

Standards for PET Image Acquisition and Quantitative Data Analysis

Ronald Boellaard

Department of Nuclear Medicine and PET Research, VU University Medical Center, Amsterdam, The Netherlands

Quantitative ^{18}F -FDG PET is increasingly being recognized as an important tool for diagnosis, determination of prognosis, and response monitoring in oncology. However, PET quantification with, for example, standardized uptake values (SUVs) is affected by many technical and physiologic factors. As a result, some of the variations in the literature on SUV-based patient outcomes are explained by differences in ^{18}F -FDG PET study methods. Various technical and clinical studies have been performed to understand the factors affecting PET quantification. On the basis of the results of those studies, several recommendations and guidelines have been proposed with the aims of improving the image quality and the quantitative accuracy of ^{18}F -FDG PET studies. In this contribution, an overview of recommendations and guidelines for quantitative ^{18}F -FDG PET studies in oncology is provided. Special attention is given to the rationale underlying certain recommendations and to some of the differences in various guidelines.

Key Words: guidelines; standards; PET; standardized uptake value

J Nucl Med 2009;50:11S–20S

DOI: 10.2967/jnumed.108.057182

PET is increasingly being used for diagnosis, staging, and therapy response evaluation (1–10). Interest in PET especially increased after the introduction of PET/CT scanners, which allowed for the collection of both anatomic information and metabolic or functional information in vivo in one scanning session. To date, the success of PET in the oncology domain still relies on the use of ^{18}F -FDG (5,11,12).

In clinical practice, visual inspection of PET or PET/CT images is the main tool for image interpretation, and for staging or restaging, this method is usually adequate (7,13). Although visual inspection may suffice in, for example, many cases of evaluation of the response of gastrointestinal stromal tumors to imatinib and lymphoma restaging, evaluation of the response of solid tumors to therapy is more challenging and requires some form of quantification. Ultimately, PET was developed as a quantitative tool, and its quantitative characteristics are increasingly being recog-

nized as providing an objective, more accurate, and less observer-dependent measure for prognosis and response monitoring purposes than visual inspection alone. Quantification of ^{18}F -FDG uptake therefore has the potential to allow an early, accurate assessment of responses to stratify responding and nonresponding patients (9,14). Moreover, recognition of the potential of quantitative ^{18}F -FDG PET for early response assessment has increased its role in anticancer drug development (15).

Various quantitative measures can be derived from ^{18}F -FDG PET studies (16). The rate of metabolism of glucose, obtained by applying a pharmacokinetic model to data derived from dynamic PET studies, may be considered the gold standard; however, its requirement for dynamic scanning, which is not feasible for whole-body scans, prohibits its routine use in many clinical settings (17). Moreover, dynamic scanning generally requires scan durations of 60–90 min, which reduce patient throughput; in addition, with current PET/CT scanners, dynamic studies cover a field of view of only up to 20 cm. However, this kind of quantification can provide valuable information regarding the validity of the use of simplified quantitative methods (18). The difficulties associated with quantitative data on the rate of metabolism of glucose have led to the development of simplified quantitative measures that can be combined with static whole-body ^{18}F -FDG PET studies.

The standardized uptake value (SUV) is an example of such a simplified measure, and it is now probably the most widely used method for the quantification of ^{18}F -FDG PET studies, although other measures have been developed as well (19–21). The SUV represents the ^{18}F -FDG uptake within a tumor, measured over a certain interval after ^{18}F -FDG administration and normalized to the dose of ^{18}F -FDG injected and to a factor (such as body weight) that takes into account distribution throughout the body (22,23). The SUV normalized to body weight is given by the following equation (SUV equation):

$$\text{SUV} = \frac{\text{AC}_{\text{voi}} (\text{kBq/ml})}{\text{FDGdose (MBq)}/\text{BW (kg)}}. \quad \text{Eq. 1}$$

In Equation 1, AC_{voi} represents the average activity concentration, in kBq/mL , in the specified volume of interest (or the maximum value); FDGdose is the dose of ^{18}F -FDG

Received Nov. 3, 2008; revision accepted Jan. 29, 2009.

For correspondence or reprints contact: Ronald Boellaard, Department of Nuclear Medicine and PET Research, VU University Medical Center, De Boelelaan 1117, 1081 HV Amsterdam, The Netherlands.

E-mail: r.boellaard@vumc.nl

COPYRIGHT © 2009 by the Society of Nuclear Medicine, Inc.

administered, in megabecquerels (corrected for physical decay); and BW is the body weight, in kilograms.

Many factors affect the outcome of the SUV. These factors can be both physiologic (9,22,24,25) and technical (26–29) and have been discussed extensively elsewhere (4,7,9,13,17,30–35). An overview of these factors is provided in Table 1. Approximate ranges and maximum effects are provided to give inexperienced readers an impression of the magnitude of potential errors. However, because the values listed were derived from published studies and unpublished data, deviations might be both larger and smaller in specific cases. Moreover, many factors have, on average, a relatively small effect (<15%) on the SUV outcome, yet the accumulation of many small errors can lead to substantial differences in SUV outcomes among sites (26,43–45). Therefore, strict standardization is of utmost importance.

Figure 1 shows an example of the effects of image reconstruction settings on the maximum SUV (SUV_{max}) in a lesion. As a consequence of such effects, the SUV has been referred to as “silly useless value” (46), a description that is partly justified because of the lack of standardization of procedures; such standardization is needed to minimize the variability of SUVs across institutes and studies. The main disadvantage of nonstandardized SUVs is that, although a proof of concept for various clinical applications has been demonstrated in several single-center studies, the results cannot be directly applied at other sites or in multicenter studies (45). Such heterogeneity partly explains why even the latest response evaluation criteria in solid tumors (47) still do not incorporate quantitative PET, although the oncologic community does recognize its potential.

The need for standardization of quantitative PET was recognized as early as 1998 by Schelbert et al. (34) and in 1999 by a European Organization for Research and Treatment of Cancer (EORTC) task force (17). Subsequently, several other studies reported the impact of various factors on PET quantification and provided recommendations for performing ^{18}F -FDG PET studies. Some of these studies focused mainly on the clinical use of or indications for ^{18}F -FDG PET, on improving PET study quality, and on providing guidelines for PET study interpretation or measurement of the response to therapy (4,7,13,30,31,33,48). Coleman et al. (31) discussed various aspects and technical issues regarding the use of integrated imaging systems, that is, PET/CT. In the present review, the focus is on recommendations and standards given specifically for quantitative ^{18}F -FDG PET oncology studies. The various factors affecting PET quantification and recommendations given in various reports are discussed.

STANDARDS AND RECOMMENDATIONS FOR ^{18}F -FDG PET

There are wide variations in PET and PET/CT scanners, each having its own characteristics, PET acquisition possibilities (e.g., acquisitions in 2-dimensional [2D] and 3-dimensional [3D] modes), image reconstruction methods, and software for visualization and data analysis. The

performance of a PET or PET/CT scanner is generally characterized with National Electrical Manufacturers Association (NEMA) NU 2 standards. The NEMA NU 2 protocol provides a standardized way of assessing the basic performance characteristics of a scanner, such as sensitivity, spatial resolution, noise equivalent count rate curves, scatter fraction, counting rate linearity, and image quality. Although scanner performance can be well characterized with NEMA NU 2 standards, there are still considerable differences in SUV outcomes among centers because of differences in patient preparation methods, PET acquisition settings, image reconstruction algorithms and settings, and data analysis software (26,45).

Differences in scanner performance, (implementation of) image reconstruction algorithms, and data analysis tools cannot be eliminated easily, as they generally are built into the PET or PET/CT scanner itself; that is, scanners from different vendors usually have different acquisition protocols, image reconstruction algorithms, and data analysis software. Moreover, default settings used within these algorithms or software may differ as well. Consequently, it may seem impossible to design guidelines that can ensure the appropriate exchange of SUV results in a multicenter setting. However, it has been shown that SUV results are determined primarily by several factors, parameters, or settings that can and should be standardized (45). Table 1 provides an overview of factors affecting SUVs and their impact on SUVs, which have been extensively discussed in numerous articles (9,14,17,22,25–28,34,45,49,50). Following is a short review of some published recommendations for quantitative PET.

OVERVIEW OF PUBLISHED ^{18}F -FDG PET RECOMMENDATIONS AND GUIDELINES

Table 2 provides an overview of published recommendations or guidelines for ^{18}F -FDG PET or PET/CT. To my knowledge, Table 2 lists articles with detailed and comprehensive recommendations or guidelines for performing and analyzing ^{18}F -FDG PET studies. These articles were obtained from a PubMed search with (a combination of) the following search terms: ^{18}F -FDG, PET, recommendation, standard, guideline, harmonization, quantification, and protocol. Some of the articles dealing specifically with quantitative PET are discussed here; however, various other studies reported on the effects of various factors on SUVs.

A decade ago, Schelbert et al. (34) proposed a procedure guideline for tumor imaging with ^{18}F -FDG. That article summarized indications for ^{18}F -FDG PET and provided recommendations for patient preparation, image acquisition and intervention procedures, and processing (reconstruction) and interpretation or reporting. It was stated that quantification might be helpful in identifying malignant tumors. Moreover, the need for quality control (QC) of radiopharmaceuticals and instrumentation was indicated. Finally, sources of error affecting PET interpretation were listed. In 2002, Bourguet et al. (30) provided guidelines on the clinical use of PET. The objective of that article was to

TABLE 1. Overview of Factors Affecting ^{18}F -FDG PET Quantification

Category	Factor	Explanation	Typical range (maximum effect)*	Reference or source
Technical errors	Relative calibration between PET scanner and dose calibrator	Systematic error in SUV is equal to error in relative calibration between PET scanner and dose calibrator	−10%–10% ($\pm 50\%$)	44,45
	Residual activity in syringe or administration system	Lower net administered dose results in incorrect lower uptake level and SUV	0%–5% (typically <15%, but can be much greater in worst-case situations)	Unpublished data
	Incorrect synchronization of clocks of PET/CT camera and dose calibrator	Incorrect decay correction results in incorrect SUV	0%–10% (21%, as seen in ongoing multicenter study)	Unpublished data
	Injection vs. calibration time	Incorrect time interval is used for decay correction of administered dose	0%–10% (NaN)	Unpublished data
	Paravenous administration of ^{18}F -FDG	Rate and quantity of delivery of ^{18}F -FDG to tumor are reduced, resulting in incorrect SUV	0%–50% or more, strongly depending on quality of administration	Estimated values based on unpublished data
Biologic factors	Blood glucose level	Lower uptake levels or SUVs occur with increasing blood glucose levels	−15%–+15% ($\pm 75\%$) [†]	14,16,54
	Uptake period	Higher SUVs occur at increasing time intervals between injection and start of PET study	+0%–+15% at 60–90 min ($\pm 30\%$)	25
	Patient motion or breathing	Image artifacts result from mismatches in positions between CT-AC and PET emission scans, and lower SUV may result from respiratory motion (resolution loss)	0%–30% ($\pm 60\%$)	36,37
	Patient comfort	Patient stress and poor waiting conditions result in uptake of ^{18}F -FDG in muscle or brown fat and affect SUV quantification	NaN, mainly giving rise to false-positive results ($\text{SUV}_{\text{BW}} = 2\text{--}12$) and possibly incorrect SUV in case of spillover	38
	Inflammation	Inflammatory processes near or at tumor result in false-positive increase in SUV	NaN, mainly giving rise to false-positive results and possibly incorrect SUV in case of spillover	39
Physical factors	Scan acquisition parameters	SNR of PET scan is affected, e.g., lower SNR results in upward bias of SUV	0%–15% ($\pm 15\%$)	26,45
	Image reconstruction parameters	Insufficient convergence and lower resolution result in lower SUV and increase in partial-volume effects; insufficient convergence makes SUV more dependent on surrounding activity distributions	−30%–0% (−30%)	26,27,43,45,49,50
	ROI	SUV outcome is strongly dependent on size and type of ROI used	0%–55% ($\pm 55\%$)	26,45
	Normalization factor for SUV	SUV outcomes are numerically different when body weight, body surface area, and lean body mass are used as normalization factors in SUV equation	Trivial [‡]	14,22
	Blood glucose level correction	Higher serum glucose level results in underestimation of SUV; use of serum glucose level correction in SUV equation therefore results in different SUV outcomes	−15%–15% ($\pm 75\%$) [†]	14,16,54
	Use of contrast agents during CT-AC	Overestimation of attenuation and therefore higher SUV (upward bias) may occur	0%–15% ($\pm 50\%$) [§]	40–42

*Values represent estimated or approximate ranges and maximum deviations derived from published studies or unpublished data. Therefore, quoted values partly reflect personal opinions of authors. However, listed references or other reviews provide more details. For unpublished data, values were estimated or derived from my own data. NaN = not a number available; SUV_{BW} = SUV normalized to body weight.

[†]Very large errors may occur when blood glucose levels are very high (>11 mmol/L). Blood glucose levels should be checked before PET study; if blood glucose levels are high, PET study should be rescheduled (43).

[‡]Use of different normalization factors (e.g., body weight [kg] and body surface area [m^2]) in SUV equation yields different SUVs.

[§]In general, use of contrast agents produces artifacts of up to about 20%. However, very large errors may occur when high-density oral contrast agents (e.g., barium) are used.

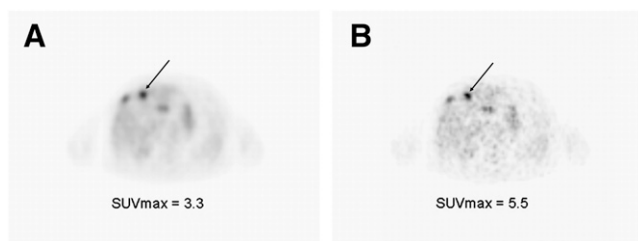


FIGURE 1. Effects of different image reconstruction settings on SUV_{max} outcome (upper liver lesion indicated by arrow) for 2 different reconstruction settings actually applied across institutes in The Netherlands. (A) PET study was reconstructed by use of ordered-subset expectation maximization (OSEM) with 2 iterations and 8 subsets, image matrix size of 128×128 , and postsmoothing with 8-mm FWHM gaussian filter. These reconstruction settings provided (estimated) image resolution of about 11-mm FWHM. Adaptive 50%-of-maximum 3D isocontour revealed estimated lesion size of 4.5 mL. (B) PET study was reconstructed by use of 4 iterations and 16 subsets, image matrix size of 256×256 , and postsmoothing with 5-mm FWHM gaussian filter. These reconstruction settings provided (estimated) image resolution of about 7-mm FWHM. Adaptive 50%-of-maximum 3D isocontour revealed estimated lesion size of 1.5 mL. Difference in SUV_{max} between images was attributable to combined effects of different numbers of iterations and subsets in OSEM, different image matrix sizes, and different smoothing filters.

review literature on the role of and indications for ^{18}F -FDG PET in oncology. In 2008, Fletcher et al. (4) provided an extensive overview of and recommendations for the use of PET for detection, diagnosis, and staging in oncology.

Although those articles suggested opportunities for the use of quantitative PET, they focused mainly on the clinical

indications for (qualitative use of) ^{18}F -FDG PET in oncology. In 1999, the EORTC PET Study Group (17) published a review of and recommendations for the measurement of clinical and subclinical tumor responses with ^{18}F -FDG PET. That article discussed various methods for ^{18}F -FDG PET data analysis, including visual inspection, use of semiquantitative indices, and full kinetic analysis. Several factors affecting ^{18}F -FDG uptake measurements, such as partial-volume effects, applied region-of-interest (ROI) definition, and blood glucose levels, were described. After a review of the assessment of tumor response, recommendations for patient preparation, timing of ^{18}F -FDG PET scans, use of attenuation correction, ^{18}F -FDG dosage, quantification methods, and ROI methodology were made. On the basis of data available at that time on the test-retest reproducibility of quantitative PET measures, quantitative criteria for assessing tumor response were proposed.

In 2005, Weber (9) reviewed the application of PET for monitoring cancer therapy and predicting outcome. That article discussed visual and quantitative response assessment with ^{18}F -FDG PET and provided a detailed overview of factors affecting SUV outcome and quantification methods. Moreover, it discussed when and whether changes in ^{18}F -FDG uptake may be considered significant and the issue of proper timing of ^{18}F -FDG PET studies before and during treatment. Finally, the need for strict adherence to protocols for data acquisition, image reconstruction, and data analysis was emphasized.

In 2005, Coleman et al. (31) summarized an intersociety dialogue on integrated imaging systems. That article provided an overview of clinical applications for PET/CT, issues affecting PET/CT image quality and quantification (such as the effects of using contrast agents and patient motion during CT-based attenuation correction [CT-AC]), the need for qualified personnel, safety issues, and regulatory and legal issues. In 2006, Delbeke et al. (32), who also participated in that intersociety dialogue, provided guidelines for ^{18}F -FDG PET/CT tumor imaging, including guidelines on patient preparation, intervention, the need for collection of other clinical information, CT and PET image acquisition procedures, uptake period, reconstruction and viewing, interpretation, QC, and qualification of personnel. That article provided an extensive point-by-point list of procedures and actions for performing PET/CT studies.

In the same year, Shankar et al. (35) published recommendations for the use of ^{18}F -FDG PET to measure treatment responses in National Cancer Institute (NCI) trials. Like earlier publications, that article described factors affecting SUVs and provided recommendations for patient preparation, image acquisition and reconstruction, timing of ^{18}F -FDG PET studies during therapy, image analysis, and ROI methodology. The authors concluded that there is no single “best” methodology for acquiring and analyzing ^{18}F -FDG PET studies and that standardized protocols needed to be developed for NCI-sponsored trials to assess when ^{18}F -FDG PET could be used as a surrogate endpoint for determining therapeutic efficacy.

TABLE 2. Overview of Studies of ^{18}F -FDG PET or PET/CT Guidelines

Study	Year	Guidelines	
		Clinical use or indications	PET quantification
Schelbert et al. (34)	1998	—	+
Young et al. (17)	1999	—	+
Bourguet et al. (30)	2003	+	—
Weber (9)	2005	—	+
Coleman et al. (31)	2005	+	+
Delbeke et al. (32)	2006	—	+
Shankar et al. (35)	2006	—	+
Kirby and Mikhaeel (33)	2007	+	—
Miller et al. (48)	2007	+	—
Juweid et al. (7)	2007	+	—
Cheson et al. (13)	2007	+	—
Fletcher et al. (4)	2008	+	—
Boellaard et al. (43)	2008	—	+

+ means that this study focused on specified topic; — means that this study did not focus on or had limited focus on specified topic (but does not mean that indications for clinical use were not discussed).

Lammertsma et al. (18) discussed various methods for analyzing ^{18}F -FDG PET studies performed to monitor tumor response. They emphasized the need for standardization. Moreover, they indicated that the relationship between SUVs and data obtained from a full kinetic analysis may be altered during (i.e., because of) treatment. In other words, the observed relative change in the SUV may under- or overestimate the response measured by a full quantitative outcome measure derived from a kinetic analysis. Consequently, the need to validate the use of simplified measures, such as the SUV, against a full kinetic analysis for response monitoring trials was stressed, as was also done by the EORTC PET Study Group (17).

In 2008, a Dutch cooperative group (43) published a protocol for the standardization and quantification of ^{18}F -FDG PET studies in multicenter trials. After a description of factors that affect SUVs, recommendations for patient preparation, PET acquisition, ^{18}F -FDG dosage, image reconstruction, data analysis, ROI procedures, SUV normalization, and QC measures were made. That article specifically focused on the interchangeability of both absolute and relative SUV results in multicenter trials.

OVERVIEW OF SPECIFIC RECOMMENDATIONS FOR QUANTITATIVE ^{18}F -FDG PET STUDIES

As may be deduced from Table 1 and the preceding literature summary, the standardization of quantitative ^{18}F -FDG PET studies is urgently needed and may be achieved by the standardization of several principles. These principles or items reflect, to some extent, the chronological order of performing PET studies and may be identified as patient preparation procedures and interventions, ^{18}F -FDG administration procedures, PET study acquisition, image quality and signal-to-noise ratio (SNR), image reconstruction, clinical image resolution, data analysis procedures and SUV normalization, and QC of instrumentation and qualification of personnel (43).

Patient Preparation Procedures

The procedures used for patient preparation affect ^{18}F -FDG uptake in both tumors and surrounding healthy tissues. Patient preparation procedures therefore should be aimed at maximizing uptake in tumors and minimizing uptake in healthy tissues, thereby optimizing image quality and reducing SUV variability among subjects. The various studies described earlier all provided guidelines for patient preparation, and there seemed to be a general consensus on the optimal preparation procedure. In general, guidelines were given for a fasting period to achieve euglycemic conditions, hydration, use of sedatives and waiting conditions, bladder voiding or use of diuretics, and limits for blood glucose levels. In most cases, additional guidelines were provided for diabetes mellitus patients. A complete delineation of patient preparation guidelines can be found in the articles listed in Table 2.

Until recently, there was still some debate on the optimal time interval between ^{18}F -FDG administration and the start of a PET study. Lowe et al. (25) and Shankar et al. (35) reported that ^{18}F -FDG uptake was still rising up to 120 min after injection, although uptake curves seemed to become flatter at 60–90 min after injection. At present, an interval of 60 min with a tolerance of 5–10 min seems to be considered acceptable in most guidelines. In other publications (17,32,34), a minimum uptake period of 30–40 min was recommended. The shift toward a longer uptake period may reflect the trend toward using PET in a quantitative rather than a qualitative manner.

Additional recommendations are still needed for the time interval between (the end of) therapy cycles and the execution of a PET study. For chemotherapy, a minimum interval of 14 d is usually applied, but more detailed recommendations are given by Juweid et al. (7). For radiotherapy, intervals between the end of treatment and the start of a PET study of even 3 mo may be required. The optimal interval may therefore be study specific, and further investigations are required (11). The appropriate timing of PET studies is one of the topics addressed further in other contributions in this supplement issue of *The Journal of Nuclear Medicine*.

It is also necessary to measure weight and, depending on the SUV normalization procedure, the height of the patient at the time of each PET study. Moreover, the net administered dose specified at the dose calibration time or injection time, which can vary, must be known with certainty. Because all of these values are entered into the SUV equation, they should be reported on a scan report form or entered into the PET system during acquisition, so that the data are stored within the DICOM (Digital Imaging and Communications in Medicine) file headers of the PET scan (43).

^{18}F -FDG Administration Procedures

The net amount of the administered dose is directly used in the SUV calculation. Consequently, the exact ^{18}F -FDG dose given to a patient must be known; that is, the dose must be corrected for residual activity in the syringe or administration system. Moreover, decay corrections must be applied to compensate for the radioactive decay of ^{18}F between the dose calibration time or injection time and the beginning of acquisition. Therefore, clocks in the PET or PET/CT system must be synchronized with those in the dose calibrators used to measure or determine the dose of ^{18}F -FDG injected. A detailed discussion of these issues can be found elsewhere (9,43).

PET Study Acquisition, Image Quality, and SNR

PET acquisition parameters, such as acquisition mode, scan duration per bed position, and amount of bed overlap in subsequent bed positions, in combination with patient weight and ^{18}F -FDG dose, affect PET image quality. It has been shown that poorer image quality (increased noise levels) may result in an upward bias of SUV measurements (26,51). To optimize image quality, recommendations are generally given for uptake period, scan duration, and ^{18}F -FDG dose. The dose can be selected from a range of generally used doses

(17,32,35); as a function of patient weight (34); or as a function of the combination of patient weight and scanner type, scanning procedure, and scan duration (43). The last approach attempts to minimize variability in image quality between different types of scanners.

For PET/CT, it is recommended that a patient be positioned with the arms above the head to improve CT-AC (31,32,43). Breath holding at midinspiration or shallow breathing is recommended to minimize artifacts attributable to mismatches in registration (and blurring) between CT-AC and PET emission scans.

Both Delbeke et al. (32) and Boellaard et al. (43) discussed the use of contrast agents during CT-AC. Delbeke et al. (32) and Coleman et al. (31) indicated that intravenous contrast agents may cause attenuation correction artifacts in PET images but that these effects are usually modest with modern PET/CT scanners. The standardization protocol described by Boellaard et al. (43), however, advises that no contrast agent be used until it has been established that attenuation correction artifacts are completely absent when oral or intravenous contrast agents are used during CT-AC. This more conservative approach has been followed because these recommendations focused mainly on the quantification of PET studies. Whether contrast agents can or cannot be used during CT-AC is, however, a topic for further discussion and evaluation (31,43).

Image Reconstruction and Image Resolution

In most articles that make recommendations, guidelines for optimal reconstruction settings are limited because image reconstruction algorithms and specifications are highly specific to manufacturer and scanner type, making generalizations difficult.

Schelbert et al. (34) indicated the need for reconstruction of PET studies with correction for decay during subsequent bed positions and both with and without attenuation correction. Obviously, attenuation correction is needed for quantification. Young et al. (17) recommended the use of attenuation correction but otherwise provided no image reconstruction parameters. However, at that time, data on the effects of image reconstruction algorithms and settings on PET quantification were sparse, and filtered backprojection was still the most commonly used reconstruction method.

Later, Shankar et al. (35) recommended the use of the same scanner and reconstruction algorithms and settings for multiple scans of a given patient. This approach is justified when ^{18}F -FDG PET studies are used for tumor response measurements based on relative changes in SUVs. It has been shown (45,49) that relative changes in SUVs are almost independent of applied methodology, provided that it is consistently used for all longitudinal PET studies of a given patient and that the scanner is continuously properly calibrated against the dose calibrator.

Delbeke et al. (32) indicated that PET image reconstruction should include corrections for detector efficiency (normalization), system dead time, random coincidences, scatter, attenuation, and sampling nonuniformity.

Similar recommendations were provided by Boellaard et al. (43) to ensure that ^{18}F -FDG PET images are fully quantitative. Apart from the application of all of the usual corrections for quantification, however, differences in PET image resolution are probably major factors contributing to variability in SUVs among centers (45). Therefore, in multicenter studies, when absolute SUV results are used, it is crucial for image resolution to be matched as much as possible across centers and scanners. This is especially true because methods for overcoming quantitative inaccuracies attributable to (differences in) resolution, so-called partial-volume effects (29), are not widely available yet and, more importantly, are still not sufficiently accurate and precise for structures smaller than 3 times the “clinical” scanner resolution (the clinical resolution is usually ~ 7 mm full width at half maximum [FWHM]). To some extent, partial-volume effects may be minimized by use of the SUV_{max} within a lesion. However, differences in image resolution still contribute to interinstitutional differences in the SUV_{max} (Fig. 1) (26,43,45). Furthermore, PET image resolution is not equal to and is usually worse than the scanner resolution measured according to NEMA NU 2 specifications, a fact that is often overlooked or not understood. Boellaard et al. (43) therefore provided reconstruction settings for various scanners but explained that reconstruction should be performed so that activity concentration recovery coefficients as a function of sphere size meet the multicenter QC specifications given in the same study. To this end, dedicated QC phantom experiments were proposed. With this approach, resolution and (iterative reconstruction) convergence matching across various scanners and institutes can be achieved, regardless of the fact that the scanners may have different reconstruction algorithms and software.

Data Analysis Procedures and SUV Normalization

After data acquisition and image reconstruction, the extraction of quantitative measures starts with segmentation of the tumor, that is, definition of an ROI. To this end, various ROI approaches have been used; these approaches include manually defined tumor boundaries, semiautomated 2D and 3D region-growing techniques that involve the application of a fixed or relative (to maximum uptake) threshold, fixed-size regions (SUV_{peak}), and maximum uptake over the entire tumor (SUV_{max}). Variations in SUVs are also caused by imprecise performance of these “simple” ROI methods, especially for small lesions (smaller than 3 times the FWHM), as can be deduced from the estimated lesion sizes in Figure 1. Therefore, there is an urgent need to develop objective and highly automated volumetric delineation methods that accurately yield true lesion size, such as those described by Geets et al. (52) and Hatt et al. (53). These and other sophisticated tumor segmentation methods are presently being developed but are still being evaluated. However, it is important to remember that different ROI methods will result in different quantitative outcomes (26).

Young et al. (17) suggested that the whole tumor uptake be reported and that the same ROI volumes be used on subsequent scans. Mean uptake and maximum uptake, expressed in MBq/L, should be recorded. Shankar et al. (35) did not recommend a specific ROI methodology but indicated that consistent ROI methodology should be used for a given patient during a longitudinal study. Both articles specifically addressed the use of ^{18}F -FDG PET for assessing the tumor response to therapy, for which consistency has been shown to be of utmost importance. Boellaard et al. (43) recommended the use of various 3D ROIs (i.e., volume of interest [VOI]) based on relative (to maximum uptake) thresholds, either with or without background corrections (adaptive thresholds). Because a larger VOI may be more precise, it was suggested that the largest VOI be used to provide stable and accurate VOIs corresponding to the tumor location and extent on all subsequent scans. However, this strategy is only applicable to the assessment of a response to therapy. Nevertheless, it was specified that in all cases, the maximum uptake should be reported and the same VOI method should be used for a given patient and across institutes participating in a multicenter trial.

Finally, SUV normalization variables and the serum glucose level affect SUVs. The SUV is usually normalized to body weight, but lean body mass and body surface area are being used as well. Young et al. (17), Weber (9), Shankar et al. (35), Lammertsma et al. (18), and Boellaard et al. (43) all indicated that SUV normalized to body surface area might be more appropriate during longitudinal studies in case of weight loss during therapy, as has been demonstrated in other studies (22,23). The most appropriate method for SUV normalization is still a matter of debate, but the method should be standardized for multicenter trials.

A higher blood glucose level results in a lower SUV. The possible effects of blood glucose levels were investigated by Eary and Mankoff (54) and discussed by Young et al. (17), Weber (9), Delbeke et al. (32), Shankar et al. (35), Lammertsma et al. (18), and Boellaard et al. (43). They stated that the blood glucose level should be checked before the administration of ^{18}F -FDG. In general, 2 strategies for minimizing the effects of blood glucose levels were recommended. First, with a strict patient preparation protocol (including, e.g., at least 4 h of fasting), a blood glucose level within the reference range (70–130 mg/dL, corresponding to approximately 4–7 mmol/L) usually can be achieved. When needed, the blood glucose level may be reduced by the administration of insulin (17,32), but rescheduling is generally recommended for patients in hyperglycemic states (32,35,43). The blood glucose level threshold for rescheduling ranges from 130 to 200 mg/dL, corresponding to 7.2–11.1 mmol/L. Delbeke et al. (32) and Shankar et al. (35) did not specify an exact threshold but indicated that, in general, a threshold of 150–200 mg/dL (8.3–11.1 mmol/L) is applied by most institutes.

The second strategy for overcoming variability in SUVs because of variations in blood glucose levels involves

incorporation of the blood glucose level as a correction factor in the SUV calculation, as suggested by, for example, Lammertsma et al. (18) and Boellaard et al. (43). However, the use of blood glucose level correction is still a matter of debate, and a possible improvement in SUV accuracy should not be counterbalanced by a disproportionate increase in SUV variability. At the very least, a standardized and validated method for measuring the blood glucose level is needed (i.e., bedside devices should not be used (43)), and the method should be calibrated among centers.

QC Measures and Qualification of Personnel

Technical prerequisites for performing quantitative PET studies are that a scanner is operating according to specifications and that it is calibrated correctly with a phantom that has known activity concentrations, as described elsewhere (9,32,35,43). In addition to QC measures, Delbeke et al. (32) provided additional guidelines for minimum PET scanner performance based on NEMA NU 2 specifications and provided guidelines for the qualification of personnel.

Correct functioning of a PET or PET/CT scanner is usually verified on a daily basis with a set of manufacturer-supplied daily QC routines. In addition, a PET scanner should be cross calibrated against a dose calibrator; that is, the activity measured by the PET camera should be directly compared with the injected activity measured by the dose calibrator being used clinically (12,43,51). Boellaard et al. (43) indicated that this relative calibration should be accurate to within $\pm 5\%$ on the basis of multicenter calibration experiences (43–45).

These QC measures for PET are usually sufficient when quantitative ^{18}F -FDG PET studies are performed to assess the response to therapy on the basis of relative changes in SUVs. However, absolute SUV results can be used for differentiating between benign and malignant lesions, determination of prognosis, and response monitoring by evaluation of residual uptake after therapy or a combination of relative changes and residual uptake during or after therapy; therefore, absolute SUV results are being used more frequently. In all of these cases, especially multicenter studies, additional QC measures are required. Takahashi et al. (55) showed that a variation in the SUV of up to 47% was observed when 5 different scanners were compared and that this variation was reduced to within 22.6% when standardized protocols were used. Boellaard et al. (43) observed that by (very) strict standardization of PET acquisition, image reconstruction, and data analysis procedures, variability in the SUV as a function of sphere size could be reduced to approximately $\pm 10\%$. Therefore, it was suggested that an assessment of activity concentration (or SUV) recovery coefficients as a function of sphere size be included in multicenter QC procedures. Specification of activity concentration or SUV recovery coefficients for each sphere (within the NEMA NU 2 2001 image quality phantom) is an attempt to achieve resolution and reconstruction convergence matching across scanners and institutes.

DISCUSSION

Is There a Need for Different Levels of Standardization?

The required level of standardization may depend on the intended use of ^{18}F -FDG PET. When oncologic ^{18}F -FDG PET studies are interpreted visually for staging and diagnostic purposes, the PET procedure should focus on optimizing image quality for tumor detection. In such a study or clinical practice, specifications for patient preparation, ^{18}F -FDG dosage, and scan duration are required, but it is likely that standardization of reconstruction settings across various institutes will be less stringent.

More strict guidelines in a multicenter setting are needed for quantitative PET studies. A distinction can be made between the use of relative changes in quantitative outcomes for measuring the response to therapy and the use of absolute quantitative measures for diagnosis, determination of prognosis, or prediction of responses (e.g., residual ^{18}F -FDG uptake during or after therapy). With relative changes in SUVs, that is, the ratio of the SUV in response studies to the SUV obtained during a baseline study, the impact of most factors affecting individual SUVs in response studies is minimized (i.e., canceled out). Indeed, it has been shown that the consistent application of a certain methodology, in addition to all other actions undertaken to optimize image quality, is likely to be sufficient (45,49). However, constant scanner performance and calibration over time must be verified (32,35,44). Moreover, if changes in metabolic volume occur, differences in ROI methodologies and image resolution will still affect the observed responses (45), so that standardization of ^{18}F -FDG PET protocols is still mandatory.

When absolute quantitative measures are used in a multicenter trial, all participating centers must follow strict guidelines for patient preparation, PET acquisition, ^{18}F -FDG dosage, image reconstruction (i.e., matching of resolution and iterative reconstruction convergence), data analysis procedures, and QC measures (43).

Finally, Mankoff et al. (56) and Lammertsma et al. (18) emphasized the need for the verification of responses based on SUVs against responses obtained from a full kinetic analysis. For this purpose, dynamic quantitative ^{18}F -FDG PET studies of a small series of patients should be done for comparison with established associations of SUV and kinetic analyses. Several centers, including my own, have large datasets for such comparisons, increasing statistical power. The execution and data analysis of these studies are much more complicated than those needed for SUV quantification. Therefore, these studies are likely to be performed at a limited number of sites. However, at least one single-center validation study should be included during phase I or II studies to correctly interpret SUV responses in larger clinical trials (17). To date, there are no guidelines for performing such quantitative dynamic ^{18}F -FDG PET studies, although some considerations have been described (17,18). It is obvious that such guidelines should incorpo-

rate the same recommendations for patient preparation, ^{18}F -FDG dosage, administration procedures, and image reconstruction settings as those used for studies based on SUVs alone (Table 2). Guidelines for data analysis have not yet been officially provided, although Young et al. (17) suggested Patlak analysis as the preferred kinetic method for analyzing dynamic ^{18}F -FDG PET studies.

In summary, different levels of standardization could be based on the intended use of ^{18}F -FDG PET, ranging from visual inspection to full kinetic analysis. Depending on the capabilities of a certain site, different levels of accreditation could be assigned, thereby allowing or not allowing a site to participate in certain clinical multicenter trials. However, PET centers should be encouraged to meet the highest feasible quality standards. It is important to remember that many small factors affect SUV results and that it is therefore of utmost importance to standardize all of these factors; that is, many small factors can add up to considerable variations in SUVs—up to 50% or more (9,26,45,49,55)—and guidelines therefore should be followed strictly.

Issues in Maintaining and Updating Future Standards

Once a standardized protocol has been established and implemented, regular verification of correct implementation of ^{18}F -FDG PET guidelines is needed. To this end, multicenter QC measurements should be obtained regularly and centrally reviewed and archived.

Because different software platforms or programs provide different SUV outcomes, even when identical datasets are used (57), standardization of ROI methodology and data analysis software is urgently needed. At present, some initiatives, such as that of the Quantitative Imaging Biomarkers Alliance (QIBA), are being undertaken to address these issues. Alternatively, in the absence of standardized ROI software and implementation, quantification of ^{18}F -FDG PET studies in a multicenter setting may need to be reviewed and analyzed centrally.

A potential threat for maintaining standard procedures across various institutes and PET/CT scanners over time is the limited flexibility in defining acquisition and reconstruction settings for most modern PET/CT scanners. This limited flexibility is some form of standardization and may, at present, be considered beneficial. Nevertheless, the impact of future software and hardware upgrades on SUV quantification needs to be monitored carefully. Any change in acquisition and reconstruction settings will have a direct effect on derived SUVs. Any software or hardware upgrade therefore should be clearly specified by the manufacturer, including a description of the effects of the upgrade on image quality, that is, on resolution recovery and SNR (55). Moreover, manufacturers should allow for the continuous use of previously implemented acquisition and reconstruction methods and settings to ensure the consistent use of methodology. This matter is of utmost importance when clinical decisions are being made on the basis of observed changes in SUVs. In other words, an artificial change in an

SUV because of a change in methodology could incorrectly be interpreted as tumor response or progression; obviously, such a situation should be avoided. An alternative strategy is for each manufacturer to assist in the implementation and maintenance of acquisition and image reconstruction protocols to meet the specifications of a multicenter standardized protocol.

Furthermore, future PET/CT scanners are likely to have higher sensitivity, improved image quality, and higher spatial resolution. Higher sensitivity and improved image quality resulting from more advanced image reconstruction methods may be used to reduce the ^{18}F -FDG dose, shorten the scan duration, or both. Changes in spatial resolution, however, have a greater impact on SUV quantification (26,27,45,49). Therefore, as mentioned earlier, it is of utmost importance that any PET study can be reconstructed in such a way that it meets the SUV recovery coefficient requirements of current ^{18}F -FDG PET standardized protocols. A simple solution could be the generation of a second (standardized) PET image by another reconstruction of the PET study with earlier software versions or settings or the simple downsampling of the higher (spatial)-resolution PET image with an appropriate filter to meet the (lower) resolution requirements of ongoing multicenter studies. An additional benefit of this strategy would be the collection of paired (high- and low-resolution) quantitative PET datasets. The use of such paired datasets would allow a proper evaluation of the clinical benefits of new technology and might facilitate the design of new ^{18}F -FDG PET standards, that is, the “translation” of existing protocols, SUV thresholds, or response criteria to newer ones.

Finally, apart from improvements in PET and PET/CT technology, other multimodality imaging devices, such as PET/MRI scanners, will be introduced (31,58–62). Both their clinical use and optimal data acquisition procedures will need to be assessed. Future recommendations will therefore need to focus on PET combined with CT or MRI. Other new developments may be the increasing use of tracers other than ^{18}F -FDG or of multitracer studies (63–65) to further enhance clinical accuracy or in other applications, such as radiation oncology (66,67). These issues are beyond the scope of the present review, but multitracer and multimodality studies are likely to play an increasingly important role in oncology in the future (62,65).

CONCLUSION

The present article reviewed a number of published standards and recommendations for quantitative ^{18}F -FDG PET and PET/CT studies. In most publications, various factors affecting SUVs or quantitative outcomes were identified; subsequently, recommendations to take these factors into account were proposed. Consequently, most recommendations made in these publications seem to be in agreement. The extent and detail of these guidelines seem to reflect mainly the level of standardization required for the intended clinical or multicenter use of

quantitative PET. The main challenges in the near future will be to implement and maintain standards in larger (international) multicenter trials and to incorporate procedures for dealing with new multimodality and multitracer studies.

ACKNOWLEDGMENTS

I would like to thank Otto S. Hoekstra and Adriaan A. Lammertsma for many fruitful discussions, review of this article, and support. I would also like to thank all members of the HOVON imaging working group for their contributions in setting up The Netherlands standardization protocol for multicenter quantitative whole-body ^{18}F -FDG PET studies. Finally, I would like to express much appreciation for many fruitful discussions with colleagues from various societies and working groups, including the EORTC, the American College of Radiology Imaging Network (ACRIN), the American Association of Physicists in Medicine (AAPM) Task Group 145, the QIBA, the European Association of Nuclear Medicine (EANM), the NCI, the American Society of Clinical Oncology (ASCO), EORTC–NCI–ASCO (ENASCO), the Dutch Society of Clinical Physics (NVKF), and the Dutch Society of Nuclear Medicine (NVNG).

REFERENCES

1. Avril NE, Weber WA. Monitoring response to treatment in patients utilizing PET. *Radiol Clin North Am*. 2005;43:189–204.
2. Bastiaannet E, Groen H, Jager PL, et al. The value of FDG-PET in the detection, grading and response to therapy of soft tissue and bone sarcomas: a systematic review and meta-analysis. *Cancer Treat Res*. 2004;30:83–101.
3. Erdi YE. The use of PET for radiotherapy. *Curr Med Res*. 2007;3:3–16.
4. Fletcher JW, Djulbegovic B, Soares HP, et al. Recommendations on the use of ^{18}F -FDG PET in oncology. *J Nucl Med*. 2008;49:480–508.
5. Geus-Oei LF, van der Heijden HF, Corstens FH, Oyen WJ. Predictive and prognostic value of FDG-PET in nonsmall-cell lung cancer: a systematic review. *Cancer*. 2007;110:1654–1664.
6. Hoekstra CJ, Stroobants SG, Smit EF, et al. Prognostic relevance of response evaluation using [^{18}F]-2-fluoro-2-deoxy-D-glucose positron emission tomography in patients with locally advanced non-small-cell lung cancer. *J Clin Oncol*. 2005;23:8362–8370.
7. Juweid ME, Stroobants S, Hoekstra OS, et al. Use of positron emission tomography for response assessment of lymphoma: consensus of the Imaging Subcommittee of International Harmonization Project in Lymphoma. *J Clin Oncol*. 2007;25:571–578.
8. Vansteenkiste JF, Stroobants SG. The role of positron emission tomography with ^{18}F -fluoro-2-deoxy-D-glucose in respiratory oncology. *Eur Respir J*. 2001;17:802–820.
9. Weber WA. Use of PET for monitoring cancer therapy and for predicting outcome. *J Nucl Med*. 2005;46:983–995.
10. Borst GR, Belderbos JS, Boellaard R, et al. Standardised FDG uptake: a prognostic factor for inoperable non-small cell lung cancer. *Eur J Cancer*. 2005;41:1533–1541.
11. Weber WA. PET for response assessment in oncology: radiotherapy and chemotherapy. *Br J Radiol*. 2005;78:42–49.
12. Larson SM, Schwartz LH. ^{18}F -FDG PET as a candidate for “qualified biomarker”: functional assessment of treatment response in oncology. *J Nucl Med*. 2006;47:901–903.
13. Cheson BD, Pfister B, Juweid ME, et al. Revised response criteria for malignant lymphoma. *J Clin Oncol*. 2007;25:579–586.
14. Hoekstra CJ, Hoekstra OS, Stroobants SG, et al. Methods to monitor response to chemotherapy in non-small cell lung cancer with ^{18}F -FDG PET. *J Nucl Med*. 2002;43:1304–1309.
15. Weber WA. Chaperoning drug development with PET. *J Nucl Med*. 2006;47:735–737.

16. Hoekstra CJ, Paglianiti I, Hoekstra OS, et al. Monitoring response to therapy in cancer using [F-18]-2-fluoro-2-deoxy-D-glucose and positron emission tomography: an overview of different analytical methods. *Eur J Nucl Med.* 2000; 27:731–743.
17. Young H, Baum R, Cremerius U, et al. Measurement of clinical and subclinical tumour response using [18F]-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.
18. Lammertsma AA, Hoekstra CJ, Giaccone G, Hoekstra OS. How should we analyse FDG PET studies for monitoring tumour response? *Eur J Nucl Med Mol Imaging.* 2006;33(suppl 1):16–21.
19. Hunter GJ, Hamberg LM, Alpert NM, Choi NC, Fischman AJ. Simplified measurement of deoxyglucose utilization rate. *J Nucl Med.* 1996;37:950–955.
20. Sadato N, Tsuchida T, Nakamura S, et al. Non-invasive estimation of the net influx constant using the standardized uptake value for quantification of FDG uptake of tumours. *Eur J Nucl Med.* 1998;25:559–564.
21. Graham MM, Peterson LM, Hayward RM. Comparison of simplified quantitative analyses of FDG uptake. *Nucl Med Biol.* 2000;27:647–655.
22. Stahl A, Ott K, Schwaiger M, Weber WA. Comparison of different SUV-based methods for monitoring cytotoxic therapy with FDG PET. *Eur J Nucl Med Mol Imaging.* 2004;31:1471–1479.
23. Kim CK, Gupta NC. Dependency of standardized uptake values of fluorine-18 fluorodeoxyglucose on body size: comparison of body surface area correction and lean body mass correction. *Nucl Med Commun.* 1996;17:890–894.
24. Lindholm P, Minn H, Leskinen-Kallio S, Bergman J, Ruotsalainen U, Joensuu H. Influence of the blood glucose concentration on FDG uptake in cancer: a PET study. *J Nucl Med.* 1993;34:1–6.
25. Lowe VJ, DeLong DM, Hoffman JM, Coleman RE. Optimum scanning protocol for FDG-PET evaluation of pulmonary malignancy. *J Nucl Med.* 1995;36:883–887.
26. 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;45:1519–1527.
27. Jaskowiak CJ, Bianco JA, Perlman SB, Fine JP. Influence of reconstruction iterations on ¹⁸F-FDG PET/CT standardized uptake values. *J Nucl Med.* 2005;46:424–428.
28. Lartizen C, Kinahan PE, Swensson R, et al. Evaluating image reconstruction methods for tumor detection in 3-dimensional whole-body PET oncology imaging. *J Nucl Med.* 2003;44:276–290.
29. Soret M, Bacharach SL, Buvat I. Partial-volume effect in PET tumor imaging. *J Nucl Med.* 2007;48:932–945.
30. Bourguet P, Blanc-Vincent MP, Boneu A, et al. Summary of the standards, options and recommendations for the use of positron emission tomography with 2-[¹⁸F]fluoro-2-deoxy-D-glucose (FDG-PET scanning) in oncology (2002). *Br J Cancer.* 2003;89(suppl 1):S84–S91.
31. Coleman RE, Delbeke D, Guiberteau MJ, et al. Concurrent PET/CT with an integrated imaging system: intersociety dialogue from the joint working group of the American College of Radiology, the Society of Nuclear Medicine, and the Society of Computed Body Tomography and Magnetic Resonance. *J Nucl Med.* 2005;46:1225–1239.
32. Delbeke D, Coleman RE, Guiberteau MJ, et al. Procedure guideline for tumor imaging with ¹⁸F-FDG PET/CT 1.0. *J Nucl Med.* 2006;47:885–895.
33. Kirby AM, Mikhael NG. The role of FDG PET in the management of lymphoma: practical guidelines. *Nucl Med Commun.* 2007;28:355–357.
34. Schelbert HR, Hoh CK, Royal HD, et al. Procedure guideline for tumor imaging using fluorine-18-FDG. Society of Nuclear Medicine. *J Nucl Med.* 1998;39:1302–1305.
35. Shankar LK, Hoffman JM, Bacharach S, et al. Consensus recommendations for the use of ¹⁸F-FDG PET as an indicator of therapeutic response in patients in National Cancer Institute trials. *J Nucl Med.* 2006;47:1059–1066.
36. Erdi YE, Nehmeh SA, Pan T, et al. The CT motion quantitation of lung lesions and its impact on PET-measured SUVs. *J Nucl Med.* 2004;45:1287–1292.
37. Hamill JJ, Bosmans G, Dekker A. Respiratory-gated CT as a tool for the simulation of breathing artifacts in PET and PET/CT. *Med Phys.* 2008;35:576–585.
38. Alkhalwaleh K, Alavi A. Quantitative assessment of FDG uptake in brown fat using standardized uptake value and dual-time-point scanning. *Clin Nucl Med.* 2008;33:663–667.
39. Gorospe L, Raman S, Echeveste J, Avril N, Herrero Y, Herna NS. Whole-body PET/CT: spectrum of physiological variants, artifacts and interpretative pitfalls in cancer patients. *Nucl Med Commun.* 2005;26:671–687.
40. Cohade C, Osman M, Nakamoto Y, et al. Initial experience with oral contrast in PET/CT: phantom and clinical studies. *J Nucl Med.* 2003;44:412–416.
41. Dizendorf E, Hany TF, Buck A, von Schulthess GK, Burger C. Cause and magnitude of the error induced by oral CT contrast agent in CT-based attenuation correction of PET emission studies. *J Nucl Med.* 2003;44:732–738.
42. Mawlawi O, Erasmus JJ, Munden RF, et al. Quantifying the effect of IV contrast media on integrated PET/CT: clinical evaluation. *AJR.* 2006;186:308–319.
43. Boellaard R, Oyen WJ, Hoekstra CJ, et al. The Netherlands protocol for standardisation and quantification of FDG whole body PET studies in multi-centre trials. *Eur J Nucl Med Mol Imaging.* 2008;35:2320–2333.
44. Geworski L, Knoop BO, de Wit M, Ivancevic V, Bares R, Munz DL. Multicenter comparison of calibration and cross calibration of PET scanners. *J Nucl Med.* 2002;43:635–639.
45. Westertep M, Pruim J, Oyen W, et al. Quantification of FDG PET studies using standardised uptake values in multi-centre trials: effects of image reconstruction, resolution and ROI definition parameters. *Eur J Nucl Med Mol Imaging.* 2007;34:392–404.
46. Keyes JW Jr. SUV: standard uptake or silly useless value? *J Nucl Med.* 1995;36:1836–1839.
47. Eisenhauer EA, Therasse P, Bogaerts J, et al. New response evaluation criteria in solid tumours: revised RECIST guideline (version 1.1). *Eur J Cancer.* 2009; 45:228–247.
48. Miller JC, Fischman AJ, Aquino SL, Blake MA, Thrall JH, Lee SI. FDG-PET CT for tumor imaging. *J Am Coll Radiol.* 2007;4:256–259.
49. Krak NC, Boellaard R, Hoekstra OS, Twisk JWR, Hoekstra CJ, Lammertsma AA. Effects of ROI definition and reconstruction method on quantitative outcome and applicability in a response monitoring trial. *Eur J Nucl Med Mol Imaging.* 2005;32:294–301.
50. Visvikis D, Cheze-LeRest C, Costa DC, Bomanji J, Gacinovic S, Ell PJ. Influence of OSEM and segmented attenuation correction in the calculation of standardised uptake values for [¹⁸F]FDG PET. *Eur J Nucl Med.* 2001;28:1326–1335.
51. Thie JA. Understanding the standardized uptake value, its methods, and implications for usage. *J Nucl Med.* 2004;45:1431–1434.
52. Geets X, Lee JA, Bol A, Lonneux M, Gregoire V. A gradient-based method for segmenting FDG-PET images: methodology and validation. *Eur J Nucl Med Mol Imaging.* 2007;34:1427–1438.
53. Hatt M, Lamare F, Boussion N, et al. Fuzzy hidden Markov chains segmentation for volume determination and quantitation in PET. *Phys Med Biol.* 2007;52:3467–3491.
54. Eary JF, Mankoff DA. Tumor metabolic rates in sarcoma using FDG PET. *J Nucl Med.* 1998;39:250–254.
55. Takahashi Y, Oriuchi N, Otake H, Endo K, Murase K. Variability of lesion detectability and standardized uptake value according to the acquisition procedure and reconstruction among five PET scanners. *Ann Nucl Med.* 2008;22:543–548.
56. Mankoff DA, Muzi M, Krohn KA. Quantitative positron emission tomography imaging to measure tumor response to therapy: what is the best method? *Mol Imaging Biol.* 2003;5:281–285.
57. Leal J, Lodge M, Wahl R. Reproducibility of SUV_{max} for oncologic PET: significant differences in quantification of the same study between PET-only and PET/CT analysis modes [abstract]. *J Nucl Med.* 2008;49(suppl):107P.
58. Beyer T, Antoch G, Muller S, et al. Acquisition protocol considerations for combined PET/CT imaging. *J Nucl Med.* 2004;45(suppl 1):25S–35S.
59. Beyer T, Antoch G, Bockisch A, Stataus J. Optimized intravenous contrast administration for diagnostic whole-body ¹⁸F-FDG PET/CT. *J Nucl Med.* 2005;46:429–435.
60. Czernin J, Schelbert HR. PET/CT in cancer patient management: introduction. *J Nucl Med.* 2007;48(suppl):2S–3S.
61. De Wever W, Stroobants S, Verschakelen JA. Integrated PET/CT in lung cancer imaging: history and technical aspects. *JBR-BTR.* 2007;90:112–119.
62. Smit EF, Hoekstra OS. New imaging techniques (PET, PET/CT, MRI, E(B)US) for lung cancer staging and response assessment. *J Thorac Oncol.* 2007;2:S144–S145.
63. Dunnwald LK, Gralow JR, Ellis GK, et al. Tumor metabolism and blood flow changes by positron emission tomography: relation to survival in patients treated with neoadjuvant chemotherapy for locally advanced breast cancer. *J Clin Oncol.* 2008;26:4449–4457.
64. Linden HM, Stekhova SA, Link JM, et al. Quantitative fluoroestradiol positron emission tomography imaging predicts response to endocrine treatment in breast cancer. *J Clin Oncol.* 2006;24:2793–2799.
65. Mankoff DA, Eary JF, Link JM, et al. Tumor-specific positron emission tomography imaging in patients: [¹⁸F]fluorodeoxyglucose and beyond. *Clin Cancer Res.* 2007;13:3460–3469.
66. Nestle U, Kremp S, Grosu AL. Practical integration of [¹⁸F]-FDG-PET and PET-CT in the planning of radiotherapy for non-small cell lung cancer (NSCLC): the technical basis, ICRU-target volumes, problems, perspectives. *Radiother Oncol.* 2006;81:209–225.
67. Ollers M, Bosmans G, van Baardwijk A, et al. The integration of PET-CT scans from different hospitals into radiotherapy treatment planning. *Radiother Oncol.* 2008;87:142–146.