Radiology

Prognostic Value of Metabolic and Volumetric Parameters of FDG PET in Pediatric Osteosarcoma: A Hypothesis-generating Study[1]

Hyung-Jun Im, MD, PhD[2]
Yi Zhang, BS
Huiyun Wu, PhD
Jianrong Wu, PhD
Najat C. Daw, MD
Fariba Navid, MD
Barry L. Shulkin, MD
Steve Y. Cho, MD

Purpose: To preliminarily assess the potential prognostic value of various fluorine 18 fluorodeoxyglucose (FDG) positron emission tomography (PET) parameters before, during, and after neoadjuvant chemotherapy (NCT).

Materials and Methods: Thirty-four patients with osteosarcoma were enrolled prospectively from 2008 to 2012 and underwent FDG PET/computed tomography (CT) imaging before (baseline scan), during (interim scan) and after NCT (posttherapy scan). The study was approved by the institutional review board and informed consent was received from patients. Maximum and peak standardized uptake value (SUV_{max} and SUV_{peak}), metabolic tumor volume (MTV), and total lesion glycolysis (TLG) were measured. Predictive value of FDG PET parameters for event-free survival (EFS) and overall survival (OS) were evaluated. Multivariable Cox regression analysis for EFS and OS was performed by using histologic response and initial presence of metastasis as covariates.

Results: At baseline scan, SUV_{peak}, MTV, and TLG were predictive of EFS (P = .006–.03) and OS (P = .001–.03) but not associated with histologic response. At interim and posttherapy scan, SUV_{max}, SUV_{peak}, MTV, and TLG were associated with histologic response (P = .0002–.04) and predictive of EFS (P = .004–.02) and OS (P = .001–.03). Multivariable Cox regression analysis revealed that the FDG PET parameters either at baseline, interim, or posttherapy were independently predictive of EFS and OS. In particular, baseline MTV was an independent predictor of EFS (hazard ratio, 5.0 [95% confidence interval {CI}: 1.5, 16.8]) and OS (hazard ratio, 29.4 [95% CI: 2.2, 392.2]).

Conclusion: SUV_{peak}, MTV, and TLG either at baseline, interim, or posttherapy were predictive of EFS and OS and may be useful prognostic biomarkers for osteosarcoma.

Online supplemental material is available for this article.

[1] From the Department of Radiology, Nuclear Medicine/PET Section, University of Wisconsin School of Medicine and Public Health, 1111 Highland Ave, Madison, WI 53705 (H.J.I., Y.Z., S.Y.C.); Departments of Biostatistics (H.W., J.W.) and Diagnostic Imaging (B.L.S.), St. Jude Children's Research Hospital, Memphis, Tenn; Department of Pediatrics, MD Anderson Cancer Center, Houston, Tex (N.C.D.); Department of Pediatrics, Children's Hospital of Los Angeles, Los Angeles, Calif (F.N.); Keck School of Medicine, University of Southern California, Los Angeles, Calif (F.N.); and University of Wisconsin Carbone Cancer Center, Madison, Wis (S.Y.C.). Received December 13, 2016; revision requested January 27, 2017; revision received August 1; accepted September 15; final version accepted November 3. **Address correspondence to** S.Y.C. (e-mail: *scho@uwhealth.org*).

Study supported by American Lebanese Syrian Associated Charities.

[2] **Current address:** Department of Transdisciplinary Studies, Graduate School of Convergence Science and Technology, Seoul National University, Seoul, Korea.

Osteosarcoma is a common primary bone malignancy in children and adolescents. The standard care of treatment for high-grade osteosarcoma includes neoadjuvant chemotherapy (NCT) and subsequent surgical resection. The most important prognostic factors are initial tumor stage and histologic response after NCT (1,2). However, even with the same stage and histologic response, survival of patients could differ substantially, and histologic response only can be evaluated after surgical resection. Because changing treatment postoperatively does not improve patient outcomes, a preoperative prognostic factor could be useful for further risk stratification before the surgical resection (3).

The quantitative parameters derived from fluorine 18 fluorodeoxyglucose (FDG) positron emission tomography (PET)/computed tomography (CT) showed prognostic value in a variety of malignancies (4–7). Recent meta-analyses showed that various FDG PET parameters including maximum standardized uptake value (SUV; SUV_{max}), metabolic tumor volume (MTV), and total lesion glycolysis (TLG) were prognostic factors in multiple types of malignancies (6,8–10). Peak SUV (SUV_{peak}) is also reported to be a prognostic factor in non–small cell lung cancer (11). In osteosarcoma, few retrospective studies assessed the prognostic value of FDG PET parameters including SUV_{max} and MTV (3,7,12,13). However, there has been no consensus on selection of an optimal parameter and time to predict survival in osteosarcoma because, to our knowledge, no systematic or prospective study has been conducted to evaluate the predictive value and prognostic significance of various FDG PET/CT parameters before, during, and after NCT.

In our prospective study, we aimed to preliminarily assess the potential prognostic value of various FDG PET parameters before, during, and after NCT.

Materials and Methods

Patients

Our study was approved by the institutional review board. Written informed consent was received from all enrolled patients. Thirty-four patients with newly diagnosed high-grade osteosarcoma were prospectively enrolled from June 2008 to May 2012 in a clinical trial (trial NCT00667342) and underwent FDG PET/CT imaging, NCT, and subsequent surgical resection. The eligibility criteria included age 30 years or younger; newly diagnosed, confirmed with histologic analysis, high-grade, resectable osteosarcoma; Karnofsky or Lansky performance score of 50 or greater or World Health Organization/Eastern Cooperative Oncology Group performance score of 2 or less; and no previous chemotherapy or radiation therapy. The exclusion criteria included osteosarcoma as a second malignancy or a major surgical procedure or significant traumatic injury within 28 days of study entry or known bleeding diathesis. Other organ-specific inclusion/exclusion criteria are provided in Appendix E1 [online].

Patients received a uniform protocol of NCT in St. Jude Children's Research Hospital (Memphis, Tenn), which consisted of the following: cisplatin; doxorubicin at weeks 0 and 5; high-dose methotrexate at weeks 3, 4, 8, and 9; and bevacizumab at weeks 0, 3, and 5. FDG PET/CT imaging was performed before (baseline scan), during (5 weeks after initiation of NCT, interim scan), and after NCT (10 weeks after initiation of NCT, posttherapy scan).

Twenty-nine of the thirty-four patients were previously reported in an abstract (14). This previous abstract evaluated the difference of primary tumor SUV_{max} between routine (~1 hour after FDG injection) and delayed (~3 hours after FDG injection) imaging times, whereas we report the prognostic value of multiple FDG PET parameters, including SUV_{max}, SUV_{peak}, MTV, and TLG. Also, in the same group of patients who participated in our study, the predictive value of SUV_{max} from FDG PET for histologic response was evaluated, and SUV_{max} at 5 and 10 weeks was found to be predictive of histologic response (15). Histologic response to NCT was evaluated by examination of postsurgical tumor specimen after surgical resection compared with the pretreatment biopsy. Patients with tumor necrosis fraction of 90% or more were considered to be responders.

All patients were regularly assessed after surgery with bone scan and contrast agent–enhanced CT. When recurrence was suspected, further examinations such as FDG PET/CT, magnetic resonance (MR) imaging, and/or biopsy were performed to confirm recurrence. Event-free survival (EFS; recurrence, progression, and death) and overall survival (OS; death from any cause) were defined as the time interval from study enrollment to date of first event, or to date of last contact for patients without events. Data were censored at the time of last follow-up if patients were alive or free of disease recurrence or progression.

FDG PET/CT Imaging and Quantification

FDG PET/CT image acquisition, reconstruction, and attenuation correction were performed as previously described (16). After patients had fasted for at least 4 hours, FDG (5.4 MBq per kilogram of body weight; maximum 12 μCi) was injected intravenously. One

https://doi.org/10.1148/radiol.2017162758

Content code: NM

Radiology 2018; 000:1–10

Abbreviations:
2SD = two SD
EFS = event-free survival
FDG = fluorodeoxyglucose
MTV = metabolic tumor volume
NCT = neoadjuvant chemotherapy
OS = overall survival
SD = standard deviation
SUV = standardized uptake value
SUV_{max} = maximum SUV
SUV_{peak} = peak SUV
TLG = total lesion glycolysis

Author contributions:
Guarantor of integrity of entire study, S.Y.C.; study concepts/study design or data acquisition or data analysis/interpretation, all authors; manuscript drafting or manuscript revision for important intellectual content, all authors; approval of final version of submitted manuscript, all authors; agrees to ensure any questions related to the work are appropriately resolved, all authors; literature research, H.J.I., Y.Z., J.W., N.C.D., S.Y.C.; clinical studies, N.C.D., F.N., B.L.S., S.Y.C.; experimental studies, B.L.S., S.Y.C.; statistical analysis, H.J.I., H.W., J.W., S.Y.C.; and manuscript editing, H.J.I., J.W., N.C.D., F.N., B.L.S., S.Y.C.

Conflicts of interest are listed at the end of this article.

hour after the injection, transmission CT and PET scans were performed from the top of the skull to the feet.

Table 1

Patient Characteristics

Characteristic	Value
Median age at diagnosis (y)*	12.2 (6.8–19.0)
Sex	
Male	17 (50)
Female	17 (50)
Location of primary tumor	
Femur	17 (50)
Tibia	9 (26)
Humerus	5 (15)
Fibula	1 (3)
Radius	1 (3)
Mandible	1 (3)
Stage at diagnosis	
Localized	25 (74)
Metastatic	9 (26)
Pathologic type	
Osteoblastic osteosarcoma	29 (85)
Chondroblastic osteosarcoma	2 (6)
Fibroblastic osteosarcoma	1 (3)
Telangiectatic osteosarcoma	2 (6)

Note.—Unless otherwise indicated, data are number of patients and data in parentheses are percentage; $n = 34$.

* Data in parentheses are range.

PET emission images were acquired for 5 minutes per bed position in two-dimensional mode. CT scans were performed (section thickness, 5 mm; tube rotation time, 0.8 seconds; table speed, 1.5 cm per rotation; pitch, 1.5:1; and 120 kVp and milliampere-second adjusted for body weight [maximum, 90 mAs]). Oral or intravenous contrast agent was not used at the CT scans. A program (Mirada XD3; Mirada Medical, Denver, Colo) was used to help analyze the FDG PET/CT images.

Spherical volume of interest was drawn to include the primary osteosarcoma lesion on the FDG PET images. SUV_{max} and SUV_{peak} within the volume of interest were measured. The tumor boundaries were defined by using various SUV thresholds that were absolute (SUV 2.0, SUV 2.5), relative percent tumor SUV_{max} (40% or 60% of the tumor SUV_{max}), and mean liver SUV + one standard deviation (SD) or two SDs (2SD). Mean liver SUV and standard deviation were measured by using a 3-cm (diameter) spherical volume of interest on the right lobe of the liver. MTVs were defined as the tumor volumes inside the tumor boundaries by using the various SUV thresholds and named as follows with their SUV thresholds inside the parenthesis: MTV (SUV 2.0), MTV (SUV 2.5), MTV (40%), MTV (60%), MTV (liver + SD), and MTV (liver + 2SD). TLGs were calculated by multiplying mean SUV by the tumor volume inside the tumor boundaries and named as follows with their SUV thresholds inside the parenthesis: TLG (SUV 2.0), TLG (SUV 2.5), TLG (40%), TLG (60%), TLG (liver + SD), and TLG (liver + 2SD). Also, percent changes between parameters on baseline and interim scans, and baseline and posttherapy scans were calculated as follows, respectively: [percentage change (baseline scan – interim scan) = 100 × (interim scan – baseline scan)/baseline scan] and [percentage change (baseline scan – posttherapy scan) = 100 × (posttherapy scan − baseline scan)/baseline scan].

Table 2

Association between Patient Characteristics and Histologic Tumor Response to Neoadjuvant Chemotherapy

Characteristic	Poor Responder ($n = 17$)	Responder ($n = 15$)	*P* Value
Mean age (y)*	12.5 ± 3.9	12.5 ± 2.5	.924
Sex			.288
Male	7	9	…
Female	10	6	…
Primary tumor			.758
Femur	8	7	…
Tibia	4	5	…
Others	5	3	…
Initial presence of metastasis			.863
No	12	11	…
Yes	5	4	…
Pathologic result			.737
Osteoblastic	14	13	…
Others	3	2	…

Note.—Data are number of patients except where otherwise indicated.

* Data are mean ± standard deviation.

Statistical Analysis

Associations between histologic response and categorical clinical and pathologic parameters were determined by using a χ^2 test (two categories) or Kruskal-Wallis test (three or more categories). Each PET parameter was divided into two groups by using an optimized cut-off value for further analysis. The optimized cut-off value was determined by the maximum Youden index from receiver operating characteristic analysis. Each FDG PET/CT parameter was compared between responders and poor responders by using the Wilcoxon rank-sum test. EFS and OS distributions between FDG PET parameter groups were compared by using the log-rank test. Multivariable Cox regression analysis was performed to study the association between individual FDG PET parameters and EFS and OS after adjusting for histologic response and initial presence of metastasis. All statistical analyses were performed by using software (SAS version 9.3 for Windows; SAS Institute, Cary, NC). *P* values were adjusted for multiple comparison on the basis of the method of false discovery rate. *P* values less than .05 were considered to indicate statistical significance.

Results

Patient Characteristics

We enrolled 34 patients; their clinical and pathologic characteristics are summarized in Table 1. The age of patients

Figure 1

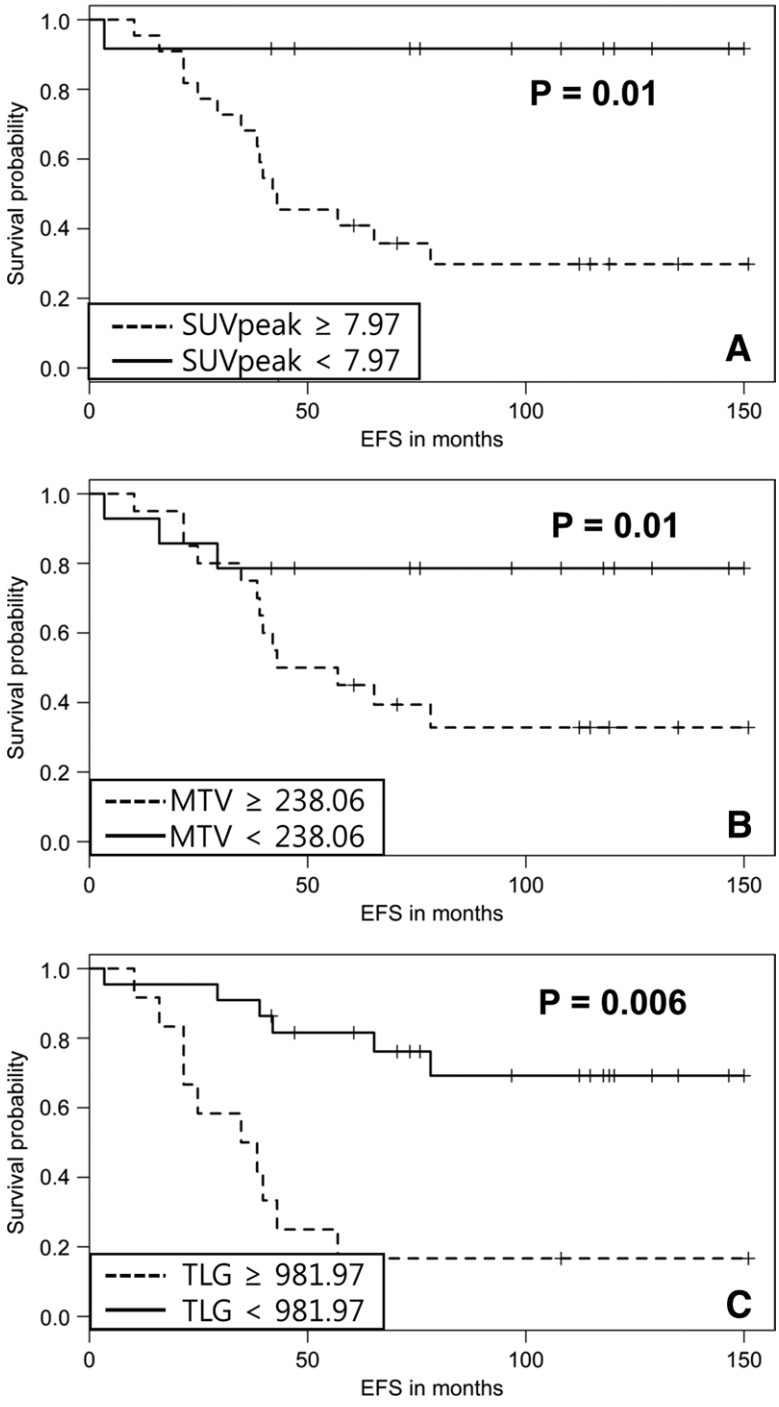

Figure 1: Kaplan-Meier Curves of event-free survival *(EFS)*. Kaplan-Meier curves of EFS for, *A*, peak standardized uptake value *(SUV; SUVpeak)*, *B*, metabolic tumor volume *(MTV)* with SUV threshold of 2.5 *(MTV [SUV, 2.5])*, and, *C*, total lesion glycolysis with SUV threshold of 2.5 *(TLG [SUV, 2.5])* at baseline fluorodeoxyglucose PET/CT scan are shown.

ranged from 6.8 to 19 years. Most of the patients had primary tumor at extremities (97%) except one at the mandible. Nine patients (26%) had metastatic lesions before initiation of treatment (seven in lung; one in bone; one in lung and bone). Of note, two patients were considered to be non-evaluable for histologic response because surgical resection was not performed at the protocol specified time (ie, week 10).

Association with Histologic Response

Histologic responses after NCT were evaluable in 32 patients, and 15 were classified as responders and the others as poor responders. Age, sex, site of primary tumor, initial presence of metastasis, and type of pathology were not associated with histologic response (Table 2). Also, none of the FDG PET parameters at baseline were associated with histologic response. However, at the interim scans, SUV_{max}, SUV_{peak}, MTV (liver + 2SD), TLG (liver + 2SD) showed statistically significant association with histologic response (P = .03–.04). Moreover, at the posttherapy scans, parameters except MTV (40%), MTV (60%), TLG (40%), and TLG (60%) were associated with histologic response (P = .0002–.02); MTV (liver + 2SD) and TLG (liver + 2SD) were the most statistically significant parameters (MTV [liver + 2SD] of poor responders vs responders: 78.7 ± 74.3 vs 2.8 ± 6.9, respectively; $P < .0002$; TLG [liver + 2SD] poor responders vs responders: 272.9 ± 319.1 vs 6.6 ± 15.6, respectively; $P < .0002$). Similarly, among percent change between baseline and interim FDG PET parameters, percent change (baseline to interim) of all parameters except MTV (40%), MTV (60%), TLG (40%), and TLG (60%) were associated with histologic response (P = .003–.04). Percent change (baseline to interim) of MTV (liver + SD) was the most statistically significant parameter (P = .003). Lastly, among percent change between baseline and posttherapy FDG PET/CT values, percentage change (baseline to posttherapy scan) of all parameters except MTV (60%), TLG (40%), and TLG (60%) were associated with histologic response (P = .0002–.01), and

Figure 2

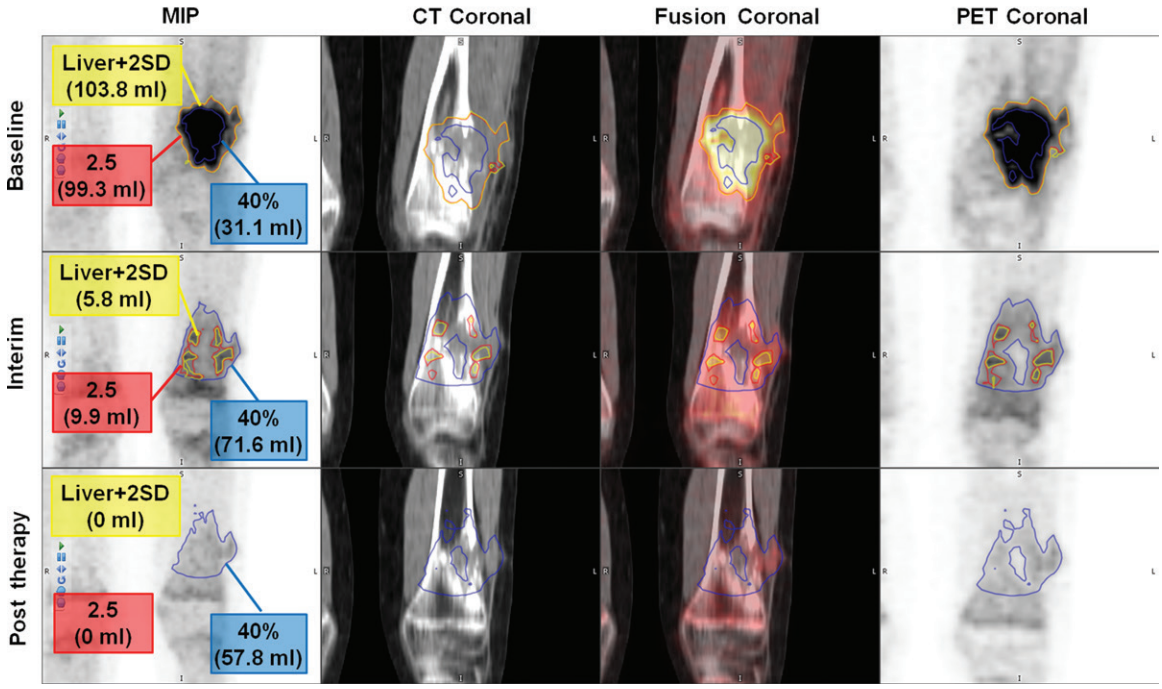

Figure 2: Representative case showing overestimation of metabolic tumor volume *(MTV)* by relative threshold at interim and posttherapy scans. MTV was measured by using fixed (standardized uptake *[SUV]* threshold of 2.5, in red) and liver-based threshold (liver + two SDs *[2SD]*, in yellow) predicts a favorable outcome at interim and posttherapy scan; however, MTV measured by using a relative threshold (40% of maximum SUV *[SUVmax]*, in blue) overestimated the tumor volume at interim and posttherapy scan, thus incorrectly predict the outcome. The patient was free of disease at 54 months. The MTV values are in the parenthesis within the boxes (yellow, red, and blue) of maximum intensity projection (*MIP*) images.

MTV (liver + 2SD) and TLG (liver + 2SD) were the most statistically significant parameters (MTV [liver + 2SD] of poor responders vs responders, −68.3% ± 29.9% vs −98.2% ± 5.3%, respectively; $P < .0002$; TLG [liver + 2SD] of poor responders vs responders, −73.8% ± 30.5% vs −98.9% ± 3.0%, respectively; $P = .00023$). To summarize, posttherapy and percent change (baseline to posttherapy) FDG PET parameters showed the greatest statistical associations with histologic response. Among the volumetric parameters using various thresholds, MTV and TLG using a threshold of 2.0, 2.5, liver + SD and liver + 2SD showed the better association with histologic response than MTV and TLG by using relative thresholds (Fig E1, Table E1 [online]).

Evaluation of Prognostic Factors for EFS

Baseline FDG PET parameters were predictive of EFS except SUV_{max}, MTV (60%), MTV (liver + 2SD), TLG (60%); MTV (SUV 2.0) and TLG (SUV, 2.5) were the most statistically significant parameters ($P = .006$, for both) (Fig E2 [online]). Also, all FDG PET parameters at interim scan were predictive of EFS except MTV (40%), MTV (60%), and TLG (40%), whereas MTV (SUV, 2.5) was the most statistically significant parameter ($P = .004$). Similarly at posttherapy scan, FDG PET parameters were predictive of EFS except MTV (40%), MTV (60%), TLG (40%), and TLG (60%), whereas MTV (liver + 2SD) was the most statistically significant parameter ($P = .01$). However, among percent changes between baseline and interim FDG PET/CT scans, only percent change (baseline to interim) of TLG (60%) was predictive of EFS ($P = .01$). Among percent changes between baseline and posttherapy scans, only percent change (baseline to posttherapy) of SUV_{peak} and MTV (SUV, 2.0) were the parameters to show predictive value for EFS ($P = .036$ and 0.033, respectively). In summary, metabolic parameters (SUV_{peak}) and volumetric parameters (MTV and TLG) were predictive of EFS at all points (baseline, interim, and posttherapy), and percent change of the parameters was not better than those from the single scan (Fig E2, Table E2 [online]).

Notably, SUV_{peak}, MTV (SUV, 2.5), and TLG (SUV, 2.5) at baseline could predict EFS ($P = .01$, .01, and .006, respectively) (Fig 1). In volumetric parameters, parameters measured by thresholds of absolute and liver-based thresholds were better than those by relative thresholds in predicting EFS. In particular, MTV measured by using relative threshold at interim and posttherapy imaging did not accurately visually represent the tumor volume, and could not predict the outcome of patients who showed a favorable response (Fig 2). Histologic

response and initial presence of metastasis tended to be predictive of EFS (P = .069 and .056, respectively).

Multivariable Cox regression analysis for EFS was performed by using SUV_{max}, SUV_{peak}, MTV (SUV, 2.5), TLG (SUV, 2.5), MTV (liver + 2SD), and TLG (liver + 2SD). Initial presence of metastasis and histologic response were used as covariates. Even at baseline scan, MTV (SUV, 2.5), TLG (SUV, 2.5), and TLG (liver + 2SD) were independent prognostic factors for EFS. At interim scan, MTV (SUV, 2.5), TLG (SUV, 2.5) were independent prognostic factors for EFS. At posttherapy scan, MTV (liver + 2SD) and TLG (liver + 2SD) were predictive of EFS independently. Notably, MTV (SUV, 2.5) and TLG (SUV, 2.5) were independent prognostic factors for EFS at baseline (MTV [SUV, 2.5] > 238.06 mL: hazard ratio, 5.0 [95% CI: 1.5, 16.8], P = .046; TLG (SUV, 2.5) > 981.97: hazard ratio, 5.7 [95% CI: 1.3, 24.5], P = .046) and interim scans (MTV [SUV, 2.5] > 35.8 mL: hazard ratio, 8.2 [95% CI: 1.5, 43.7], P = .046; and TLG [SUV, 2.5] > 117.7: hazard ratio, 8.2 [95% CI: 1.5, 43.7], P = .046) (Table 3).

Table 3

Multivariable Cox Regression Analysis by Using FDG PET/CT Parameters for Event-free Survival

Parameter	Hazard Ratio	*P* Value
SUV_{max}		
Baseline	4.457 (0.945, 21.014)	.071
Interim	4.675 (1.231, 17.749)	.054
Posttherapy	5.473 (0.937, 31.982)	.071
SUV_{peak}		
Baseline	9.926 (1.245, 79.156)	.06
Interim	3.723 (1.078, 12.857)	.068
Posttherapy	5.473 (0.937, 31.982)	.071
MTV (SUV 2.5)		
Baseline	5.024 (1.505, 16.770)	.046
Interim	8.155 (1.522, 43.693)	.046
Posttherapy	4.946 (1.008, 24.275)	.071
TLG (SUV 2.5)		
Baseline	5.740 (1.344, 24.509)	.046
Interim	8.155 (1.522, 43.693)	.046
Posttherapy	3.464 (0.849, 14.134)	.088
MTV (liver + 2SD)		
Baseline	3.163 (0.997, 10.037)	.071
Interim	7.349 (0.802, 67.353)	.088
Posttherapy	11.774 (1.567, 88.465)	.046
TLG (liver + 2SD)		
Baseline	4.433 (1.334, 14.737)	.046
Interim*		.995
Posttherapy	13.121 (2.137, 80.571)	.046

Note.—Data in parentheses are 95% confidence intervals (CIs). These parameters were adjusted for histologic response and initial presence of metastasis. All P values were adjusted based on the method of false discover rate. SUV_{max} = maximum standardized uptake value, SUV_{peak}= peak standardized uptake value, MTV (SUV 2.5) = MTV with threshold of SUV 2.5, MTV (liver + 2SD) = MTV with relative threshold of mean liver SUV + 2 standard deviation, TLG (SUV 2.5) = TLG with threshold of SUV 2.5, TLG (liver + 2SD) = TLG with liver based threshold of mean liver SUV + 2 standard deviation.

* Hazard ratio and CI were not stated when the CI tended toward infinity.

Evaluation of Prognostic Factors for OS

FDG PET parameters at baseline were predictive of OS except SUV_{max}, MTV (60%), MTV (liver + 2SD), and TLG (60%), whereas MTV (SUV, 2.0), TLG (SUV, 2.0), and TLG (SUV, 2.5) were the most statistically significant parameters (P = .001 for all). At interim FDG PET/CT scan, FDG PET parameters were predictive of OS except MTV (40%), MTV (60%), TLG (40%), and TLG (60%), whereas MTV (SUV, 2.0), MTV (SUV, 2.5), and TLG (SUV, 2.5) were the most significant parameters (P = .001, for all). Also at posttherapy scan, FDG PET parameters were predictive of OS except MTV (40%), MTV (60%), TLG (40%), and TLG (60%); SUV_{max}, SUV_{peak}, MTV (SUV, 2.5), and TLG (SUV, 2.5) were the most statistically significant parameters (P = .002, for all).

Among percent changes between baseline and interim FDG PET/CT scans, only percent change (baseline to interim) of TLG (60%) was predictive of OS (P = .01). Among percent change between baseline and posttherapy FDG PET/CT scans, only percent change (baseline to after therapy) of MTV (SUV, 2.0), MTV (40%), and TLG (SUV, 2.5) were parameters to show predictive value for OS (P = .03 ~ 0.04). Overall, FDG PET parameters from a single-point scan (baseline, interim, or posttherapy) were predictive of OS, and percent changes of the parameters were less predictive than those from a single-point scan (Fig E3, Table E3 [online]). In particular, SUV_{peak}, MTV (SUV, 2.5), and TLG (SUV, 2.5) at baseline could predict OS (P = .02, .004, and .001, respectively) (Fig 3). For volumetric parameters, absolute or liver-based thresholds were better than relative thresholds to predict OS. Histologic response and initial presence of metastasis were also predictive of OS (P = .047 and .026, respectively).

Multivariable Cox regression analysis for OS was performed by using SUV_{max}, SUV_{peak}, MTV (SUV, 2.5), TLG (SUV, 2.5), MTV (liver + 2SD), and TLG (liver + 2SD). Initial presence of metastasis and histologic response were used as covariates. At baseline scan, MTV (SUV, 2.5), TLG (SUV, 2.5), and TLG (liver + 2SD) were independent prognostic factors for OS, whereas SUV_{max} and SUV_{peak} were not. For example, a patient with small MTV (SUV, 2.5) (24.9 mL) (Fig 4a)

Figure 3

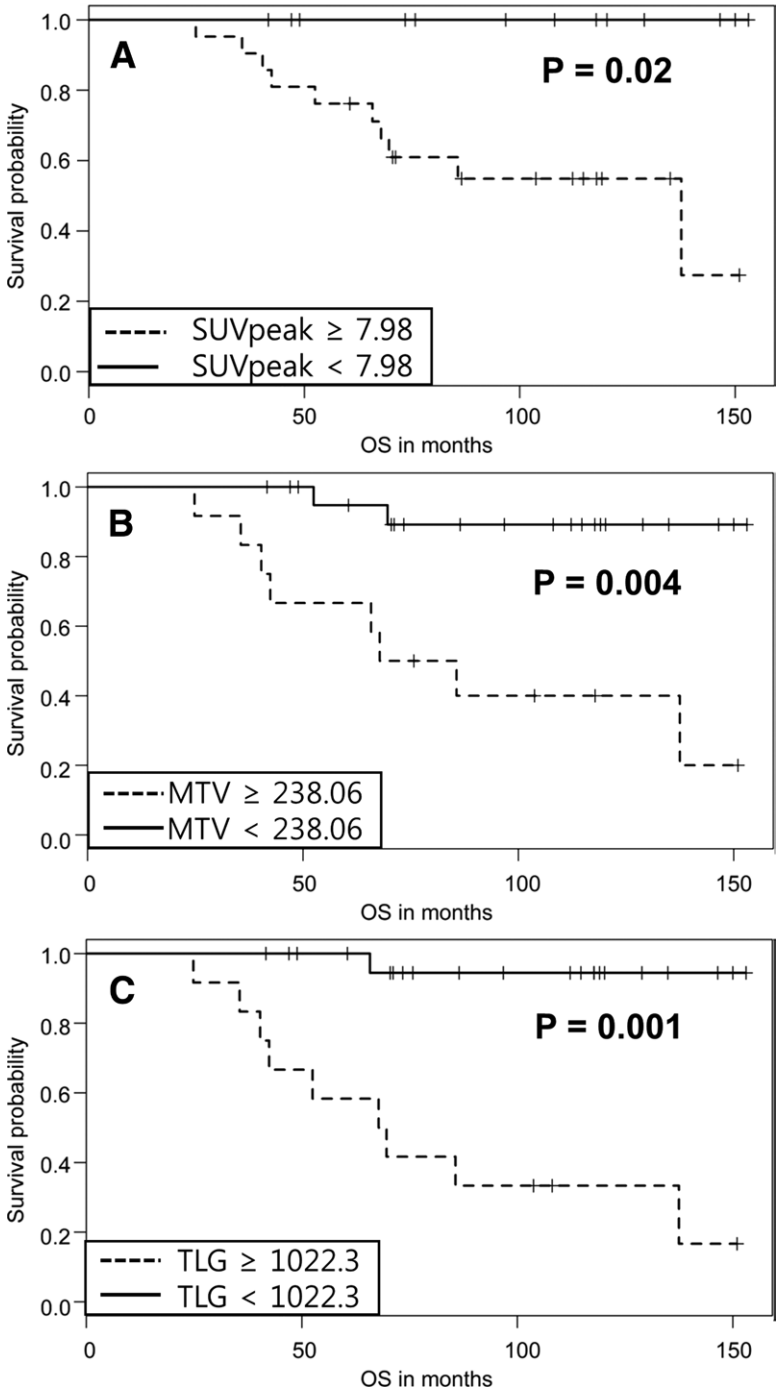

Figure 3: Kaplan-Meier Curves of overall survival *(OS)*. Kaplan-Meier Curves of OS for, *A*, peak standardized uptake value *(SUV; SUVpeak)*, *B*, metabolic tumor volume with SUV threshold of 2.5 *(MTV [SUV, 2.5])*, and, *C*, total lesion glycolysis with SUV threshold of 2.5 *(TLG [SUV, 2.5])* at baseline fluorodeoxyglucose PET/CT scan are presented.

showed a good outcome despite a poor histologic response after NCT. On the other hand, patients with large MTV (SUV, 2.5) experienced worse outcomes independent of histologic response after NCT (Figs 4b, 4c). At interim scan, only TLG (SUV, 2.5) and TLG (liver + 2SD) were independent prognostic factors for OS. Finally, at posttherapy scan, SUV_{max}, SUV_{peak}, MTV (SUV, 2.5), and TLG (SUV, 2.5) were independently predictive of OS. SUV_{max} and SUV_{peak} were independent prognostic factors for OS only at posttherapy scans. However, TLG (SUV, 2.5) was an independent prognostic factor for OS at all points (baseline TLG [SUV, 2.5] > 1022.3: hazard ratio, 29.4 [95% CI: 2.2, 392.3], P = .033; interim TLG [SUV, 2.5] > 120.4: hazard ratio, 34.8 [95% CI: 2.5, 483.8], P = .033; posttherapy TLG [SUV, 2.5] > 72.0: hazard ratio, 32.2 [95% CI: 2.1, 498.4], P = .033) (Table 4).

Discussion

We evaluated prognostic value of various FDG PET parameters before, during, and after NCT. We demonstrated that SUV_{max}, SUV_{peak}, MTV, and TLG at interim and posttherapy scans are associated with histologic response but not at baseline scan. MTV and TLG at baseline, interim, and posttherapy scans were predictive of EFS and OS after adjusting known prognostic factors.

Prognostic value of FDG PET in osteosarcoma was reported in some retrospective studies (3,7,12,13). Franzius et al (12) reported that higher tumor-to-nontumor ratio of FDG uptake correlated with poor outcome. Sato et al also reported that tumor SUV_{max} after NCT is associated with poor prognosis (13). Costelloe et al (7) showed that SUV_{max} was predictive of EFS both before and after NCT, and also SUV_{max} before NCT was predictive of OS. Additionally, TLG before NCT was predictive of OS but not of EFS. Additionally, MTV before NCT was an independent predictor of metastasis-free survival (3). To our knowledge, our study was the first prospective study to evaluate the prognostic value of FDG PET/CT by systematically

Figure 4

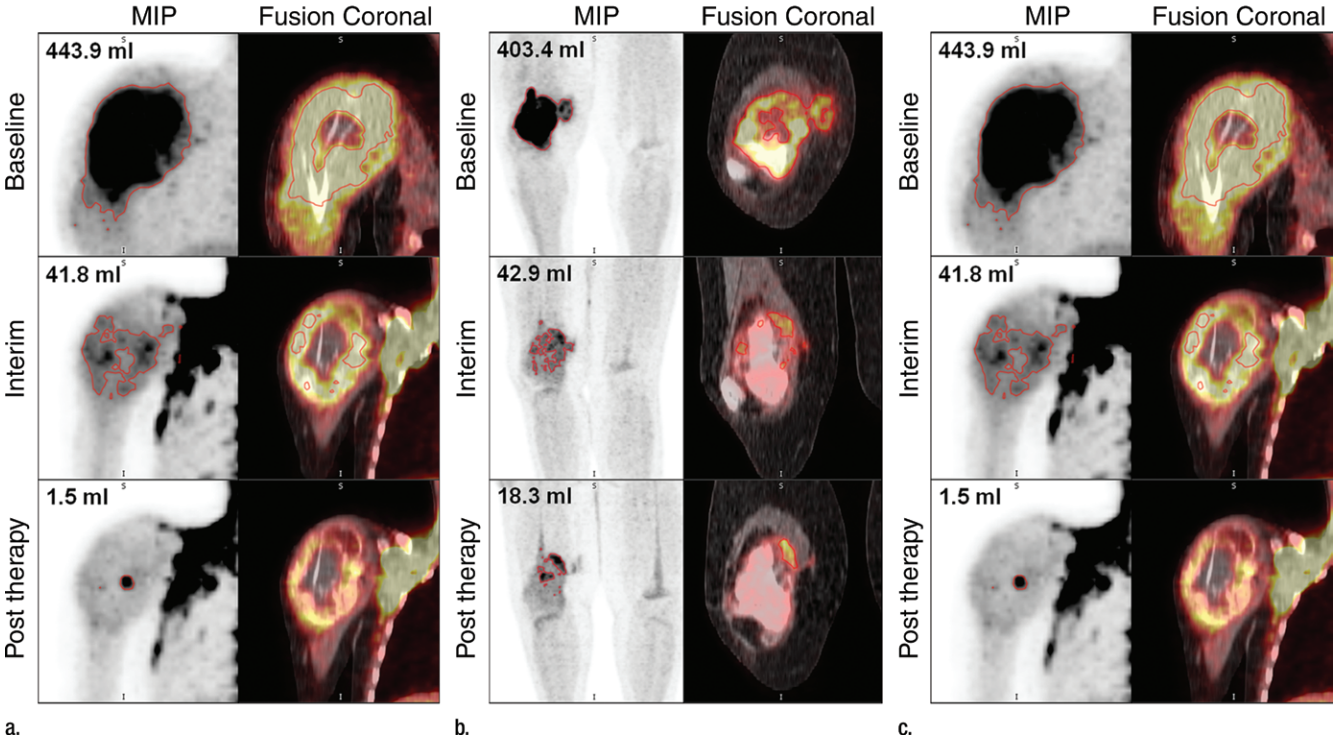

Figure 4: Representative cases with small and large metabolic tumor volume *(MTV)*. **(a)** A patient with small baseline MTV (24.9 mL) with poor histologic response after neoadjuvant chemotherapy *(NCT)* showed good clinical outcome (free of disease at 52 months). **(b)** A patient with large baseline MTV (403.4 mL) and poor histologic response after NCT experienced recurrence at 22 months after the surgery. **(c)** Another patient with a large baseline MTV (443.9 mL) and favorable histologic response after NCT experienced recurrence at 15.4 months and died at 55 months after the surgery. Red volume of interest indicates MTV by using standardized uptake value *(SUV)* threshold of 2.5 *(MTV [SUV, 2.5])*. MTV (SUV, 2.5) value at each point is on the upper-left part of maximum intensity projection (*MIP*) image. The optimized cut-off by Youden index for baseline MTV (SUV, 2.5) was 238.06 mL.

testing the predictive value of various parameters (SUV_{max}, SUV_{peak}, MTV, and TLG) with three scans at various points (at baseline, interim, and posttherapy) for three different types of end points (histologic response, EFS, and OS). Our study reproduced the individual results of the previous retrospective studies and further confirmed the prognostic value of FDG PET/CT. Additional information gained from our study is as follows: *(a)* SUV_{peak} was associated with histologic response (interim and posttherapy scan) and predictive of EFS and OS (at all points); *(b)* at the interim scan, SUV_{max}, SUV_{peak}, MTV, and TLG were predictive of EFS and OS; *(c)* MTV and TLG have prognostic value for EFS and OS at all points; and *(d)* percent change of the parameters did not provide better prognostic value compared with those from single-point scan. In a previous study (15) from our group that used the same patient data, we reported that there was no significant correlation between EFS and SUV_{max} at baseline, interim, and posttherapy scan. However, in the prior study, we only analyzed the correlation between the continuous values of EFS and SUV_{max}, which may lead to underestimation of the association. However, we dichotomized the patients according to various FDG PET parameters so that we could find the predictive values of the parameters in our study.

MTV was reported (3) to be an independent predictor of metastasis-free survival in patients with osteosarcoma, after adjustment for American Joint Committee on Cancer stage and histologic response, which are known powerful predictors of survival. In our study, various FDG PET parameters remained independent of stage and histologic response as additional prognostic factors for clinical outcome (EFS and OS).

Although to our knowledge SUV_{max} is the most widely used FDG PET parameter, the value is measured from only one voxel, and thus could be affected by noise or artifact (17). SUV_{peak} has been proposed to overcome the shortcoming of SUV_{max}, defined as the average SUV within a small, fixed size region of interest centered on a high uptake part of the tumor (18). SUV_{peak} has been reported (11,19,20) to be prognostic in several types of malignancy. In patients with osteosarcoma, we revealed that SUV_{peak} at the posttherapy scan is an independent prognostic factor for OS.

Multiple threshold methods of deriving MTV have been used in studies to predict survival in various malignancies,

Table 4

Multivariable Cox Regression Analysis by Using FDG PET Parameters for Overall Survival

Parameter	Hazard Ratio	*P* Value
SUV_{max}		
Baseline	4.696 (0.548, 40.207)	.19
Interim	3.462 (0.664, 18.045)	.186
Posttherapy	32.512 (2.254, 468.962)	.033
SUV_{peak}		
Baseline*		.945
Interim	4.921 (0.577, 41.960)	.186
Posttherapy	32.512 (2.254, 468.962)	.033
MTV (SUV 2.5)		
Baseline	29.447 (2.211, 392.266)	.033
Interim*		.945
Posttherapy	32.217 (2.083, 498.397)	.033
TLG (SUV 2.5)		
Baseline	29.447 (2.211, 392.266)	.033
Interim	34.789 (2.501, 483.842)	.033
Posttherapy	32.217 (2.083, 498.397)	.033
MTV (liver + 2SD)		
Baseline	7.948 (1.212, 52.125)	.056
Interim*		.945
Posttherapy	6.889 (0.743, 63.864)	.146
TLG (liver + 2SD)		
Baseline	9.331 (1.535, 56.723)	.034
Interim	9.616 (1.368, 67.608)	.046
Posttherapy	5.592 (0.650, 48.096)	.176

Note.—Data in parentheses are 95% confidence intervals (CIs). The parameters were adjusted for histologic response and initial presence of metastasis. All *P* values were adjusted on the basis of the method of false discover rate. SUV_{max} = maximum standardized uptake value, SUV_{peak} = peak standardized uptake value, MTV (SUV 2.5) = MTV with threshold of SUV 2.5, MTV (liver + 2SD) = MTV with relative threshold of mean liver SUV + 2 standard deviation, TLG (SUV 2.5) = TLG with threshold of SUV 2.5, TLG (liver + 2SD) = TLG with liver based threshold of mean liver SUV + 2 standard deviation.

* Hazard ratio and CI were not stated when the CI tended toward infinity.

including absolute SUV threshold methods such as SUV 2.0 or SUV 2.5 (21), relative threshold methods such as 40%–60% of SUV_{max} of the lesion (22), and liver-based threshold methods (23). Among osteosarcoma studies, Costelloe et al (7) used 45% of SUV_{max}, while Byun et al (3) used SUV 2.0 and 2.5, and 45% of SUV_{max}. In our study, we used multiple thresholds to define MTV including absolute thresholds, relative thresholds, and liver-based thresholds, and the predictive values of these volumetric parameters were comparatively assessed. We found that prognostic values of volumetric parameters measured by using absolute and liver-based thresholds were comparable and better than those measured by using relative thresholds to accurately represent the PET tumor volume. MTV by using a relative threshold with 40% and 60% tumor SUV_{max} tended to underestimate the PET tumor volume at baseline scans in tumors with very high SUV_{max}, and overestimate PET tumor volume at interim and posttherapy scans because of decreasing tumor SUV_{max} after NCT. This result also has been reported in a previous study by Erdi et al (24), which showed underestimation of the volume by using relative threshold method in a tumor with low signal to background. Thus, our data indicate that measuring MTV by using fixed or liver-based thresholds is superior for prediction of outcome than relative threshold for high tumor SUV_{max} baseline scans and especially at interim or posttherapy scans.

Our study has several limitations. First, our study was an exploratory study and thus was not planned with power analysis. Second, we analyzed relatively small sample size from a single institute which may limit generalizability. However, our study is a prospective study of patients uniformly treated and evaluated at fixed intervals to determine the prognostic value of metabolic and volumetric parameters of FDG PET/CT in pediatric patients with osteosarcoma. Another limitation of our study is the adequate but somewhat limited clinical follow-up time for the survived patients (median, 44 months; range, 17–61 months) and exclusion of two patients for histologic analysis.

In summary, a single-point FDG PET/CT scan either at baseline, interim, or posttherapy was predictive of EFS and OS even after adjusting for histologic response and initial presence of metastasis. Volumetric parameters by using absolute SUV threshold or liver-based threshold were comparable to predict EFS and OS; however, relative threshold was not consistently predictive. These findings need further validation in larger prospective multi-institutional studies before these parameters could be used as a prognostic and predictive biomarker during treatment for risk stratification and therapy adjustment.

Disclosures of Conflicts of Interest: H.J.I. disclosed no relevant relationships. **Y.Z.** disclosed no relevant relationships. **H.W.** disclosed no relevant relationships. **J.W.** disclosed no relevant relationships. **N.C.D.** disclosed no relevant relationships. **F.N.** disclosed no relevant relationships. **B.L.S.** disclosed no relevant relationships. **S.Y.C.** disclosed no relevant relationships.

References

1. Bielack SS, Kempf-Bielack B, Delling G, et al. Prognostic factors in high-grade osteosarcoma of the extremities or trunk: an analysis of 1,702 patients treated on neoadjuvant cooperative osteosarcoma study group protocols. J Clin Oncol 2002;20(3):776–790.

2. Davis AM, Bell RS, Goodwin PJ. Prognostic factors in osteosarcoma: a critical review. J Clin Oncol 1994;12(2):423–431.

3. Byun BH, Kong CB, Park J, et al. Initial metabolic tumor volume measured by 18F-FDG PET/CT can predict the outcome of os-

teosarcoma of the extremities. J Nucl Med 2013;54(10):1725–1732.

4. O JH, Choi WH, Han EJ, et al. The Prognostic Value of (18)F-FDG PET/CT for Early Recurrence in Operable Breast Cancer: Comparison with TNM Stage. Nucl Med Mol Imaging 2013;47(4):263–267.

5. Hyun SH, Choi JY, Shim YM, et al. Prognostic value of metabolic tumor volume measured by 18F-fluorodeoxyglucose positron emission tomography in patients with esophageal carcinoma. Ann Surg Oncol 2010;17(1):115–122.

6. Im HJ, Kim TS, Park S-Y, et al. Prediction of tumour necrosis fractions using metabolic and volumetric 18F-FDG PET/CT indices, after one course and at the completion of neoadjuvant chemotherapy, in children and young adults with osteosarcoma. Eur J Nucl Med Mol Imaging 2012;39(1):39–49.

7. Costelloe CM, Macapinlac HA, Madewell JE, et al. 18F-FDG PET/CT as an indicator of progression-free and overall survival in osteosarcoma. J Nucl Med 2009;50(3):340–347.

8. Na F, Wang J, Li C, Deng L, Xue J, Lu Y. Primary tumor standardized uptake value measured on F18-Fluorodeoxyglucose positron emission tomography is of prediction value for survival and local control in non-small-cell lung cancer receiving radiotherapy: meta-analysis. J Thorac Oncol 2014;9(6):834–842.

9. Sarker A, Im HJ, Cheon GJ, et al. Prognostic Implications of the SUVmax of Primary Tumors and Metastatic Lymph Node Measured by 18F-FDG PET in Patients With Uterine Cervical Cancer: A Meta-analysis. Clin Nucl Med 2016;41(1):34–40.

10. Pak K, Cheon GJ, Nam HY, 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):884–890.

11. Machtay M, Duan F, Siegel BA, et al. Prediction of survival by [18F]fluorodeoxyglucose positron emission tomography in patients with locally advanced non-small-cell lung cancer undergoing definitive chemoradiation therapy: results of the ACRIN 6668/RTOG 0235 trial. J Clin Oncol 2013;31(30):3823–3830.

12. Franzius C, Bielack S, Flege S, Sciuk J, Jürgens H, Schober O. Prognostic significance of (18)F-FDG and (99m)Tc-methylene diphosphonate uptake in primary osteosarcoma. J Nucl Med 2002;43(8):1012–1017.

13. Sato J, Yanagawa T, Dobashi Y, Yamaji T, Takagishi K, Watanabe H. Prognostic significance of 18F-FDG uptake in primary osteosarcoma after but not before chemotherapy: a possible association with autocrine motility factor/phosphoglucose isomerase expression. Clin Exp Metastasis 2008;25(4):427–435.

14. Shulkin B, Davis J, Navid F, et al. FDG PET/CT in pediatric osteosarcoma: Comparison of standard & delayed scanning. J Nucl Med 2013;54(Suppl 2):478.

15. Davis JC, Daw NC, Navid F, et al. FDG Uptake During Early Adjuvant Chemotherapy Predicts Histologic Response in Pediatric and Young Adult Patients with Osteosarcoma. J Nucl Med 2017 Jun 13. [Epub ahead of print].

16. Hurley C, McCarville MB, Shulkin BL, et al. Comparison of (18) F-FDG-PET-CT and Bone Scintigraphy for Evaluation of Osseous Metastases in Newly Diagnosed and Recurrent Osteosarcoma. Pediatr Blood Cancer 2016;63(8):1381–1386.

17. Boellaard R, Krak NC, Hoekstra OS, Lammertsma AA. Effects of noise, image resolution, and ROI definition on the accuracy of standard uptake values: a simulation study. J Nucl Med 2004;45(9):1519–1527.

18. 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–150S.

19. Hoshikawa H, Yamamoto Y, Mori T, et al. Predictive value of SUV-based parameters derived from pre-treatment (18)F-FLT PET/CT for short-term outcome with head and neck cancers. Ann Nucl Med 2014;28(10):1020–1026.

20. Lee HS, Kim HO, Hong YS, et al. Prognostic value of metabolic parameters in patients with synchronous colorectal cancer liver metastasis following curative-intent colorectal and hepatic surgery. J Nucl Med 2014;55(4):582–589.

21. Im HJ, Kim YK, Kim YI, Lee JJ, Lee WW, Kim SE. Usefulness of Combined Metabolic-Volumetric Indices of (18)F-FDG PET/CT for the Early Prediction of Neoadjuvant Chemotherapy Outcomes in Breast Cancer. Nucl Med Mol Imaging 2013;47(1):36–43.

22. Fonti R, Larobina M, Del Vecchio S, et al. Metabolic tumor volume assessed by 18F-FDG PET/CT for the prediction of outcome in patients with multiple myeloma. J Nucl Med 2012;53(12):1829–1835.

23. Dholakia AS, Chaudhry M, Leal JP, et al. Baseline metabolic tumor volume and total lesion glycolysis are associated with survival outcomes in patients with locally advanced pancreatic cancer receiving stereotactic body radiation therapy. Int J Radiat Oncol Biol Phys 2014;89(3):539–546.

24. Erdi YE, Mawlawi O, Larson SM, et al. Segmentation of lung lesion volume by adaptive positron emission tomography image thresholding. Cancer 1997;80(12 Suppl):2505–2509.