

---

---

# Does PET SUV Harmonization Affect PERCIST Response Classification?

Elske Quak<sup>1</sup>, Pierre-Yves Le Roux<sup>2</sup>, Charline Lasnon<sup>3-5</sup>, Philippe Robin<sup>2</sup>, Michael S. Hofman<sup>6</sup>, David Bourhis<sup>2</sup>, Jason Callahan<sup>6</sup>, David S. Binns<sup>6</sup>, Cédric Desmots<sup>3</sup>, Pierre-Yves Salaun<sup>2</sup>, Rodney J. Hicks<sup>6,7\*</sup>, and Nicolas Aide<sup>3-5\*</sup>

<sup>1</sup>Nuclear Medicine Department, François Baclesse Cancer Centre, Caen, France; <sup>2</sup>Nuclear Medicine Department and EA 3878 IFR 148, University Hospital, Brest, France; <sup>3</sup>Nuclear Medicine Department, University Hospital, Caen, France; <sup>4</sup>Normandy University, Caen, France; <sup>5</sup>INSERM 1199, Caen University, Caen, France; <sup>6</sup>Cancer Imaging, Peter MacCallum Cancer Institute, East Melbourne, Australia; and <sup>7</sup>The Sir Peter MacCallum Department of Oncology, The University of Melbourne, Melbourne, Australia

---

Pre- and posttreatment PET comparative scans should ideally be obtained with identical acquisition and processing, but this is often impractical. The degree to which differing protocols affect PERCIST classification is unclear. This study evaluates the consistency of PERCIST classification across different reconstruction algorithms and whether a proprietary software tool can harmonize SUV estimation sufficiently to provide consistent response classification.

**Methods:** Eighty-six patients with non-small cell lung cancer, colorectal liver metastases, or metastatic melanoma who were scanned for therapy monitoring purposes were prospectively recruited in this multicenter trial. Pre- and posttreatment PET scans were acquired in protocols compliant with the Society of Nuclear Medicine and Molecular Imaging and the European Association of Nuclear Medicine (EANM) acquisition guidelines and were reconstructed with a point spread function (PSF) or PSF + time-of-flight (TOF) for optimal tumor detection and also with standardized ordered-subset expectation maximization (OSEM) known to fulfill EANM harmonizing standards. After reconstruction, a proprietary software solution was applied to the PSF ± TOF data (PSF ± TOF.EQ) to harmonize SUVs with the OSEM values. The impact of differing reconstructions on PERCIST classification was evaluated. **Results:** For the OSEM<sub>PET1</sub>/OSEM<sub>PET2</sub> (OSEM reconstruction for pre- and posttherapeutic PET, respectively) scenario, which was taken as the reference standard, the change in SUL was  $-41\% \pm 25$  and  $+56\% \pm 62$  in the groups of tumors showing a decrease and an increase in <sup>18</sup>F-FDG uptake, respectively. The use of PSF reconstruction affected classification of tumor response. For example, taking the PSF ± TOF<sub>PET1</sub>/OSEM<sub>PET2</sub> scenario increased the apparent reduction in SUL in responding tumors ( $-48\% \pm 22$ ) but reduced the apparent increase in SUL in progressing tumors ( $+37\% \pm 43$ ), as compared with the OSEM<sub>PET1</sub>/OSEM<sub>PET2</sub> scenario. As a result, variation in reconstruction methodology (PSF ± TOF<sub>PET1</sub>/OSEM<sub>PET2</sub> or OSEM<sub>PET1</sub>/PSF ± TOF<sub>PET2</sub>) led to 13 of 86 (15%) and 17 of 86 (20%) PERCIST classification discordances, respectively. Agreement was better for these scenarios with application of the propriety filter, with  $\kappa$  values of 1 and 0.95 compared with 0.79 and 0.72, respectively. **Conclusion:** Reconstruction algorithm-dependent variability in PERCIST classification is a significant issue but can be overcome by harmonizing SULs using a proprietary software tool.

**Key Words:** PET; <sup>18</sup>F-FDG; therapy response; PERCIST; harmonization

**J Nucl Med 2016; 57:1699–1706**

DOI: 10.2967/jnumed.115.171983

---

**P**ET with <sup>18</sup>F-FDG is increasingly being used for response evaluation in cancer patients, especially in clinical trials (1). For pragmatic reasons, SUV is the most frequently used quantitative parameter. In an effort to bring consistency to the classification of response across trials, emulating the use of the RECIST for radiologic examinations (2), various schema based on the degree of SUV change after treatment have been proposed. The European Organization for Research and Treatment of Cancer criteria (3) have been largely supplanted by PERCIST (4).

However, many sources of error in SUV measurement exist (5,6). In particular, technologic improvements can lead to significant device-dependent and reconstruction-dependent variations in quantitative values. For instance, point spread function (PSF) reconstruction, which improves spatial resolution throughout the entire field of view (7) and which is commercially available from all major PET vendors, has been shown to increase SUV by up to 66% compared with conventional OSEM reconstruction (8). Although this might be advantageous for the detection of small lesions and increase clinical confidence when interpreting scans, this could lead to classification errors by exceeding thresholds used for discriminating between responding and nonresponding tumors unless acquisition and processing of pre- and posttreatment scans are performed on the same scanner and processed identically. However, in busy nuclear medicine departments, which may have several scanners or which periodically update equipment, and with patient mobility requiring scanning at different sites, this may not always be practical.

Therefore, there is a growing interest in SUV harmonization strategies such as the European Association Research Ltd. accreditation program (9), the North American Quantitative Imaging Biomarkers Alliance (10), and the Uniform Protocols for Imaging in Clinical Trials (11), which aim at minimizing the variability in SUV measurements by harmonizing patient preparation and scan acquisition and processing (12). Although many sources of error in SUV measurements can be overcome by complying with EANM or Society of Nuclear Medicine and Molecular Imaging guidelines for PET tumor imaging (13–15), reconstruction-dependent variations require

---

Received Jan. 14, 2016; revision accepted May 6, 2016.

For correspondence or reprints contact: Nicolas Aide, Nuclear Medicine Department, Caen University Hospital, Avenue Côte de Nacre, 14000 Caen, France.

E-mail: aide-n@chu-caen.fr

\*Contributed equally to this work.

Published online Jun. 9, 2016.

COPYRIGHT © 2016 by the Society of Nuclear Medicine and Molecular Imaging, Inc.

either the use of an additional filtering step (16) or the generation of 2 sets of images, 1 set to provide optimal diagnostic quality and 1 to meet quantitative harmonization standards (17). However, the impact of reconstruction methodology on classification of treatment response is unclear.

To assess this impact, we reconstructed the same PET raw data with an OSEM algorithm known to meet EANM requirements and also with PSF with or without TOF reconstruction (PSF ± TOF). Postreconstruction filtering was then applied to the PSF or PSF plus TOF reconstruction with EQ.PET, a proprietary software solution allowing visualization of optimized images while simultaneously obtaining harmonized SUVs (16).

The primary aim of this multicenter prospective study was to evaluate the impact of SUV reconstruction-dependency on PERCIST classification and the ability of EQ.PET technology to minimize variability in response assessment.

## MATERIALS AND METHODS

### Patients

Over an 18-mo period, patients with non-small cell lung cancer, colorectal liver metastases, or metastatic melanoma and scanned for monitoring efficacy of chemotherapy, molecularly targeted therapies, or radiotherapy were prospectively included in 3 PET centers. Informed consent was waived for this type of study by the local ethics committee (Ref A12-D24-VOL13, *Comité de protection des personnes Nord-Ouest III*) because the scans were obtained for clinical indications and the trial procedures were performed independently without influencing clinical reporting.

### PET Systems

Data from the following 3 PET systems were used for this study: a Biograph 6 TrueV with PSF reconstruction, an mCT with PSF + TOF, and a Biograph 64 TrueV with PSF reconstruction (Siemens Medical Solutions). Both the Biograph systems were equipped with an extended axial field of view.

### Patient Preparation

All patients were requested to fast for 6 h before the <sup>18</sup>F-FDG injection. Patient height, weight, and blood glucose levels were recorded. Patients were injected intravenously with <sup>18</sup>F-FDG, followed by a 60-min rest in a warm room. The injected activity and the exact delay between injection and the start of the acquisition were extracted from the DICOM standard headers and were recorded for each patient for baseline scan (PET1) and posttherapeutic scan (PET2).

### PET Acquisition and Reconstruction Parameters

A daily calibration of each PET system was performed with a <sup>68</sup>Ge source according to the manufacturer's protocol. A quarterly cross-calibration of each PET system was performed according to the EANM guidelines, as described elsewhere (13,14), and all clocks were synchronized weekly.

The PET acquisition was performed in 3-dimensional (3D) mode. Scatter and attenuation corrections were applied on all PET acquisitions. Patients were scanned from skull vertex or base to the mid thighs, with acquisition extended to legs in melanoma patients with primary site of disease in the lower limbs.

All raw PET data were reconstructed with the local PSF ± TOF settings for optimal lesion detection and a 3D OSEM reconstruction algorithm fulfilling the EANM guidelines regarding recovery coefficients (Table 1).

### EQ.PET Technology

For each PET system, the EQ.PET filter was calculated on the phantom data of each PSF ± TOF reconstruction as described in detail elsewhere (17). Briefly, the recovery coefficients (defined as the ratio between the measured and true activity concentration for each sphere) of a National Electrical Manufacturers Association NU2 phantom scanned per EANM guidelines were aligned to the EANM reference recovery coefficients by applying a gaussian filter.

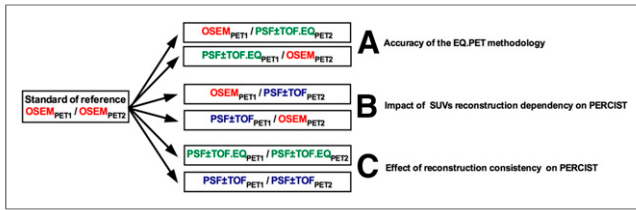
### PERCIST Evaluation

All PET examinations were analyzed with Syngo.via software equipped with EQ.PET (Siemens Medical Solutions). For interpretation purposes, both the reconstruction for optimal lesion detection

**TABLE 1**  
PET/CT Acquisition and Reconstruction Parameters for the 3 Participating Centers

Parameter	Center 1		Center 2		Center 3			
	Biograph 6	Biograph 6	Biograph mCT	Biograph mCT	Biograph 64	Biograph 64	Biograph 64	Biograph 64
PET acquisition	2 min, 40 s	3 min, 40 s (BMI > 25)	2 min, 00 s		2 min, 30 s (≤65 kg)	3 min (65–85 kg)	3 min, 30 s (85–100 kg)	4 min, 00 s (>100 kg)
PET reconstruction								
Details	—	—	—	—		≤65kg	65–100 kg	>100 kg
Reconstruction	OSEM3D	PSF	OSEM3D	PSF + TOF	OSEM3D	PSF	PSF	PSF
Iterations/subsets	4i 8s	3i 21s	2i 24s	2i 21s	4i 8s	3i 21s	3i 21s	3i 21s
Postprocessing filter (mm)	5	0	4.4	2	3.5	6	5	4
Matrix	168 × 168	168 × 168	200 × 200	200 × 200	168 × 168	168 × 168	168 × 168	168 × 168
Pixel spacing	4.07 × 4.07	4.07 × 4.07	4.07 × 4.07	4.07 × 4.07	3.39 × 3.39	3.39 × 3.39	3.39 × 3.39	3.39 × 3.39
Slice thickness (mm)	5	5	2.027	2.027	3	3	3	3
EQ filter (mm)	0	6.9	0	6.3	0	2.4	3.9	4.9

BMI = body mass index.



**FIGURE 1.** Flow chart describing different scenarios for determining accuracy of EQ.PET technology in setting of therapy assessment with PERCIST.

(PSF  $\pm$  TOF) and the OSEM reconstruction were displayed on the screen together with the EQ.PET-filtered harmonized SUV results for the tumor regions of interest and the liver background. The EQ.PET-filtered images were not displayed on the screen.

As defined in PERCIST (4), the measurable target lesion is the single most intense tumor site on pre- and posttreatment scans, which means that the target lesion is not necessarily the same before and after treatment.

To evaluate the impact of reconstruction methods on assessment of therapeutic response, we first assessed the accuracy of EQ.PET in correcting PSF-reconstructed lesions by assigning the same lesion (i.e., the most intense lesion determined on OSEM pre- and posttreatment scans) as a target for all reconstructions (Fig. 1A). These results are presented in Figure 2. Subsequently, we evaluated the impact of reconstruction inconsistency on PERCIST classification in real life by assessing whether the use of a different reconstruction algorithm for pre- and posttreatment scans would lead to a change in the most intense lesion location with consequent impact on PERCIST classification. For this analysis, we selected the most intense lesion on each reconstruction (Fig. 1B). Finally, we studied whether using PSF  $\pm$  TOF reconstruction for both the pre- and the posttreatment scans would give results similar to those using the former-generation OSEM algorithm (Fig. 1C).

In practice, the most intense lesion was located by scaling the 3D maximum-intensity-projection view on both the OSEM and the

PSF  $\pm$  TOF reconstructions. The location of this lesion for both reconstructions was noted. Then, volumes of interest were drawn on one reconstruction and automatically propagated to the second set of reconstructions (propagation from OSEM to PSF  $\pm$  TOF and vice versa). Within these volumes of interest, lean body mass  $SUV_{peak}$  ( $SUL_{peak}$ ) was measured. Background activity ( $SUL_{mean}$ ) was measured in an automatically placed 3-cm-diameter sphere in the right liver lobe avoiding metastases, especially in colorectal cancer patients, and in a 1-cm-diameter and 2-cm-height cylinder in the descending thoracic aorta.

Additionally, a volume of interest (1-cm-diameter and 2-cm-height cylinder) was placed next to the target lesion to assess tumor-to-background ratios.

On the basis of  $SUL_{peak}$  variation between the pre- and posttreatment scan, patients were classified according to PERCIST as follows: complete metabolic response (CMR)—complete resolution of  $^{18}F$ -FDG uptake in the tumor volume, with tumor  $SUL$  lower than liver  $SUL$  and background blood pool, and disappearance of all lesions if multiple; partial metabolic response (PMR)—at least 30% reduction in tumor uptake; stable metabolic disease (SMD)—less than 30% increase or less than 30% decrease in tumor  $^{18}F$ -FDG  $SUL_{peak}$  and no new lesions; and progressive metabolic disease (PMD)—greater than 30% increase in  $^{18}F$ -FDG tumor  $SUL_{peak}$  within the tumor or appearance of new lesions.

### Statistical Analysis

Quantitative data from clinical PET/CT examinations are presented as mean  $\pm$  SD. The relationship between PSF  $\pm$  TOF, PSF  $\pm$  TOF.EQ, and OSEM quantitative values was assessed with Bland–Altman plots. The PSF  $\pm$  TOF/OSEM  $SUL_{peak}$  ratios and PSF  $\pm$  TOF.EQ/OSEM  $SUL_{peak}$  ratios were compared using the Wilcoxon test for paired samples. To assess potential confounding factors, the ratios between PSF  $\pm$  TOF.EQ and OSEM  $SUL_{peak}$  according target lesion size (<10, 10–20, and >20 mm), tumor-to-background ratio (<5, 5–10, and >10), and reconstruction algorithm (PSF or PSF + TOF) were compared using the Kruskal–Wallis test (with a post hoc Dunn test) for multiple-groups comparison or the Mann–Whitney test for unpaired samples when appropriate. A 2-tailed  $P$  value of less than 0.05 was considered statistically significant.

Levels of agreement between the different types of reconstruction were evaluated using the  $\kappa$ -statistic. OSEM reconstruction both for pre- and for posttherapeutic PET examination (OSEM<sub>PET1</sub>/OSEM<sub>PET2</sub>) was used as the current standard to classify the therapeutic response of each lesion and compared with other scenarios described in Figure 1.  $\kappa$ -values were reported using the benchmarks of Landis and Koch (18).

Graphs and analyses were performed using Prism GraphPad and the Vassar University website for statistical computation (<http://vassarstats.net>).

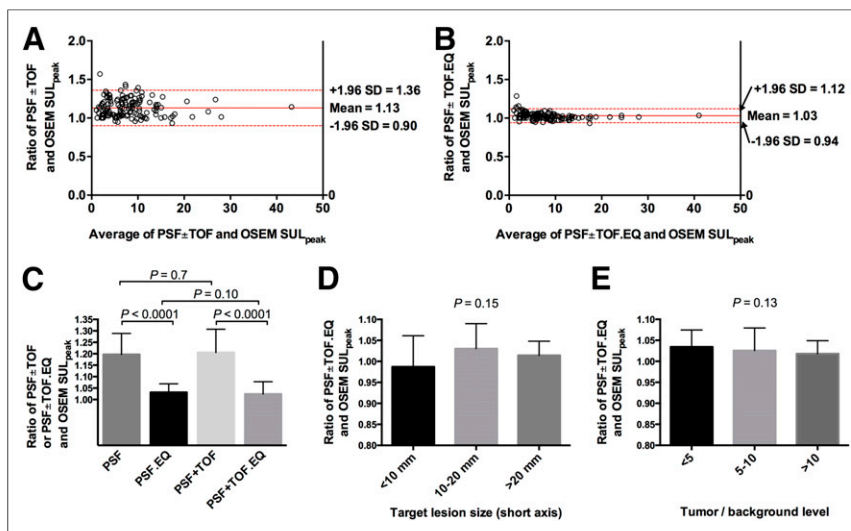
## RESULTS

### Patients' Demographics

The patients' sex ratio (male-to-female) was 2.4, and mean age ( $\pm$ SD) was  $63 \pm 9$  y. Details about the different treatment modalities for the 3 different cancer types can be found in Table 2. The interval between the pre- and posttreatment PET scans was  $114 \pm 55$  d.

### Compliance to PERCIST Acquisition Requirements

The injected activity (MBq/kg) was  $3.96 \pm 0.29$  and  $3.94 \pm 0.24$  for pre- and



**FIGURE 2.** (A and B) Relationship between  $SUL_{mean}$  in tumor lesions extracted from PSF  $\pm$  TOF or PSF  $\pm$  TOF.EQ and OSEM images, assessed using Bland–Altman plots. (C) Mean  $\pm$  SD ratio of  $SUL_{peak}$  obtained with conventional OSEM algorithm and those obtained with PSF or PSF + TOF reconstructions are shown before and after application of EQ.PET technology. (D and E) Ratio of PSF  $\pm$  TOF.EQ and OSEM  $SUL_{peak}$  depending on target lesion size and tumor/background.

**TABLE 2**  
Treatment Modalities

Cancer type	Treatment type	No. of patients
Colorectal cancer ( <i>n</i> = 24)	Chemotherapy	23
	Chemotherapy + radiotherapy	1
Melanoma ( <i>n</i> = 11)	Tyrosine kinase inhibitor	6
	Immunotherapy	3
	Radiotherapy	2
Non-small cell lung cancer ( <i>n</i> = 51)	Chemotherapy	49
	Immunotherapy	1
	Radiotherapy	1

posttreatment scans, respectively. The percentage variation between injected doses for the pre- and posttreatment scans ranged from 0% to 38% and exceeded the 20% limit in only 3 cases. PERCIST requirements were met in 83 of 86 patients (97%).

The uptake time was  $64 \pm 5$  and  $64 \pm 7$  min for the pre- and posttreatment scans, respectively. The absolute difference in uptake time between the pre- and posttreatment scans ranged from 0 to 42 min and exceeded the 20-min limit in only 3 cases. PERCIST requirements were met in 83 of 86 patients (97%).

In each individual patient, the baseline and the follow-up scans were always acquired on the same PET/CT system. When the 2 centers in which acquisition parameters in terms of time per bed position and postfiltering differed depending on patient weight or body mass index were considered, interscan acquisition consistency was achieved in 81 of 86 (94%) patients.

#### Ability of EQ.PET Technology to Harmonize SUL Assessments

The mean  $SUL_{peak}$  ( $\pm$ SD) for OSEM, PSF  $\pm$  TOF, and PSF  $\pm$  TOF.EQ reconstructions were  $7.5 \pm 5.6$ ,  $8.4 \pm 6.2$ , and  $7.6 \pm 5.7$ , respectively.

The mean ratio between PSF  $\pm$  TOF and OSEM reconstructions for  $SUL_{peak}$  was 1.13 (95% confidence interval, 0.90–1.36) (Fig. 2A). After application of the EQ.PET filter, this ratio was reduced to 1.03 (95% confidence interval, 0.94–1.12) (Fig. 2B).

The impact of potential confounders on the ratios between PSF  $\pm$  TOF.EQ and OSEM reconstructions for  $SUL_{peak}$  are shown in Figures 2C–2E. Because no difference in the ratios of  $SUL_{peak}$  was observed between the PSF and PSF + TOF reconstructions, we decided not to separate these results in subsequent analyses. The ratios of  $SUL_{peak}$  were not influenced by the size of the target lesion or by the background surrounding the target lesion.

#### Impact of Reconstruction-Dependent Variation of SUL on PERCIST Evaluation

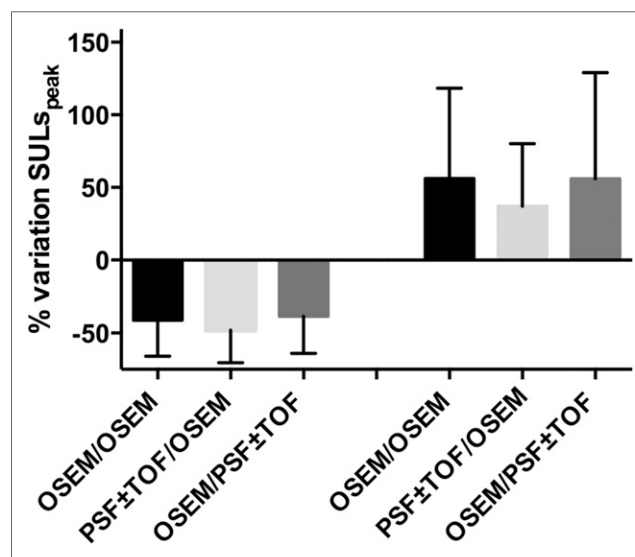
The target lesion selected was the same for OSEM and PSF  $\pm$  TOF reconstructions for 83 (97%) of the pretreatment scans and 85 (99%) of the posttreatment scans. The mean size of the target lesion was  $31 \pm 21$  mm for the baseline scan and  $27 \pm 22$  mm for the posttreatment scan.

The variation in  $SUL_{peak}$  between the pre- and posttreatment scans is shown in Figure 3. For the OSEM/OSEM scenario, which was taken as the reference standard, the change in SUL was  $-41\% \pm 25$  and  $+56\% \pm 62$  in the groups of tumors showing a decrease and an increase in  $^{18}F$ -FDG uptake, respectively. The use of PSF reconstruction affected classification of tumor response, depending on

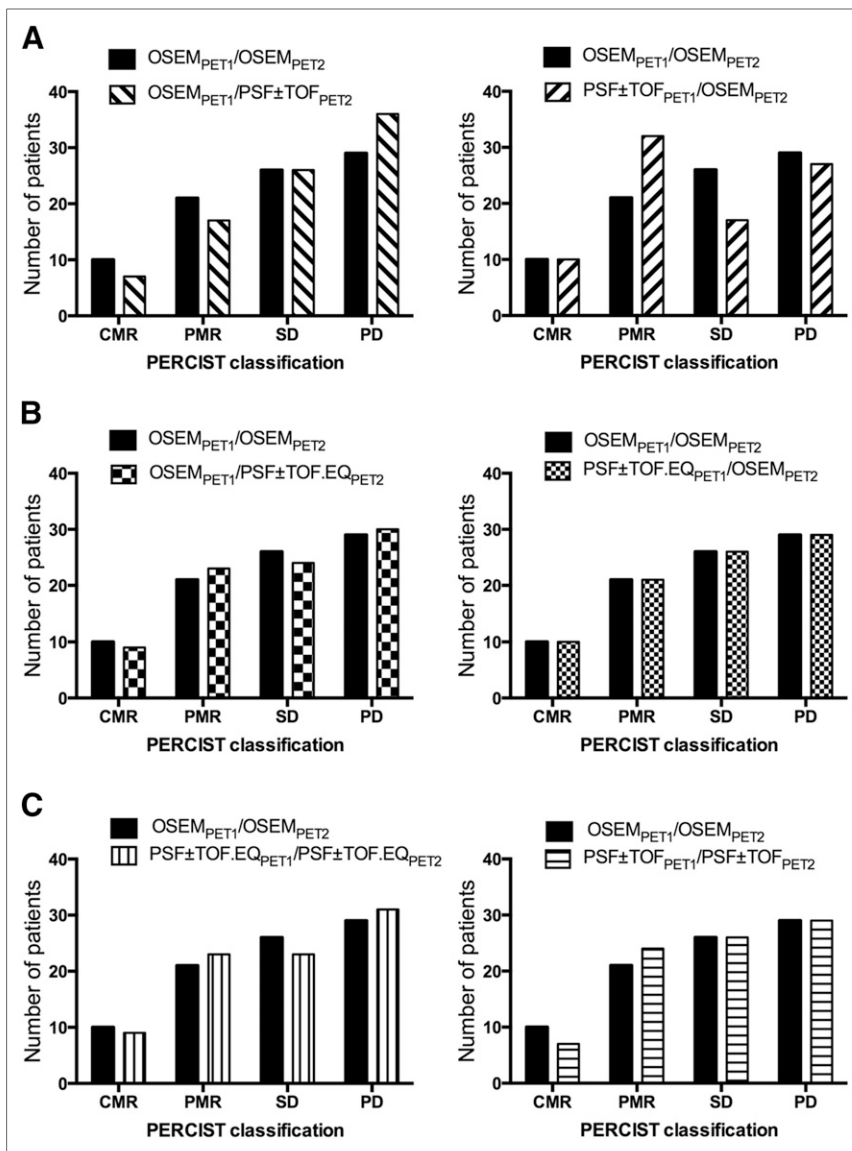
whether this reconstruction was used for the pre- or posttreatment scans. For example, the PSF  $\pm$  TOF<sub>PET1</sub>/OSEM<sub>PET2</sub> scenario increased the apparent reduction in SUL in responding tumors ( $-48\% \pm 22$ ) and reduced the apparent increase in SUL in progressing tumors ( $+37\% \pm 43$ ), as compared with the OSEM<sub>PET1</sub>/OSEM<sub>PET2</sub> scenario described above. Accordingly, inconsistent reconstructions induced discordant response classifications among the different scenarios.

When OSEM for the pre- and posttreatment scans was used, PET classified 10 patients as CMR, 21 as PMR, 26 as SMD, and 29 as PMD (Fig. 4; Supplemental Table 1 [supplemental materials are available at <http://jnm.snmjournals.org>]). CMR occurred in 6 patients with a decrease in  $SUL_{peak}$  to a level below the liver and blood-pool background and in 4 patients to complete disappearance of the target lesions. PMD occurred in 7 patients with an increase in tumor  $SUL_{peak}$  greater than 30% and in 22 patients with new lesions on the posttreatment scan.

Agreement levels between the OSEM<sub>PET1</sub>/OSEM<sub>PET2</sub> scenario and other scenarios involving reconstruction inconsistency (Table 3) were found to be almost perfect, with narrow confidence intervals for the scenarios using EQ.PET-filtered data either before or after treatment and the reconstruction-consistent scenario PSF  $\pm$  TOF<sub>PET1</sub>/PSF  $\pm$  TOF<sub>PET2</sub>.



**FIGURE 3.** Impact of reconstruction consistency on percentage variation in  $SUL_{peak}$  in responding (left) and progressing (right) tumors. Data are mean  $\pm$  SD.



**FIGURE 4.** Impact of reconstruction consistency on PERCIST classification. PERCIST classification is shown for standard of reference and for other scenarios: reconstruction inconsistency between baseline and posttreatment scans (A), use of EQ.PET methodology either for baseline or for posttreatment scan (B), and reconstruction consistency but use of different generations of reconstruction algorithms (C).

Agreement levels were fair to substantial for the scenario  $OSEM_{PET1}/PSF \pm TOF_{PET2}$  and  $PSF \pm TOF_{PET1}/OSEM_{PET2}$ , with wide confidence intervals.

Table 4 shows the number of discordances in the PERCIST classification that occurred for the different scenarios tested. Discordances ( $n = 42$ ) were most frequent for the scenarios  $OSEM_{PET1}/PSF \pm TOF_{PET2}$  ( $n = 17$ ) and  $PSF \pm TOF_{PET1}/OSEM_{PET2}$  ( $n = 13$ ). For  $OSEM_{PET1}/PSF \pm TOF_{PET2}$ , these led to 7 patients being classified as SMD instead of PMR, 3 as PMR instead of CMR, and 7 as PMD instead of SMD. For  $PSF \pm TOF_{PET1}/OSEM_{PET2}$ , 11 patients were classified as PMR instead of SMD and 2 as SMD instead of PMD. Figure 5 illustrates a patient classified as SMD according to the  $OSEM_{PET1}/OSEM_{PET2}$  standard of reference, whereas reconstruction inconsistency ( $OSEM_{PET1}/PSF \pm TOF_{PET2}$ ) led to a PMD classification.

### Impact of Reconstruction-Dependent Variation of SUL on Liver Reference Background

As shown in Figure 6, the PERCIST requirement that normal liver SUL must be within 20% and  $<0.3$  SUL mean units for the pre- and posttreatment PET scans to be assessable was not met for 15 (17%) for  $OSEM_{PET1}/OSEM_{PET2}$  and up to 17 (20%) of the 86 patients for the other scenarios of reconstruction inconsistency.

In none of the patients was this due to a difference in uptake time  $> 20$  min, or variation in injected dose  $> 20\%$  between the pre- and posttreatment PET scans. In 3 patients, there was a change in blood glucose level of  $> 1$  mmol/l, and 2 patients experienced a weight loss  $> 5$  kg.

### DISCUSSION

$^{18}F$ -FDG PET is being increasingly used for therapy monitoring in cancer patients. In the context of constant technologic evolutions that affect PET quantitation, harmonization strategies have emerged (9,10,12). In a previous study, we investigated the capability of the EQ.PET technology to harmonize  $PSF \pm TOF$  and OSEM SUV metrics while optimizing tumor detection (16). In this prospective multicenter study, we evaluated the impact of inconsistent reconstruction on PERCIST response classification, demonstrating variation in up to 20% of cases. Further, we showed that proprietary software, such as the EQ.PET technology, provided more consistent response classification. The EQ.PET technology is not affected by the type of reconstruction, the tumor size, or the tumor-to-background level. This has practical advantages when use of the same scanner for both scans is impractical, or when there is inadvertent variation of acquisition/reconstruction settings. The latter situation seems relatively common even in centers running the same

PET system, as recently reported by the Clinical Trials Network of the Society of Nuclear Medicine and Molecular Imaging (19).

In centers running 2 or more PET systems, this problem is compounded as illustrated by the study by Skougaard et al. (20), in which 12 of 81 (14%) patients undergoing pre- and posttreatment PET in the same department were excluded for analysis because they were scanned on 2 different generation PET systems.

We observed compliance in more than 90% of cases to PERCIST requirements regarding the injected dose and the uptake time between the pre- and posttreatment scans in a series of patients scanned in routine clinical practice. Despite good acquisition compliance, inconsistent reconstruction led to variability in PERCIST classification compared with OSEM as the reference standard. Taking, for example, the scenario of a system upgrade during a trial, the use of OSEM for the pretreatment scan while using

**TABLE 3**

Agreement Levels Between OSEM<sub>PET1</sub>/OSEM<sub>PET2</sub> Scenario and Other Scenarios Involving Reconstruction Inconsistency

Scenario	κ (95% confidence interval)			
	Overall (n = 86)	NSCLC (n = 51)	Colorectal liver metastases (n = 24)	Melanoma metastases (n = 11)
OSEM <sub>PET1</sub> /OSEM <sub>PET2</sub> vs. OSEM <sub>PET1</sub> /PSF ± TOF.EQ <sub>PET2</sub>	0.95 (0.90–1.00)	0.94 (0.87–1.00)	1.00 (1.00–1.00)	0.85 (0.59–1.00)
OSEM <sub>PET1</sub> /OSEM <sub>PET2</sub> vs. PSF ± TOF.EQ <sub>PET1</sub> /OSEM <sub>PET2</sub>	1.00 (1.00–1.00)	1.00 (1.00–1.00)	1.00 (1.00–1.00)	1.00 (1.00–1.00)
OSEM <sub>PET1</sub> /OSEM <sub>PET2</sub> vs. OSEM <sub>PET1</sub> /PSF ± TOF <sub>PET2</sub>	0.72 (0.60–0.84)	0.74 (0.60–0.90)	0.71 (0.48–0.94)	0.58 (0.18–0.99)
OSEM <sub>PET1</sub> /OSEM <sub>PET2</sub> vs. PSF ± TOF <sub>PET1</sub> /OSEM <sub>PET2</sub>	0.79 (0.69–0.90)	0.75 (0.60–0.90)	0.76 (0.55–0.97)	1.00 (1.00–1.00)
OSEM <sub>PET1</sub> /OSEM <sub>PET2</sub> vs. PSF ± TOF.EQ <sub>PET1</sub> /PSF ± TOF.EQ <sub>PET2</sub>	0.93 (0.87–1.00)	0.91 (0.83–1.00)	1.00 (1.00–1.00)	0.86 (0.59–1.00)
OSEM <sub>PET1</sub> /OSEM <sub>PET2</sub> vs. PSF ± TOF <sub>PET1</sub> /PSF ± TOF <sub>PET2</sub>	0.92 (0.85–0.99)	0.92 (0.83–1.00)	0.94 (1.83–1.00)	0.86 (0.59–1.00)

NSCLC = non-small cell lung cancer.

Data in parentheses are 95% confidence intervals.

PSF ± TOF for the posttreatment scan led to discordant response assessments in 17 of 86 (20%) (Table 4). It is noteworthy that a change in selected PERCIST target lesion occurred in only 3 of 172 scans (2%) and that among patients classified as PMD because of the appearance of new lesions, OSEM and PSF ± TOF performed equally in detecting these new lesions despite the potential for PSF reconstruction to detect smaller cancer lesions and to significantly increase SUV metrics as compared with OSEM reconstruction (8).

We found that with the appropriate EQ.PET filter for each center's PSF ± TOF reconstruction, we were able to harmonize PET quantitative data for tumors with a mean ratio of 1.02 for SUL<sub>peak</sub>, with narrow confidence interval. The use of the EQ.PET methodology for either the pre- or the posttreatment scans gave almost perfect agreement levels in comparison with the OSEM<sub>PET1</sub>/OSEM<sub>PET2</sub> reference standard, with narrow confidence intervals. We observed only 3 discordances for the OSEM<sub>PET1</sub>/PSF ± TOF.EQ<sub>PET2</sub> versus OSEM<sub>PET1</sub>/OSEM<sub>PET2</sub> scenario, and no discordance occurred

for the PSF ± TOF.EQ<sub>PET1</sub>/OSEM<sub>PET2</sub> versus OSEM<sub>PET1</sub>/OSEM<sub>PET2</sub> scenario. In the melanoma group, the OSEM<sub>PET1</sub>/OSEM<sub>PET2</sub> versus OSEM<sub>PET1</sub>/PSF ± TOF<sub>PET2</sub> scenario led to a 0.58 κ-value that can be explained by 3 discordances (patients classified as SMD for the OSEM<sub>PET1</sub>/OSEM<sub>PET2</sub> scenario and classified as PMD for the OSEM<sub>PET1</sub>/PSF ± TOF<sub>PET2</sub> scenario). This involved a group of only 11 patients, contributing to the wide confidence interval.

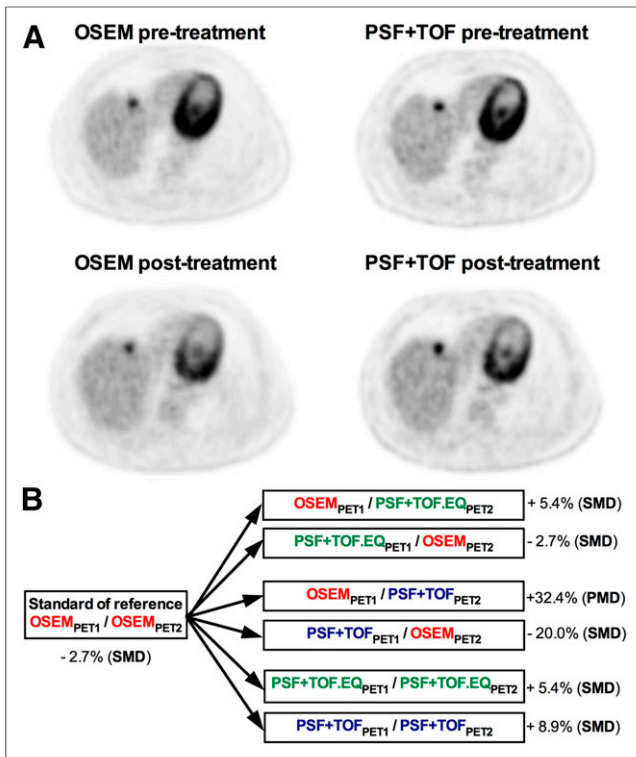
Consistent reconstruction (i.e., the PSF ± TOF<sub>PET1</sub>/PSF ± TOF<sub>PET2</sub> and PSF ± TOF.EQ<sub>PET1</sub>/PSF ± TOF.EQ<sub>PET2</sub>) did not, however, give perfect agreement compared with the OSEM<sub>PET1</sub>/OSEM<sub>PET2</sub> standard of reference. These discordances were due to differences in background level (liver and blood pool) on posttreatment scan among the different reconstructions, leading to CMR being changed to PMR and vice versa (Supplemental Fig. 1), and to a percentage change in SUL<sub>peak</sub> close to +30% or –30% for the OSEM<sub>PET1</sub>/OSEM<sub>PET2</sub> scenario, resulting in changes

**TABLE 4**

Number of Discordances in PERCIST Classification That Occurred for Different Scenarios Tested

Scenario	No. of discordances			
	Overall	NSCLC	Colorectal liver metastases	Melanoma metastases
OSEM <sub>PET1</sub> /OSEM <sub>PET2</sub> vs. OSEM <sub>PET1</sub> /PSF ± TOF.EQ <sub>PET2</sub>	3	2	0	1
OSEM <sub>PET1</sub> /OSEM <sub>PET2</sub> vs. PSF ± TOF.EQ <sub>PET1</sub> /OSEM <sub>PET2</sub>	0	0	0	0
OSEM <sub>PET1</sub> /OSEM <sub>PET2</sub> vs. OSEM <sub>PET1</sub> /PSF ± TOF <sub>PET2</sub>	17	9	5	3
OSEM <sub>PET1</sub> /OSEM <sub>PET2</sub> vs. PSF ± TOF <sub>PET1</sub> /OSEM <sub>PET2</sub>	13	9	4	0
OSEM <sub>PET1</sub> /OSEM <sub>PET2</sub> vs. PSF ± TOF.EQ <sub>PET1</sub> /PSF ± TOF.EQ <sub>PET2</sub>	4	3	0	1
OSEM <sub>PET1</sub> /OSEM <sub>PET2</sub> vs. PSF ± TOF <sub>PET1</sub> /PSF ± TOF <sub>PET2</sub>	5	3	1	1





**FIGURE 5.** A 50-year-old man with liver metastasis (right lobe) from colon cancer treated by chemotherapy, classified as SMD according to OSEM<sub>PET1</sub>/OSEM<sub>PET2</sub> standard of reference whereas reconstruction inconsistency (OSEM<sub>PET1</sub>/PSF + TOF<sub>PET2</sub>, a scenario mimicking system upgrade during a trial) led to PMD classification. Use of EQ.PET technology correctly classified patient as SMD. (A) Transverse slices at level of the liver metastasis for OSEM and PSF + TOF reconstructions and for baseline and posttreatment scans. (B) Percentage change in SUL<sub>peak</sub> and PERCIST classification according to different scenarios.

from SMD to either PMR or PMD and vice versa for other scenarios.

When strictly applying the PERCIST criterion about normalization of liver uptake (normal liver SUL must be within 20% and <0.3 SUL mean units for PET1 and PET2 to be assessable),

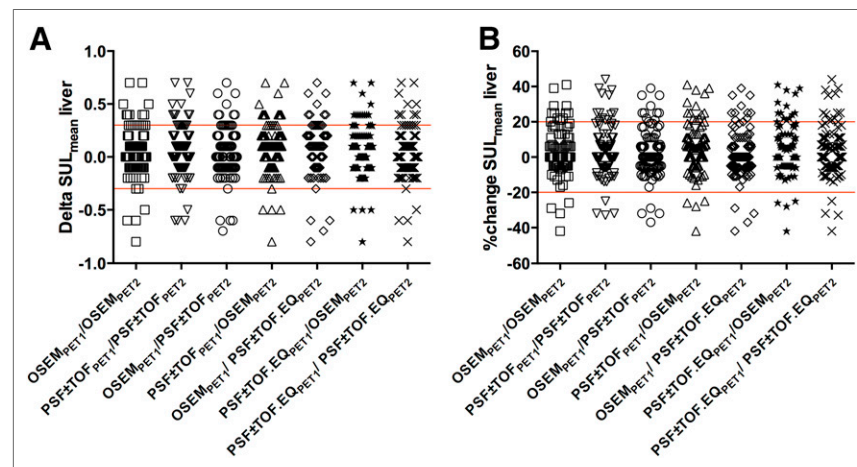
despite excellent compliance with acquisition consistency, a somewhat unexpected finding of this study was that 17% of the response evaluations in this study would not have been considered assessable. As this was observed for the OSEM<sub>PET1</sub>/OSEM<sub>PET2</sub> scenario and consistent in all but 2 patients for the other scenarios, we don't think that this is due to inconsistent reconstruction and may warrant reconsideration of this criterion, particularly for therapies that may alter hepatic metabolism. Factors influencing hepatic <sup>18</sup>F-FDG uptake have been found to be chemotherapy, patient weight/BMI, blood glucose level, and hepatic steatosis (21). In this study, most of the 86 patients were treated with chemotherapy. However, only a few of the patients suffered from severe weight loss or unstable glycaemia.

A limitation of this study is that EQ.PET is a software solution developed for, and applied only to, scanners and reconstruction algorithms of the company that developed this product (including 2 PET/CT systems using a similar PET component) and has not been validated for equipment from other manufacturers.

The alternative approach of obtaining a second reconstruction dataset, as recommended by the European Association Research Ltd. accreditation program for quantitation, can be easily implemented in any PET unit irrespective of their equipment. Using the EQ technology to process images acquired on non-Siemens PET systems would require the vendor-neutral capacity of this software to be validated, using clinical data and other tools such as the digital reference object technique recently published by Pierce et al. (22). However, the EQ.PET software does not require a second standardized reconstruction and could be applied to oldest examinations, acquired and stored before the era of PET standardization programs, provided other sources of SUL variability are controlled and data regarding calibration of the PET system are available. Also, the EQ.PET filter could be adapted to meet any given standard, which is important in the context of evolving guidelines.

## CONCLUSION

Reconstruction algorithm-dependent variability in PERCIST classification is a significant issue but can be overcome by harmonizing SULs using a proprietary software tool. Other manufacturers are encouraged either to emulate this solution or to produce a vendor-neutral approach.



**FIGURE 6.** Absolute (A) and relative (B) variation in liver SUL<sub>mean</sub> within pre- and posttreatment scans. Red lines illustrate PERCIST requirements regarding these parameters.

## DISCLOSURE

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. Prof. Nicolas Aide received a research grant from Siemens Molecular Imaging. No other potential conflict of interest relevant to this article was reported.

## REFERENCES

1. Weber WA, Gatsonis CA, Mozley PD, et al. Repeatability of <sup>18</sup>F-FDG PET/CT in advanced non-small cell lung cancer: prospective assessment in 2 multicenter trials. *J Nucl Med.* 2015;56:1137-1143.

2. Therasse P, Arbuuck 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.
3. Young H, Baum R, Cremerius U, et al. Measurement of clinical and subclinical tumour response using [<sup>18</sup>F]-fluorodeoxyglucose and positron emission tomography: review and 1999 EORTC recommendations. European Organization for Research and Treatment of Cancer (EORTC) PET Study Group. *Eur J Cancer.* 1999;35:1773–1782.
4. 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.
5. Boellaard R. Mutatis mutandis: harmonize the standard! *J Nucl Med.* 2012;53:1–3.
6. Boellaard R. Standards for PET image acquisition and quantitative data analysis. *J Nucl Med.* 2009;50(suppl 1):11S–20S.
7. Panin VY, Kehren F, Michel C, Casey M. Fully 3-D PET reconstruction with system matrix derived from point source measurements. *IEEE Trans Med Imaging.* 2006;25:907–921.
8. Bellevue D, Blanc Fournier C, Switers O, et al. Staging the axilla in breast cancer patients with <sup>18</sup>F-FDG PET: how small are the metastases that we can detect with new generation clinical PET systems? *Eur J Nucl Med Mol Imaging.* 2014;41:1103–1112.
9. European Association of Nuclear Medicine. EARL FDG-PET/CT accreditation. EANM website. [http://earl.eanm.org/cms/website.php?id=/en/projects/fdg\\_pet\\_ct\\_accreditation.htm](http://earl.eanm.org/cms/website.php?id=/en/projects/fdg_pet_ct_accreditation.htm). 2015. Accessed July 6, 2016.
10. FDG-PET/CT Technical Committee. FDG-PET/CT as an imaging biomarker measuring response to cancer therapy, version 1.05, Publicly Reviewed Version. QIBA. RSNA website. [https://www.rsna.org/uploadedFiles/RSNA/Content/Science\\_and\\_Education/QIBA/QIBA\\_FDG-PET\\_Profile\\_v105\\_Publicly\\_Reviewed\\_Version\\_FINAL\\_11Dec2013.pdf](https://www.rsna.org/uploadedFiles/RSNA/Content/Science_and_Education/QIBA/QIBA_FDG-PET_Profile_v105_Publicly_Reviewed_Version_FINAL_11Dec2013.pdf). 2013. Accessed July 6, 2016.
11. Graham MM, Wahl RL, Hoffman JM, et al. Summary of the UPICT protocol for <sup>18</sup>F-FDG PET/CT imaging in oncology clinical trials. *J Nucl Med.* 2015;56:955–961.
12. Makris NE, Huisman MC, Kinahan PE, Lammertsma AA, Boellaard R. Evaluation of strategies towards harmonization of FDG PET/CT studies in multicentre trials: comparison of scanner validation phantoms and data analysis procedures. *Eur J Nucl Med Mol Imaging.* 2013;40:1507–1515.
13. Boellaard R, Delgado-Bolton R, Oyen WJ, et al. FDG PET/CT: EANM procedure guidelines for tumour imaging: version 2.0. *Eur J Nucl Med Mol Imaging.* 2015;42:328–354.
14. Boellaard R, O'Doherty MJ, Weber WA, et al. FDG PET and PET/CT: EANM procedure guidelines for tumour PET imaging: version 1.0. *Eur J Nucl Med Mol Imaging.* 2010;37:181–200.
15. Delbeke D, Coleman RE, Guiberteau MJ, et al. Procedure guideline for tumor imaging with <sup>18</sup>F-FDG PET/CT 1.0. *J Nucl Med.* 2006;47:885–895.
16. Quak E, Le Roux PY, Hofman MS, et al. Harmonizing FDG PET quantification while maintaining optimal lesion detection: prospective multicentre validation in 517 oncology patients. *Eur J Nucl Med Mol Imaging.* 2015;42:2072–2082.
17. Lasnon C, Desmots C, Quak E, et al. Harmonizing SUVs in multicentre trials when using different generation PET systems: prospective validation in non-small cell lung cancer patients. *Eur J Nucl Med Mol Imaging.* 2013;40:985–996.
18. Landis JR, Koch GG. The measurement of observer agreement for categorical data. *Biometrics.* 1977;33:159–174.
19. Sunderland JJ, Christian PE. Quantitative PET/CT scanner performance characterization based upon the society of nuclear medicine and molecular imaging clinical trials network oncology clinical simulator phantom. *J Nucl Med.* 2015;56:145–152.
20. Skougaard K, Nielsen D, Jensen BV, Hendel HW. Comparison of EORTC criteria and PERCIST for PET/CT response evaluation of patients with metastatic colorectal cancer treated with irinotecan and cetuximab. *J Nucl Med.* 2013;54:1026–1031.
21. Boktor RR, Walker G, Stacey R, Gledhill S, Pitman AG. Reference range for intrapatient variability in blood-pool and liver SUV for <sup>18</sup>F-FDG PET. *J Nucl Med.* 2013;54:677–682.
22. Pierce LA, 2nd, Elston BF, Clunie DA, Nelson D, Kinahan PE. A digital reference object to analyze calculation accuracy of PET standardized uptake value. *Radiology.* 2015;277:538–545.