Eur J Nucl Med Mol Imaging (2016) 43:2155–2165
DOI 10.1007/s00259-016-3433-2

ORIGINAL ARTICLE

TLG-S criteria are superior to both EORTC and PERCIST for predicting outcomes in patients with metastatic lung adenocarcinoma treated with erlotinib

Kung-Chu Ho[1,2] · Yu-Hua Dean Fang[3] · Hsiao-Wen Chung[1] · Yuan-Chang Liu[4] · John Wen-Cheng Chang[5] · Ming-Mo Hou[5] · Cheng-Ta Yang[6] · Nai-Ming Cheng[2] · Tzu-Pei Su[7] · Tzu-Chen Yen[2]

Received: 11 April 2016 / Accepted: 26 May 2016 / Published online: 3 June 2016

Abstract

Purpose In this retrospective review of prospectively collected data, we sought to investigate whether early FDG-PET assessment of treatment response based on total lesion glycolysis measured using a systemic approach (TLG-S) would be superior to either local assessment with EORTC (European Organization for Research and Treatment of Cancer) criteria or single-lesion assessment with PERCIST (PET Response Criteria in Solid Tumors) for predicting clinical outcomes in patients with metastatic lung adenocarcinoma treated with erlotinib. We also examined the effect of bone flares on tumor response evaluation by single-lesion assessment with PERCIST in patients with metastatic bone lesions.

Methods We performed a retrospective review of prospectively collected data from 23 patients with metastatic lung adenocarcinoma treated with erlotinib. All participants underwent FDG-PET imaging at baseline and on days 14 and 56 after completion of erlotinib treatment. In addition, diagnostic CT scans were performed at baseline and on day 56. FDG-PET response was assessed with TLG-S, EORTC, and PERCIST criteria. Response assessment based on RECIST 1.1 (Response Evaluation Criteria in Solid Tumors) from diagnostic CT imaging was used as the reference standard. Two-year progression-free survival (PFS) and overall survival (OS) served as the main outcome measures.

Results We identified 13 patients with bone metastases. Of these, four (31 %) with persistent bone uptake due to bone flares on day 14 were erroneously classified as non-responders according to the PERCIST criteria, but they were correctly classified as responders according to both the EORTC and TLG-S criteria. Patients who were classified as responders on day 14 based on TLG-S criteria had higher rates of 2-year PFS (26.7 % vs. 0 %, $P = 0.007$) and OS (40.0 % vs. 7.7 %, $P = 0.018$). Similar rates were observed in patients who showed a response on day 56 based on CT imaging according to the RECIST criteria. Patients classified as responders on day 14 according to the EORTC criteria on FDG-PET imaging had better rates of 2-year OS than did non-responders (36.4 % vs. 8.3 %, $P = 0.015$).

Conclusions TLG-S criteria may be of greater help in predicting survival outcomes than other forms of assessment. Bone flares, which can interfere with the interpretation of treatment response based on PERCIST criteria, are not uncommon in patients with metastatic lung adenocarcinoma treated with erlotinib.

✉ Tzu-Chen Yen
yen1110@adm.cgmh.org.tw

1 Graduate Institute of Biomedical Electronics and Bioinformatics, National Taiwan University, Taipei, Taiwan

2 Department of Nuclear Medicine and Center for Advanced Molecular Imaging and Translation, Chang Gung Memorial Hospital and Chang Gung University, 5 Fu-Shin Street, Kueishan, Taoyuan 333, Taiwan

3 Department of Biomedical Engineering, National Cheng Kung University, Tainan, Taiwan

4 Department of Medical Imaging and Intervention, Chang Gung Memorial Hospital and Chang Gung University, Taoyuan, Taiwan

5 Division of Hematology-Oncology, Department of Internal Medicine, Chang Gung Memorial Hospital and Chang Gung University, Taoyuan, Taiwan

6 Department of Thoracic Medicine, Chang Gung Memorial Hospital and Chang Gung University, Taoyuan, Taiwan

7 Department of Nuclear Medicine, Chang Gung Memorial Hospital, Keelung, Taiwan

Keywords Lung cancer · Erlotinib · FDG-PET · Tumor response · Survival · Outcomes

Introduction

Erlotinib (Tarceva®; Roche Products Ltd., Welwyn Garden City, UK) is a small-molecule inhibitor of epidermal growth factor receptor (EGFR) tyrosine kinase enzymatic activity. Although patients carrying mutations of the *EGFR* gene generally respond better to tyrosine kinase inhibitors (TKIs) [1], erlotinib may improve survival even in subjects with *EGFR* wild-type tumors [2]. Several studies have shown that FDG-PET is a useful imaging modality for predicting response to erlotinib in patients with non-small cell lung cancer (NSCLC) [3–19]. However, most of these studies used only the primary tumor as the target lesion for sequential imaging [3–12], and treatment response was largely assessed using the European Organization for Research and Treatment of Cancer (EORTC) criteria [3–9, 20]. Other studies have evaluated treatment response using the Positron Emission Tomography Response Criteria in Solid Tumors (PERCIST) [21], based on the measurement of a specific focus on the hottest single lesion, including the metastatic tumor [15, 16]. Finally, other studies have used total lesion glycolysis measured using a systemic approach (TLG-S) based either on the sum of up to five measurable target lesions [18] or on all measurable lesions [19].

Growing evidence indicates that significant differences in tumor biology exist between primary malignant cells and their metastasized progeny. In NSCLC, a significant discordance in *EGFR* and *K-RAS* mutation status has been reported between primary tumors and their corresponding lymph node metastases [22] or distant metastases [23, 24]. In light of the discrepancies in genetic alterations between primary and metastatic tumors [25–27], a deeper understanding of their specific metabolic phenotype on FDG-PET scans can aid in the clinical investigation of therapeutic response to TKIs in patients with advanced NSCLC. Accordingly, researchers have reported a correlation between *EGFR* mutation heterogeneity and a mixed FDG-PET response in patients with lung adenocarcinoma treated with TKIs [28]. Unfortunately, the question remains as to whether systemic assessment (i.e., including sites of distant metastases) of tumor response by FDG-PET would be superior to the exclusive focus on primary tumor response in this setting.

Starting from this premise, we designed the current study to investigate whether early FDG-PET assessment of treatment response using TLG-S would be superior to either local assessment with EORTC criteria or single-lesion assessment with PERCIST for predicting clinical outcomes in patients with metastatic lung adenocarcinoma treated with erlotinib. The study hypothesis originated from the differences in tumor biology between primary malignant cells and their metastasized progeny. In addition, bone flares in NSCLC patients treated with TKIs have been reported in both CT [29] and bone scan [30] studies. During the evaluation of the hottest single lesion with PERCIST criteria, bone flares can interfere with the assessment of tumor response when FDG-PET is used in patients with metastatic bone disease. Therefore, we examined the effect of bone flares on the assessment of tumor response using the PERCIST criteria.

Materials and methods

Patients

Eligibility criteria for the study were as follows: 1) stage IIIB-IV lung adenocarcinoma or recurrent adenocarcinoma of the lung that failed to respond to frontline chemotherapy or relapsed thereafter; 2) complete recovery from any toxic effects of previous antitumor therapy; and 3) no chemotherapy within 1 month of enrollment. Patients were excluded if they met any of the following criteria: 1) symptomatic brain metastases, 2) severe comorbidities, 3) the presence of malignant pleural effusion without other measurable lesions, or 4) active infections. The study protocol was approved by the institutional review board of the Chang Gung memorial hospital. Written informed consent was obtained from all participants.

Study design

This was a single-center, single-arm, open-label study. All patients received oral erlotinib at a fixed dose of 150 mg (one tablet per day). Baseline FDG-PET examinations (day 0 FDG-PET) were performed in the 2 weeks preceding the start of erlotinib therapy. Follow-up FDG-PET scans were performed on days 14 and 56 after beginning erlotinib treatment, in order to assess early and late treatment response, respectively. Contrast-enhanced diagnostic CT scans were performed both at baseline and on day 56.

The primary aim of the study was twofold: 1) to investigate whether early FDG-PET could predict late tumor response, and 2) to examine the prognostic value of early FDG-PET for survival outcomes. Because several papers addressing these issues have been published [3–19], we also performed a retrospective review of our cohort data. The first goal of the retrospective analysis was to investigate whether early FDG-PET assessment of treatment response using TLG-S would be superior to either local assessment with EORTC criteria or hottest-single-lesion assessment with PERCIST criteria for predicting 2-year survival outcomes. The second objective was to analyze the impact of bone flare on tumor response assessment with PERCIST criteria.

FDG-PET/CT image acquisition

Patients were asked to fast for 4 h before examination, and blood glucose levels were <200 mg/dL in all participants. No intravenous contrast enhancement was used. Images were acquired 50 min after intravenous injection of 370–555 MBq ^{18}F-FDG (depending on body weight). Whole-body PET emission scans were performed from the base of the skull to the mid-thigh, with no position changes. FDG-PET/CT was performed in 18 patients using a Discovery ST 16 scanner (GE Healthcare, Milwaukee, WI, USA), whereas imaging in the five remaining patients was performed on a Biograph mCT scanner (Siemens Medical Solutions, Malvern, PA, USA). Low-dose CT images were used for attenuation correction of PET data. PET images were reconstructed using a CT-based attenuation correction with an ordered-subset expectation maximization iterative reconstruction algorithm (4 iterations and 10 subsets for the Discovery ST16 scanner; 2 iterations and 21 subsets for the Biograph mCT scanner). Using these reconstruction parameters, axial spatial resolutions of PET at the center of the gantry were 4.80 mm and 2.16 mm for the Discovery ST16 and the Biograph mCT scanners, respectively. We performed cross-calibration of the Siemens Biograph mCT scanner every 3 months. In addition, the GE Discovery ST scanner at our site undergoes 3D normalization and well counter correction every 3 months to ensure optimal quantitative accuracy. Our site receives a certificate of validation from the Society of Nuclear Medicine and Molecular Imaging (SNMMI) Clinical Trials Network program every time a scanner is validated. Consequently, there were no significant differences in terms of standardized uptake value (SUV) quantitation between the two scanners.

Imaging analysis and assessment of treatment response

FDG-PET images were obtained in transaxial planes using a dedicated workstation (*syngo*; Siemens Medical Solutions). The SUV for each tumor volume was calculated with the following formula: (measured activity concentration [Bq/mL])/(injected activity [Bq]/body weight [kg] × 1000). Rather than using the peak SUV utilized by the PERCIST criteria, we measured the maximum SUV (SUV_{max}) within a region of interest (ROI) [16]. An SUV_{max} >2.5 was used as the threshold for target volume delineation of the metabolic tumor volume (MTV) [31]. TLG-S was calculated as follows: TLG-S = mean SUV × MTV (cm^3) [32].

The metabolic response according to the EORTC criteria is based on the same ROI volumes sampled on subsequent scans. Although EORTC criteria define partial metabolic response (PMR) as a minimum of 15–25 % reduction in SUV_{max} after one cycle of therapy, we adopted a value of ≥25 % for defining PMR. Stable metabolic disease (SMD) was diagnosed in the presence of either an increase or a decrease of <25 %. Finally, progressive metabolic disease (PMD) was defined as an increase in SUV_{max} of >25 % [20].

In line with the standard procedures recommended by the PERCIST criteria, we measured the change in SUV_{max} between the hottest single tumor lesion on the baseline scan and on the subsequent scan. The target lesions could differ between the two scans. Complete metabolic response (CMR) was defined as complete abrogation of tumor FDG uptake; PMR was defined as a reduction in SUV_{max} of at least 30 %, and PMD as either an increase in SUV_{max} of at least 30 % or development of a new lesion. Finally, SMD was considered to be present when CMR, PMR, and PMD did not occur [21].

According to the PERCIST recommendations [21], the measurement of TLG-S was based on the delineation of target lesions (two or fewer lesions per organ, with a maximum of five lesions). PMR was defined as a reduction of at least 45 % in TLG-S, whereas PMD was diagnosed in the presence of a 75 % or greater increase in this parameter [21]. SMD was considered to be present when PMR or PMD did not occur [21].

Standard CT response was assessed through an independent review of diagnostic CT images obtained on day 56 compared with baseline scans. All diagnostic CT images were analyzed by investigators blinded to PET results. Target lesions (two or fewer lesions per organ, with a maximum of five lesions) were identified. Tumor response was classified as complete response (CR), partial response (PR), stable disease (SD), or progressive disease (PD) according to the Response Evaluation Criteria in Solid Tumors (RECIST), version 1.1 [33].

Based on FDG-PET results, patients with CMR or PMR were considered as responders, whereas those with SMD or PMD were classified as non-responders. Similarly, patients with CR or PR on CT images were classified as responders, and those showing SD or PD were considered non-responders.

Statistical analysis

Progression-free survival (PFS) and overall survival (OS) served as the main outcome measures. PFS was defined as the time from the date of inclusion in the study to disease recurrence or progression. OS was defined as the time from the date of inclusion in the study to the date of death from any cause or last follow-up. Survival curves were plotted using the Kaplan–Meier method and compared using the log-rank test. Cox regression analysis was used to calculate the adjusted hazard ratios (HRs) and their corresponding 95 % confidence intervals (CIs). All calculations were performed using the

SPSS 18.0 statistical package (SPSS Inc., Chicago, IL, USA). Two-tailed P values < 0.05 were considered statistically significant.

Results

Patients

Between April 2009 and May 2012, we identified a total of 23 patients (16 women and 7 men) with advanced lung adenocarcinoma who were treated with erlotinib (Table 1). The median age at enrollment was 57 years. Most of the study patients (87 %) had an ECOG performance status of 0 or 1. The median follow-up time in the study cohort was 14 months (range, 1–51 months). At the end of the follow-up period, two patients survived and 21 had died. The two patients who survived had follow-up periods of 51 and 39 months, respectively.

Response on FDG-PET versus CT imaging according to the RECIST criteria

All patients underwent FDG-PET imaging on day 14; data for day 56 FDG-PET scans were missing for three participants. Based on FDG-PET imaging on day 14 and according to the EORTC criteria, 11 patients (48 %) had PMR, 11 (48 %) had SMD, and one (4 %) had PMD. Based on FDG-PET imaging on day 56, we identified five patients (25 %) with CMR, seven (35 %) with PMR, five (25 %) with SMD, and three (15 %) with PMD. Based on FDG-PET imaging on day 14 and according to the PERCIST criteria, six patients (26 %) had PMR, 15 (65 %) had SMD, and two (9 %) had PMD. According to FDG-PET imaging on day 56, we identified one patient (5 %) with CMR, eight patients (40 %) with PMR, seven patients (35 %) with SMD, and four patients (20 %) with PMD. Based on FDG-PET imaging on day 14 and according to the TLG-S criteria (summing the five hottest lesions), ten patients (26 %) had PMR, 11 (65 %) had SMD, and two (9 %) had PMD. According to FDG-PET imaging on day 56, we identified one patient (5 %) with CMR, eight patients (40 %) with PMR, eight patients (35 %) with SMD, and three patients (20 %) with PMD. Based on CT imaging on day 56 and according to the RECIST criteria, ten patients (26 %) had PR, five (65 %) had SD, and eight (9 %) had PD. Two patients classified as having PD and one patient who had PR did not undergo day 56 FDG-PET imaging.

The overall response according to early FDG-PET findings versus the standard CT response is summarized in Table 2. Four patients who were classified as responders based on CT imaging on day 56 and according to the RECIST criteria were considered non-responders when the PERCIST criteria were applied to early FDG-PET findings (Fig. 1). The overall response rate (43.5 %, 10 of 23 patients) obtained by applying the TLG-S system to early FDG-PET results was identical to that calculated by applying the RECIST criteria to CT data obtained on day 56. Eight patients with PD according to the RECIST criteria on day 56 were classified as non-responders when the PERCIST and TLG-S criteria were applied on early FDG-PET findings; however, one of these subjects was classified as a responder based on the EORTC criteria. Taking into account the missing FDG-PET results for three patients, tumor response based on the PERCIST and TLG-S criteria using day 56 FDG-PET data (9 responders and 11 non-responders) was the same as that observed when applying the RECIST criteria to CT findings (10 responders and 13 non-responders).

Table 1 General characteristics of the study patients

Characteristic	Patients, *n* (%)
Number of patients	23 (100)
Age (years)	
Median	57
Range	38–81
Sex	
Male	7 (30)
Female	16 (70)
Performance status	
0	7 (30)
1	13 (57)
2	3 (13)
AJCC clinical stage[a]	
IIIB	1 (4)
IV	17 (74)
Post-operative recurrence	5 (22)

[a] Seventh edition

AJCC American Joint Committee on Cancer

Table 2 Overall response according to early FDG-PET findings on day 14 versus standard CT response on day 56

		Day 56 RECIST criteria	
		Responder	Non-responder
Day 14 EORTC criteria	Responder	9	2
	Non-responder	1	11
Day 14 PERCIST criteria	Responder	6	0
	Non-responder	4	13
Day 14 TLG-S criteria	Responder	10	0
	Non-responder	0	13

RECIST Response Evaluation Criteria in Solid Tumors version 1.1, *EORTC* European Organization for Research and Treatment of Cancer, *PERCIST* PET Response Criteria in Solid Tumors, *TLG-S* Total Lesion Glycolysis-Systemic approach

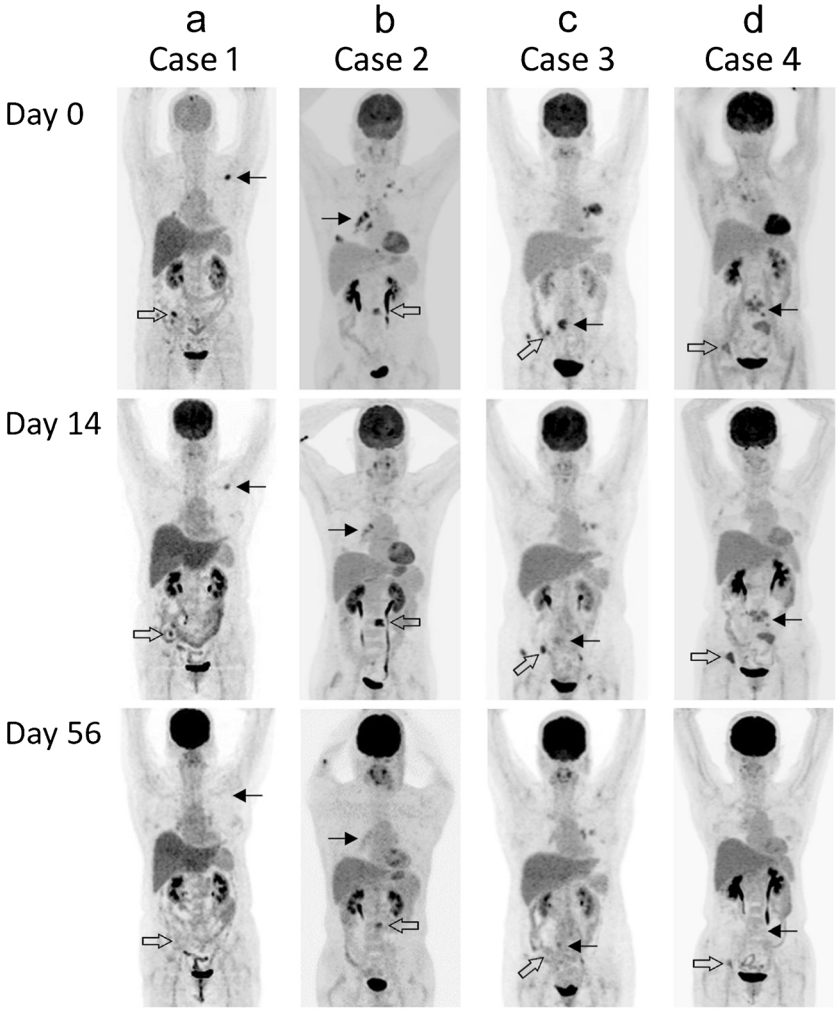

Fig. 1 Illustrative images of four non-responders according to PERCIST criteria on day 14 PET who had persistent bone uptake due to the bone flare effect during erlotinib treatment (case numbers in Fig. 1 correspond to those in Table 3). (**a**) In case 1, the hottest lesion was identified at the scapula (SUV_{max} 8.1; *arrow*) on day 0, and at the ilium (SUV_{max} 7.3; *hollow arrow*) on day 14. A complete metabolic response was observed on day 56. (**b**) In case 2, the hottest lesion was located at the mediastinal lymph nodes (SUV_{max} 15.3; arrow) on day 0, and at the L3 vertebra (SUV_{max} 11.5; *hollow arrow*) on day 14. On day 56, a partial metabolic response was observed, with tracer uptake decreased at the L3 vertebra (SUV_{max} 5.3; *hollow arrow*). (**c**) In case 3, the hottest lesion was identified at the L5 vertebra (SUV_{max} 10.3; *arrow*) on day 0, and at the sacroiliac junction (SUV_{max} 8.2; *hollow arrow*) on day 14. On day 56, decreased activity was observed at the L5 vertebra (SUV_{max} 3.2) (*arrow*), and the lesion located at the sacroiliac junction was not measurable (*hollow arrow*). (**d**) In case 4, the hottest lesion was identified at the lumbosacral spine (SUV_{max} 6.6; *arrow*) on day 0, and at the acetabulum (SUV_{max} 6.2; *hollow arrow*) on day 14. On day 56, a partial metabolic response was observed, with tracer uptake decreased at the acetabulum (SUV_{max} 3.5; *hollow arrow*)

Impact of bone flares on early assessment of treatment response from FDG-PET images using the PERCIST criteria

A total of 13 study patients had bone metastases (Table 3). Persistent bone uptake due to bone flares occurred in four patients (31 %), with the highest tracer uptake in the bone identified on day 14 FDG-PET. Persistent bone uptake was defined as a reduction in SUV_{max} of less than 30 % or an increase in SUV_{max} value. Such bone flares led to erroneous classification of these patients as non-responders when PERCIST criteria were applied. All of the bone flares regressed on day 56 FDG-PET images (Fig. 1). Notably, all four patients were correctly classified as responders according to either the EORTC or TLG-S criteria on day 14 (Fig. 2). All four patients identified as non-responders on day 14 according to the PERCIST criteria were classified as responders based on day 56 CT findings (group A in Table 3). Of the remaining nine patients who did not have bone flares, three patients were classified as responders (group B) and six as non-responders (group C) based on the PERCIST criteria applied to day 14 FDG-PET results and the RECIST approach applied to day 56 CT findings (Table 3).

Table 3 Changes in FDG uptake observed in bone lesions and in the hottest single lesions identified during erlotinib treatment among patients with lung cancer and skeletal metastases

Group	Case no.	Day 0 (SUV)	Day 14 (SUV)	Day 56 (SUV)
A[a]	1	Bone (8.1)[b]	Bone (7.3)[b]	Bone (−)
Day 14 PERCIST criteria: non-responders	2	Bone (11.3), MLN (15.3)[b]	Bone (11.5)[b]	Bone (5.3)[b]
Day 56 RECIST criteria: responders	3	Bone (10.3)[b]	Bone (8.2)[b]	Bone (3.2), MLN (5.0)[b]
	4	Bone (6.6)[b]	Bone (6.2)[b]	Bone (3.5)[b]
B	5	Bone (21.8)[b]	Bone (8.2)[b]	Bone (2.8), MLN (5.3)[b]
Day 14 PERCIST criteria: responders	6	Bone (9.2)[b]	Bone (5.3)[b]	Bone (3.6)[b]
Day 56 RECIST criteria: responders	7	Bone (6.6), liver (7.7)[b]	Bone (4.2), MLN (4.9)[b]	Bone (−), MLN (4.2)[b]
C	8	Bone (5.6), lung (7.1)[b]	Bone (8.8)[b]	Bone (11.9)[b]
Day 14 PERCIST criteria: non-responders	9	Bone (6.8)[b]	Bone (7.8)[b]	Bone (3.9), lung (6.1)[b]
Day 56 RECIST criteria: non-responders	10	Bone (9.4)[b]	Bone (9.0)[b]	Bone (6.8)[b]
	11	Bone (13.0)[b]	Bone (9.6)[b]	Bone (9.3)[b]
	12	Bone (8.4), lung (10.5)[b]	Bone (7.4), lung (11.3)[b]	N/A
	13	Bone (8.0), lung (13.4)[b]	Bone (8.8), lung (15.5)[b]	Bone (8.3), lung (16.4)[b]

PERCIST PET Response Criteria in Solid Tumors, *RECIST* Response Evaluation Criteria in Solid Tumors version 1.1, *MLN* mediastinal lymph node, *N/A* not available

[a] Group A was the bone flare group

[b] Hottest lesion identified during a whole-body PET scan

Prediction of progression-free survival

Patients who were classified as responders on day 14 based on TLG-S criteria had higher 2-year PFS (26.7 % vs. 0 %; $P = 0.007$ [log-rank test, Kaplan–Meier analysis], Fig. 3; HR = 0.28, 95 % CI = 0.10–0.76, $P = 0.012$). However, the assessment of early response on FDG-PET scans based on either EORTC or PERCIST criteria was not significantly associated with PFS. Using FDG-PET images obtained at day 56, we identified significant univariate associations between 2-year PFS and response according to both PERCIST (16.7 % vs. 0 %; $P = 0.044$ [log-rank test, Kaplan–Meier analysis]; HR = 0.37, 95 % CI = 0.13–1.03, $P = 0.057$) and TLG-S PERCIST (16.7 % vs. 0 %; $P = 0.044$ [log-rank test, Kaplan–Meier analysis]; HR = 0.37, 95 % CI = 0.13–1.03, $P = 0.057$) criteria. On day 56, CT response according to the RECIST was significantly associated with a higher 2-year PFS rate (26.7 % vs. 0 %; $P = 0.007$ [log-rank test, Kaplan–Meier analysis]; HR = 0.28, 95 % CI = 0.10–0.76, $P = 0.012$). However, FDG-PET response according to the EORTC criteria did not show a statistically significant association with PFS.

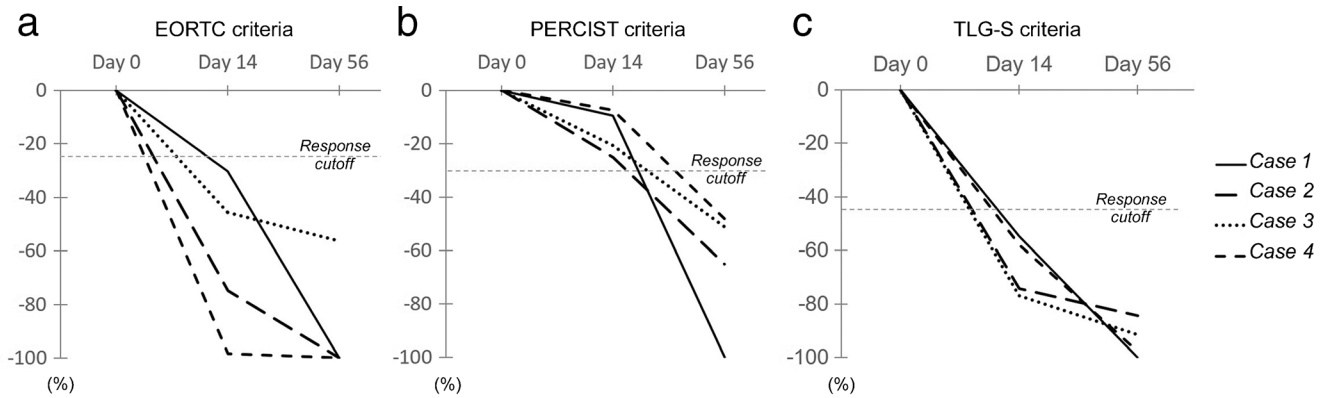

Fig. 2 (**a**) Percentage change in FDG uptake in the four patients with skeletal metastases who were wrongly classified as non-responders based on FDG-PET imaging at day 14 using the PERCIST criteria. (**b**) All patients were correctly classified as early responders according to the EORTC criteria. (**c**) The use of a systemic approach that included both primary and metastatic tumors (TLG-S method) was similarly effective in classifying these patients as early responders. The cutoff values for defining a reduction in FDG uptake as significant were 25, 3, and 45 % of baseline values for EORTC, PERCIST, and TLG-S criteria, respectively

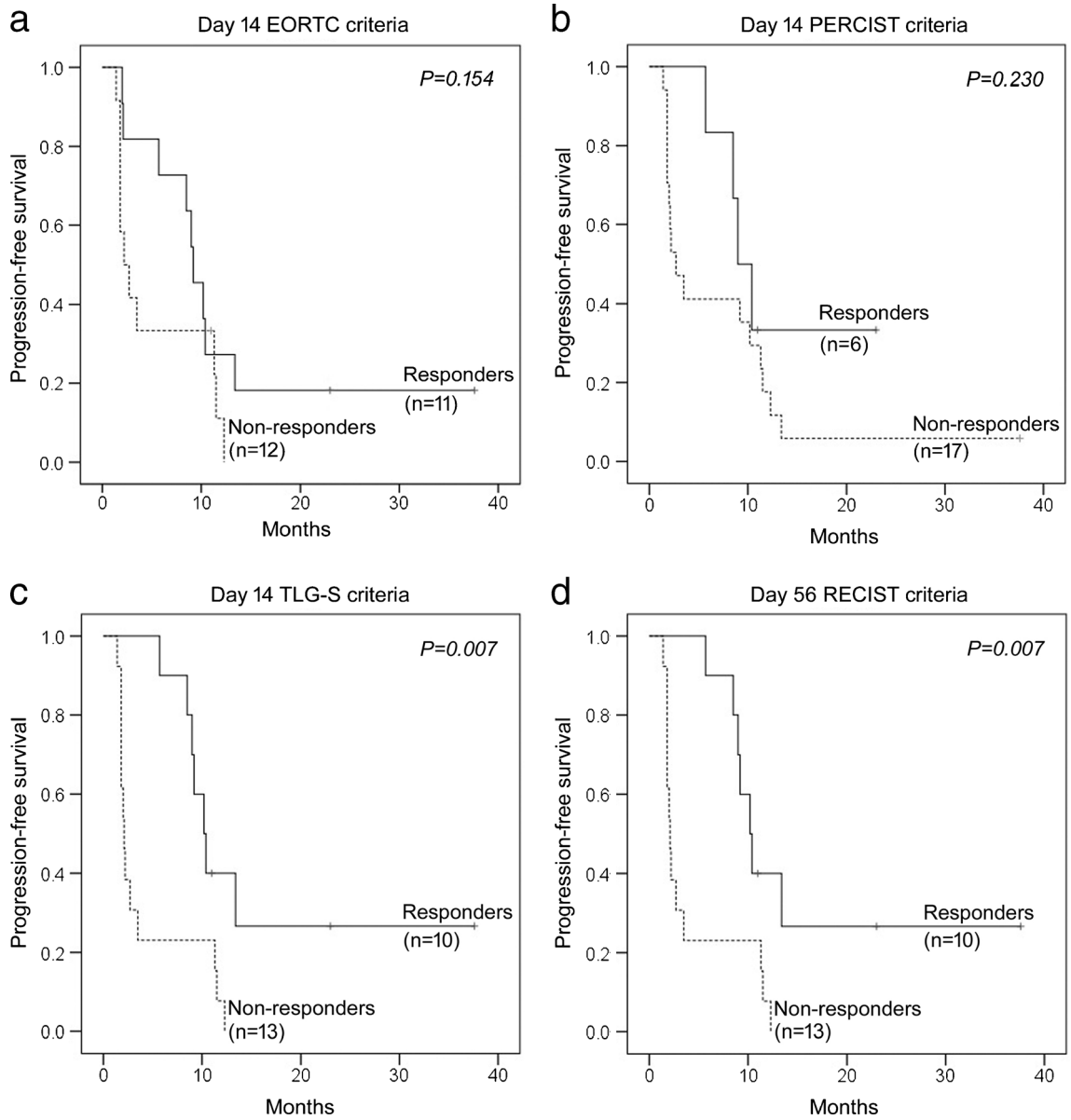

Fig. 3 Kaplan–Meier estimates of progression-free survival (PFS) according to different criteria used for assessing FDG-PET response. No significant differences in PFS were identified between responders and non-responders defined according to FDG-PET imaging at day 14 using either the EORTC criteria (**a**) or the PERCIST criteria (**b**). The higher rate of PFS for patients classified as responders according to FDG-PET imaging at day 14 using the TLG-S criteria (**c**) was identical to that according to CT imaging at day 56 using the RECIST criteria (**d**) ($P = 0.007$)

Prediction of overall survival

Patients who were classified as responders on day 14 based on EORTC criteria had higher 2-year OS (36.4 % vs. 8.3 %; $P = 0.015$ [log-rank test, Kaplan–Meier analysis]; HR = 0.32, 95 % CI = 0.12–0.83, $P = 0.020$). Similar findings were obtained when responders were identified using the TLG-S method (40.0 % vs. 7.7 %; $P = 0.018$ [log-rank test, Kaplan–Meier analysis], Fig. 4; HR = 0.32, 95 % CI = 0.12–0.86, $P = 0.024$). Early FDG-PET response according to the PERCIST criteria was not significantly associated with OS. On day 56, CT response based on the RECIST criteria was the only variable significantly associated with 2-year OS (40.0 % vs. 7.7 %; $P = 0.018$ [log-rank test, Kaplan–Meier analysis], Fig. 4; HR = 0.32, 95 % CI = 0.12–0.86, $P = 0.024$). Although patients classified as responders or non-responders according to PERCIST and TLG-S on day 56 were the same as those identified using the RECIST criteria, the association between FDG-PET response and OS was not significant because of missing data in three participants.

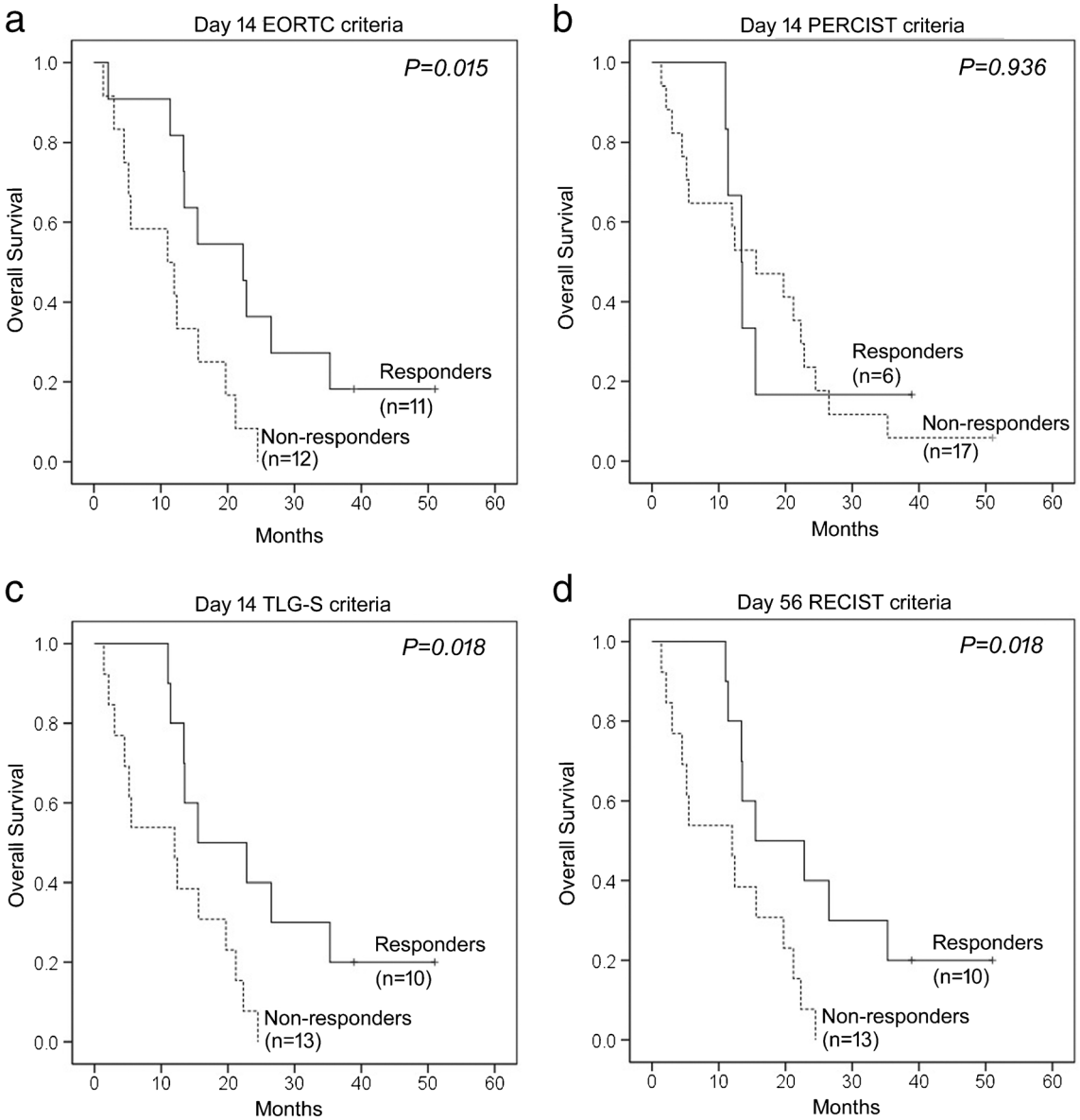

Fig. 4 Kaplan–Meier estimates of overall survival (OS) according to different criteria used for assessing FDG-PET response. OS rates were significantly better for patients classified as responders according to FDG-PET imaging on day 14 using the EORTC criteria (**a**). No significant differences in OS were identified between responders and non-responders defined according to FDG-PET imaging at day 14 using the PERCIST criteria (**b**). The higher rate of OS in patients classified as responders according to FDG-PET imaging at day 14 using the TLG-S criteria (**c**) was identical to that according to CT imaging at day 56 using the RECIST criteria (**d**) ($P = 0.018$)

Discussion

The PERCIST criteria use the hottest lesion on FDG-PET as the target, considering both the primary tumor and its distant metastases. In a study involving 22 patients, Benz et al. [16] reported that patients with PMD according to FDG-PET results obtained 2 weeks after the start of erlotinib treatment displayed a significantly shorter time to progression and poorer OS compared with those showing either SMD or PMR. A report by Zander et al. [15] demonstrated that the PERCIST criteria obtained using FDG-PET data acquired after 1 week of erlotinib therapy predicted both PFS and OS in patients with advanced NSCLC, independent of *EGFR* mutation status. Similar results were obtained when the EORTC criteria were applied [15]. However, the results of our study indicate that FDG-PET response on day 14 according to the PERCIST criteria was not significantly associated with PFS or OS. One possible explanation for these findings is that some patients classified as responders according to CT imaging on day 56 using the RECIST criteria were erroneously considered non-responders based on early FDG-PET results. Notably, patient classification errors were caused mainly by the presence of high skeletal tracer uptake on day 14, ultimately resulting in a much smaller reduction in SUV than that in other

study participants. Indeed, it should be noted that 1) all of these bone lesions disappeared by day 56, and 2) the four patients incorrectly classified by the PERCIST criteria were correctly identified as early responders according to both the EORTC and TLG-S criteria. Starting from these premises, bone flares are a plausible explanation for misclassification when PERCIST criteria are used.

In this study, we observed that the presence of "persistent bone uptake" in patients who showed primary tumor response resulted in erroneous categorization of four patients (group A in Table 3). Although the discrepancy between primary tumor (response) and bone metastasis (no-response) on day 14 may be caused by tumor heterogeneity, bone uptake was either greatly reduced or absent on day 56 in group A patients. Based on these findings, we reasoned that the occurrence of bone flares would be the most plausible mechanism to explain "discordant persistent" tracer uptake in the bone. However, the peak time of bone flare can be influenced by several factors (e.g., tracer, tumor type, and drugs). Numerous data on bone flares are available from bone scintigraphy studies, but less information is available on their occurrence in FDG-PET images. In this study, we defined persistent bone uptake as a reduction in SUV_{max} of less than 30 % or an increase in SUV_{max} value. Persistent bone uptake observed on day 14 imaging may have occurred before or after peak time. Consequently, non-peak persistent bone uptake was attributed to the bone flare phenomenon. The clinical significance of this phenomenon is still a matter of debate [34]. Osteoblastic bone flares were previously described as transiently worsening bone lesions on FDG-PET scans in a case series of four NSCLC patients treated with bevacizumab [35]. A study using CT imaging and the RECIST criteria identified the occurrence of osteoblastic bone flares in three NSCLC patients who received erlotinib [29]. Another study reported that 21 % of NSCLC patients undergoing bone scintigraphy developed bone flares during therapy with TKIs [30]. In the present study, bone flares were observed in 31 % of patients with skeletal metastases on FDG-PET scans performed on day 14. However, a case report of the use of FDG-PET for assessing response to erlotinib indicated that disease progression may be misdiagnosed as a bone flare as well [36]. In our study, six non-responders with persistent bone lesions on day 14 had stable disease on day 56.

Consistent with previous reports [3–9], the results of our study demonstrate that assessment of early FDG-PET response using the EORTC criteria predicts OS in NSCLC patients treated with erlotinib. In our report, the number and timing of FDG-PET scans (at baseline and on days 14 and 56) were in line with the protocol utilized by Mileshkin et al. [3]. Interestingly, these authors reported that FDG-PET response according to the EORTC criteria on day 14 was significantly associated with a better OS, whereas the same response on day 56 was not. Nonetheless, the biological heterogeneity between the primary tumor and its metastatic progeny, as well as the intermetastatic heterogeneity [37], had not previously been adequately taken into account. We thus reasoned that a systemic assessment that included the metastatic sites could be superior to the exclusive local assessment of primary tumor response according to the EORTC criteria. Our findings support the original study hypothesis. Accordingly, a systemic approach based on the TLG-S method (including both primary and metastatic tumors up to a total of five target lesions) identified a significant association between early FDG-PET response and survival endpoints (PFS and OS).

In this study, in line with the PERCIST criteria, we defined PMR as a reduction of at least 45 % in TLG-S, whereas PMD was diagnosed in the presence of a 75 % or higher increase in this parameter. Kahraman et al. [18] had previously shown that the TLG-S percentage change was a strong predictor of survival outcomes in NSCLC patients treated with erlotinib. In that study, the authors defined TLG-S as the sum of up to five measurable target lesions; different cutoff values for defining the metabolic response were calculated as well [18]. Another recent report demonstrated that high TLG-S values were an independent predictor of survival in patients with advanced NSCLC who received erlotinib [19]. In that study, TLG-S was calculated by taking into account all of the measurable lesions in whole-body scans; in addition, TLG-S was dichotomized according to the median value [19]. Together, these findings indicate strong prognostic significance of TLG-S, although neither the extent of target lesions nor the definition of metabolic response have yet been standardized.

Some controversy still exists regarding the discordance in *EGFR* and *K-RAS* mutation status between primary and metastatic tumors among NSCLC patients [22–24]. Therefore, local imaging assessment of the primary tumor has been largely supported by reports showing that a heterogeneous distribution of *EGFR* mutations rarely occurs [38, 39]. In our study, we demonstrate that FDG-PET response based on the EORTC criteria was associated with OS but not PFS. It should be noted, however, that small core biopsies may not reflect the clonal heterogeneity of the entire tumor [40]. Moreover, intratumor heterogeneity (consisting of a mixed population of *EGFR*-mutated and wild-type cells) may reduce the response to TKIs [41]. Significant heterogeneity in *EGFR* mutation status between primary lung tumors and their metastases can also cause a mixed response to TKIs in certain patients [28, 40]. At the imaging level, intratumor heterogeneity of FDG uptake has been associated with tumor response and clinical outcomes in patients with NSCLC treated with erlotinib [11]. Based on these findings, we believe that systemic approaches including distant metastases will be superior to single-site assessments when this patient group undergoes FDG-PET imaging. Our current findings using the TLG-S method support this contention quite strongly.

Conclusions

The results of the present study demonstrate that TLG-S criteria may be superior to other forms of assessment in predicting PFS and OS based on early FDG-PET response. Bone flares, which can interfere with the interpretation of treatment response using PERCIST criteria, are not uncommon in patients with metastatic lung adenocarcinoma treated with erlotinib.

Acknowledgments This study was financially supported by grants from the Chang Gung Memorial Hospital (CMRPG380271) and the National Science Council (NSC 99-2314-B-182A-082-MY2), Taiwan.

Compliance with ethical standards

Conflict of interest The authors declare that they have no conflict of interest.

Ethical approval All procedures performed in studies involving human participants were in accordance with the ethical standards of the institutional and/or national research committee and with the 1964 Declaration of Helsinki and its later amendments or comparable ethical standards.

Informed consent Informed consent was obtained from all participants included in the study.

References

1. Lynch TJ, Bell DW, Sordella R, Gurubhagavatula S, Okimoto RA, Brannigan BW, et al. Activating mutations in the epidermal growth factor receptor underlying responsiveness of non-small-cell lung cancer to gefitinib. N Engl J Med. 2004;350:2129–39.
2. Cappuzzo F, Ciuleanu T, Stelmakh L, Cicenas S, Szczesna A, Juhasz E, et al. Erlotinib as maintenance treatment in advanced non-small-cell lung cancer: a multicentre, randomised, placebo-controlled phase 3 study. Lancet Oncol. 2010;11:521–9.
3. Mileshkin L, Hicks RJ, Hughes BG, Mitchell PL, Charu V, Gitlitz BJ, et al. Changes in 18F-Fluorodeoxyglucose and 18F-Fluorodeoxythymidine Positron Emission Tomography Imaging in Patients with Non-Small Cell Lung Cancer Treated with Erlotinib. Clin Cancer Res. 2011;17:3304–15.
4. Binns DS, Pirzkall A, Yu W, Callahan J, Mileshkin L, Conti P, et al. Compliance with PET acquisition protocols for therapeutic monitoring of erlotinib therapy in an international trial for patients with non-small cell lung cancer. Eur J Nucl Med Mol Imaging. 2011;38:642–50.
5. Aukema TS, Kappers I, Olmos RA, Codrington HE, van Tinteren H, van Pel R, et al. Is 18F-FDG PET/CT useful for the early prediction of histopathologic response to neoadjuvant erlotinib in patients with non-small cell lung cancer? J Nucl Med. 2010;51:1344–8.
6. O'Brien ME, Myerson JS, Coward JI, Puglisi M, Trani L, Wotherspoon A, et al. A phase II study of (1)(8)F-fluorodeoxyglucose PET-CT in non-small cell lung cancer patients receiving erlotinib (Tarceva); objective and symptomatic responses at 6 and 12 weeks. Eur J Cancer. 2012;48:68–74.
7. Schaake EE, Kappers I, Codrington HE, Valdes Olmos RA, Teertstra HJ, van Pel R, et al. Tumor response and toxicity of neo-adjuvant erlotinib in patients with early-stage non-small-cell lung cancer. J Clin Oncol. 2012;30:2731–8.
8. van Gool MH, Aukema TS, Schaake EE, Rijna H, Valdes Olmos RA, van Pel R, et al. Timing of metabolic response monitoring during erlotinib treatment in non-small cell lung cancer. J Nucl Med. 2014;55:1081–6.
9. van Gool MH, Aukema TS, Schaake EE, Rijna H, Codrington HE, Valdes Olmos RA, et al. (18)F-fluorodeoxyglucose positron emission tomography versus computed tomography in predicting histopathological response to epidermal growth factor receptor-tyrosine kinase inhibitor treatment in resectable non-small cell lung cancer. Ann Surg Oncol. 2014;21:2831–7.
10. Kobe C, Scheffler M, Holstein A, Zander T, Nogova L, Lammertsma AA, et al. Predictive value of early and late residual 18F-fluorodeoxyglucose and 18F-fluorothymidine uptake using different SUV measurements in patients with non-small-cell lung cancer treated with erlotinib. Eur J Nucl Med Mol Imaging. 2012;39:1117–27.
11. Cook GJ, O'Brien ME, Siddique M, Chicklore S, Loi HY, Sharma B, et al. Non-small cell lung cancer treated with erlotinib: heterogeneity of F-FDG Uptake at PET-association with treatment response and prognosis. Radiology. 2015;276:883–93.
12. de Langen AJ, van den Boogaart V, Lubberink M, Backes WH, Marcus JT, van Tinteren H, et al. Monitoring response to antiangiogenic therapy in non-small cell lung cancer using imaging markers derived from PET and dynamic contrast-enhanced MRI. J Nucl Med. 2011;52:48–55.
13. Hachemi M, Couturier O, Vervueren L, Fosse P, Lacoeuille F, Urban T, et al. [(1)(8)F]FDG positron emission tomography within two weeks of starting erlotinib therapy can predict response in non-small cell lung cancer patients. PLoS One. 2014;9:e87629.
14. Bengtsson T, Hicks RJ, Peterson A, Port RE. 18F-FDG PET as a surrogate biomarker in non-small cell lung cancer treated with erlotinib: newly identified lesions are more informative than standardized uptake value. J Nucl Med. 2012;53:530–7.
15. Zander T, Scheffler M, Nogova L, Kobe C, Engel-Riedel W, Hellmich M, et al. Early prediction of nonprogression in advanced non-small-cell lung cancer treated with erlotinib by using [(18)F]fluorodeoxyglucose and [(18)F]fluorothymidine positron emission tomography. J Clin Oncol. 2011;29:1701–8.
16. Benz MR, Herrmann K, Walter F, Garon EB, Reckamp KL, Figlin R, et al. (18)F-FDG PET/CT for monitoring treatment responses to the epidermal growth factor receptor inhibitor erlotinib. J Nucl Med. 2011;52:1684–9.
17. Suleiman AA, Frechen S, Scheffler M, Zander T, Kahraman D, Kobe C, et al. Modeling tumor dynamics and overall survival in advanced non-small-cell lung cancer treated with erlotinib. J Thorac Oncol. 2015;10:84–92.
18. Kahraman D, Holstein A, Scheffler M, Zander T, Nogova L, Lammertsma AA, et al. Tumor lesion glycolysis and tumor lesion proliferation for response prediction and prognostic differentiation in patients with advanced non-small cell lung cancer treated with erlotinib. Clin Nucl Med. 2012;37:1058–64.
19. Winther-Larsen A, Fledelius J, Sorensen BS, Meldgaard P. Metabolic tumor burden as marker of outcome in advanced EGFR wild-type NSCLC patients treated with erlotinib. Lung Cancer. 2016;94:81–7.
20. Young H, Baum R, Cremerius U, Herholz K, Hoekstra O, Lammertsma AA, et al. Measurement of clinical and subclinical tumour response using [18F]-fluorodeoxyglucose and positron emission tomography: review and 1999 EORTC recommendations. Eur J Cancer. 1999;35:1773–82.

21. Wahl RL, Jacene H, Kasamon Y, Lodge MA. From RECIST to PERCIST: evolving considerations for PET response criteria in solid tumors. J Nucl Med. 2009;50:122S–50.
22. Schmid K, Oehl N, Wrba F, Pirker R, Pirker C, Filipits M. EGFR/KRAS/BRAF mutations in primary lung adenocarcinomas and corresponding locoregional lymph node metastases. Clin Cancer Res. 2009;15:4554–60.
23. Kalikaki A, Koutsopoulos A, Trypaki M, Souglakos J, Stathopoulos E, Georgoulias V, et al. Comparison of EGFR and K-RAS gene status between primary tumours and corresponding metastases in NSCLC. Br J Cancer. 2008;99:923–9.
24. Takahashi K, Kohno T, Matsumoto S, Nakanishi Y, Arai Y, Yamamoto S, et al. Clonal and parallel evolution of primary lung cancers and their metastases revealed by molecular dissection of cancer cells. Clin Cancer Res. 2007;13:111–20.
25. Marusyk A, Almendro V, Polyak K. Intra-tumour heterogeneity: a looking glass for cancer? Nat Rev Cancer. 2012;12:323–34.
26. Stoecklein NH, Klein CA. Genetic disparity between primary tumours, disseminated tumour cells, and manifest metastasis. Int J Cancer. 2010;126:589–98.
27. Klein CA. Parallel progression of primary tumours and metastases. Nat Rev Cancer. 2009;9:302–12.
28. Chen ZY, Zhong WZ, Zhang XC, Su J, Yang XN, Chen ZH, et al. EGFR mutation heterogeneity and the mixed response to EGFR tyrosine kinase inhibitors of lung adenocarcinomas. Oncologist. 2012;17:978–85.
29. Lind JSW, Postmus PE, Smit EF. Osteoblastic bone lesions developing during treatment with erlotinib indicate major response in patients with non-small cell lung cancer: a brief report. J Thorac Oncol. 2010;5:554–7.
30. Chao HS, Chang CP, Chiu CH, Chu LS, Chen YM, Tsai CM. Bone scan flare phenomenon in non-small-cell lung cancer patients treated with gefitinib. Clin Nucl Med. 2009;34:346–9.
31. Hong R, Halama J, Bova D, Sethi A, Emami B. Correlation of PET standard uptake value and CT window-level thresholds for target delineation in CT-based radiation treatment planning. Int J Radiat Oncol Biol Phys. 2007;67:720–6.
32. Larson SM, Erdi Y, Akhurst T, Mazumdar M, Macapinlac HA, Finn RD, et al. 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. 1999;2:159–71.
33. Eisenhauer EA, Therasse P, Bogaerts J, Schwartz LH, Sargent D, Ford R, et al. New response evaluation criteria in solid tumours: revised RECIST guideline (version 1.1). Eur J Cancer. 2009;45:228–47.
34. Fogelman I. The flare phenomenon: still learning after 35 years. Eur J Nucl Med Mol Imaging. 2011;38:5–6.
35. Krupitskaya Y, Eslamy HK, Nguyen DD, Kumar A, Wakelee HA. Osteoblastic bone flare on F18-FDG PET in non-small cell lung cancer (NSCLC) patients receiving bevacizumab in addition to standard chemotherapy. J Thorac Oncol. 2009;4:429–31.
36. Al-Nabhani K, Syed R, Haroon A, Almukhailed O, Bomanji J. Flare response versus disease progression in patients with non-small cell lung cancer. J Radiol Case Rep. 2012;6:34–42.
37. Jamal-Hanjani M, Quezada SA, Larkin J, Swanton C. Translational Implications of Tumor Heterogeneity. Clin Cancer Res. 2015;21:1258–66.
38. Yatabe Y, Matsuo K, Mitsudomi T. Heterogeneous distribution of EGFR mutations is extremely rare in lung adenocarcinoma. J Clin Oncol. 2011;29:2972–7.
39. Vignot S, Frampton GM, Soria JC, Yelensky R, Commo F, Brambilla C, et al. Next-generation sequencing reveals high concordance of recurrent somatic alterations between primary tumor and metastases from patients with non-small-cell lung cancer. J Clin Oncol. 2013;31:2167–72.
40. Thomas A, Rajan A, Lopez-Chavez A, Wang Y, Giaccone G. From targets to targeted therapies and molecular profiling in non-small cell lung carcinoma. Ann Oncol. 2013;24:577–85.
41. Taniguchi K, Okami J, Kodama K, Higashiyama M, Kato K. Intratumor heterogeneity of epidermal growth factor receptor mutations in lung cancer and its correlation to the response to gefitinib. Cancer Sci. 2008;99:929–35.