The utility of ^{18}F-FDG PET/CT for evaluation of tumor response to immune checkpoint inhibitor therapy and prognosis prediction in patients with non-small-cell lung cancer

Kazuhiro Kitajima[1] MD,
Yusuke Kawanaka[1] MD,
Hisashi Komoto[1] MD,
Toshiyuki Minami[2] MD,
Takashi Yokoi[2] MD,
Kozo Kuribayashi[2] MD,
Takashi Kijima[2] MD,
Akifumi Nakamura[3] MD,
Masaki Hashimoto[3] MD,
Nobuyuki Kondo[3] MD,
Seiki Hasegawa[3] MD,
Koichiro Yamakado[1] MD

1. Department of Radiology, Hyogo College of Medicine, Hyogo, Japan
2. Division of Respiratory Medicine and Hematology, Department of Internal Medicine, Hyogo College of Medicine, Nishinomiya, Hyogo, Japan
3.Department of Thoracic Surgery, Hyogo College of Medicine, Nishinomiya, Hyogo, Japan

Keywords: NSCLC -^{18}F-FDG - PET/CT - EORTC - PERCIST - RECIST

Corresponding author:
Kazuhiro Kitajima MD,
Department of Radiology, Hyogo College of Medicine,
Nishinomiya, Hyogo, Japan
1-1 Mukogawa-cho, Nishinomiya,
Hyogo 663-8501 Japan
Phone:+81-798-45-6883,
Fax:+81-798-45-6262
kazu10041976@yahoo.co.jp

Received:
10 August 2021
Accepted revised:
27 September 2021

Abstract

Objective: To compare three fluorine-18-fluorodeoxyglucose positron emission tomography (^{18}F-FDG PET) (EORTC criteria and PERCIST) and computed tomography (CT) (RECIST1.1) for response evaluation and prognosis prediction in non-small-cell lung cancer (NSCLC) patients treated with immune checkpoint inhibitor (ICI) monotherapy. **Subjects and Methods:** Forty NSCLC patients underwent ^{18}F-FDG PET/CT scans at baseline and after 4 to 8 cycles of nivolumab or pembrolizumab. Therapeutic response was evaluated according to EORTC criteria, PERCIST, and RECIST1.1,then concordance among those was assessed using Cohen's κ coefficient. Progression-free survival (PFS) and overall survival (OS) was examined using log-rank and Cox methods. **Results:** The number of complete metabolic response (CMR)/partial metabolic response (PMR)/stable metabolic disease (SMD)/progressive metabolic disease (PMD) were 8/10/4/18 for EORTC criteria and 9/9/4/18 for PERCIST. Using RECIST1.1, those of CR/PR/SD/PD were 4/10/12/14. Although there was high concordance between PERCIST and EORTC (92.5% of patients; κ=0.924), that between PERCIST and RECIST1.1 was substantial (65.0%; κ=0.560) and that between EORTC and RECIST1.1 (65.0%; κ=0.574). After a median 23.2 months (range 7.2 to 51.8 months), 32 patients had documented progression and 24 patients died from NSCLC. According to both PET and CT, patients with no progression (CMR/PMR/SMD or CR/PR/SD) showed significantly longer PFS and OS than PMD or PD patients (EORTC: P<0.0001 and P<0.0001, respectively, PERCIST: P<0.0001 and P=0.0001, respectively, RECIST1.1: P<0.0001 and P<0.0001, respectively). In a univariate analysis total MTV (P=0.042) on pre-ICI treatment ^{18}F-FDGPET/CT scans was significantly associated with progression. Highest SUVmax (P<0.0001), total MTV (P=0.0062), total TLG (P<0.0001), highest SULpeak (P<0.0001), and total TLGL (P<0.0001) on post-ICI treatment ^{18}F-FDG PET/CT scans were also were significantly associated with progression. Moreover, the change rate of highest SUVmax (P<0.0001), total metabolic tumor volume (MTV) (P<0.0001), total lesion glycolysis (TLG) (P<0.0001), highest SULpeak (P<0.0001), total TLGL (P<0.0001), size (P=0.0012), EORTC (P<0.0001), PERCIST (P<0.0001), and RECIST 1.1 (P<0.0001) on two PET/CT scans were significantly associated with progression. A multivariate analysis confirmed the change rate of total MTV (P=0.034), and total TLGL (P=0.0027), EORTC (P=0.018), PERCIST (P=0.045), and RECIST1.1 (P=0.0037) as independent negative PFS predictors. **Conclusion:** Both ^{18}F-FDG PET (EORTC criteria and PERCIST) and CT (RECIST1.1) after 4 to 8 ICI monotherapy cycles are accurate for evaluation of tumor response and predicting prognosis in NSCLC patients.

Hell J Nucl Med 2021; 24(3): 186-198 *Epub ahead of print: 17 December 2021* *Published online: 28 December 2021*

Introduction

Non-small-cell lung cancer (NSCLC) is accounts for approximately 85% of all lung cancer cases and has the highest incidence amongst malignancies. It is the leading cause of cancer-related death worldwide. Treatment options available for patients with inoperable or recurrent NSCLC who are candidates for systemic therapy include platinum-based cytotoxic chemotherapy regimens, molecular targeted agents for patients who carry specific driver mutations and, more recently, immune checkpoint inhibitors (ICI). Specifically, immunotherapy with antibodies that prevent the interaction of the programmed death ligand-1 (PD-L1) with the programmed cell death-1 (PD-1) receptor, thus releasing T cells to eliminate tumor cells, has led to significantly improved survival in patients with NSCLC [1-3].

An adequate assessment of response to systemic treatment is crucial for effective cancer treatment management, as efficient monitoring of tumor responsiveness to systemic therapy is essential to mitigate high mortality risk and the cytotoxic effects of systemic therapeutics. Present techniques for monitoring therapeutic response are typically based on anatomical changes seen with computed tomography (CT) imaging or other anatomical imaging methods, and the response evaluation criteria in solid tumors (RECIST) was updated by the World Health Organization in 2009 (version 1.1) [4]. However,

anatomical imaging may be limited in regard to its applicability to distinguish viable residual tumors from reactive changes, such as those related to edema and scar tissue, or killed cells and shrunken tumors. Functional evaluations of metabolic activity can be performed with fluorine-18-fluorodeoxyglucose (^{18}F-FDG)-positron emission tomography (PET), and its use is considered helpful to overcome these limitations and believed to be more suitable for therapeutic response examinations in patients receiving cytotoxic chemotherapy and molecular targeted agents, which causes direct reduction in the viability of tumor cells. However, the mechanism of action of immunotherapy is different from that of classical cytotoxic drugs, and is based on stimulation of the host's immune response against cancer cells, which may result in the development of inflammation at the tumor site and subsequent antitumor response [5]. The assessment of ICI therapeutic efficacy is challenging, and the role of ^{18}F-FDG PET is not well established.

Quantitative assessment of treatment response using ^{18}F-FDG PET is based on differences in the standardized uptake value (SUV) between baseline and follow-up examinations. The European Organization for Research and Treatment of Cancer (EORTC) developed in 1999 recommends using SUV normalized to body surface area (SUVbsa) to reduce the influence of body weight on SUV, usually maximum SUV (SUVmax) [6], though SUV can be affected by various aspects, such as technical, physical, and biological factors. To improve reproducibility for comparisons of results from separate trials, a widely accepted standardized protocol is needed. For this reason, positron emission tomography response criteria in solid tumors (PERCIST) 1.0, which uses peak lean body mass SUV (SULpeak), was developed in 2009 [7].The purpose of this retrospective study was to compare two functional ^{18}F-FDG PET criteria (EORTC criteria and PERCIST) with morphological CT criteria (RECIST1.1) after 4 to 8 ICI therapy cycles for response evaluation and prognosis prediction using serial ^{18}F-FDGPET/CT in patients with NSCLC who were treated with ICI therapy.

Subjects and Methods

Patients

This retrospective study was conducted after receiving approval from a local review board and waived the requirement for patient-informed consent. Between January 2016 and December 2018,all patients with a histological diagnosis of NSCLC and who were treated with ICI therapy were retrospectively selected from our database. A total of 40 patients (mean age 69.1±7.9 years, range 50-88 years) who underwent two ^{18}F-FDG PET/CT examinations in our institution at the baseline and again after 4-8 cycles of anti-PD-1 antibody therapy (i.e. nivolumab or pembrolizumab monotherapy) for the evaluation of treatment response were included in the present analysis (4 cycles in 8 patients, 5 cycles in 8 patients, 6 cycles in 7 patients, 7 cycles in 11 patients, and 8 cycles in 6 patients). Baseline ^{18}F-FDG PET/CT scans were obtained at a median of 1.4 months (range, 0.3-2.7 months) before initiation of anti-PD-1 antibody therapy. The median interval from initiation of ICI therapy to second ^{18}F-FDG PET/CT scan was 4.0 months (range, 2.1-5.6 months). Patient and tumor characteristics are shown in Table 1.

Table 1. *Patient characteristics.*

Character	N	%
Sex		
Male	32	80.0%
Female	8	20.0%
Age		
Mean	69.1±7.9	
Range	50-88	
Histological subtypes		
Adenocarcinoma	19	47.5%
Squamous cell carcinoma	17	42.5%
Adenosquamous cell carcinoma	2	5.0%
Plemorphiccarcinioma	2	5.0%
EGFR mutation		
Presence/none/unknown	7/21/12	17.5%/52.5%/30.0%
ALK mutation		
Presence/none/unknown	1/33/6	2.5%/82.5%/15.0%
PD-L1 expression (Tumor proportion score)		
0%	3	7.5%
1~50%	13	32.5%
51~100%	16	40.0%
Unknown	8	20.0%
Initial stage		
I	6	15.0%
II	7	17.5%
III	11	27.5%
IV	16	40.0%

(continued)

Previous treatment		
None	4	10.0%
Surgery and chemotherapy	10	25.0%
Surgery and chemotherapy plus radiotherapy	4	10.0%
Chemotherapy	13	32.5%
Chemotherapy and radiotherapy	4	10.0%
Chemoradiotherapy and chemotherapy	5	12.5%
ICI therapy regimen		
Nivolumab	30	75.0%
Pembrolizumab	10	25.0%

EGFR: Epidermal Growth Factor Receptor, ICI: immune checkpoint inhibitor

The ICI therapy regimen used was nivolumab (n=30 patients) orpembrolizumab (n=10). Nivolumab was administered intravenously at a dose of 3mg/kg every 2 weeks and pembrolizumab was administered intravenously at a dose of 2mg/kg every 3 weeks. This anti-PD-1 antibody therapy was administered until apparent disease progression or unacceptable toxicity occurred, or the decision to discontinue was made by the patient or attending physician. The total number of cycles of ICI treatment was 4 to 8 for 13 patients, 9 to 18 for 12 patients, 19 to 36 for 9 patients, and more than 36 for 6 patients.

^{18}F-FDG PET/CT

The ^{18}F-FDG PET/CT examinations were performed using one of four PET/CT scanners (Gemini GXL16, Gemini TF64, or Ingenuity TF; Philips Medical Systems, Eindhoven, The Netherlands and Discovery IQ; GE Healthcare, Waukesha, WI, USA). Briefly, patients were instructed to fast for 5 hours before the scan, then blood glucose was measured immediately prior to injection of ^{18}F-FDG at 4.0MBq/kg body weight for GXL16 and 3.0MBq/kg for TF64, and 3.7MBq/kg body weight for Ingenuity TF and Discovery IQ. All in the present cohort had a blood glucose level lower than 160 mg/dL. Approximately 60 minutes after the injection, static emission images were obtained. Helical CT scan images from the top of the head to the mid-thigh were obtained for attenuation correction and anatomic localization using the following parameters: tube voltage 120kV (all four scanners), effective tube current auto-mA up to 120mA (GXL16) 100mA (TF64), 155mA (Ingenuity TF) or 15~390mA [Smart mA : Noise Index: 25] (Discovery IQ), gantry rotation speed 0.5 seconds, detector configuration 16×1.5mm (GXL16) or 64×0.625mm (TF64 and Ingenuity TF), 16x1.25mm (Discovery IQ), slice thickness 2mm, and transverse field of view 600mm (GXL16, TF64, Ingenuity TF) or 700mm (Discovery IQ). Immediately upon completion of the CT examination, PET images from the head to mid-thigh were acquired for 90 seconds (GXL16, TF64, and Ingenuity TF) or 180 seconds (Discovery IQ) per bed position obtained in 3-dimensional mode. During PET scanning, the patient was allowed to breathe normally. Attenuation-corrected PET images were reconstructed with a line-of-response row-action maximum likelihood algorithm for GXL16, and an ordered-subset expectation maximization (OSEM) iterative reconstruction algorithm (33 subsets, 3 iterations) was used for TF64 and Ingenuity, while Q.Clear (block sequential regularized expectation maximization (BSREM) (β=400) was used for Discovery IQ.

Image analysis

Fluorine-18-FDG PET/CT images were retrospectively reviewed by one experienced physician board-certified in both diagnostic radiology and nuclear medicine with 12 years of experience with oncologic ^{18}F-FDG PET/CT, without knowledge of the other imaging results, or clinical and histopathologic data of the present patients. The commercial software package GI-PET (AZE Co., Ltd., Tokyo, Japan), capable of harmonizing SUV obtained with different PET/CT systems using phantom data was used to evaluate the treatment PET response (EORTC and PERCIST) [8,9]. Maximum SUV (SUVmax) was defined as the maximum concentration in the target lesion (injected dose/body weight). Peak SUV was calculated using a 1.2-cm diameter volume region of interest (ROI) placed on the hottest site of the tumor, then normalized to SULpeak (SUVpeak×[lean body mass]/[total body mass]). Metabolic tumor volume (MTV) was defined as ^{18}F-FDG-avid tumor volume, with the margin threshold set at 40% of SUVmax. Total lesion glycolysis (TLG) was then calculated as follows: SUVmean × MTV, with SUVmean representing the mean SUV value. Total lesion glycolysis lean (TLGL) was then calculated as follows: SULmean × MTV, with SULmean representing the mean SUL value.

Highest SUVmax and highest SULpeak were defined as the highest SUVmax and SULpeak of all tumors per patient, respectively. By summing the corresponding values for each lesion in the body, total MTV, total TLG and total TLGL measurements were computed.

RECIST1.1

RECIST 1.1 was used for morphological response evaluation [4]. Target lesion was defined as ≥1cm well-defined lesion for soft tissue in longest axis and ≥1.5cm in shortest axis for lymph node. The largest sum of diameter of five target lesions with maximum two lesions per organ was evaluated. Sclerotic or lytic/sclerotic (mixed type) bone metastases were considered non-measurable lesions. Greater than or equal to 30% decrease in the largest sum of diameter was considered as partial response (PR) while ≥20% increase was considered as PD. Change in-between PR and PD (<−30% and <+20%) was considered as stable disease (SD). The appearance of new lesion was considered as PD.

EORTC

According to the EORTC criteria [6], complete resolution of ^{18}F-FDG uptake within the tumor volume so that it is indistinguishable from surrounding normal tissue was conside-

red to be complete metabolic response (CMR). On the other hand, the appearance of new ^{18}F-FDG uptake in another region on the second ^{18}F-FDG PET scan was classified as progression metabolic disease (PMD). EORTC recommends using the pre-treatment scan to define regions on high ^{18}F-FDG uptake that represent viable tumor, and also recommends to use the same ROI volumes on subsequent scans, positioned as close to original tumor as possible and to measure maximal tumor ROI counts per pixel per second calibrated as MBq/L. As EORTC gives no information about the right number of lesions to measure, we chose up to 5 of the lesions with the highest ^{18}F-FDG uptake and up to two lesions per organ and measured the same lesions on the subsequent follow-up scan [10]. All 5 targets SUVmax measurements were summed on each scan, giving ΣSUVmax. A percentage change in baseline and second summed SUVmax was calculated. The patients who underwent more than one cycle of systemic treatment were then classified into 3 response groups defined in EORTC. Partial metabolic response (PMR) was defined as a 25% or greater reduction in highest SUVmax. An increase in tumor SUVmax of 25% or more within the ROI defined with the baseline scan was classified as PMD, while stable metabolic disease (SMD) was classified as an increase in highest SUVmax of less than 25% or a decrease of less than 25%.

PERCIST

To determine therapeutic response with PERCIST [7], SUL values were calculated using a 1.2-cm diameter volume ROI placed on the target lesion. We also determined whether the SULpeak value of the tumor was 1.5 times or more than that of the liver SUL (mean + 2 standard deviations) in a 3cm diameter spherical ROI on the normal right lobe. The following classifications were used: CMR, complete resolution of ^{18}F-FDG uptake within the target lesion that was lower than mean liver activity and indistinguishable from background blood-pool level. In patients with metabolically active lesions on the follow-up scan, the SULpeak of up to 5 lesions on the baseline and follow-up scan was summed (maximum of 2 per organ). Since the hottest lesions were selected in each scan, target lesions on follow-up scans were not necessarily the same as target lesions at baseline. If the sum of SULpeak decreased by at least 30%, tumor response was classified as PMR. Conversely, PMD was defined as an increase of the sum of SULpeak by at least 30% or the appearance of new hypermetabolic lesions on follow-up ^{18}F-FDG PET/CT scan. Cases not meeting the definitions for CMR, PMR, or PMD were classified as SMD.

Statistical analysis

Assessment of concordance between 2 criteria methods was done using Cohen's κ coefficient [11], with agreement noted as slight (κ<0.21), fair (κ=0.21-0.40), moderate (κ= 0.41-0.60), substantial (κ=0.61-0.80), or nearly perfect (κ> 0.80).

Progression-free survival (PFS) was defined as the time elapsed from the start of anti-PD-1 antibody therapy to date of disease progression (shown by radiological and/or clinical examination findings) or death from any cause. Patients with no evidence of progressive disease were censored at the date of the last follow-up examination. Overall survival (OS) was defined as the time from start of anti-PD-1 antibody therapy until death from any cause. Patients alive on the date of last follow-up were censored, and classified as alive with disease or no evidence of progression. Actuarial survival curves were generated using the Kaplan-Meier method and differences between groups were tested using a log-rank test. Receiver operating characteristic (ROC) curve analysis was also performed to determine the cut-off values for predicting recurrence or death. Univariate and multivariate analyses of potential prognostic factors were performed using the Cox proportional hazards regression model. The results from the Cox models were expressed as hazard ratios with 95% confidence intervals, and P values <0.05 were considered statistically significant. Variables with a P-value of < 0.1 in the univariate analysis were entered into the multivariate analysis.

Differences of parameters determined between two groups were assessed using a Student's t test.

Statistical analyses were performed with SAS, version 9.3 (SAS Institute Inc., Cary, NC, USA), with P<0.05 considered to indicate significance.

Results

Treatment response assessment

Using the EORTC criteria with ^{18}F-FDG PET/CT findings, CMR was noted in 8 patients (20.0%), PMR in 10 (25.0%), SMD in 4 (10.0%), and PMD in 18 (45.0%). Using PERCIST with 18F-FDG PET/CT findings, CMR was noted in 9 patients (22.5%), PMR in 9 (22.5%), SMD in 4 (10.0%), and PMD in 18 (45.0%). Using RECIST1.1 with CT, 4 patients (10.0%) had CR, 10 (25.0%) PR, 12 (30.0%) SD, and 14 (35.0%) PD, respectively. Two representative cases are shown in Figures 1,2.

Comparisons of treatment response assessment among criteria methods

Concordance between the PERCIST and EORTC criteria response classifications was seen in 37 (92.5%) cases and discordance in 3 (7.5%), with nearly perfect agreement (κ=0.963) demonstrated between them for response classification (Table 2). Concordance between the PERCIST and RECIST1.1 response classifications was seen in 26 (65.0%) cases and discordance in 14 (35.0%), with substantial agreement (κ=0.560) demonstrated between them for response classification. Concordance between the EORTC and RECIST1.1 response classifications was seen in 26 (65.0%) cases and discordance in 14 (35.0%), with substantial agreement (κ=0.574) demonstrated between them for response classification.

One patient who was classified as CR in RECIST1.1 and PMR in EORTC and PERCIST was a case in which metastatic lymph node decreased to the size of 8mm on CT and showed remaining ^{18}F-FDG uptake on PET. One patient who was classified as PR in RECIST1.1 and PMD in EORTC and PERCIST was a case in which tiny bone metastasis could not be detected by CT and could be correctly diagnosed by PET.

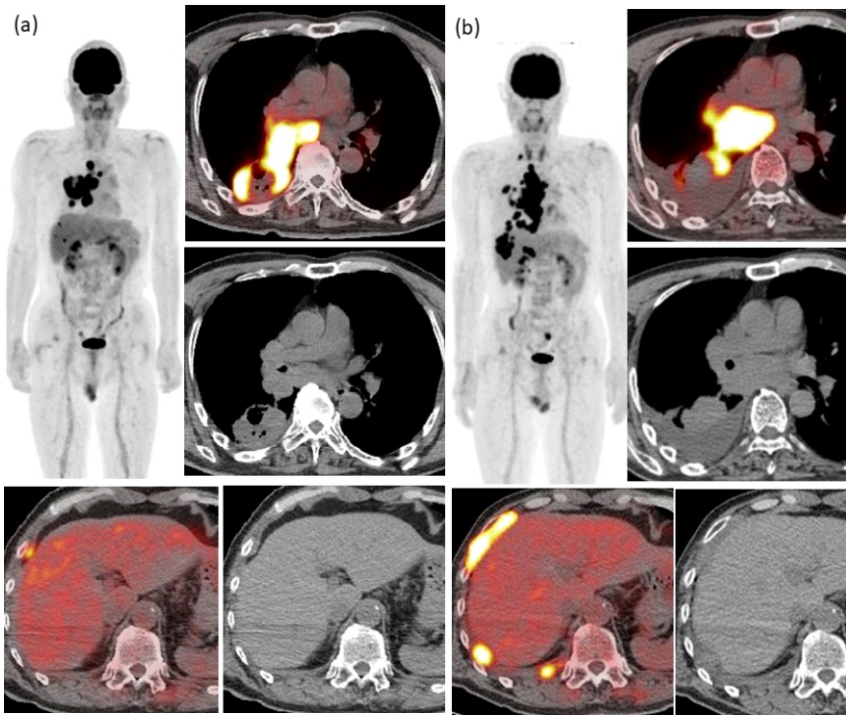

Figure 1. A 65-year-old man with squamous cell carcinoma lung cancer, cT3N2M1, negative EGFR mutation, negative ALK mutation, and high PD-L1 expression (90%), treated by pembrolizumab. (a) Pre-treatment ^{18}F-FDGPET/CT shows the strong ^{18}F-FDG uptakes of primary lung cancer with ipsilateral hilar and mediastinal lymph nodal metastases and pleural dissemination. (b) Post-treatment ^{18}F-FDG PET/CT after five cycles of pembrolizumab shows the improvement of primary tumor, the progression of ipsilateral hilar and mediastinal lymph nodal metastases and pleural dissemination, and appearance of and supraclavicular nodal metastasis and pleural disseminations.The classification of EORTC, PERCIST, and RECIST1.1 are all PMD or PD. Progression was observed at 4.5 months and he died 18.7 months after pembrolizumab initiation. This is the "non-responder" case.

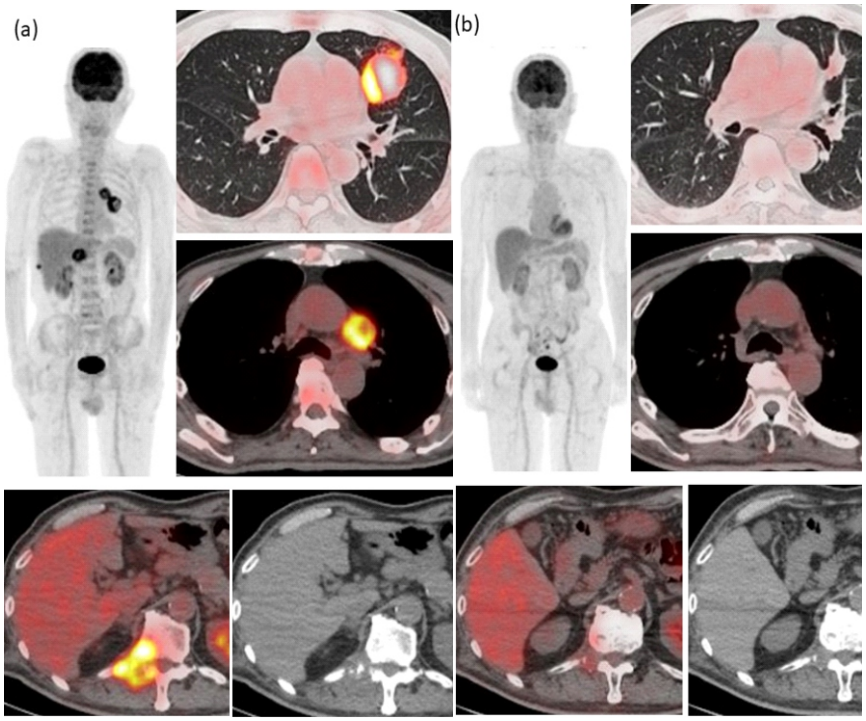

Figure 2. A 73-year-old man with squamous cell carcinoma lung cancer, cT3N1M1, negative EGFR mutation, negative ALK mutation, and high PD-L1 expression (90%), treated by pembrolizumab. (a) Pre-treatment ^{18}F-FDGPET/CT shows the strong ^{18}F-FDG uptakes of primary lung cancer with ipsilateral hilar lymph nodal metastasis and Th12 bone metastasis. The ^{18}F-FDG uptake of right lower rib was physiological uptake of the fracture. (b) Post-treatment ^{18}F-FDGPET/CT after six cycles of pembrolizumab shows the remarkable improvement of primary tumor, ipsilateral hilar lymph nodal metastasis and Th12 bone metastasis with no ^{18}F-FDG uptake less than surrounding normal tissue or liver activity. At the interpretation of RECIST1.1, the sum of three lesion's size decreased from 93mm to 39mm with 58.1% decreasing. The classification of EORTC and PERCIST are CMR, and that of RECIST1.1 is PR. No recurrence was seen at 21.2 months after pembrolizumab initiation. This is the "responder" case.

Table 2. *Comparison of treatment response assessments among three criteria.*

	EORTC criteria					RECIST1.1				
	CMR	PMR	SMD	PMD	Total	CR	PR	SD	PD	Total
PERCIST										
CMR	8	1	0	0	9	3	4	2	0	9
PMR	0	9	0	0	9	1	5	3	0	9
SMD	0	0	3	1	4	0	0	4	0	4
PMD	0	0	1	17	18	0	1	3	14	18
Total	8	10	4	18	40	4	10	12	14	40

	RECIST1.1				
	CR	PR	SD	PD	Total
EORTC					
CMR	3	4	1	0	8
PMR	1	5	4	0	10
SMD	0	0	4	0	4
PMD	0	1	3	14	18
Total	4	10	12	14	40

PERCIST: Positron Emission Tomography Response Criteria in Solid Tumors, EOPRTC: European Organization for Research and Treatment of Cancer, RECIST1.1: Response Evaluation Criteria in Solid Tumors Version 1.1, CMR: complete metabolic response, PMR: partial metabolic response, SMD: stable metabolic disease, PMD: progressive metabolic disease, CR: complete response, PR: partial response, SD: stable disease, PD: progressive disease

Progression-free survival (PFS)

Progressive disease was noted in 32 (80.0%) of the 40 patients after a median period of 9.1 months (range 2.6-51.8 months). In the 8 patients without progression, the overall median follow-up duration was 36.8 months (8.4-51.8 months), while that duration was 5.1 months (2.6-24.7 months) in the 32 with progression during the follow-up period.

According to both PET criteria (EORTC criteria and PERCIST) and CT criteria (RECIST1.1), patients with no progression (CMR/PMR/SMD or CR/PR/SD) showed significantly longer PFS than PMD or PD patients (EORTC criteria: P< 0.0001, PERCIST:P<0.0001, RECIST1.1: P<0.0001) (Figure3a).

Similarly, according to both PET criteria (EORTC criteria and PERCIST) and CT criteria (RECIST1.1), responders (CMR/PMR or CR/PR) showed significantly longer PFS than non-responders (SMD/PMD or SD/PD) (EORTC criteria: P< 0.0001, PERCIST:P<0.0001, RECIST1.1: P=0.0002 (Figure3b).

Based on the ROC curve analysis and log-rank tests on PFS, the patients could be divided into two groups according to highest SUVmax (cut-off value: 8.57), total MTV (15.5 g), total TLG (87.7), highest SULpeak (6.6), and total TLGL (59.1) on pre-ICI treatment ^{18}F-FDG PET/CT scans, highest SUVmax (cut-off value: 6.80), total MTV (13.2 g), total TLG (47.3), highest SULpeak (4.8), and total TLGL (56.4) on post-ICI treatment ^{18}F-FDG PET/CT scans, and change rate of highest SUVmax (cut-off value: -26.0%), total MTV (-14.4%), total TLG (-43.2%), highest SULpeak (-6.3%), total TLGL (-42.5%), size (13.1%) on two PET/CT scans.

In a univariate analysis showed, total MTV (P=0.042) on pre-ICI treatment ^{18}F-FDG PET/CT scans was significantly associated with progression. Highest SUVmax (P<0.0001), total MTV (P=0.0062), total TLG (P<0.0001), highest SULpeak (P<0.0001), and total TLGL (P<0.0001) on post-ICI treatment ^{18}F-FDG PET/CT scans were also significantly associated with progression. Moreover, the change rate of highest SUVmax (P<0.0001), total MTV (P<0.0001), total TLG (P< 0.0001), highest SULpeak (P<0.0001), total TLGL (P<0.0001), and size (P=0.0012) on two PET/CT scans were significantly associated with progression (Table 3). On the other hand, highest SUVmax, total TLG, highest SULpeak, and total TLGL

on pre-ICI treatment ^{18}F-FDG PET/CT scans were not significantly associated with progression.

Several variables (pre-treatment total MTV on pre-ICI treatment ^{18}F-FDG PET/CT scans, post-treatment highest SUVmax, highest SULpeak, total TLG, and total TLGL on post-ICI treatment ^{18}F-FDG PET/CT scans, and change rate of highest SUVmax, highest SULpeak, total MTV, total TLG, total TLGL, size, EORTC criteria, PERCIST, and RECIST1.1) in the univariate analysis were entered into the multivariate analysis. A multivariate analysis confirmed the change rate of total MTV (hazard ration [HR]:35.4, 95% confidence interval [CI]:3.28-491.67, P=0.034), and total TLGL (HR:2108.7, 95% CI:11.7-740179.9, P=0.0027), EORTC criteria (HR:44.7, 95% CI:5.19-1241.7, P=0.018), PERCIST (HR:20.3, 95% CI:2.05-127.9, P=0.045), and RECIST 1.1 (HR:128.9, 95% CI:9.37-25616.9, P=0.0037) on two PET/CT scans as independent predictors of PFS (Table 3).

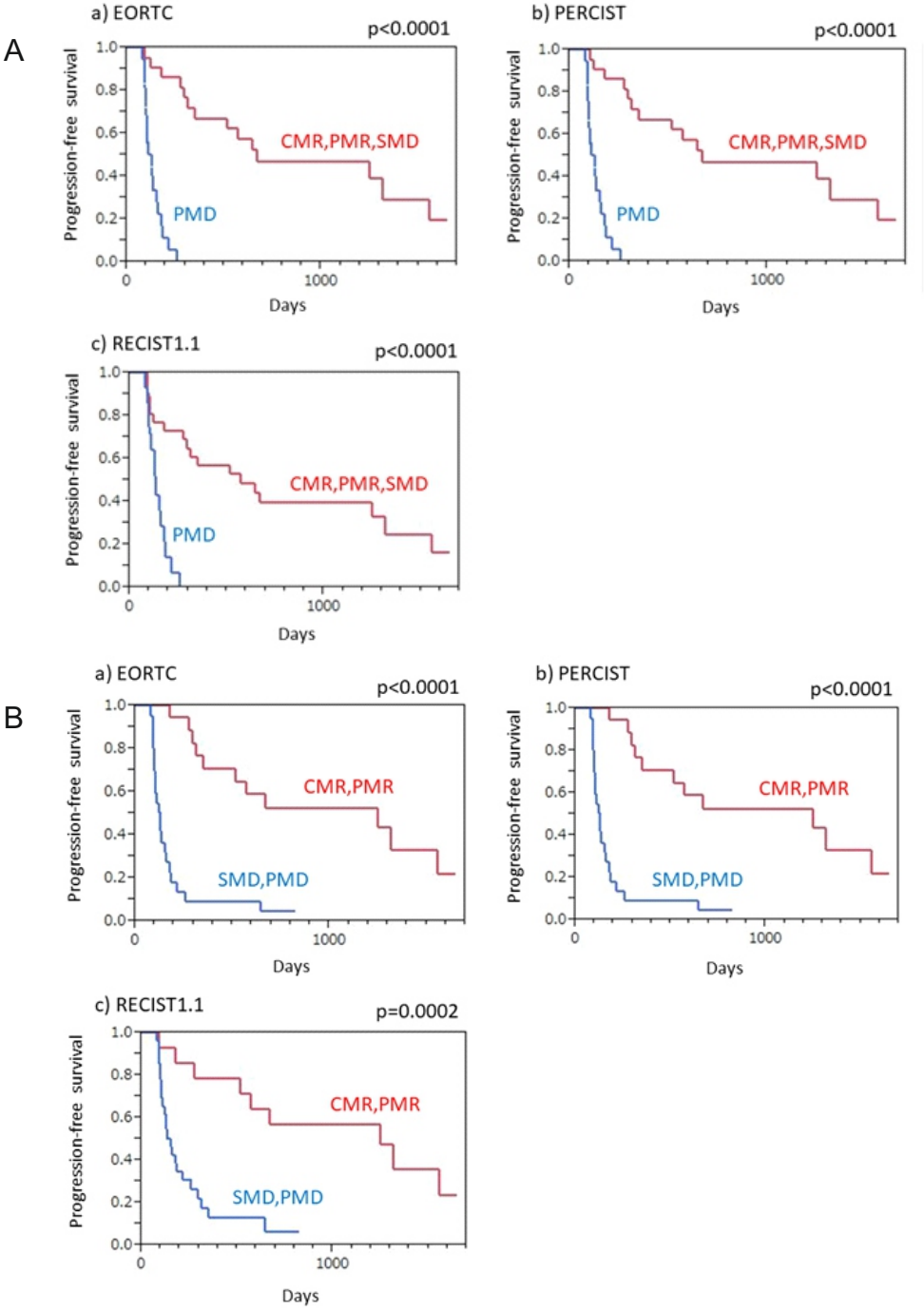

Figure 3. Kaplan-Meier curve of progression free survival. In part A we dichotomized the patients between PMD and all other groups and in part B between PMD and SMD and all other groups. **A.** According to both PET criteria (EORTC criteria and PERCIST) and CT criteria (RECIST1.1), patients with no progression (CMR/PMR/SMD or CR/PR/SD) showed significantly longer PFS than PMD or PD patients (EORTC criteria: P<0.0001, PERCIST:P<0.0001, RECIST1.1: P<0.0001). **B.** According to both PET criteria (EORTC criteria and PERCIST) and CT criteria (RECIST1.1), responders (CMR/PMR or CR/PR) showed significantly longer PFS than non-responders (SMD/PMD or SD/PD) (EORTC criteria: P<0.0001, PERCIST:P<0.0001, RECIST1.1: P=0.0002).

Table 3. *Factors associated with progression free survival (PFS).*

		Univariate analysis		Multivariate analysis	
Factors		**P (log-rank)**	**Hazard ration (95% CI)**	**P (log-rank)**	**Hazard ration (95% CI)**
Pre-ICI treatment					
SUVmax	<8.57	0.92	1.04 (0.49-2.18)		
	>8.57				
MTV	<15.5	0.042	2.15 (1.03-4.73)	0.74	0.82 (0.25-2.66)
	>15.5				
TLG	<87.7	0.70	1.15 (0.55-2.42)		
	>87.7				
SULpeak	<6.6	0.62	1.20 (0.58-2.53)		
	>6.6				
TLGL	<59.1	0.17	1.68 (0.79-3.78)		
	>59.1				
Post-ICI treatment					
SUVmax	<6.8	<0.0001	5.14 (2.29-12.39)	0.13	2.51 (0.21-15.83)
	>6.8				
MTV	<13.2	0.0062	2.86 (1.35-6.17)	0.34	1.93 (0.39-2.25)
	>13.2				
TLG	<47.3	<0.0001	6.82 (2.83-19.13)	0.87	1.29 (0.05-2.13)
	>47.3				
SULpeak	<4.8	<0.0001	7.11 (3.01-18.83)	0.77	1.38 (0.58-3.38)
	>4.8				
TLGL	<56.4	<0.0001	8.34 (3.37-23.93)	0.33	1.97 (0.12-4.83)
	>56.4				
Change rate of two values					
SUVmax	<-26.0%	<0.0001	5.50 (2.43-13.68)	0.49	2.74 (0.19-4.12)
	>-26.0%				
MTV	<-14.4%	<0.0001	5.23 (2.30-12.6)	0.034	35.4 (3.28-491.67)
	>-14.4%				
TLG	<-43.2%	<0.0001	28.7 (7.82-186.7)	0.32	1.44 (0.14-3.92)
	>-43.2%				
SULpeak	<-6.3%	<0.0001	5.34 (2.34-12.62)	0.14	2.49 (0.45-14.74)
	>-6.3%				

(continued)

TLGL	<-42.5%	<0.0001	16.66 (5.36-74.02)	0.0027	2108.7 (11.7-740179.9)
	>-42.5%				
Size	<13.1%	0.0012	3.55 (1.64-8.17)	0.36	3.19 (0.81-5.82)
	>13.1%				
EORTC	CMR/PMR/SMD	<0.0001	18.99 (6.04-84.76)	0.018	44.77 (5.19-1241.7)
	PMD				
PERCIST	CMR/PMR/SMD	<0.0001	8.48 (3.57-22.52)	0.045	20.3 (2.05-127.9)
	PMD				
RECIST1.1	CR/PR/SD	<0.0001	6.29 (2.52-17.26)	0.0037	128.9 (9.37-25616.9)
	PD				

SUVmax: maximum standardized uptake value, MTV: metabolic tumor volume, TLG: total lesion glycolysis, SULpeak: peak lean body mass standardized uptake value, TLGL: total lesion glycolysis lean, PERCIST: Positron Emission Tomography Response Criteria in Solid Tumors, EOPRTC: European Organization for Research and Treatment of Cancer, RECIST1.1: Response Evaluation Criteria in Solid Tumors Version 1.1, CMR: complete metabolic response, PMR: partial metabolic response, SMD: stable metabolic disease, PMD: progressive metabolic disease, CR: complete response, PR: partial response, SD: stable disease, PD: progressive disease

Overall survival (OS)

Among all 40 patients, 24 (60.0%) died from NSCLC after a median 23.2 months (range 7.2-51.8 months). For the 16 surviving patients, the median overall follow-up period was 33.5 months (10.6-51.8 months) as compared to 18.2 months (7.2-46.0 months) for the 24 who died during follow-up.

According to both PET criteria (EORTC criteria and PERCIST) and CT criteria (RECIST1.1), patients with no progression (CMR/PMR/SMD or CR/PR/SD) showed significantly longer OS than PMD or PD patients (EORTCcriteria: P< 0.0001, PERCIST: P=0.0001, RECIST1.1: P<0.0001) (Figure 4a).

Similarly, according to both PET criteria (EORTC criteria and PERCIST) and CT criteria (RECIST1.1), responders (CMR/PMR or CR/PR) showed significantly longer OS than non-responders (SMD/PMD or SD/PD) (EORTCcriteria: P=0.0004, PERCIST: P=0.0004, RECIST1.1: P=0.0003) (Figure4b).

A univariate analysis showed that total MTV (P=0.042) on pre-ICI treatment ^{18}F-FDG PET/CT scans, highest SUVmax (P<0.0001), total MTV (P=0.013), total TLG (P<0.0001), highest SULpeak (P=0.0001), and total TLGL (P<0.0001) on post-ICI treatment ^{18}F-FDG PET/CT scans, and change rate of highest SUVmax (P=0.0008), total MTV (P=0.0004), total TLG (P<0.0001), highest SULpeak (P=0.0020), total TLGL (P< 0.0001), and size (P=0.0028) on two PET/CT scans were significantly associated with death (Table 4). On the other hand, highest SUVmax, total TLG, highest SULpeak, and total TLGL on pre-ICI treatment ^{18}F-FDG PET/CT scans were not significantly associated with progression.

Several variables (pre-treatment total MTV on pre-ICI treatment ^{18}F-FDG PET/CT scans, post-treatment highest SUVmax, highest SULpeak, total MTV, total TLG, and total TLGL on post-ICI treatment ^{18}F-FDG PET/CT scans, and change rate of highest SUVmax, highest SULpeak, total MTV, total TLG, and total TLGL, larger size, EORTC criteria, PERCIST, and RECIST1.1) in the univariate analysis were entered into the multivariate analysis. A multivariate analysis confirmed the pre-ICI treatment total MTV (HR:37.99, 95% CI:4.24-457.3, P=0.0035), post-ICI treatment highest SUVmax (HR:138.1, 95% CI:7.64-6612.1, P=0.0005),total MTV (HR:29.04, 95% CI:1.65-360.6, P=0.0070), SULpeak (HR:15.5, 95% CI:1.67-44.3, P=0.011), change rate of highest SULpeak (HR:14.93, 95% CI:1.11-26.28 P=0.013), and total TLGL (HR:35.74, 95% CI:2.48-402.7, P=0.0053) on two PET/CT scans as independent predictors of OS (Table 4).

Discussion

To the best of our knowledge, this is the first study to compare two ^{18}F-FDG PET criteria (EORTC criteria and PERCIST) and CT (RECIST1.1) after 4 to 8 ICI therapy cycles for evaluation of tumor response to ICI therapy and prediction of prognosis in patients with NSCLC. All ^{18}F-FDG PET (EORTC criteria and PERCIST) and CT (RECIST1.1) after 4 to 8 ICI therapy cycles are accurate for evaluation of tumor response and predicting prognosis in NSCLC patients. Although metabolic changes normally precede anatomic changes, ^{18}F-FDG is superior to CT in the early assessment of response. This study was performed after at least 4 cycles and in the majority (n = 26, 65%) after 6 ICI cycles. This may explain the relatively good CT results.

Assessing early (≦4 cycles) tumor response to ICI is ideal for effective cancer treatment management. Although several CT response criteria such as immune-related response criteria (irRC) [12], immune-related RECIST (irRECIST) [13] and immune RECIST (iRECIST) [14] and ^{18}F-FDG PET response criteria such as PET/CT criteria for early prediction of

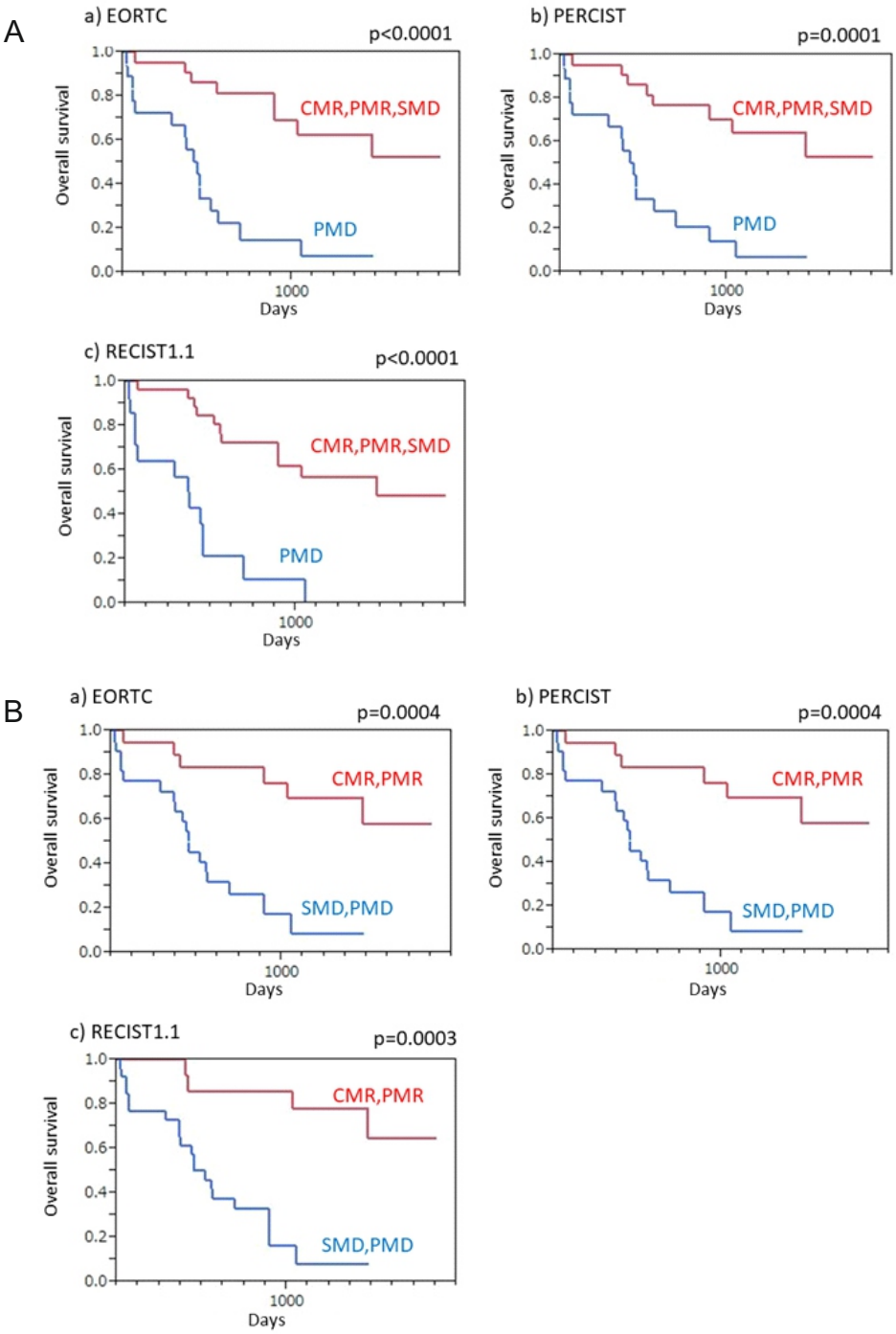

Figure 4. Kaplan-Meier curve of overall survival. In part A we dichotomized the patients between PMD and all other groups and in part B between PMD and SMD and all other groups. A. According to both PET criteria (EORTC criteria and PERCIST) and CT criteria (RECIST1.1), patients with no progression (CMR/PMR/SMD or CR/PR/SD) showed significantly longer OS than PMD or PD patients (EORTC criteria: P<0.0001, PERCIST: P=0.0001, RECIST1.1: P<0.0001). B. According to both PET criteria (EORTC criteria and PERCIST) and CT criteria (RECIST1.1), responders (CMR/PMR or CR/PR) showed significantly longer OS than non-responders (SMD/PMD or SD/PD) (EORTC criteria: P=0.0004, PERCIST: P=0.0004, RECIST1.1: P=0.0003).

Response to Immune checkpoint inhibitor Therapy (PECRIT) [15], PET Response Evaluation Criteria for Immunotherapy (PERCIMT) [16], immunotherapy-modified PERCIST imPERCIST) [17], and immune PERCIST (iPERCIST) [18] for ICI treatment have been proposed, an optimal evaluation method has yet to be determined. Most of these immune related response criteria have been developed in melanomas, due to the fact that melanomas were first treated with ICI and in particular with ipilimumab monotherapy. Melanomas have own characteristics and tend to give a lot of new metastatic lesions everywhere in the body. This is different than in NSCLC.

Table 4. *Factors associated with overall survival (OS).*

		Univariate analysis		Multivariate analysis	
Factors		**P (log-rank)**	**Hazard ration (95% CI)**	**P (log-rank)**	**Hazard ration (95% CI)**
Pre-ICI treatment					
SUVmax	<8.57	0.30	1.56 (0.67-3.69)		
	>8.57				
MTV	<15.5	0.042	2.15 (1.03-4.73)	0.0035	37.99 (4.24-457.3)
	>15.5				
TLG	<87.7	0.47	1.35 (0.59-3.13)		
	>87.7				
SULpeak	<6.6	0.39	1.45 (0.62-3.44)		
	>6.6				
TLGL	<59.1	0.80	1.12 (0.46-2.55)		
	>59.1				
Post-ICI treatment					
SUVmax	<6.8	<0.0001	10.54 (3.43-46.1)	0.0005	138.1 (7.64-6612.1)
	>6.8				
MTV	<13.2	0.013	2.88 (1.25-7.01)	0.0070	29.04 (1.65-360.6)
	>13.2				
TLG	<47.3	<0.0001	11.65 (3.33-73.77)	0.12	2.68 (0.21-28.71)
	>47.3				
SULpeak	<4.8	0.0001	6.39 (2.32-22.50)	0.011	15.5 (1.67-44.3)
	>4.8				
TLGL	<56.4	<0.0001	6.93 (2.52-24.42)	0.64	1.14 (0.51-7.1)
	>56.4				

(continued)

Change rate of two values					
SUVmax	<-26.0%	0.0008	4.74 (1.85-14.52)	0.39	1.82 (0.77-14.19)
	>-26.0%				
MTV	<-14.4%	0.0004	4.87 (1.99-13.67)	0.13	3.57 (0.90-25.29)
	>-14.4%				
TLG	<-43.2%	<0.0001	6.28 (2.43-19.41)	0.056	5.19 (0.64-18.7)
	>-43.2%				
SULpeak	<-6.3%	0.0020	3.94 (1.66-9.79)	0.013	14.93 (1.11-26.28)
	>-6.3%				
TLGL	<-42.5%	<0.0001	6.40 (2.48-19.75)	0.0053	35.74 (2.48-402.7)
	>-42.5%				
Size change rate	<13.1%	0.0028	3.87 (1.57-10.92)	0.36	1.12 (0.25-4.56)
	>13.1%				
EORTC	CMR/PMR/SMD	0.0001	5.77 (2.33-16.33)	0.073	5.07 (0.41-81.1)
	PMD				
PERCIST	CMR/PMR/SMD	0.001	4.27 (1.78-11.30)	0.44	1.59 (0.14-21.4)
	PMD				
RECIST1.1	CR/PR/SD	0.0001	5.68 (2.39-14.12)	0.40	1.87 (0.20-25.4)
	PD				

SUVmax: maximum standardized uptake value, MTV: metabolic tumor volume, TLG: total lesion glycolysis, SULpeak: peak lean body mass standardized uptake value, TLGL: total lesion glycolysis lean, PERCIST: Positron Emission Tomography Response Criteria in Solid Tumors, EOPRTC: European Organization for Research and Treatment of Cancer, RECIST1.1: Response Evaluation Criteria in Solid Tumors Version 1.1, CMR: complete metabolic response, PMR: partial metabolic response, SMD: stable metabolic disease, PMD: progressive metabolic disease, CR: complete response, PR: partial response, SD: stable disease, PD: progressive disease

Cho et al. (2017) [15] introduced PECRIT, which combined change in lesions' size with the change in ^{18}F-FDG avidity at ^{18}F-FDG PET/CT 1 cycle after ICI monotherapy (ipilimumab, nivolumab, or BMS-936559) onset in 20 advanced melanoma patients and reported that a criteria with SD by RECIST1.1 and an increase >15.5% in SULpeak of the hottest lesion by ^{18}F-FDG PET/CT was accurate for predicting treatment response at 4 months with a sensitivity/specificity/accuracy of 100%/93%/95%, respectively. Anwar et al. (2018) [16] introduced PERCIMT, which use the absolute number of new lesions rather than changes in metabolic parameters (i.e. SUV) at ^{18}F-FDG PET/CT 4 cycles after ipilmumab onset in 41 metastatic melanoma patients and reported that a criteria with 4 or more new lesions of less than 1cm in functional diameter was accurate for predicting clinical benefit with a sensitivity/specificity of 84%/100%, respectively. Ito et al. (2019) [17] introduced imPERCIST, in which the appearance of new lesions do not configure PMD,

and the sum of SULpeak for up to 5 measured lesions increased by at least 30% reflect PMD at ^{18}F-FDG PET/CT 2-4 cycles after ipilimumab onset in 60 metastatic melanoma and reported that 2-year OS for responders versus non-responders according to PERCIST and imPERCIST was 61% vs. 33% (P=0.028) and 66% vs 29% (P=0.003). Goldfarb et al. (2019) [18] used iPERCIST by introducing two new categories of response: unconfirmed PMD (UPMD) and confirmed PMD (CPMD), indicating that all metabolic progression observed at 8 weeks (4 cycles) should be confirmed by another ^{18}F-FDG PET/CT study 4 weeks later in 28 NSCLC patients receiving nivolumab and reported that iPERCIST was useful to differentiate responders from non-responders and predict OS (P=0.0003).

This study has some limitations. First, it was retrospectively performed at a single center with a small sample size, thus generalization of the findings is limited and statistical errors are possible. A prospective multicenter trial with a larger cohort would help to clarify the exact roles of ^{18}F-FDG PET/CT and CT for decision-making and predicting long-term outcome in clinical settings. Second, the enrolled population was heterogeneous, as it included patients with pre-treatment and post-treatment, as well as those with from 4 to 8 cycles. Such heterogeneity likely introduced confounding factors into the analysis. Third, for the interpretation of RECIST1.1, we used not diagnostic contrast-enhanced CT but low-dose and non-contrast CT. Ideally, diagnostic contrast-enhanced CT is desirable.

In conclusion, all two^{18}F-FDGPET criteria (EORTC criteria and PERCIST) and one CT criteria (RECIST1.1) after 4 to 8 ICI therapy cycles were found to be significantly predictive of PFS and OS in NSCLC patients treated by ICI therapy. The classical functional ^{18}F-FDG PET/CT criteria seemed to be suitable in late (≧4 cycles) assessment of ICI-response classification in NSCLC patients. In early (<4 cycles) assessment of ICI treatment response in NSCLC patients, ^{18}F-FDG PET may be superior to CT. This needs to be confirmed in a larger cohort prospective study.

Funding Source
This work was supported by JSPS KAKENHI grant numbers 19K08187.

The authors declare that they have no conflicts of interest.

Bibliography

1. Borghaei H, Paz-Ares L, Horn L et al. Nivolumab versus docetaxel in advanced nonsquamous non-small-cell lung cancer. *N Engl J Med* 2015; 373: 1627-39.
2. Reck M, Rodríguez-Abreu D, Robinson AG et al. KEYNOTE-024 Investigators. Pembrolizumab versus chemotherapy for PD-L1-Positive non-Small-Cell lung cancer. *N Engl J Med* 2016; 375: 1823-33.
3. Rittmeyer A, Barlesi F, Waterkamp D et al. OAK Study Group. Atezolizumab versus docetaxel in patients with previously treated non-small-cell lung cancer (OAK): a phase 3, open-label, multicentrer andomised controlled trial. *Lancet* 2017; 389: 255-65.
4. Eisenhauer EA, Therasse P, Bogaerts J et al. New response evaluation criteria in solid tumours: revised RECIST guideline (version 1.1).*Eur J Cancer* 2009; 45: 228-47.
5. Chiou VL, Burotto M. Pseudoprogression and immune-related response in solid tumors. *J Clin Oncol* 2015; 33: 3541-3.
6. Young H, Baum R, Cremerius U et al. Measurement of clinical and subclinical tumour response using ^{18}F-fluorodeoxyglucose and positron emission tomography: review and 1999 EORTC recommendations. *Eur J Cancer* 1999; 35: 1773-82.
7. 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 (Suppl 1): 122S-50S.
8. Tsuchitani T, Takahashi Y, Maeda Y et al. Investigation of appropriate semi-quantitative index for assessment of esophageal and breast cancer treatment response in Japanese patients using ^{18}F-FDG PET/CT findings. *Hell J Nucl Med* 2019; 22: 20-4.
9. Kitajima K, Yamano T, Miyoshi Y et al. Prognostic value of ^{18}F-FDG PET/CT prior to breast cancer treatment. Comparison with magnetic resonance spectroscopy and diffusion weighted imaging. *Hell J Nucl Med* 2019; 22: 25-35.
10. Depardon E, Kanoun S, Humbert O et al. FDG PET/CT for prognostic stratification of patients with metastatic breast cancer treated with first line systemic therapy: Comparison of EORTC criteria and PERCIST. *PLoS One* 2018; 13: e0199529.
11. Kundel HL, Polansky M. Measurement of observer agreement. *Radiology* 2003; 228: 303-8.
12. Wolchok JD, Hoos A, O'Day S et al. Guidelines for the evaluation of immune therapy activity in solid tumors: immune-related response criteria. *Clin Cancer Res* 2009; 15: 7412-20.
13. Chiou VL, Burotto M. Pseudoprogression and Immune-Related Response in Solid Tumors. *J Clin Oncol* 2015; 33: 3541-3.
14. Seymour L, Bogaerts J, Perrone A et al. RECIST working group. iRECIST: guidelines for response criteria for use in trials testing immunotherapeutics. *Lancet Oncol* 2017; 18: e143-e52.
15. Cho SY, Lipson EJ, Im HJ et al. Prediction of response to immune checkpoint inhibitor therapy using early-time-point ^{18}F-FDG PET/CT imaging in patients with advanced melanoma. *J Nucl Med* 2017; 58: 1421-8.
16. Anwar H, Sachpekidis C, Winkler J et al. Absolute number of new lesions on ^{18}F-FDG PET/CT is more predictive of clinical response than SUV changes in metastatic melanoma patients receiving ipilimumab. *Eur J Nucl Med Mol Imaging* 2018; 45: 376-83.
17. Ito K, Teng R, Schöder H, Humm JL et al. ^{18}F-FDG PET/CT for monitoring of Ipilimumab therapy in patients with metastatic melanoma. *J Nucl Med* 2019; 60: 335-41.
18. Goldfarb L, Duchemann B, Chouahnia K et al. Monitoring anti-PD-1-based immunotherapy in non-small cell lung cancer with FDG PET: introduction of iPERCIST. *EJNMMI Res* 2019; 9: 8.