3	2.52	1.45 – 4.35	0	Fixed effect
		Others	4	2.24	1.30 – 3.88	82	Random effects
		Tumour delineation					
		SUV2.5	3	1.82	1.06 – 3.13	71	Random effects
		Others	4	2.97	2.14 – 4.13	0	Fixed effect
		TNM stage					
		I/II	3	3.07	2.15 – 4.38	0	Fixed effect
		III/IV	2	1.62	1.04 – 2.52	71	Random effects
	TLG	Cut-off values					
		Median	2	2.33	1.22 – 4.45	0	Fixed effect
		Others	4	2.48	1.51 – 4.07	60	Random effects
		Tumour delineation					
		SUV2.5	3	1.86	1.32 – 2.63	0	Fixed effect
		Others	3	3.25	2.27 – 4.64	0	Fixed effect
		TNM stage					
		I/II	2	3.42	2.31 – 5.06	12	Fixed effect
		III/IV	2	3.42	2.31 – 5.06	12	Fixed effect

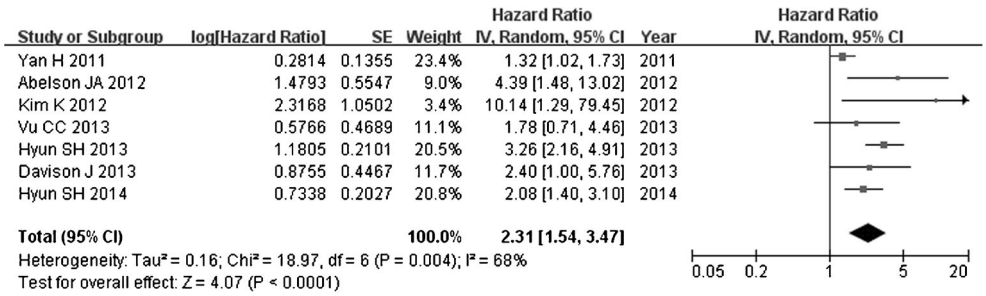

Fig. 4 Forest plots of HR for death in studies investigating the prognostic value of MTV. HRs for death in individual studies together with the pooled result are shown (*error bars* 95 % CI, *SE* standard error)

MTV had a 2.71-fold higher risk of adverse events and a 2.31-fold higher risk of death than patients with low MTV. Patients with a high TLG had a 2.35-fold higher risk of adverse events and a 2.43-fold higher risk of death. HRs of MTV and TLG for OS were higher than those of SUVmax for OS without overlapping 95 % CI (Table 4). In addition, SUVmax was not a significant prognostic factor either for EFS (three of seven studies) or for OS (two of four studies) in most studies included in this meta-analysis. In contrast, a single study [3] (1 of 13 studies) showed that MTV and TLG cannot predict EFS and OS. However, for adverse events, we could not confirm if pooled HRs of MTV and TLG are higher than that of SUVmax because of overlapping 95 % CIs (Table 4).

MTV and TLG are combined volumetric and metabolic parameters that reflect both properties of the tumour. More precisely, MTV is affected by tumour size and the distribution of the SUV and TLG is affected by MTV and also SUV. Also, SUV itself can vary according to blood glucose level, fasting time, uptake time, and methods of attenuation correction and reconstruction. We reviewed these factors in the studies included using a quality assessment form (Supplementary Table 1). In the quality assessment of the PET studies, four studies scored 5/8 (62.5 %) and the other nine studies scored 6/8 (75 %). In all studies, blood sugar levels were determined and imaging was done when the patient had a blood sugar level lower than their upper limit (blood sugar range 120 – 200 mg/dL). Fasting time was also well documented in all studies except one [36], and ranged from 4 to 8 h. Uptake time after injection of ^{18}F-FDG was well reported in all studies and ranged from 45 to about 60 min, except in one study with an uptake time of 84±32 min [36] (Table 2). The procedure for measuring SUV was acceptable in all studies except one which had relatively long uptake periods with a wide range [36]. However, exclusion of this study did not affect the pooled HRs: the HR for OS in patients with a high MTV changed from 2.31 (95 % CI 1.54 – 3.47) to 2.33 (95 % CI 1.48 – 3.66), and the HR for OS in patients with a high TLG changed from 2.49 (95 % CI 1.94 – 3.18) to 2.48 (95 % CI 1.92 – 3.21).

Although an SUVmax threshold of 2.5 is widely used for tumour delineation, Abelson et al. [32] found in their patient population that an SUVmax threshold of 7 was better than a threshold of 2 or 4 for predicting prognosis. Thus, to find specific cut-off MTV and TLG values for a worse prognosis for further research, the measurement of SUV should be well controlled and the SUV for tumour delineation should also be standardized. However, regardless of the method of tumour delineation or the MTV and TLG cut-off values selected in each study, high values of MTV and TLG were associated with a higher risk of adverse events and/or death.

The search for previous meta-analyses evaluating the utility of PET or PET/CT in lung cancer identified 20 articles (Table 5). Of these 20 studies, 15 evaluated PET for detecting lymph node metastasis [41–49] or distant metastasis [50–55], 2 evaluated the accuracy of PET for diagnosing a solitary pulmonary nodule [56, 57], one determined the predictive value of PET after neoadjuvant therapy [58], and two evaluated PET for determining disease-free survival and OS using the HR effect size [11, 12]. In a meta-analysis, Berghmans et al. [11] determined the prognostic value of SUVmax in NSCLC patients. These authors subsequently conducted another meta-analysis [12] which showed that SUVmax was associated with a 2.08-fold higher risk of death (95 % CI 1.69 – 2.56), which is similar to the pooled HR found in the current study (2.33, 95 % CI 1.51 – 3.61), even though there was no overlap in the studies between the two meta-analyses.

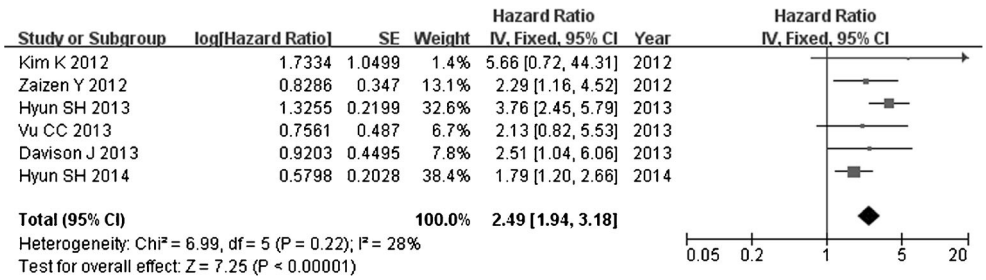

Fig. 5 Forest plots of HR for death in studies investigating the prognostic value of TLG. HRs for death in individual studies together with the pooled result are shown (*error bars* 95 % CI, *SE* standard error)

Table 4 Pooled hazard ratios of ^{18}F-FDG PET parameters

Endpoint	Parameter	HR	95 % CI of HR	*p* value
Event-free survival	SUVmax	2.12	1.30 – 3.47	0.003
	MTV	2.71	1.82 – 4.02	<0.00001
	TLG	2.35	1.91 – 2.89	<0.00001
Overall survival	SUVmax	1.20	1.05 – 1.38	0.008
	MTV	2.31	1.54 – 3.47	<0.0001
	TLG	2.43	1.89 – 3.11	<0.00001

In one study in patients with advanced stage NSCLC, high SUVmax was not a significant risk factor [16]. This might be explained by the fact that if the cancer becomes advanced, SUVmax can neither represent the whole tumour burden nor predict prognosis. Interestingly, the subgroup analysis in this study according to TNM stage showed that both MTV and TLG were significant risk factors for EFS and OS in patients with stage I/II and III/IV NSCLC.

In 11 of the included studies multivariate analysis was performed using the Cox proportional hazards model [2, 24, 25, 33–35, 37–40] or logistic regression model [36] to

Table 5 Previous meta-analyses of ^{18}F-FDG PET in patients with lung cancer

Reference	Year of publication	Country	No. of studies	No. of patients (lesions)	Classification	Effect size	Performance measure
[56]	2001	USA	40	(1,474)	Diagnosis	Selection of malignant lesion among focal pulmonary lesions	Sensitivity/ specificity
[57]	2013	China	8	415	Diagnosis	Selection of malignant lesion among focal pulmonary lesions	Sensitivity/ specificity
[41]	1999	USA	14	514	Staging	Detection of mediastinal lymph node metastasis	Sensitivity/ specificity
[42]	2003	USA	32	1,959	Staging	Detection of mediastinal lymph node metastasis	Sensitivity/ specificity
[43]	2005	Netherlands	17	833	Staging	Detection of mediastinal lymph node metastasis	Sensitivity/ specificity
[44]	2006	Italy	13	2,912	Staging	Detection of mediastinal lymph node metastasis	Sensitivity/ specificity
[50]	2011	China	8	2,446	Staging	Detection of bone metastasis	Sensitivity/ specificity
[45]	2011	China	14	2,550	Staging	Detection of mediastinal lymph node metastasis	Sensitivity/ specificity
[52]	2012	China	7	1,746	Staging	Detection of bone metastasis in comparison with bone scan	Sensitivity/ specificity
[46]	2012	Taiwan	7	1,248	Staging	Detection of mediastinal lymph node metastasis by comparison between regions	Sensitivity/ specificity
[51]	2012	China	17	2,940	Staging	Detection of bone metastasis in comparison with bone scan and MRI	Sensitivity/ specificity
[47]	2011	China	10	1,122	Staging	Negative predictive value for mediastinal lymph node metastasis	Negative predictive value
[48]	2012	China	19	2,845	Staging	Detection of hilar and mediastinal lymph node metastasis in comparison with DWI	Sensitivity/ specificity
[53]	2012	China	5	578	Staging	Detection of distant malignancy	Sensitivity/ specificity
[49]	2012	China	20	3,028	Staging	Detection of mediastinal lymph node metastasis	Sensitivity/ specificity
[54]	2012	China	56	8,699	Staging	Detection of metastasis	Sensitivity/ specificity
[55]	2013	China	9	780	Staging	Detection of distant metastasis	Sensitivity/ specificity
[58]	2013	China	13	414	Follow-up	Prediction of pathological response after neoadjuvant chemotherapy	Sensitivity/ specificity
[11]	2008	Belgium	13	1,474	Prognosis	DFS/OS	Hazard ratio
[12]	2010	Belgium	21	2,637	Prognosis	DFS/OS	Hazard ratio

evaluate the independence of MTV and TLG as prognostic markers with covariates including TNM stage and/or tumour size. Of seven studies in which multivariate analysis for EFS was performed [23, 25, 31, 33, 39, 45, 52], three [25, 38, 40] of six [25, 33, 35, 38–40] and four [2, 35, 38, 40] of seven showed that MTV and TLG, respectively, are independent prognostic markers for EFS. On the other hand, SUVmax was found to be an independent prognostic marker in only one study [40] of seven [2, 25, 33, 35, 38–40]. Of six studies in which multivariate analysis was performed for OS, five of five [24, 25, 34, 36, 37] and three [2, 24, 37] of five [2, 24, 25, 36, 37] showed that MTV and TLG, respectively, are independent prognostic markers for OS. However, SUVmax was found to be an independent prognostic marker in only one study [37] of six [2, 24, 25, 34, 36, 37]. These results indicate that, unlike SUVmax, MTV and TLG might be independent prognostic markers regardless of TNM stage and tumour size. However, since the results are heterogeneous and all included studies had a retrospective design, a further large-scale prospective study is warranted to assess whether MTV and TLG could be independent prognostic factors for clinical outcome.

Heterogeneity was detected in the present meta-analysis. In pooled data, significant heterogeneity was found for MTV in predicting EFS [38] and OS [34], and thus a random effect model was used to derive a pooled HR. In each analysis of the value of MTV in predicting EFS, studies that showed heterogeneity were identified [34, 38]. The study by Yan et al. [34] was the only study that used PET rather than PET/CT in analysis of the value of MTV in predicting EFS, and the study by Yoo et al. [38] was the only study that included only patients with stage IV lung cancer. Excluding these two studies reduced the heterogeneity (I^2, from 56 % to 42 % for EFS, and from 68 % to 11 % for OS) with HR of 2.34 (95 % CI 1.64 – 3.34) for EFS and 2.64 (95 % CI 1.99 – 3.50) for OS.

This is the first meta-analysis investigating the prognostic value of volumetric parameters in patients with lung cancer; however, the study had several limitations. We were unable to determine an optimal cut-off value to categorize volumetric parameters as high or low. Different cut-off values and delineation strategies, and various histological methods were applied in the studies, which might have affected the occurrence of events and survival. Further studies with data from individual patients are needed to determine standard cut-off values and delineation methods for predicting prognosis using volumetric PET parameters. Although we found that patients with a high MTV or TLG had a higher risk of adverse events or death than patients with a low MTV or TLG, there was difficulty in interpreting the HRs for MTV and TLG because exact incidence rates for the events were unknown. Further prospective studies are needed which also include incidence rates. The included studies were all retrospective in design and thus the results could have been underpowered. There was a single study with a prospective design, but we could not extract survival data [7]. A publication bias cannot be excluded even though funnel plots showed no clear evidence of it. In addition, language bias could have been present because articles in languages other than English were excluded. In addition, although two reviewers independently extracted data from each study, the complete accuracy of the data could not be ensured by the strategy.

Conclusion

Volumetric parameters from ^{18}F-FDG PET are significant prognostic factors for outcome in patients with NSCLC. Patients with a high MTV or TLG are at higher risk of adverse events or death. In addition, volumetric parameters may be used as incremental predictors of EFS rather than SUVmax even in patients with advanced NSCLC.

Acknowledgments This study was supported by the National Research Fund (grant no. HI13C1299) of the Korea Health Industry Development Institute (KHIDI) and an Intramural Research Grant from Seoul National University.

Conflicts of interest None.

References

1. Siegel R, Naishadham D, Jemal A. Cancer statistics, 2013. CA Cancer J Clin. 2013;63(1):11–30.
2. Zaizen Y, Azuma K, Kurata S, Sadashima E, Hattori S, Sasada T, et al. Prognostic significance of total lesion glycolysis in patients with advanced non-small cell lung cancer receiving chemotherapy. Eur J Radiol. 2012;81(12):4179–84.
3. Vu CC, Matthews R, Kim B, Franceschi D, Bilfinger TV, Moore WH. Prognostic value of metabolic tumor volume and total lesion glycolysis from (18)F-FDG PET/CT in patients undergoing stereotactic body radiation therapy for stage I non-small-cell lung cancer. Nucl Med Commun. 2013;34(10):959–63.
4. Zarogoulidis K, Zarogoulidis P, Darwiche K, Boutsikou E, Machairiotis N, Tsakiridis K, et al. Treatment of non-small cell lung cancer (NSCLC). J Thorac Dis. 2013;5 Suppl 4:S389–S96.
5. Budiawan H, Cheon GJ, Im HJ, Lee SJ, Paeng JC, Kang KW, et al. Heterogeneity analysis of ^{18}F-FDG uptake in differentiating between metastatic and inflammatory lymph nodes in adenocarcinoma of the lung: comparison with other parameters and its application in a clinical setting. Nucl Med Mol Imaging. 2013;47(4):232–41.
6. Goldstraw P, Crowley J, Chansky K, Giroux DJ, Groome PA, Rami-Porta R, et al. 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. 2007;2(8):706–14.
7. Lee P, Weerasuriya DK, Lavori PW, Quon A, Hara W, Maxim PG, et al. Metabolic tumor burden predicts for disease progression and death in lung cancer. Int J Radiat Oncol Biol Phys. 2007;69(2):328–33.
8. De Leyn P, Lardinois D, Van Schil PE, Rami-Porta R, Passlick B, Zielinski M, et al. ESTS guidelines for preoperative lymph node

staging for non-small cell lung cancer. Eur J Cardiothorac Surg. 2007;32(1):1–8.
9. Stroobants SG, D'Hoore I, Dooms C, De Leyn PR, Dupont PJ, De Wever W, et al. Additional value of whole-body fluorodeoxyglucose positron emission tomography in the detection of distant metastases of non-small-cell lung cancer. Clin Lung Cancer. 2003;4(4):242–7.
10. Paidpally V, Chirindel A, Lam S, Agrawal N, Quon H, Subramaniam RM. FDG-PET/CT imaging biomarkers in head and neck squamous cell carcinoma. Imaging Med. 2012;4(6):633–47.
11. Berghmans T, Dusart M, Paesmans M, Hossein-Foucher C, Buvat I, Castaigne C, et al. Primary tumor standardized uptake value (SUVmax) measured on fluorodeoxyglucose positron emission tomography (FDG-PET) is of prognostic value for survival in non-small cell lung cancer (NSCLC): a systematic review and meta-analysis (MA) by the European Lung Cancer Working Party for the IASLC Lung Cancer Staging Project. J Thorac Oncol. 2008;3(1):6–12.
12. Paesmans M, Berghmans T, Dusart M, Garcia C, Hossein-Foucher C, Lafitte JJ, et al. Primary tumor standardized uptake value measured on fluorodeoxyglucose positron emission tomography is of prognostic value for survival in non-small cell lung cancer: update of a systematic review and meta-analysis by the European Lung Cancer Working Party for the International Association for the Study of Lung Cancer Staging Project. J Thorac Oncol. 2010;5(5):612–9.
13. Edge S, Byrd DR, Compton CC, Fritz AG, Greene FL, Trotti A (eds) AJCC cancer staging handbook. From the AJCC cancer staging manual, 7th ed. New York: Springer; 2010.
14. National Comprehensive Cancer Network. Non-Small Cell Lung Cancer (version 3.2014). http://www.nccn.org/professionals/physician_gls/pdf/nscl.pdf. Accessed 13 March 2014.
15. Lim R, Eaton A, Lee NY, Setton J, Ohri N, Rao S, et al. ^{18}F-FDG PET/CT metabolic tumor volume and total lesion glycolysis predict outcome in oropharyngeal squamous cell carcinoma. J Nucl Med. 2012;53(10):1506–13.
16. Hoang JK, Hoagland LF, Coleman RE, Coan AD, Herndon 2nd JE, Patz Jr EF. Prognostic value of fluorine-18 fluorodeoxyglucose positron emission tomography imaging in patients with advanced-stage non-small-cell lung carcinoma. J Clin Oncol. 2008;26(9):1459–64.
17. Romesser PB, Qureshi MM, Shah BA, Chatburn LT, Jalisi S, Devaiah AK, et al. Superior prognostic utility of gross and metabolic tumor volume compared to standardized uptake value using PET/CT in head and neck squamous cell carcinoma patients treated with intensity-modulated radiotherapy. Ann Nucl Med. 2012;26(7):527–34.
18. Van de Wiele C, Kruse V, Smeets P, Sathekge M, Maes A. Predictive and prognostic value of metabolic tumour volume and total lesion glycolysis in solid tumours. Eur J Nucl Med Mol Imaging. 2013;40(2):290–301.
19. Dewalle-Vignion AS, Yeni N, Petyt G, Verscheure L, Huglo D, Beron A, et al. Evaluation of PET volume segmentation methods: comparisons with expert manual delineations. Nucl Med Commun. 2012;33(1):34–42.
20. Moon SH, Hyun SH, Choi JY. Prognostic significance of volume-based PET parameters in cancer patients. Korean J Radiol. 2013;14(1):1–12.
21. Arslan N, Miller TR, Dehdashti F, Battafarano RJ, Siegel BA. Evaluation of response to neoadjuvant therapy by quantitative 2-deoxy-2-[^{18}F] fluoro-D-glucose with positron emission tomography in patients with esophageal cancer. Mol Imaging Biol. 2002;4(4):301–10.
22. Rahim MK, Kim SE, So H, Kim HJ, Cheon GJ, Lee ES, et al. Recent trends in PET image interpretations using volumetric and texture-based quantification methods in nuclear oncology. Nucl Med Mol Imaging. 2014;48(1):1–15
23. Chung HH, Kim JW, Han KH, Eo JS, Kang KW, Park NH, et al. Prognostic value of metabolic tumor volume measured by FDG-PET/CT in patients with cervical cancer. Gynecol Oncol. 2011;120(2):270–4.
24. Hyun SH, Ahn HK, Kim H, Ahn MJ, Park K, Ahn YC, et al. Volume-based assessment by (18)F-FDG PET/CT predicts survival in patients with stage III non-small-cell lung cancer. Eur J Nucl Med Mol Imaging. 2014;41(1):50–8.
25. Kim K, Kim SJ, Kim IJ, Kim YS, Pak K, Kim H. Prognostic value of volumetric parameters measured by F-18 FDG PET/CT in surgically resected non-small-cell lung cancer. Nucl Med Commun. 2012;33(6):613–20.
26. Pan L, Gu P, Huang G, Xue H, Wu S. Prognostic significance of SUV on PET/CT in patients with esophageal cancer: a systematic review and meta-analysis. Eur J Gastroenterol Hepatol. 2009;21(9):1008–15.
27. Pak K, Cheon GJ, Nam HY, Kim SJ, Kang KW, Chung JK, et al. Prognostic value of metabolic tumor volume and total lesion glycolysis in head and neck cancer: a systematic review and meta-analysis. J Nucl Med. 2014;55(6):884–90.
28. Zhao Q, Feng Y, Mao X, Qie M. Prognostic value of fluorine-18-fluorodeoxyglucose positron emission tomography or PET-computed tomography in cervical cancer: a meta-analysis. Int J Gynecol Cancer. 2013;23(7):1184–90.
29. Parmar MK, Torri V, Stewart L. Extracting summary statistics to perform meta-analyses of the published literature for survival endpoints. Stat Med. 1998;17(24):2815–34.
30. Higgins JP, Thompson SG, Deeks JJ, Altman DG. Measuring inconsistency in meta-analyses. BMJ. 2003;327(7414):557–60.
31. Egger M, Davey Smith G, Schneider M, Minder C. Bias in meta-analysis detected by a simple, graphical test. BMJ. 1997;315(7109):629–34.
32. Abelson JA, Murphy JD, Trakul N, Bazan JG, Maxim PG, Graves EE, et al. Metabolic imaging metrics correlate with survival in early stage lung cancer treated with stereotactic ablative radiotherapy. Lung Cancer. 2012;78(3):219–24.
33. Lin Y, Lin WY, Kao CH, Yen KY, Chen SW, Yeh JJ. Prognostic value of preoperative metabolic tumor volumes on PET-CT in predicting disease-free survival of patients with stage I non-small cell lung cancer. Anticancer Res. 2012;32(11):5087–91.
34. Yan H, Wang R, Zhao F, Zhu K, Jiang S, Zhao W, et al. Measurement of tumor volume by PET to evaluate prognosis in patients with advanced non-small cell lung cancer treated by non-surgical therapy. Acta Radiol. 2011;52(6):646–50.
35. Melloni G, Gajate AM, Sestini S, Gallivanone F, Bandiera A, Landoni C, et al. New positron emission tomography derived parameters as predictive factors for recurrence in resected stage I non-small cell lung cancer. Eur J Surg Oncol. 2013;39(11):1254–61.
36. Davison J, Mercier G, Russo G, Subramaniam RM. PET-based primary tumor volumetric parameters and survival of patients with non-small cell lung carcinoma. AJR Am J Roentgenol. 2013;200(3):635–40.
37. Hyun SH, Choi JY, Kim K, Kim J, Shim YM, Um SW, et al. Volume-based parameters of (18)F-fluorodeoxyglucose positron emission tomography/computed tomography improve outcome prediction in early-stage non-small cell lung cancer after surgical resection. Ann Surg. 2013;257(2):364–70.
38. Yoo SW, Kim J, Chong A, Kwon SY, Min JJ, Song HC, et al. Metabolic tumor volume measured by F-18 FDG PET/CT can further stratify the prognosis of patients with stage IV non-small cell lung cancer. Nucl Med Mol Imaging. 2012;46(4):286–93.
39. Chen HH, Chiu NT, Su WC, Guo HR, Lee BF. Prognostic value of whole-body total lesion glycolysis at pretreatment FDG PET/CT in non-small cell lung cancer. Radiology. 2012;264(2):559–66.
40. Kim DH, Son SH, Kim CY, Hong CM, Oh JR, Song BI, et al. Prediction for recurrence using F-18 FDG PET/CT in pathologic N0 lung adenocarcinoma after curative surgery. Ann Surg Oncol. 2014:21(2):589–96.

41. Dwamena BA, Sonnad SS, Angobaldo JO, Wahl RL. Metastases from non-small cell lung cancer: mediastinal staging in the 1990s – meta-analytic comparison of PET and CT. Radiology. 1999;213(2): 530–6.
42. Gould MK, Kuschner WG, Rydzak CE, Maclean CC, Demas AN, Shigemitsu H, et al. Test performance of positron emission tomography and computed tomography for mediastinal staging in patients with non-small-cell lung cancer: a meta-analysis. Ann Intern Med. 2003;139(11):879–92.
43. Birim O, Kappetein AP, Stijnen T, Bogers AJ. Meta-analysis of positron emission tomographic and computed tomographic imaging in detecting mediastinal lymph node metastases in nonsmall cell lung cancer. Ann Thorac Surg. 2005;79(1):375–82.
44. Alongi F, Ragusa P, Montemaggi P, Bona CM. Combining independent studies of diagnostic fluorodeoxyglucose positron-emission tomography and computed tomography in mediastinal lymph node staging for non-small cell lung cancer. Tumori. 2006;92(4):327–33.
45. Lv YL, Yuan DM, Wang K, Miao XH, Qian Q, Wei SZ, et al. Diagnostic performance of integrated positron emission tomography/computed tomography for mediastinal lymph node staging in non-small cell lung cancer: a bivariate systematic review and meta-analysis. J Thorac Oncol. 2011;6(8):1350–8.
46. Liao CY, Chen JH, Liang JA, Yeh JJ, Kao CH. Meta-analysis study of lymph node staging by ^{18}F-FDG PET/CT scan in non-small cell lung cancer: comparison of TB and non-TB endemic regions. Eur J Radiol. 2012;81(11):3518–23.
47. Wang J, Welch K, Wang L, Kong FM. Negative predictive value of positron emission tomography and computed tomography for stage T1-2N0 non-small-cell lung cancer: a meta-analysis. Clin Lung Cancer. 2012;13(2):81–9.
48. Wu LM, Xu JR, Gu HY, Hua J, Chen J, Zhang W, et al. Preoperative mediastinal and hilar nodal staging with diffusion-weighted magnetic resonance imaging and fluorodeoxyglucose positron emission tomography/computed tomography in patients with non-small-cell lung cancer: which is better? J Surg Res. 2012;178(1):304–14.
49. Zhao L, He ZY, Zhong XN, Cui ML. (18)FDG-PET/CT for detection of mediastinal nodal metastasis in non-small cell lung cancer: a meta-analysis. Surg Oncol. 2012;21(3):230–6.
50. Liu T, Xu JY, Xu W, Bai YR, Yan WL, Yang HL. Fluorine-18 deoxyglucose positron emission tomography, magnetic resonance imaging and bone scintigraphy for the diagnosis of bone metastases in patients with lung cancer: which one is the best? – a meta-analysis. Clin Oncol (R Coll Radiol). 2011;23(5):350–8.
51. Qu X, Huang X, Yan W, Wu L, Dai K. A meta-analysis of (18)FDG-PET-CT, (18)FDG-PET, MRI and bone scintigraphy for diagnosis of bone metastases in patients with lung cancer. Eur J Radiol. 2012;81(5):1007–15.
52. Chang MC, Chen JH, Liang JA, Lin CC, Yang KT, Cheng KY, et al. Meta-analysis: comparison of F-18 fluorodeoxyglucose-positron emission tomography and bone scintigraphy in the detection of bone metastasis in patients with lung cancer. Acad Radiol. 2012;19(3):349–57.
53. Xu G, Zhao L, He Z. Performance of whole-body PET/CT for the detection of distant malignancies in various cancers: a systematic review and meta-analysis. J Nucl Med. 2012;53(12):1847–54.
54. Wu Y, Li P, Zhang H, Shi Y, Wu H, Zhang J, et al. Diagnostic value of fluorine-18 fluorodeoxyglucose positron emission tomography/computed tomography for the detection of metastases in non-small-cell lung cancer patients. Int J Cancer. 2013;132(2):E37–47.
55. Li J, Xu W, Kong F, Sun X, Zuo X. Meta-analysis: accuracy of ^{18}FDG PET-CT for distant metastasis staging in lung cancer patients. Surg Oncol. 2013;22(3):151–5.
56. Gould MK, Maclean CC, Kuschner WG, Rydzak CE, Owens DK. Accuracy of positron emission tomography for diagnosis of pulmonary nodules and mass lesions: a meta-analysis. JAMA. 2001;285(7):914–24.
57. Zhang L, Wang Y, Lei J, Tian J, Zhai Y. Dual time point ^{18}FDG-PET/CT versus single time point 18FDG-PET/CT for the differential diagnosis of pulmonary nodules: a meta-analysis. Acta Radiol. 2013;54(7):770–7.
58. Zhang C, Liu J, Tong J, Sun X, Song S, Huang G. ^{18}F-FDG-PET evaluation of pathological tumour response to neoadjuvant therapy in patients with NSCLC. Nucl Med Commun. 2013;34(1):71–7.