Impact of Partial-Volume Effect Correction on the Predictive and Prognostic Value of Baseline ^{18}F-FDG PET Images in Esophageal Cancer

Mathieu Hatt[1], Adrien Le Pogam[1,2], Dimitris Visvikis[1], Olivier Pradier[1,3], and Catherine Cheze Le Rest[1]

[1]INSERM, U650 LaTIM, CHRU Morvan, Brest, France; [2]MRC Clinical Sciences Centre, Hammersmith Hospital, London, United Kingdom; and [3]Department of Radiotherapy, CHRU Morvan, Brest, France

The objective of this study was to investigate the clinical impact of partial-volume effect (PVE) correction on the predictive and prognostic value of metabolically active tumor volume (MATV) measurements on ^{18}F-FDG PET baseline scans for therapy response and overall survival in esophageal cancer patients. **Methods:** Fifty patients with esophageal cancer treated with concomitant radiochemotherapy between 2004 and 2008 were retrospectively considered. PET baseline scans were corrected for PVE with iterative deconvolution incorporating wavelet denoising. MATV delineation on both original and corrected images was performed using the automatic fuzzy locally adaptive Bayesian methodology. Several parameters were extracted considering the original and corrected images: maximum and peak standardized uptake value (SUV), mean SUV, MATV, and total lesion glycolysis (TLG) (TLG = MATV × mean SUV). The predictive value of each parameter with or without correction was investigated using Kruskal–Wallis tests, and the prognostic value was determined with Kaplan–Meier curves. **Results:** Whereas PVE correction had a significant quantitative impact on the absolute values of the investigated parameters, their clinical value within the clinical context of interest was not significantly modified—a result that was observed for both overall survival and response to therapy. The hierarchy between parameters was the same before and after correction. SUV measurements (maximum, peak, and mean) had nonsignificant ($P > 0.05$) predictive or prognostic value, whereas functional tumor-related measurements (MATV and TLG) were significant ($P < 0.002$) predictors of response and independent prognostic factors. **Conclusion:** PVE correction does not improve the predictive and prognostic value of baseline PET image–derived parameters in esophageal cancer patients.

Key Words: esophageal cancer; response to therapy; overall survival; PET; partial volume effects; SUV; tumor volume; total lesion glycolysis

J Nucl Med 2012; 53:12–20
DOI: 10.2967/jnumed.111.092775

With a worldwide estimated 5-y survival of only 15% (*1*), esophageal cancer is the third most common malignancy of the digestive tract and is a leading cause of cancer mortality. Its incidence is still increasing, and there is a growing concern regarding its effective management (*2*). Surgical resection remains the most effective treatment; however, many patients have a locally advanced esophageal carcinoma at diagnosis and neoadjuvant therapy before surgery has demonstrated improved survival in such cases (*3*). The maximum improvement in terms of increased overall survival from neoadjuvant treatment is observed for patients who achieve a complete pathologic response (only 15%–30% of cases), with no residual cancer cells in the primary tumor or lymph nodes (*4*). On the other hand, nonresponders (NRs) may be unnecessarily affected by toxicity (*5*). The development of an early diagnostic test offering noninvasive prediction of the response to therapy or survival is therefore of great interest. For tumors that cannot be surgically removed, combined radiochemotherapy is the preferred treatment. In this case too, early assessment of response to therapy would allow a modification in the management of nonresponding patients early during treatment. Such a response assessment becomes even more critical when one considers the availability of new targeted drugs that could be tested with higher efficiency if applied early (*6*).

Along with the standardized uptake values (SUVs) (maximum SUV [SUV_{max}] or peak SUV [SUV_{peak}]) usually considered in clinical practice, other parameters describing functional lesions—such as metabolically active tumor volume (MATV, defined as the tumor volume that can be seen and delineated on an ^{18}F-FDG PET image) (*7*), mean SUV (SUV_{mean}), and total lesion glycolysis (TLG, defined as the product of MATV and its associated SUV_{mean}) (*8*)—have been investigated. The prognostic value of these parameters in esophageal cancer patients for overall or disease-free survival has been demonstrated (*9–12*). On the other hand regarding therapy prediction, several studies on different cancer models have recently suggested using the baseline scan only, instead of the comparison of pretreatment and posttreatment scans (late assessment) or

Received Jul. 6, 2011; revision accepted Sep. 6, 2011.
For correspondence or reprints contact: Mathieu Hatt, LaTIM, INSERM U650, CHRU Morvan, 5 Ave. Foch, 29609 Brest, France.
E-mail: hatt@univ-brest.fr

during-treatment scans (early assessment) (*13*). Such investigations were, for instance, performed in pleural mesothelioma (*14*), non-Hodgkin lymphoma (*15*), and esophageal cancer (*7,16*), demonstrating higher statistical value for MATV-based parameters than SUV measurements, whose predictive value has been found to be conflicting (*17*).

However, in most of these studies, no partial-volume effect (PVE) correction was applied, possibly explaining the observed limited value of SUV. The impact of PVE correction on the clinical value of SUV measurements has been investigated by a limited number of authors. Hoetjes et al. (*18*) investigated the impact of 4 PVE correction strategies on 15 breast cancer patients, regarding the early metabolic PET response after 1 cycle of chemotherapy. The SUV decrease between the pretreatment scan and the scan early during treatment was found to be lower after PVE correction (26%–27% vs. 31%) for the first 3 methods but not for the fourth one based on binary tumor masks (30%). Van Heijl et al. (*19*) recently demonstrated a nonsignificant impact of PVE correction on the correlation between disease-free survival and ^{18}F-FDG PET SUV measurements in 52 esophageal cancer patients. In this study, a PVE correction method based on binary tumor masks generated with adaptive thresholding delineation was used, and disease-free survival was the only clinical endpoint investigated. Both the use of adaptive thresholding and the PVE correction method based on tumor masks assume a homogeneous tracer distribution in both tumor and background and are therefore likely to provide only approximate correction (*20*). On the other hand, no data are currently available regarding the impact of PVE correction on the value of baseline ^{18}F-FDG PET–based measurements for the prediction of overall survival and response to therapy in esophageal cancer.

The current study was therefore performed to investigate the impact of an advanced PVE correction methodology and the use of an accurate MATV delineation approach on both the predictive and the prognostic value of baseline ^{18}F-FDG PET scan–derived parameters.

MATERIALS AND METHODS

Patients

Fifty consecutive patients with newly diagnosed esophageal cancer were included and retrospectively analyzed. The characteristics of the patients are given in Table 1. Most of the patients (45 of 50) were men, aged 65 ± 9 y at the time of diagnosis. Seventy-four percent of the tumors originated from the middle and lower esophagus, and 72% were squamous cell carcinoma. None of the patients underwent surgery, and all were treated with concomitant radiochemotherapy between 2004 and 2009. The therapy regime included 3 courses of 5-fluorouracil and cisplatin and a median radiation dose of 60 Gy given in 180-cGy fractions delivered once daily, 5 d a week for 6–7 wk. As part of the routine procedure for the initial staging in esophageal cancer, each patient was referred for an ^{18}F-FDG PET study before treatment, and these baseline scans were used in this study.

TABLE 1
Patient Demographics and Clinical Characteristics

Parameter	No. of patients (n = 50)
Sex	
Male	45 (90)
Female	5 (10)
Site	
Upper esophagus	13 (26)
Middle esophagus	20 (40)
Lower esophagus	17 (34)
Histology type	
Adenocarcinoma	14 (28)
Squamous cell carcinoma	36 (72)
Histologic differentiation	
Well differentiated	14 (28)
Moderately differentiated	12 (24)
Poorly differentiated	5 (10)
Unknown	19 (38)
TNM stage	
T1	7 (14)
T2	8 (16)
T3	24 (48)
T4	11 (22)
N0	20 (40)
N1	30 (60)
M0	34 (68)
M1	16 (32)
American Joint Committee on Cancer stage	
I	4 (8)
IIA	8 (16)
IIB	6 (12)
III	16 (32)
IVA	16 (32)

Age range of patients was 45–84 y, and median was 69 y. Data in parentheses are percentages.

Overall survival was determined as the time between initial diagnosis and last follow-up or death. Response to therapy was evaluated 1 mo after the completion of the concomitant radiochemotherapy using conventional thoracoabdominal CT and endoscopy. Patients were classified as NRs (including stable and progressive disease), partial responders (PRs), or complete responders (CRs). Response evaluation was based on CT evolution between pretreatment and posttreatment scans using response evaluation criteria in solid tumors (*21*). Patients also underwent fibroscopy in the case of partial or complete response. Complete response was confirmed by the absence of visible disease in the endoscopy and no viable tumor on biopsy. Partial CT response was confirmed by macroscopic residual (disease >10% viable) on biopsy. No discordance was observed between pathologic, when available, and CT evaluation. The current analysis was performed after an approval by the institutional ethics review board.

^{18}F-FDG PET Acquisitions

^{18}F-FDG PET studies were performed before the treatment. Patients were instructed to fast for at least 6 h before an injection of ^{18}F-FDG (5 MBq/kg). Static emission images were acquired from head to thigh beginning 60 min after injection and with 2 min per bed position, on a Gemini PET/CT system (Philips). Images were reconstructed using the row-action

maximum-likelihood 3-dimensional algorithm according to standard clinical protocol: 2 iterations, relaxation parameter of 0.05, 5-mm 3-dimensional gaussian postfiltering, a 4 × 4 × 4 mm voxel grid sampling, and attenuation correction based on a low-dose CT scan.

PET Image PVE Correction and Image Analysis

Images were corrected for PVE using an iterative deconvolution methodology that has been previously validated (*22*). Its principle is to iteratively estimate the inversion of the scanner's point spread function, which is assumed to be known and spatially invariant in the field of view. The considered lesions were all in the same body region, and this approximation should therefore not significantly affect the applied correction on a patient-by-patient comparison basis. Iterative deconvolution methods, such as those of Lucy-Richardson (L-R) (*23,24*) or Van Cittert (*25*), are known for the amplification of noise associated with an increasing number of iterations. To solve this issue, wavelet-based denoising of the residual was introduced within the iterative L-R deconvolution using Bayeshrink filtering (*26*), leading to images corrected for PVE without significant noise addition. The following are advantages of this methodology: it is able to generate entire whole-body corrected images independently of any manual or automatic segmentation of regions of interest, and it is voxel-based and therefore does not assume homogeneous regional radiotracer distributions for the tumor or surrounding background.

Tumor Delineation and Parameter Extraction

For each patient, the tumor was identified on the baseline pretreatment PET images by an experienced nuclear physician. It was then delineated using the fuzzy locally adaptive Bayesian algorithm (*20,27*) on both the original (without PVE correction) and the PVE-corrected images. This segmentation approach has been shown to give both robust and reproducible functional volume delineations under variable image noise characteristics (*28,29*).

The following parameters were subsequently extracted from each baseline image with or without correction for PVE: SUV_{max}, SUV_{peak} (defined as the mean of SUV_{max} and its 26 neighbors [roughly corresponding to a 1-cm region of interest]), SUV_{mean} within the volume, MATV, and TLG (determined by multiplying SUV_{mean} with the corresponding MATV).

Statistical Analysis

Pearson coefficients were used to estimate correlation between the image-derived parameters, and paired *t* tests were used to characterize the differences between uncorrected and corrected parameters. The correlation between response to therapy and each parameter was investigated using the Kruskal–Wallis test as a nonparametric statistic allowing the comparison of parameter distributions associated with each category of response (CR, PR, and NR). This test does not assume a normal distribution of variables, and the computation of its statistic H is based on ranks instead of absolute values of variables (*30*). Regarding survival, for each considered parameter, Kaplan–Meier survival curves were generated (*31*) for which the most discriminating threshold value allowing differentiation of the groups of patients was identified using receiver-operating-characteristic methodology (*32*). The prognostic value of each parameter in terms of overall survival was assessed by the log-rank test.

The significance of the following factors (with or without correction) was tested: SUV_{max}, SUV_{peak}, MATV, SUV_{mean}, and TLG. All tests were performed 2-sided using the MedCalc statistical software (MedCalc Software), and *P* values below 0.05 were considered statistically significant.

RESULTS

Impact of PVE Correction on Image-Derived Parameters

The PVE correction affected the images that could be assessed visually, with a higher contrast between the tumor and the surrounding tissues, as can be seen in Figure 1 and is illustrated using profiles in Figure 2. Table 2 provides the distributions of volumes and associated parameters measured in original and corrected images.

MATVs delineated on original images and images corrected for PVE were highly correlated ($r > 0.998$; confidence

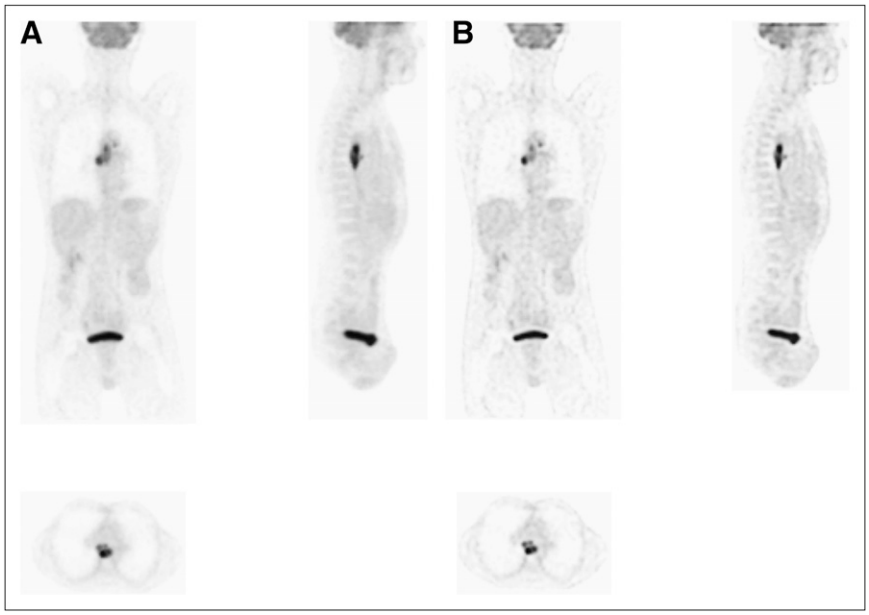

FIGURE 1. Illustration of iterative deconvolution PVE correction on whole-body ^{18}F-FDG PET image, with original image (A) and corrected image (B).

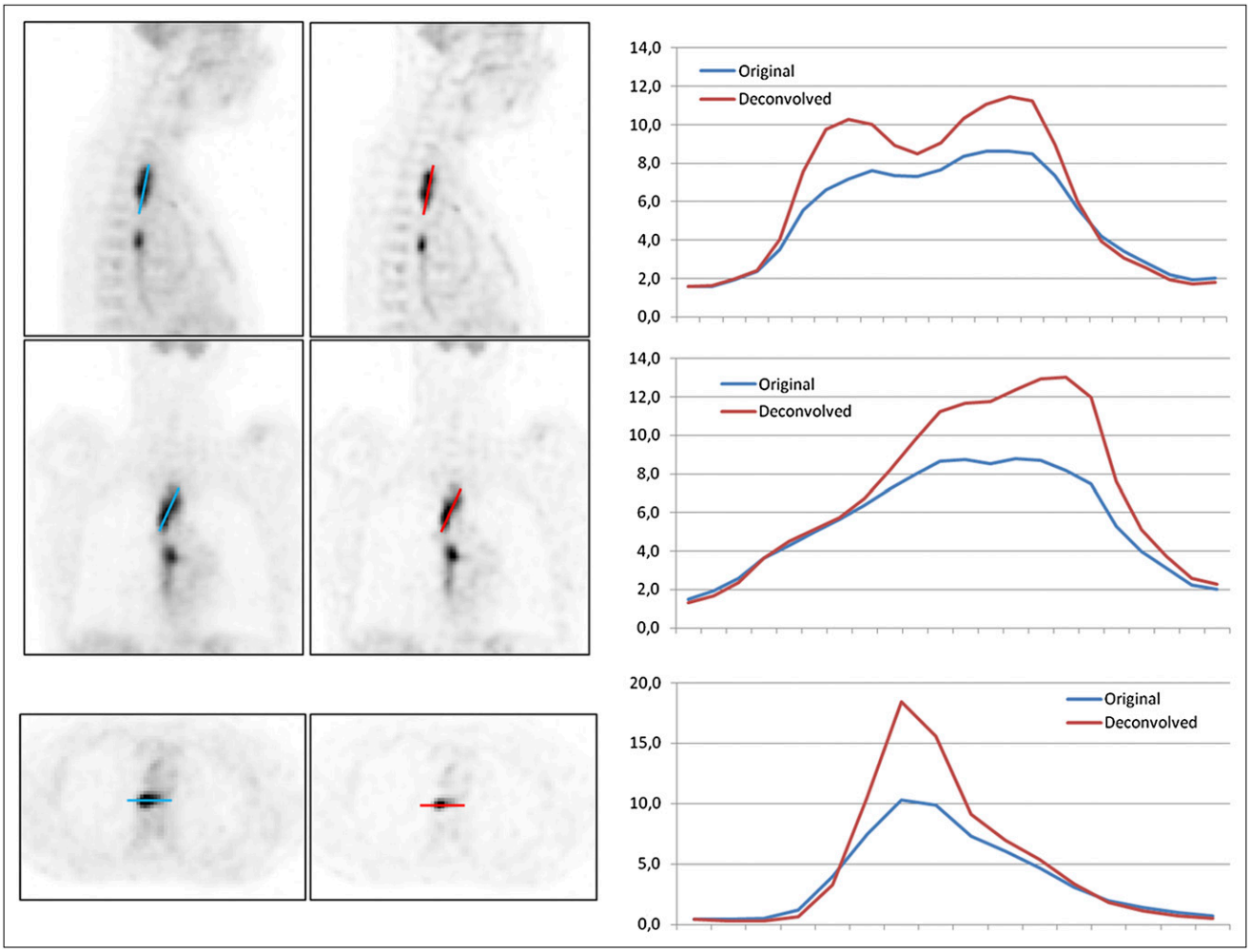

FIGURE 2. Qualitative differences between original and corrected PET images of esophageal lesion of MATV above 25 cm^3 using profiles on axial, sagittal, and coronal planes.

interval, 0.997–0.999; $P < 0.0001$). However, MATVs delineated on PVE-corrected images were systematically smaller ($P < 0.001$) by on average $-10\% \pm 5\%$ (range, -1.5% to -22.4%), resulting in a mean volume difference of -4 ± 3 cm^3 (40 ± 36 cm^3 vs. 36 ± 34 cm^3). This difference is illustrated on 3 different tumors in Figure 3. There was no significant correlation between these differences and the PET lesion volumes ($r < 0.2$, $P > 0.18$).

All primary lesions were detected by ^{18}F-FDG PET and exhibited a rather high uptake with a mean SUV_{max} of 10 ± 4. As expected, SUV_{peak} and SUV_{mean} measurements were comparatively lower (8 ± 3 and 6 ± 2, respectively). All SUV measurements are summarized in Table 2. After iterative deconvolution, SUV_{max}, SUV_{peak}, and SUV_{mean} were 15 ± 6, 10 ± 4, and 7 ± 3, respectively. All were significantly higher than noncorrected values ($P < 0.05$). SUV_{max} increased by $54\% \pm 23\%$ (range, 18%–157%), whereas the impact on SUV_{peak} and SUV_{mean} was lower, with a mean increase of $27\% \pm 10\%$ (range, 8%–51%) and $28\% \pm 11\%$ (range, 9%–59%), respectively. Considering the PVE correction–induced decrease of MATV ($-10\% \pm 5\%$) and increase of corresponding SUV_{mean}

TABLE 2
Distributions of Parameters With and Without PVE Correction

Definition	Notation	Original mean ± SD	PVE correction mean ± SD
Highest SUV	SUV_{max}	9.7 ± 3.9	14.9 ± 6.1
Mean of SUV_{max} and its 26 neighbors	SUV_{peak}	8.0 ± 3.3	10.1 ± 4.0
SUV_{mean} within MATV	SUV_{mean}	5.8 ± 2.4	7.4 ± 3.1
MATV (cm^3)	MATV	39.9 ± 36.1	36.2 ± 33.7
Total lesion glycolysis (g)	TLG	218.1 ± 208.3	235.8 ± 218.1

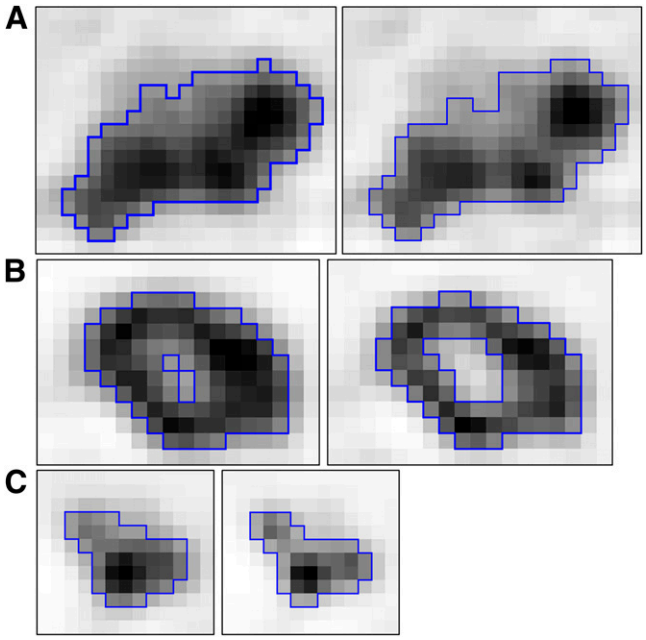

FIGURE 3. Examples of fuzzy locally adaptive Bayesian delineation results (blue contours) on original (left) and corrected (right) PET images with large, slightly heterogeneous MATV (A); MATV with necrotic core (B); and small, homogeneous MATV (C).

(+28% ± 11%), PVE correction resulted in significantly higher TLG values (+14% ± 12%; range, −2 to +50%) ($P < 0.0001$).

The increases of SUV_{max} and SUV_{peak} after PVE correction did not correlate with MATV ($r < 0.2$, $P > 0.2$), whereas the increase of SUV_{mean} correlated inversely with MATV ($r = -0.79$, $P < 0.0001$), with higher increases observed for smaller volumes.

Impact of PVE Correction on Predictive and Prognostic Values

Twenty-five patients were classified as PR, 11 were CR, and 14 were NR (including stable and progressive disease). With a median follow-up of 60 mo (range, 10–84 mo), the median overall survival was 12 mo and the 1-y and 2-y survival rates were 60% and 35%, respectively. At the time of last follow-up, 10 patients were alive with no evidence of disease, 9 were alive with recurrent disease, and 31 had died. Survival was significantly correlated with response, as overall survival was 24 ± 15 (median, 21), 22 ± 20 (median, 14), and 9 ± 4 (median, 10) months for CR, PR, and NR, respectively ($P < 0.01$). Results concerning the prognostic and predictive values of all considered parameters with and without PVE correction are summarized in Tables 3 and 4.

Initial SUV_{max}, whether corrected for PVE or not, was not predictive of response to therapy ($P = 0.2$ and 0.3 for SUV_{max} and SUV_{max} with PVE correction, respectively), although CRs tend to have a smaller SUV_{max} (7.8 ± 4.2 and 12.2 ± 6.6 after PVE correction) than PRs and NRs (10.2 ± 3.7 and 10.3 ± 3.8 for PR and NR, respectively, and 15.9 ± 6.0 and 15.5 ± 5.7, respectively, after PVE correction) (Fig. 4A). SUV_{peak} led to slightly more differentiated groups of response without reaching statistical significance ($P = 0.08$), with a mean value of 6.2 ± 3.6 in CRs, whereas both PRs and NRs were characterized by a similarly higher SUV_{peak} (8.5 ± 3.1 and 8.5 ± 3.2 for PRs and NRs, respectively). After PVE correction, the results using SUV_{peak} were similar, with 7.8 ± 4.4, 10.7 ± 3.7, and 10.8 ± 3.9 for CRs, PRs, and NRs, respectively ($P = 0.1$). The SUV_{mean} measurements could not significantly predict response ($P = 0.07$), and the differentiation between the 3 groups of response considered on the basis of SUV_{mean} was still not possible after PVE correction ($P > 0.14$).

None of the SUV measurements was a significant prognostic factor in the univariate analysis, despite a trend for longer survival associated with lower SUV (maximum, peak, or mean). For instance, an SUV_{max} below a threshold of 8 or an SUV_{mean} under 6.5 tend to be associated with a better outcome and a median survival of 20 versus 13 mo

TABLE 3
Kruskal–Wallis Test Results

Parameter	H	*P*	Response differentiation? ($P < 0.05$)		
			CR ($n = 11$)/NR ($n = 14$)	CR ($n = 11$)/PR ($n = 25$)	PR ($n = 25$)/NR ($n = 14$)
SUV_{max}	3.6	0.17	No	No	No
SUV_{max} with PVE correction	2.4	0.31	No	No	No
SUV_{peak}	5.1	0.08	No	No	No
SUV_{peak} with PVE correction	4.7	0.10	No	No	No
SUV_{mean}	5.5	0.07	No	No	No
SUV_{mean} with PVE correction	3.9	0.14	No	No	No
MATV (cm^3)	20.7	<0.0001	Yes	Yes	Yes
MATV with PVE correction (cm^3)	20.7	<0.0001	Yes	Yes	Yes
TLG (g)	25.1	<0.0001	Yes	Yes	Yes
TLG with PVE correction (g)	25.2	<0.0001	Yes	Yes	Yes

H statistic and associated *P* value are given for each parameter, with ability to differentiate ($P < 0.05$) each pair of response groups among patients.

TABLE 4
Univariate Analysis Results Using Kaplan–Meier Survival Curves

Parameter	Threshold	HR	HR 95% confidence interval	*P*	Median survival (mo)
SUV_{max}	8	1.5	0.7–3.1	0.28	20 vs. 13
SUV_{max} with PVE correction	11	1.6	0.7–3.2	0.26	20 vs. 13
SUV_{peak}	7	1.4	0.7–2.8	0.31	16 vs. 10
SUV_{peak} with PVE correction	9	1.8	0.9–3.6	0.11	20 vs. 11
SUV_{mean}	6.5	1.7	0.8–3.6	0.15	16 vs. 10
SUV_{mean} with PVE correction	7.5	1.7	0.8–3.5	0.12	20 vs. 10
MATV (cm^3)	85	3.9	1.0–15.2	0.0004	20 vs. 6
MATV with PVE correction (cm^3)	80	3.4	0.9–11.7	0.0024	16 vs. 10
TLG (g)	260	2.9	1.2–6.8	0.0012	21 vs. 10
TLG with PVE correction (g)	280	3.2	1.3–7.6	0.0004	21 vs. 10

($P = 0.3$) and 16 versus 10 mo ($P = 0.15$), respectively. Similarly, after PVE correction no threshold value could significantly differentiate groups of patients regarding their survival (Figs. 5A and 5B).

Contrary to SUV measurements with or without PVE correction, the parameters related to functional volume (MATV and TLG) allowed significant ($P < 0.0001$) differentiation of the 3 response groups and were significant prognostic factors ($P < 0.002$), as illustrated in Figure 4C. No significant differences were found using the original or PVE-corrected values.

The parameter that allowed for the best differentiation of patient groups was the TLG ($P < 0.0001$). CRs were characterized by a TLG of 55 ± 45 g, whereas PRs and NRs had a TLG of 178 ± 143 and 416 ± 238 g, respectively. After PVE correction, the absolute values of each group rose to 62 ± 45, 200 ± 155, and 437 ± 249 g for CRs, PRs, and NRs, respectively, leading to the same discrimination between groups of response ($P < 0.0001$). Although slightly less efficient than TLG, the use of MATV allowed a statistically significant differentiation of the 3 response groups ($P < 0.0001$). Use of the MATV values extracted from PVE correction images led to exactly the same discriminating power ($P < 0.0001$).

MATV and TLG were also good prognostic factors, with high MATV and TLG values being significantly associated with shorter survival, with hazard ratios between 3 and 4 (Table 3). A MATV above 85 cm^3 was identified as a predictor of poor outcome, with a median survival of only 6 mo, versus 20 mo for patients with a smaller MATV ($P = 0.0004$), as illustrated in Figure 5C. In addition, a MATV below 15 cm^3 was associated ($P = 0.009$) with longer survival (49 mo) than a larger MATV (11 mo). Similar results were obtained using the MATVs measured on the PVE-corrected images, with a median survival of 20 mo for patients with tumor volume with PVE correction below 80 cm^3 versus 10 mo for patients with MATV above 80 cm^3 ($P < 0.002$). Regarding TLG, a threshold of 260 g was found to be a good discriminating factor for outcome (21 vs. 10 mo, $P = 0.0012$), whereas using PVE-corrected TLG led to similar results, with a slightly higher threshold (TLG with PVE correction = 280 g, 21 vs. 10 mo, $P = 0.0004$).

DISCUSSION

Our study investigated the impact of PVE correction on the predictive and prognostic values of different parameters derived using the baseline pretreatment PET images. Our results confirmed that PVE correction significantly affects quantitative SUVs, with an average increase of above 50% for SUV_{max}, in agreement with previous studies (*18,19*), and a lower increase (<30%) for SUV_{peak} and SUV_{mean}. The lower increase observed for SUV_{peak} and SUV_{mean} is related to the fact that the L-R deconvolution is a voxel-by-voxel process leading to enhancement of contrasts between subvolumes within the MATV and both lower- and higher-voxels SUVs included in the averaging associated with the calculation of SUV_{mean} and SUV_{peak}. PVE correction did not significantly affect the delineation of the MATV. Overall, MATVs delineated on the corrected images were only slightly smaller than those determined on the original images. The mean reduction of 10% was within the reproducibility limits of confidence intervals regarding tumor volume measurements on double-baseline PET scans using fuzzy locally adaptive Bayesian algorithm method (±30%) (*29*). This limited impact of PVE correction on MATV can be explained by the fact that PVE is dependent on tumor size and is more pronounced on small lesions (*33*). In our group of patients, the tumors were rather large (40 ± 30 cm^3); therefore, the relative variation of volumes with respect to the entire volume is small. Twelve patients (25%) had an MATV of around 10 cm^3 or smaller. In addition, the use of a robust delineation approach instead of threshold-based methods in various configurations of blur and noise (*28,34*) ensured a limited variability in the MATV delineation results between original and corrected images.

As previously demonstrated (*7,12*), MATV and TLG extracted from noncorrected ^{18}F-FDG PET pretreatment acquisitions had high clinical value. In contrast, none of the usual SUV measurements (maximum, peak, or mean) considered in clinical practice was significantly associated with therapy response or survival, as also reported in the 2 largest available prospective trials (*35,36*).

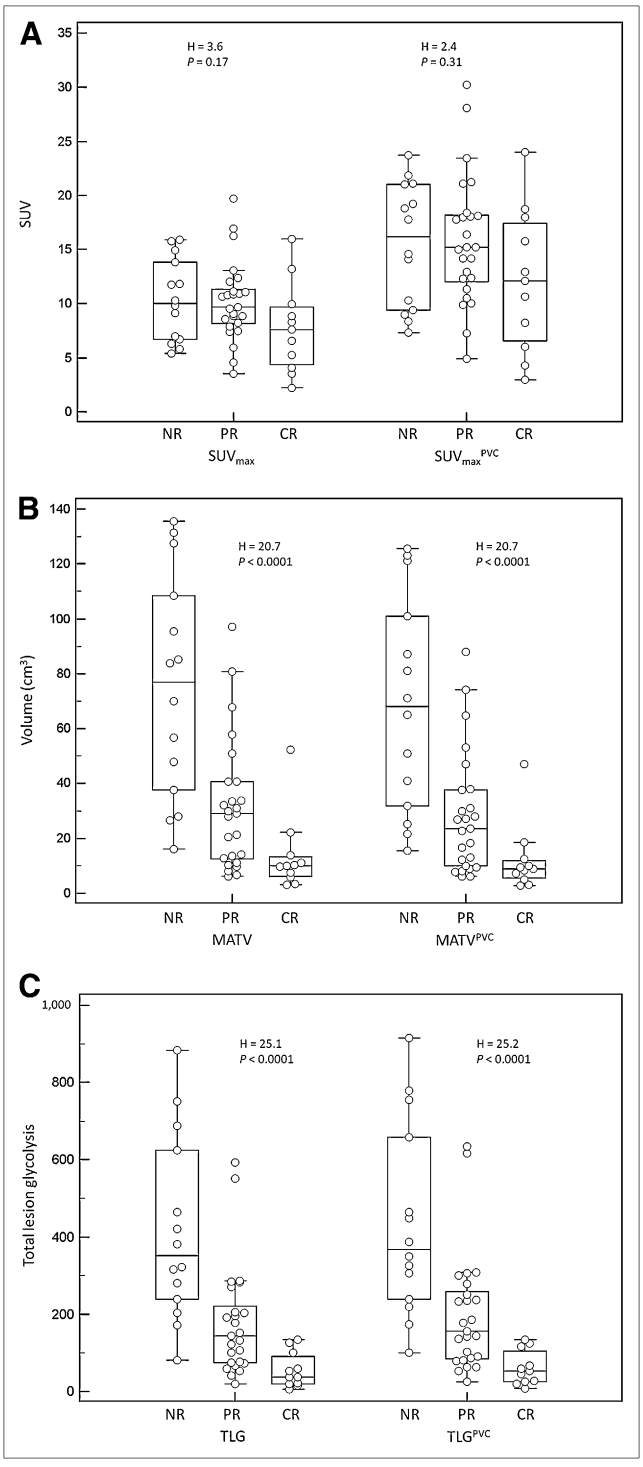

FIGURE 4. Examples of distributions of NRs, PRs, and CRs and associated Kruskal–Wallis test results: SUV_{max} and SUV_{max} with PVE correction (A), MATV and MATV with PVE correction (B), and TLG and TLG with PVE correction (C). $MATV^{PVC}$ = MATV with PVE correction; SUV_{max}^{PVC} = SUV with PVE correction; TLG^{PVC} = TLG with PVE correction.

Regarding response to therapy prediction using SUVs, we found that PVE correction did not improve the already demonstrated low discriminating power of any of the SUV measurements considered (*7*). This can be explained by the combination of several factors. First, without PVE correction, the trend of low SUV being associated with better outcome may have been exaggerated by an underestimation of SUV, because CRs had also smaller volumes in addition to low SUV_{max}. Second, after PVE correction all 3 response groups had increased SUV_{max} but with still no significant

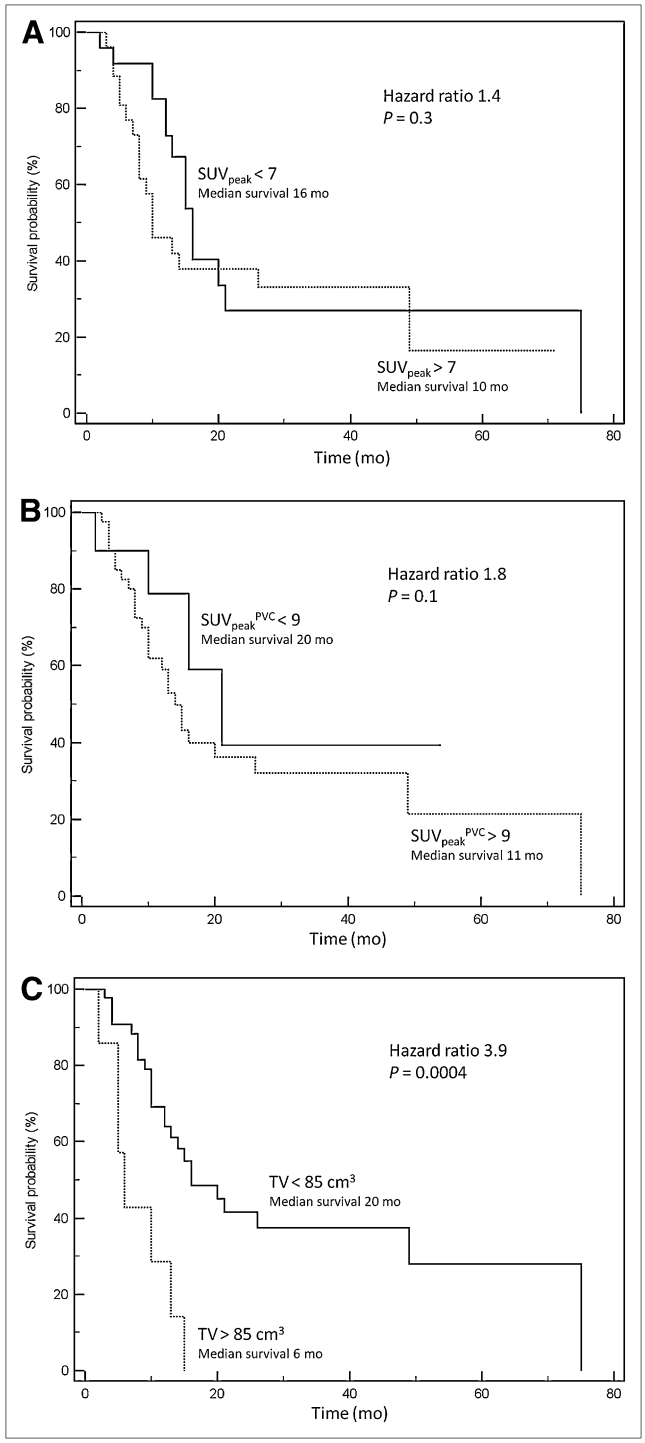

FIGURE 5. Examples of Kaplan–Meier survival curves obtained using SUV_{peak} (A), SUV_{peak} with PVE correction (B), and MATV (C). SUV_{peak}^{PVC} = SUV_{peak} with PVE correction.

difference between the groups. We have demonstrated that SUV_{mean} increase after PVE correction was inversely correlated with tumor volume ($r = 0.8$, $P < 0.0001$), with smaller volumes being characterized by higher SUV_{mean} increases after PVE correction than larger volumes. The SUV_{mean} within the MATV of PRs and NRs was therefore increased by a smaller amount (+20% ± 9%) than those within the MATV of CRs (+34% ± 13%), which were associated with smaller tumor volumes. The mean tumor SUVs of CRs were therefore closer to the SUV_{mean} of PRs and NRs after correction. Hence, the discriminating power of SUV_{mean} was reduced by PVE correction. A similar trend was observed for SUV_{max} and SUV_{peak}, although it was less significant because their respective increase was not correlated with the MATV. Therefore, PVE correction might have further reduced the clinical value of SUV measurements in this context. This effect has been previously suggested as a limitation to the prognostic value of SUV_{max} in early-stage non–small cell lung cancer (*37*).

Similar conclusions can be drawn from the results regarding the impact of PVE correction on the prognostic value of the SUV parameters. Indeed, as already demonstrated (*12*), extreme MATV values were significantly associated with longer or shorter overall survival for very small (49 mo for MATV below 15 cm^3 vs. 11 mo for MATV above 15 cm^3) or very large MATV (6 mo for tumor volume above 85 cm^3 vs. 20 mo for MATV below 80 cm^3), respectively. On the other hand, SUV measurements without correction cannot significantly differentiate between the patients with longer or shorter survival ($P > 0.05$ for all SUV measurements), although a trend for longer survival was associated with lower SUVs. After correction, this differentiation was not significantly improved, because SUVs associated with the smaller volumes were closer to SUVs associated with larger volumes. Therefore, the discrimination was again reduced by PVE correction. To our knowledge there are no similar data available on the impact of PVE correction on SUV predictive value in the literature, but our results are in agreement with previous findings that demonstrated no significant changes in disease-free survival correlation between original and corrected SUVs in esophageal cancer using alternative less accurate methodologies for both PVE correction and functional volume segmentation (*19*).

As previously demonstrated (*7,12*), MATV and associated TLG values were good predictors of response (*7*) and independent prognostic factors of overall survival (*12*). After PVE correction, the already high clinical value of MATV and TLG was not significantly altered. Considering the thresholds used to differentiate patient groups, there was no need for adjustment regarding MATV measurements because MATVs were not significantly modified by PVE correction. On the other hand, TLG thresholds needed to be adjusted, considering that PVE correction led to significantly increased SUV_{mean} and resulting TLG values. The determined threshold values for each parameter regarding prognosis or prediction of response were found using receiver-operating-characteristic analysis on the current patient cohort and would therefore require larger prospective studies to be validated.

The rather large tumor volumes (40 ± 30 cm^3) in our patient dataset might be considered as a limitation of this study, because PVEs are usually considered significant for volumes around or below 10 cm^3 (*33*). First, 25% of the tumors in this dataset were within this volume range. In addition, the shape of the primary esophageal lesions is not spheric but mostly cylindric, with a small diameter (<2 cm) in the transaxial direction. Therefore, esophageal lesions can be significantly affected by PVEs despite the overall large metabolic volumes, as can be seen in Figure 2 for a lesion with a MATV above 25 cm^3. Finally, the patient population used in this study was typical of routine clinical practice and was not selected on the basis of the overall primary MATVs.

CONCLUSION

The results of this study demonstrate that PVE correction does not add any value to parameters derived from MATVs such as MATV and TLG measured on ^{18}F-FDG PET baseline acquisitions. PVE correction did not alter the already demonstrated clinical value of both parameters as predictive factors of the response to concomitant radiochemotherapy or as prognostic factors of overall survival in locally advanced esophageal cancer. Similarly, although PVE correction led to increases in all SUV measurements (maximum, peak, or mean) considered in clinical practice, the corrected values were still not significantly associated with either therapy response or prognosis. Finally, our study is in agreement with previous investigations using simpler tools, showing limited interest in PVE correction in this specific context. However, the potential impact of PVE correction in other applications such as diagnosis or lesion detectability remains to be evaluated. In addition, the value of PVE correction in patient follow-up using serial PET scans needs to be further demonstrated.

DISCLOSURE STATEMENT

The costs of publication of this article were defrayed in part by the payment of page charges. Therefore, and solely to indicate this fact, this article is hereby marked "advertisement" in accordance with 18 USC section 1734.

ACKNOWLEDGMENT

No potential conflict of interest relevant to this article was reported.

REFERENCES

1. Parkin DM, Bray F, Ferlay J, Pisani P. Global cancer statistics, 2002. *CA Cancer J Clin.* 2005;55:74–108.
2. Hayat MJ, Howlader N, Reichman ME, Edwards BK. Cancer statistics, trends, and multiple primary cancer analyses from the Surveillance, Epidemiology, and End Results (SEER) Program. *Oncologist.* 2007;12:20–37.
3. Gebski V, Burmeister B, Smithers BM, Foo K, Zalcberg J, Simes J. Survival benefits from neoadjuvant chemoradiotherapy or chemotherapy in oesophageal carcinoma: a meta-analysis. *Lancet Oncol.* 2007;8:226–234.

4. Kelsen DP, Winter KA, Gunderson LL, et al. Long-term results of RTOG trial 8911 (USA Intergroup 113): a random assignment trial comparison of chemotherapy followed by surgery compared with surgery alone for esophageal cancer. *J Clin Oncol.* 2007;25:3719–3725.
5. Stahl M, Wilke H, Stuschke M, et al. Clinical response to induction chemotherapy predicts local control and long-term survival in multimodal treatment of patients with locally advanced esophageal cancer. *J Cancer Res Clin Oncol.* 2005;131:67–72.
6. Aklilu M, Ilson DH. Targeted agents and esophageal cancer: the next step? *Semin Radiat Oncol.* 2007;17:62–69.
7. Hatt M, Visvikis D, Pradier O, Cheze-le Rest C. Baseline ^{18}F-FDG PET image-derived parameters for therapy response prediction in oesophageal cancer. *Eur J Nucl Med Mol Imaging.* 2011;38:1595–1606.
8. Larson SM, Erdi Y, Akhurst T, 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;2:159–171.
9. Hyun SH, Choi JY, Shim YM, et al. Prognostic value of metabolic tumor volume measured by ^{18}F-fluorodeoxyglucose positron emission tomography in patients with esophageal carcinoma. *Ann Surg Oncol.* 2010;17:115–122.
10. Roedl JB, Prabhakar HB, Mueller PR, Colen RR, Blake MA. Prediction of metastatic disease and survival in patients with gastric and gastroesophageal junction tumors: the incremental value of PET-CT over PET and the clinical role of primary tumor volume measurements. *Acad Radiol.* 2009;16:218–226.
11. Mamede M, Abreu ELP, Oliva MR, Nose V, Mamon H, Gerbaudo VH. FDG-PET/CT tumor segmentation-derived indices of metabolic activity to assess response to neoadjuvant therapy and progression-free survival in esophageal cancer: correlation with histopathology results. *Am J Clin Oncol.* 2007;30:377–388.
12. Hatt M, Visvikis D, Albarghach NM, Tixier F, Pradier O, Cheze-le Rest C. Prognostic value of ^{18}F-FDG PET image-based parameters in oesophageal cancer and impact of tumour delineation methodology. *Eur J Nucl Med Mol Imaging.* 2011;38:1191–1202.
13. 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.
14. Lee HY, Hyun SH, Lee KS, et al. Volume-based parameter of ^{18}F-FDG PET/CT in malignant pleural mesothelioma: prediction of therapeutic response and prognostic implications. *Ann Surg Oncol.* 2010;17:2787–2794.
15. Cazaentre T, Morschhauser F, Vermandel M, et al. Pre-therapy ^{18}F-FDG PET quantitative parameters help in predicting the response to radioimmunotherapy in non-Hodgkin lymphoma. *Eur J Nucl Med Mol Imaging.* 2010;37:494–504.
16. Tixier F, Le Rest CC, Hatt M, et al. Intratumor heterogeneity characterized by textural features on baseline ^{18}F-FDG PET images predicts response to concomitant radiochemotherapy in esophageal cancer. *J Nucl Med.* 2011;52:369–378.
17. Kwee RM. Prediction of tumor response to neoadjuvant therapy in patients with esophageal cancer with use of ^{18}F FDG PET: a systematic review. *Radiology.* 2010;254:707–717.
18. Hoetjes NJ, van Velden FH, Hoekstra OS, et al. Partial volume correction strategies for quantitative FDG PET in oncology. *Eur J Nucl Med Mol Imaging.* 2010;37:1679–1687.
19. van Heijl M, Omloo JM, van Berge Henegouwen MI, van Lanschot JJ, Sloof GW, Boellaard R. Influence of ROI definition, partial volume correction and SUV normalization on SUV-survival correlation in oesophageal cancer. *Nucl Med Commun.* 2010;31:652–658.
20. Hatt M, Cheze le Rest C, Descourt P, et al. Accurate automatic delineation of heterogeneous functional volumes in positron emission tomography for oncology applications. *Int J Radiat Oncol Biol Phys.* 2010;77:301–308.
21. Therasse P, Arbuck SG, Eisenhauer EA, et al. New guidelines to evaluate the response to treatment in solid tumors. European Organization for Research and Treatment of Cancer, National Cancer Institute of the United States, National Cancer Institute of Canada. *J Natl Cancer Inst.* 2000;92:205–216.
22. Boussion N, Cheze Le Rest C, Hatt M, Visvikis D. Incorporation of wavelet-based denoising in iterative deconvolution for partial volume correction in whole-body PET imaging. *Eur J Nucl Med Mol Imaging.* 2009;36:1064–1075.
23. Lucy LB. An iteration technique for the rectification of observed distributions. *Astron J.* 1974;79:745–754.
24. Richardson WH. Bayesian-based iterative method of image restoration. *J Opt Soc Am.* 1972;62:55–59.
25. Van Cittert PH. Zum einfluss der spaltbreite auf die intensita tsverteilung in spektrallinien [German]. *Z Physik.* 1931;69:298.
26. Chang SG, Yu B, Vetterli M. Adaptive wavelet thresholding for image denoising and compression. *IEEE Trans Image Process.* 2000;9:1532–1546.
27. Hatt M, Cheze le Rest C, Turzo A, Roux C, Visvikis D. A fuzzy locally adaptive Bayesian segmentation approach for volume determination in PET. *IEEE Trans Med Imaging.* 2009;28:881–893.
28. Hatt M, Cheze Le Rest C, Albarghach N, Pradier O, Visvikis D. PET functional volume delineation: a robustness and repeatability study. *Eur J Nucl Med Mol Imaging.* 2011;38:663–672.
29. Hatt M, Cheze-Le Rest C, Aboagye EO, et al. Reproducibility of ^{18}F-FDG and 3′-deoxy-3′-^{18}F-fluorothymidine PET tumor volume measurements. *J Nucl Med.* 2010;51:1368–1376.
30. Kruskal W, Wallis W. Use of ranks in one-criterion variance analysis. *J Am Stat Assoc.* 1952;47:583–621.
31. Kaplan E, Meyer P. Non parametric estimation from incomplete observation. *J Am Stat Assoc.* 1958;53:457–481.
32. Metz CE. Basic principles of ROC analysis. *Semin Nucl Med.* 1978;8:283–298.
33. Soret M, Bacharach SL, Buvat I. Partial-volume effect in PET tumor imaging. *J Nucl Med.* 2007;48:932–945.
34. Hatt M, Visvikis D. Defining radiotherapy target volumes using ^{18}F-fluoro-deoxy-glucose positron emission tomography/computed tomography: still a Pandora's box?– in regard to Devic et al. (Int J Radiat Oncol Biol Phys 2010). *Int J Radiat Oncol Biol Phys.* 2010;78:1605.
35. Omloo JM, Sloof GW, Boellaard R, et al. Importance of fluorodeoxyglucose-positron emission tomography (FDG-PET) and endoscopic ultrasonography parameters in predicting survival following surgery for esophageal cancer. *Endoscopy.* 2008;40:464–471.
36. Chatterton BE, Ho Shon I, Baldey A, et al. Positron emission tomography changes management and prognostic stratification in patients with oesophageal cancer: results of a multicentre prospective study. *Eur J Nucl Med Mol Imaging.* 2009;36:354–361.
37. Hanin FX, Lonneux M, Cornet J, et al. Prognostic value of FDG uptake in early stage non-small cell lung cancer. *Eur J Cardiothorac Surg.* 2008;33:819–823.

JNM The Journal of NUCLEAR MEDICINE

Impact of Partial-Volume Effect Correction on the Predictive and Prognostic Value of Baseline ^{18}F-FDG PET Images in Esophageal Cancer

Mathieu Hatt, Adrien Le Pogam, Dimitris Visvikis, Olivier Pradier and Catherine Cheze Le Rest

J Nucl Med. 2012;53:12-20.
Doi: 10.2967/jnumed.111.092775

The Journal of Nuclear Medicine is published monthly.
SNMMI | Society of Nuclear Medicine and Molecular Imaging
1850 Samuel Morse Drive, Reston, VA 20190.
(Print ISSN: 0161-5505, Online ISSN: 2159-662X)

SNMMI SOCIETY OF NUCLEAR MEDICINE AND MOLECULAR IMAGING