

Interobserver Agreement of Qualitative Analysis and Tumor Delineation of ^{18}F -Fluoromisonidazole and 3'-Deoxy-3'- ^{18}F -Fluorothymidine PET Images in Lung Cancer

Sébastien Thureau*¹, Philippe Chaumet-Riffaud*², Romain Modzelewski¹, Philippe Fernandez³, Laurent Tessonnier⁴, Laurent Vervueren⁵, Florent Cachin⁶, Alina Berriolo-Riedinger⁷, Pierre Olivier⁸, H el ene Kolesnikov-Gauthier⁹, Oleg Blagosklonov¹⁰, Boumedi ene Bridji¹¹, Anne Devillers¹², Laurent Collombier¹³, Fr ed eric Courbon¹⁴, Eric Gremillet¹⁵, Claire Houzard¹⁶, Jean Marc Caignon¹⁷, Julie Roux¹⁸, Nicolas Aide¹⁹, Isabelle Brenot-Rossi²⁰, Kaya Doyeux¹, Bernard Dubray¹, and Pierre Vera¹

¹Nuclear Medicine and Radiotherapy, Henri Becquerel Cancer Center and Rouen University Hospital, and QuantIF-LITIS (EA [Equipe d'Accueil] 4108), Faculty of Medicine, University of Rouen, Rouen, France; ²Department of Nuclear Medicine, AP-HP Bic etre and University Paris-Sud, IR4M UMR8081 CNRS, Orsay, France; ³Bordeaux University Hospital, Bordeaux, France; ⁴University Hospital of the Timone, Marseille, France; ⁵Angers University Hospital, Angers, France; ⁶Jean Perrin Cancer Center, Clermont Ferrand, France; ⁷Georges-Fran ois Leclerc Cancer Center, Dijon, France; ⁸Nancy Brabois University Hospital, Nancy, France; ⁹Oscar Lambret Cancer Center, Lille, France; ¹⁰Besan on University Hospital, Besan on, France; ¹¹Ren e Gauducheau Cancer Center, Nantes, France; ¹²Eugene Marquis Cancer Center, Rennes, France; ¹³Nimes University Hospital, Nimes, France; ¹⁴Rangueil University Hospital, Toulouse, France; ¹⁵Nuclear Imaging Center, Saint Etienne, France; ¹⁶Lyon-Sud University Hospital, Lyon, France; ¹⁷Catherine De Siennec Center, Nantes, France; ¹⁸Grenoble University Hospital, Grenoble, France; ¹⁹Fran ois Baclesse Cancer Center, Caen, France; and ²⁰Institut Paoli Calmettes, Marseille, France

As the preparation phase of a multicenter clinical trial using ^{18}F -fluoro-2-deoxy-D-glucose (^{18}F -FDG), ^{18}F -fluoromisonidazole (^{18}F -FMISO), and 3'-deoxy-3'- ^{18}F -fluorothymidine (^{18}F -FLT) in non-small cell lung cancer (NSCLC) patients, we investigated whether 18 nuclear medicine centers would score tracer uptake intensity similarly and define hypoxic and proliferative volumes for 1 patient and we compared different segmentation methods. **Methods:** Ten ^{18}F -FDG, ten ^{18}F -FMISO, and ten ^{18}F -FLT PET/CT examinations were performed before and during curative-intent radiotherapy in 5 patients with NSCLC. The gold standards for uptake intensity and volume delineation were defined by experts. The between-center agreement (18 nuclear medicine departments connected with a dedicated network, SFMN-net [French Society of Nuclear Medicine]) in the scoring of uptake intensity (5-level scale, then divided into 2 levels: 0, normal; 1, abnormal) was quantified by κ -coefficients (κ). The volumes defined by different physicians were compared by overlap and κ . The uptake areas were delineated with 22 different methods of segmentation, based on fixed or adaptive thresholds of standardized uptake value (SUV). **Results:** For uptake intensity, the κ values between centers were, respectively, 0.59 for ^{18}F -FDG, 0.43 for ^{18}F -FMISO, and 0.44 for ^{18}F -FLT using the 5-level scale; the values were 0.81 for ^{18}F -FDG and 0.77 for both ^{18}F -FMISO and ^{18}F -FLT using the 2-level scale. The mean overlap and mean κ between observers were 0.13 and 0.19, respectively, for ^{18}F -FMISO and 0.2 and 0.3, respectively, for ^{18}F -FLT. The segmentation methods yielded significantly different volumes for ^{18}F -FMISO and ^{18}F -FLT ($P < 0.001$). In comparison with physicians, the best method found was $1.5 \times$ maximum SUV (SUVmax) of the aorta for ^{18}F -FMISO and $1.3 \times$ SUVmax of the muscle for

^{18}F -FLT. The methods using the SUV of 1.4 and the method using $1.5 \times$ the SUVmax of the aorta could be used for ^{18}F -FMISO and ^{18}F -FLT. Moreover, for ^{18}F -FLT, 2 other methods (adaptive threshold based on 1.5 or $1.6 \times$ muscle SUVmax) could be used. **Conclusion:** The reproducibility of the visual analyses of ^{18}F -FMISO and ^{18}F -FLT PET/CT images was demonstrated using a 2-level scale across 18 centers, but the interobserver agreement was low for the ^{18}F -FMISO and ^{18}F -FLT volume measurements. Our data support the use of a fixed threshold (1.4) or an adaptive threshold using the aorta background to delineate the volume of increased ^{18}F -FMISO or ^{18}F -FLT uptake. With respect to the low tumor-on-background ratio of these tracers, we suggest the use of a fixed threshold (1.4).

Key Words: PET; ^{18}F -FMISO; FLT; FDG; lung cancer; agreement

J Nucl Med 2013; 54:1543-1550

DOI: 10.2967/jnumed.112.118083

The low efficacy of radiochemotherapy in the treatment of lung cancer can be improved without significant toxicity by increasing the dose of radiotherapy, as demonstrated by Bradley et al. in a phase I/II study (1). In parallel, PET with ^{18}F -fluoro-2-deoxy-D-glucose (^{18}F -FDG), is safe and improves the staging of tumors and target volume definition for lung radiotherapy (2). These improvements increased the life expectancy of patients treated for lung cancer in a phase III study (3), and they allow clinicians to boost the dose of radiotherapy in the hypermetabolic areas identified by ^{18}F -FDG PET (4). However, the best segmentation method for delineating the gross tumor volume (GTV) with ^{18}F -FDG PET/CT remains controversial (5), particularly during treatment in the case of ^{18}F -FDG uptake alteration and tumor shrinkage (6). Several groups have proposed defining target volumes using tracers other than ^{18}F -FDG. New tracers, such as ^{18}F -fluoromisonidazole (^{18}F -FMISO) for tumor hypoxia or

Received Dec. 6, 2012; revision accepted Apr. 9, 2013.

For correspondence or reprints contact: Pierre Vera, Nuclear Medicine, Henri Becquerel Cancer Center, 1 Rue d'Amiens, 76038, Rouen, France.

E-mail: pierre.vera@chb.unicancer.fr

*Contributed equally to this work.

Published online Aug. 5, 2013.

COPYRIGHT   2013 by the Society of Nuclear Medicine and Molecular Imaging, Inc.

3'-deoxy-3'-¹⁸F-fluorothymidine (¹⁸F-FLT) for tumor proliferation, have provided additional data. Radioresistance is a well-known phenomenon in hypoxic tumors (7). Therefore, ¹⁸F-FMISO has been proposed to achieve such increases in the dose of radiotherapy in head and neck cancers (8,9). GTV was assessed on ¹⁸F-FLT images obtained from cases of esophagus cancer (10) and on ¹⁸F-FLT and ¹⁸F-FMISO images obtained from cases of rectal cancer (11). In these studies, the GTV measurements were different between ¹⁸F-FDG and ¹⁸F-FLT or ¹⁸F-FMISO. GTV using ¹⁸F-FLT to define the esophagus could change the treatment planning and doses delivered to organs at risk, according to Han et al. (10). In rectal tumors, ¹⁸F-FMISO was not specific to the tumor volume, and the volume—as defined by ¹⁸F-FLT—decreased more rapidly during chemoradiation than did the ¹⁸F-FDG volume (11).

A major limitation of these tracers, compared with ¹⁸F-FDG, is the low contrast of images using them. Some authors (9,10,12) have proposed boosting radiotherapy on these hypoxia PET images. However, the reproducibility of the visual analysis of images and delineation of volume are not known. Moreover, few data exist regarding the segmentation methods for ¹⁸F-FMISO and ¹⁸F-FLT PET/CT images. Nevertheless, to change the volume or radiotherapy dose using these tracers, it is essential to test the reproducibility of image visual analysis, the reproducibility of volume delineation, and the best method for defining ¹⁸F-FMISO and ¹⁸F-FLT volumes, particularly for multicenter clinical studies.

The objectives of this study were to determine whether image analysis and volume delineation with ¹⁸F-FMISO and ¹⁸F-FLT PET images were reproducible among different nuclear medicine physicians and to find the best method for volume delineation of hypoxia and of hyperproliferative volumes.

MATERIALS AND METHODS

Patient Population

Thirty PET/CT examinations were prospectively acquired from 5 patients before and during curative-intent radiochemotherapy. The inclusion and exclusion criteria were described in a previous paper (13). The patients gave informed written consent. The institutional review board for human studies approved the protocol and the patient information/consent form on April 25, 2008, and the study was submitted to the ClinicalTrials.gov Protocol Registration System (<http://register.clinicaltrials.gov>) in August 2008 (ClinicalTrials.gov RTEP4 study, NCT 01261585). Three PET/CT scans were obtained within 1 wk in each patient before and during radiotherapy (~46 Gy). All of the PET/CT scans were acquired using a Biograph LSO Sensation 16 Hirez (Siemens Medical Solutions).

PET/CT Protocol

For the ¹⁸F-FDG scan, the tracer was injected after 30 min of rest (5 MBq/kg). Sixty minutes later (± 5 min), acquisition began with a CT scan in the craniocaudal direction. The CT scan parameters were set to 120 kV and 100–150 mAs (based on the patient's weight), using dose-reduction software (CareDose; Siemens Medical Solutions). These parameters yielded a mean effective value of 89.1 ± 6.7 mAs. The patient's arms were positioned over the head, and scans were acquired with free breathing and a 16×0.75 -mm primary collimation. The duration of the CT scan was 20 s. No contrast medium was used in this study. PET image acquisition immediately followed in the caudocranial direction, and the scan time was based on 3 min per bed position (without respiratory gating). Six to 8 positions were acquired (whole body); the axial field of view for 1 bed position was 162 mm, with a bed overlap of 25% (plane spacing, 2 mm). The transverse spatial resolution reached 4.4 mm (centered point source in air).

For the ¹⁸F-FLT scan, acquisition started 60 ± 5 min after the injection of tracer (3 MBq/kg). The duration of ¹⁸F-FLT image acquisition was adapted to the injected dose, using the ¹⁸F-FDG sequence as a reference to obtain comparable counting rates for both the ¹⁸F-FLT and the ¹⁸F-FDG PET/CT scans. Three bed positions were centered on the thorax (without gating). The other parameters were similar to those used for ¹⁸F-FDG PET/CT acquisition.

And for the F-FMISO scan, the injected dose of tracer was 2 MBq/kg, and the delay after injection was 180 ± 10 min. Three bed positions were centered on the thorax (without gating). The other parameters were similar to those used for ¹⁸F-FDG PET/CT acquisition.

In a second time, CT images were reconstructed for attenuation correction (5-mm contiguous slices) and for anatomic localization (3-mm slices every 2 mm) on a 512×512 matrix. PET images were iteratively reconstructed (images, 168×168 matrix) using Fourier rebinning and attenuation-weighted ordered-subset expectation maximization software on a clinical Leonardo workstation (Siemens Medical Solutions) with 4 iterations and 8 subsets. Images were corrected for random coincidences, scatter, and attenuation using the CT data. The PET images were finally smoothed using a gaussian filter (full width at half of the maximum, 5 mm).

Image Analysis

Eighteen French centers participated in this study and were connected through a French network (SFMN-net [French Society of Nuclear Medicine]) created and supported by the SFMN, using Keosys workstations (Keosys). Digital Imaging and Communications in Medicine (DICOM) images were transmitted via the Imagys Interface, which is compliant with 21 CFR, part 11. The medical devices of the Keosys Company are ISO 9001- and ISO 13485-compliant. The storage and archiving of images was performed on a central server medical device that was certified IIA class.

In a first step, the 30 PET/CT images (ten ¹⁸F-FDG, ten ¹⁸F-FMISO, and ten ¹⁸F-FLT) were transferred to a dedicated server. Personal access was granted to each of the 18 reviewers who participated for all of the analyses of volumes of interest (VOIs) initially defined by a nuclear physician from the main investigator center (Rouen). Each reviewer had access to the 30 PET/CT images. Nuclear physicians analyzed the 30 PET/CT images and classified the uptake of the 16 VOIs previously defined (13). All 18 nuclear physicians analyzed the 96 lesions (16 lesions in 5 patients \times 3 types of PET/CT examinations [¹⁸F-FDG + ¹⁸F-FMISO + ¹⁸F-FLT] \times 2 times [before and during radiotherapy]).

The scoring was performed according to Hicks et al. (14) and Rischin et al. (15) for the ¹⁸F-FLT and ¹⁸F-FMISO images, respectively. This score initially had 5 classes: 0, uptake less than background; 1, no regions of focal uptake greater than background; 2, focal uptake mildly greater than background; 3, focal uptake moderately greater than background; and 4, focal uptake markedly greater than background. In a second step, these 5 classes were grouped into 2 classes (a score ≥ 2 was considered positive). In another step, the 18 physicians delineated an ¹⁸F-FMISO and ¹⁸F-FLT VOI for 1 patient (patient 003). All centers had already conducted a limited number of examinations with ¹⁸F-FMISO and ¹⁸F-FLT tracers (at least 10 tests per center). The volume delineation was performed using previously proposed threshold methods: the hypoxia volume (¹⁸F-FMISO) was the group of voxels superior to 1.3 times the background maximum standardized uptake value (SUV_{max}) of the muscle (and inside the metabolic GTV defined on the ¹⁸F-FDG PET/CT scan) (8), and the proliferative volume (¹⁸F-FLT) was the group of voxels superior to 1.4 (and inside the metabolic GTV defined on the ¹⁸F-FDG PET/CT scan) (10). This definition was provided to the physicians before the delineation of subvolumes. All of the segmentation procedures were recorded in DICOM-RT and were transferred to the Artiview Aquilab workstation (version 2.3.3; Parc Eurasanté).

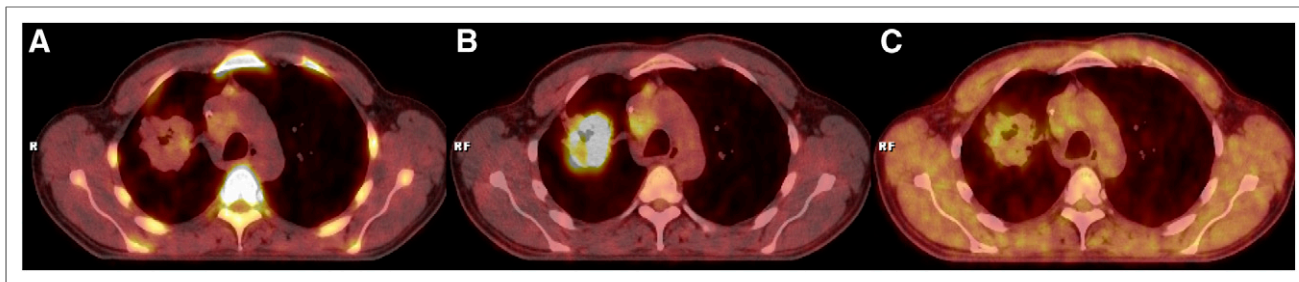


FIGURE 1. ^{18}F -FLT (A), ^{18}F -FDG (B), and ^{18}F -FMISO (C) images.

Finally, all VOIs were defined on a Leonardo workstation using TrueD software (Siemens). To determine the best methods for ^{18}F -FMISO and ^{18}F -FLT volume delineation, the 16 VOIs identified in 5 patients were segmented by the investigator (Fig. 1). Twenty-two segmentation methods, previously defined for ^{18}F -FMISO and ^{18}F -FLT tracers, were used: 2 methods using a fixed threshold with a standardized uptake value (SUV) equal to 1.2 (^{18}F -FMISO_{1,2} and ^{18}F -FLT_{1,2}) and 1.4 (^{18}F -FMISO_{1,4} and ^{18}F -FLT_{1,4}), 2 automatic delineations with a fixed 40% (^{18}F -FMISO_{40%} and ^{18}F -FLT_{40%}) and 50% (^{18}F -FMISO_{50%} and ^{18}F -FLT_{50%}) SUVmax threshold, and 18 methods relative to the background (volumes described for a threshold defined for 1.3-, 1.5-, and 1.6-fold the SUVmax or mean [SUVmean] of the background [identified successively in the aorta, lung, and muscle, respectively]). The ^{18}F -FMISO and ^{18}F -FLT volumes were systematically placed inside the ^{18}F -FDG volume. The whole process generated a database of 1,408 VOIs (16 VOIs \times 2 tracers [^{18}F -FMISO + ^{18}F -FLT] \times 2 times [before and during radiochemotherapy] \times 22 methods).

For ^{18}F -FDG, the 16 VOIs were segmented using 3 segmentation methods (16 VOIs \times 2 times \times 3), leading to a total of 96 VOIs. These 3 segmentation methods were the segmentation methods using thresholds of 40% (^{18}F -FDG_{40%}) and 50% (^{18}F -FDG_{50%}) of SUVmax and a method using a fixed threshold with the SUV equal to 2.5 (^{18}F -FDG_{2,5}).

Statistical Analysis

Descriptive data are expressed as the mean \pm SEM.

The various coefficients of interobserver reliability and agreement were obtained using specialized R packages (R version 2.15.1; CRAN, 2012 version 0.84 [http://www.r-project.org]). The interobserver agreement of specialists for the visual analysis of images was evaluated using κ -values for the different sets of data (16,17).

The reproducibility of volume delineation, compliance, and overlap were compared among the nuclear medicine readers. The quality of the agreement was defined as follows, according to the Cohen κ -test: <0.2, poor agreement; 0.21–0.40, fair agreement; 0.41–0.60, moderate agreement; 0.61–0.80, good agreement; and 0.81–1.00, very good agreement.

The correlation volume defined among the different physicians was obtained using Artiview software, and the overlap and κ -index of compliance were calculated. The higher the index was, the higher the agreement was among nuclear medicine observers for the segmentation of functional volumes (1 = the same volumes in the same site and 0 = no agreement). \cup represents the union and \cap the intersection:

$$\text{Overlap} = (C1 \cap C2 \cap \dots \cap Cn) / (C1 \cup C2 \cup \dots \cup Cn),$$

$$\kappa \text{ index of compliance} = N \times ((C1 \cap C2 \cap \dots \cap Cn) / (C1 + C2 + \dots + Cn)),$$

where C_i is the number of voxels in the VOI determined by physician i , and n is the number of physicians.

First, the comparison of the results of the volume delineation, obtained by the numerous methods, was performed with ANOVA for repeated measures (MedCalc software, version 11.6.0.0; MedCalc Software BVBA). In the second step, the different segmentation methods were compared with a reference standard (gold standard) defined by expert consensus. A lesion was considered positive (^{18}F -FMISO or ^{18}F -FLT) if more than half of the physicians considered it positive using the method of Rischin et al. and if the segmentation method determined a volume (15). A lesion was considered negative (^{18}F -FMISO or ^{18}F -FLT) if more than half of physicians considered it negative using the method of Rischin et al. and if the segmentation method did not determine a volume. The results of this classification and the reference standard established by consensus enabled us to calculate the sensitivity and specificity of each segmentation method.

RESULTS

Reproducibility of Image Analysis (Qualitative Analysis)

Table 1 shows the global agreements of the visual analyses of ^{18}F -FDG, ^{18}F -FMISO, and ^{18}F -FLT PET/CT images in 5 and 2 classes. The interobserver agreement of the 5-level grading was moderate for images acquired with ^{18}F -FDG (0.59), ^{18}F -FMISO (0.43), and ^{18}F -FLT (0.44) (slightly superior for ^{18}F -FDG). The quality of the reproducibility increased to very good for ^{18}F -FDG (0.81) and to good for ^{18}F -FMISO (0.77) and ^{18}F -FLT (0.77) when estimated using 2-level scoring. The reproducibility of the visual analysis of the PET/CT images regarding the primary tumor and nodes is shown in Table 2. With 5 classes, only the ^{18}F -FDG PET/CT primary tumors showed very good reproducibility (0.84), whereas the reproducibility was moderate for nodes (0.51). For primary tumors and nodes, the κ -values were moderate with ^{18}F -FMISO (0.45 and 0.44) and ^{18}F -FLT PET (0.49 and 0.41). With a 2-level classification system, the interobserver agreement of specialists rose from good to very good (from 0.65 to 0.87) for primary tumors, and the nodes' reproducibility became good (0.74 to 0.81). The reproducibility of the analyses of PET/CT images at both times of acquisition (before and during radiochemotherapy) is shown in Table 3. On the basis of a 5-level classification, only the ^{18}F -FDG PET/CT primary tumors showed good reproducibility (0.65). The agreements for ^{18}F -FMISO and ^{18}F -FLT PET images were, respectively, fair and moderate (0.34 to 0.52). In

TABLE 1

Global κ -Test for Visual Analysis of PET/CT Images for 18 Nuclear Medicine Physicians at 18 Centers

κ	^{18}F -FDG	^{18}F -FMISO	^{18}F -FLT
5 classes	0.59	0.43	0.44
2 classes	0.81	0.77	0.77

TABLE 2
κ-Values for PET/CT Images Regarding Primary Tumors and Nodes

κ	¹⁸ F-FDG		¹⁸ F-FMISO		¹⁸ F-FLT	
	Tumor	Node	Tumor	Node	Tumor	Node
5 classes	0.84	0.51	0.45	0.44	0.49	0.41
2 classes	0.87	0.78	0.65	0.81	0.84	0.74

2 classes, reproducibility rose from good to very good (from 0.73 to 0.88) for all tracers. When the reproducibility of tracer analysis was compared, the ¹⁸F-FDG PET/CT images had the best reproducibility, particularly for tumors and before treatment (in 5 and 2 classes).

Concordance is usually used as a scalar measurement of agreement. However, the tables for the probability of *z* were analyzed, and 95% confidence intervals of κ were calculated, resulting in the conclusion that significant concordance existed in all of the datasets (the 95% confidence intervals of κ did not cross the zero value).

Reproducibility of Volume Delineation (Quantitative Analysis)

All ¹⁸F-FMISO and ¹⁸F-FLT VOIs (patient 003) sent by the 18 centers were superimposed on the ¹⁸F-FMISO and ¹⁸F-FLT PET/CT images, respectively (Figs. 2A and 2B). The mean ± SEM volumes were 46.47 ± 147.51 mL for ¹⁸F-FMISO and 22.23 ± 80.15 mL for ¹⁸F-FLT. The mean overlap and κ-index of compliance among the observers were 0.13 (0 to 0.75) and 0.19 (0 to 0.86) for ¹⁸F-FMISO, respectively, and 0.2 (0.004 to 0.68) and 0.3 (0.13 to 0.73) for ¹⁸F-FLT, respectively.

Comparison Methods for Volume Delineation

The comparisons of ¹⁸F-FDG volumes were not significantly different using the 3 methods (FDG_{40%} = 14.5 ± 2.56 mL; FDG_{50%} = 9.8 ± 1.73 mL; and FDG_{2.5} = 13.1 ± 2.3 mL). The 22 delineation methods were significantly different for ¹⁸F-FMISO (Fig. 3; *P* < 0.0001) and ¹⁸F-FLT (Fig. 4; *P* < 0.0001). For both tracers, some methods provided volumes significantly greater than the ¹⁸F-FDG_{40%} volume. The comparisons between the different methods of delineation and the experts are expressed in Table 4. Methods that yielded an ¹⁸F-FMISO or ¹⁸F-FLT volume superior to the ¹⁸F-FDG volume were excluded from the analysis (Table 4). The best sensitivity and specificity on a receiver-operating characteristic curve were obtained, respectively, for ¹⁸F-FMISO_{1.5×SUVmaxAorta} (area under the curve for receiver-operating-characteristic curve analysis [AUC], 0.84) and ¹⁸F-FLT_{1.3×SUVmaxMuscle} (AUC, 0.90). ¹⁸F-FMISO_{1.4} and ¹⁸F-FLT_{1.4}, ¹⁸F-FMISO_{1.5×SUVmaxAorta}, ¹⁸F-FLT_{1.5×SUVmaxAorta}, ¹⁸F-FLT_{1.6×SUVmaxAorta}, ¹⁸F-FLT_{1.3×SUVmaxMuscle}, ¹⁸F-FLT_{1.5×SUVmaxMuscle}, and ¹⁸F-FLT_{1.6×SUVmaxMuscle} yielded satisfactory ranges of values for sensitivity and specificity, with an AUC greater than 0.79.

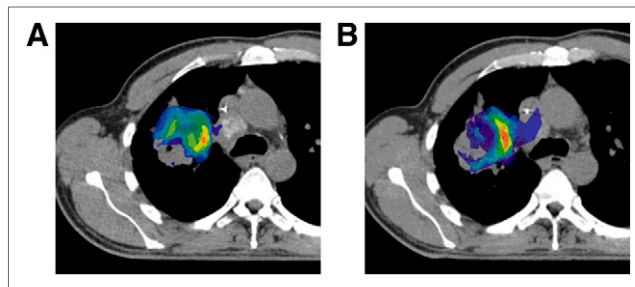


FIGURE 2. (A) Reproducibility of ¹⁸F-FMISO volume delineation for 1 patient (orange, good agreement; green, moderate agreement; blue, poor agreement). (B) Reproducibility of ¹⁸F-FLT volume delineation for 1 patient (orange, good agreement; green, moderate agreement; blue, poor agreement).

Therefore, the methods using the SUV of 1.4 and the method using 1.5 × SUVmax of the aorta could be used for the 2 tracers (¹⁸F-FMISO and ¹⁸F-FLT). In Figure 3, the mean hypoxic fraction of the lesions is almost zero for the best delineation. The ratio between metabolic and hypoxic volume is approximately 15%–10% for head and neck cancer (8,9) and probably less in lung cancer. Our study focus was on primary tumors and mediastinal lymph disease.

DISCUSSION

This study was performed in 18 nuclear medicine centers in France connected with a dedicated network. A specific network for the transfer and exchange of PET/CT images and radiotherapy VOIs in multicenter studies was specifically built by the SFMN. Its objective is to favor and foster the implementation of studies requiring the transfer of images (e.g., scintigraphy, radiotherapy images, and delineation of boundaries). This network used Web access and dedicated software and workstations. This study demonstrated the robustness of this network and its ease to use in interobserver agreement studies and multicenter clinical research in the fields of nuclear medicine and radiotherapy. There are 2 major issues to take into consideration with respect to the interpretation of hypoxia data. The first, is the tumor hypoxic (or proliferative)? The second issue is, what is the true hypoxic (or proliferative) subvolume? Both points were assessed subsequently.

This research demonstrated that visual analysis of low-contrast PET images (¹⁸F-FMISO and ¹⁸F-FLT) is feasible and reproducible when the images are analyzed using a binary scale (positive or negative). The interobserver agreement was slightly higher for ¹⁸F-FDG images, but it could also be partly explained by the relatively limited experience of physicians with respect to ¹⁸F-FMISO or ¹⁸F-FLT image analysis. In contrast, the volume delineation was different from one physician to another for these tracers. Among the physicians, the best method found was 1.5 × SUVmax of the

TABLE 3
κ-Values Regarding Times of PET/CT Images (Before or During Radiotherapy)

κ	¹⁸ F-FDG		¹⁸ F-FMISO		¹⁸ F-FLT	
	Before radio-chemotherapy	During radio-chemotherapy	Before radio-chemotherapy	During radio-chemotherapy	Before radio-chemotherapy	During radio-chemotherapy
5 classes	0.65	0.53	0.45	0.43	0.52	0.34
2 classes	0.88	0.73	0.74	0.80	0.79	0.73

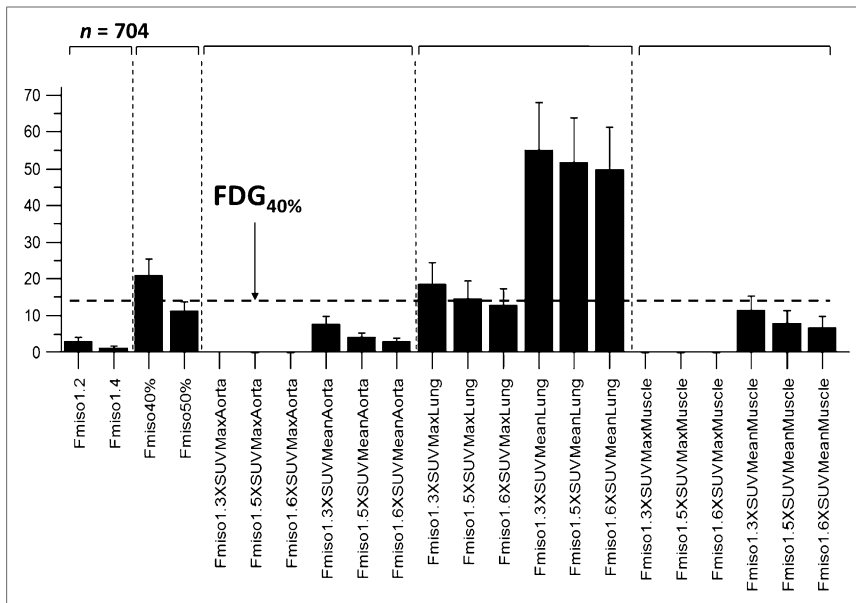


FIGURE 3. Data are mean \pm SEM volumes (mL) for 22 methods of volume delineation for ^{18}F -FMISO.

aorta for ^{18}F -FMISO and $1.3 \times \text{SUVmax}$ of the muscle for ^{18}F -FLT. The methods using an SUV of 1.4 and the method using $1.5 \times \text{SUVmax}$ of the aorta could be used for both tracers (^{18}F -FMISO and ^{18}F -FLT).

The κ -value is a chance-adjusted index of agreement, which can be used in medical research and especially to assess the reproducibility of visual analysis among physicians. It is well known that Cohen's κ -values do not assume equal classification proportions for different observers, but they depend on the characteristics of marginal distributions. A limitation of our study is that the rate of true-positives in the study population and the bias of one reader relative to another could have significantly affected our results (18,19).

However, the data analysis doesn't show any evidence of bias, because the different observers systematically differed in their

classification. Most studies assessing multirater agreement involve between 3 and 8 readers. We believe that the large number of independent reviewers in this study should provide confidence in our results.

This study, performed in 18 different nuclear medicine centers, demonstrated high reproducibility of ^{18}F -FMISO or ^{18}F -FLT PET images when the images were interpreted using a binary scale. The reproducibility of the visual analysis was greater for ^{18}F -FDG images than for ^{18}F -FMISO and ^{18}F -FLT acquisitions, no matter whether the images were interpreted using a 2- or 5-point grading scale. These results could be explained by a much steeper learning curve in the reading of ^{18}F -FDG PET/CT images by physicians and by the lower contrast of ^{18}F -FMISO and ^{18}F -FLT images than of ^{18}F -FDG. In our population, the SUVmax of ^{18}F -FDG images was 6.1 (95% confidence interval, 4.0–8.3) before radiochemotherapy and 4.3 (95% confidence interval, 3.2–5.4) during radiochemotherapy. For

^{18}F -FLT acquisitions, the mean SUVmax was 4.7 (95% confidence interval, 3.6–5.7) before radiochemotherapy and 2.2 (95% confidence interval, 1.5–2.8) during radiochemotherapy. For ^{18}F -FMISO, the SUVmax was lower, at 1.5 (95% confidence interval, 1.3–1.7) before radiotherapy and 1.4 (95% confidence interval, 1.2–1.7) during radiochemotherapy. The low contrast of ^{18}F -FMISO and ^{18}F -FLT PET images could most likely explain the best reproducibility of the analysis of PET/CT images occurring before treatment, compared with images obtained during treatment (Table 3). The interpretation appeared to be easier for primary tumors, compared with nodes, for ^{18}F -FDG, but it was also difficult for ^{18}F -FMISO and ^{18}F -FLT when using a 5-grade scale (Table 2). In all of the analyses, the reproducibility dramatically increased using a binary scale, compared with a 5-point scale. This is an important result for the

robustness and future development of these tracers in nuclear medicine. Similar results were previously shown by our group for the reproducibility of ^{18}F -FDG PET images in a series of patients with squamous cell esophageal carcinoma (20).

This study clearly demonstrated the lack of reproducibility by physicians of the delineation of hypoxic or proliferative tumoral volumes on PET/CT images. Many studies (21,22) demonstrated increases in the reproducibility of GTV delineation using ^{18}F -FDG PET images, compared with GTV delineation with CT alone. Moreover, the pretherapeutic volume is important to determine precisely because some authors have suggested that the site of highest ^{18}F -FDG uptake could be a site of recurrence (23). With ^{18}F -FMISO PET/CT, our group designed a study to boost the hypoxic volume with external maximal-dose radiotherapy (clinicaltrials.gov NCT01576796). Other researchers have proposed boosting cervical

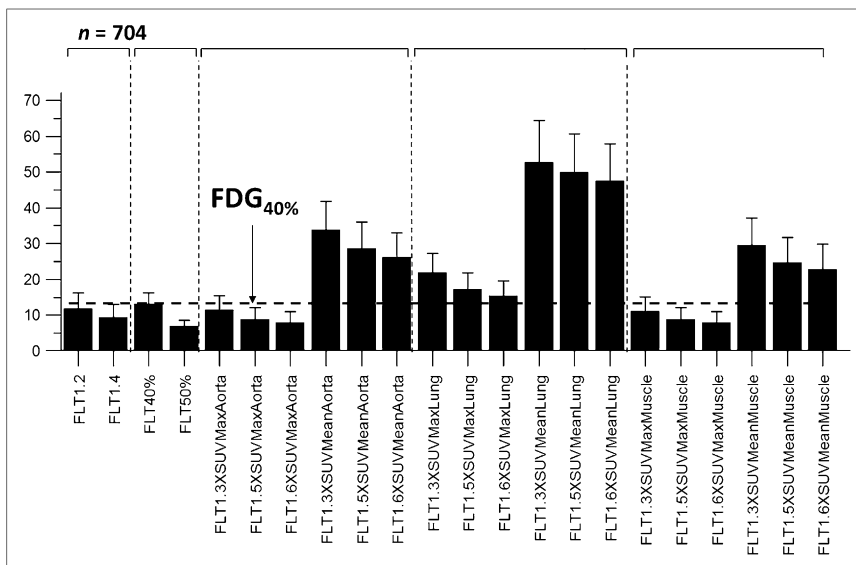


FIGURE 4. Data are mean \pm SEM volumes (mL) for 22 methods of volume delineation for ^{18}F -FLT.

TABLE 4
Sensitivity and Specificity of Different Methods of Volume Delineation

Segmentation	¹⁸ F-FMISO			¹⁸ F-FLT		
	Sensitivity	Specificity	AUC	Sensitivity	Specificity	AUC
1.2	1	0.56	0.78	0.92	0.5	0.71
1.4	0.86	0.72	0.79	0.92	0.83	0.88
SUVmax						
40%	1	0	0.50	1	0	0.50
50%	1	0	0.50	1	0	0.50
SUVmaxAorta						
1.3 ×	1	0.56	0.78	0.96	0.5	0.73
1.5 ×	0.71	0.96	0.84	0.92	0.83	0.88
1.6 ×	0.14	1	0.57	0.92	0.83	0.88
SUVmeanAorta						
1.3 ×	1	0.36	0.68	1	0.33	0.67
1.5 ×	1	0.4	0.70	1	0	0.50
1.6 ×	1	0.56	0.78	1	0	0.50
SUVmaxLung						
1.3 ×	1	0	0.50	1	0	0.50
1.5 ×	1	0.1	0.55	1	0.33	0.67
1.6 ×	1	0.1	0.55	1	0.33	0.67
SUVmeanLung						
1.3 ×	1	0	0.50	1	0	0.50
1.5 ×	1	0	0.50	1	0	0.50
1.6 ×	1	0	0.50	1	0	0.50
SUVmaxMuscle						
1.3 ×	0.43	0.96	0.70	0.96	0.83	0.90
1.5 ×	0.29	1	0.65	0.92	0.83	0.88
1.6 ×	0.29	1	0.65	0.87	0.83	0.85
SUVmeanMuscle						
1.3 ×	0	0.96	0.48	1	0.17	0.59
1.5 ×	1	0.14	0.57	1	0.17	0.59
1.6 ×	1	0.52	0.76	0.96	0.17	0.57

For sensitivity and specificity reference standard defined by consensus group.

radiotherapy in patients with head and neck cancer with the doses then escalated in the hypoxic GTV, as measured by ¹⁸F-FMISO PET/CT images (12). Similar approaches were performed (8,9) with head and neck cancers. Hendrickson et al. (9) increased the external radiotherapy dose to 60 Gy on cervical nodes.

Up until now, there have been no data published regarding the reproducibility of visual analysis of ¹⁸F-FMISO and ¹⁸F-FLT PET/CT images. Because this study demonstrated the poor reproducibility of analyses with these tracers (Figs. 2A and 2B), it appears important that each nuclear medicine department sets up a learning phase and evaluates the reproducibility among nuclear physician readers. A multidisciplinary meeting, with nuclear physicians and radiation oncologists, for volume delineation could be an interesting way to proceed, as suggested by Nestle et al. (5).

In ¹⁸F-FDG PET/CT, a universal method for GTV delineation has not been definitively validated for all situations (e.g., pretherapeutic vs. midtreatment, high vs. low contrast, and large vs. small tumors). The visual (manual) segmentation method is dependent on the experience of the nuclear physician and on the contouring protocol used, which can lead to significant variation in the assessment of GTV with ¹⁸F-FDG PET (24). To reduce these discrepancies, many authors favor the use of semiautomatic segmentation methods, which are considered to be more reproducible, taking into account the lesion volume, the level of SUV distribution heterogeneity, the tumor-to-background ratio, the background activity

level, and the region-growing initialization (25,26). The use of these segmentation methods has revealed significant variations between CT-based and PET/CT-based radiotherapy treatment plans for GTV determination. Because of a lack of consistency in tumor contour delineation on PET, interpretation of the available data is difficult, and how PET imaging should be integrated into treatment planning remains uncertain, in the absence of a consensus on the use of a definitive segmentation method (27,28). Many studies have compared ¹⁸F-FDG PET/CT segmentation methods with CT and with MR imaging (29), and only few of these studied used pathologic examination data as references (10,29). Moreover, lung PET/CT imaging applications in thoracic or abdominal pathologies can be affected by motion artifacts related to respiratory movement (signal smearing). This effect can lead to the underestimation of a lesion's SUVmax or overestimation of GTVs (30), even if different methods are applied to overcome respiratory motion (31). Carlin et al. recalled in a recent review the difficulties in defining a hypoxic volume with ¹⁸F-FMISO (32).

With ¹⁸F-FMISO and ¹⁸F-FLT, few methods have been tested to determine hypoxia and proliferative volumes. Manual methods and methods less sophisticated than ¹⁸F-FDG have been used with these tracers. Been et al. determined a proliferative volume of 40% of the SUVmax of ¹⁸F-FLT images. Nehmeh et al. (33) used a volume obtained with a value of the SUV at 1.2 × background measured in blood plasma. Grosu et al. (34) used, for ¹⁸F-fluoroazomycin

arabinoside (FAZA), a volume determined by $1.5 \times \text{SUV}_{\text{max}}$ of the cervical muscle. The issue to resolve for muscle is the variability of determination, thus Choi et al. preferred using cerebellar tissue (8). However, cerebellar tissue is not systematically in the field of view. For the esophagus, Han et al. obtained the best volume with an SUV of 1.4 for ^{18}F -FLT acquisitions, as Tian et al. and Xu et al. (35,36) did for lung cancer with the same tracer. These results are in accordance with our results.

These methods, which have been studied previously, were tested and compared in this study of lung tumors. For methods using background comparison, 3 backgrounds were tested (aorta, contralateral lung, and muscle) and were used successively with SUV_{mean} and SUV_{max} . The method of Lee et al. (9,12), using comparison to the background and using blood samples, has not been tested. More than 1,400 VOIs were drawn in this preliminary study, and no other complex methods were tested.

Our study clearly showed that all of these methods yielded different results for the volumes of ^{18}F -FMISO (Fig. 3) and ^{18}F -FLT (Fig. 4). It was difficult to validate the best method definitively because we did not have a gold standard for in vivo measurement of hypoxic or proliferative volumes. The invasive method, using Eppendorf electrodes, was insufficient because it tested only part of the tumor (37). A comparison with pathology (pimonidazole immunohistochemistry, for example) was not possible because it would have been available only after surgery. Therefore, we used the physicians' validation as the gold standard. The ability of the methods to find hypoxic or proliferative volumes was compared with the panel of physicians. Excluded were all methods that yielded a hypoxic or proliferative volume when more than 50% of the physicians did not find it and all of the methods that did not determine a hypoxic or proliferative volume when more than 50% of the physicians found it. This approach was similar to calculating sensitivity and specificity (Table 4). We postulated that the hypoxic or proliferative volume could not be superior to the ^{18}F -FDG GTV. Therefore, all of the methods yielding hypoxia or proliferative volumes superior to the ^{18}F -FDG GTV, measured at a threshold of 40%, were excluded from the analysis. Thus, for ^{18}F -FMISO, ^{18}F -FMISO_{40%}, and methods using the SUV_{mean} of lung background, comparisons were excluded (Fig. 3). For ^{18}F -FLT, methods using comparison with background SUV_{mean} (aorta, lung, and muscle) and lung $\text{SUV}_{\text{max,max}}$ were excluded (Fig. 4). For the methods using the lung for the background, the SUVs were low (SUV_{mean} and SUV_{max}) and thus responsible for large volumes. For ^{18}F -FMISO, the methods using comparison of the background with aorta SUV_{max} and muscle SUV_{max} were most likely not efficient because they resulted in the absence of a hypoxic volume (volume equal to zero). Among physicians, the best method found was $1.5 \times \text{SUV}_{\text{max}}$ of the aorta for ^{18}F -FMISO and $1.3 \times \text{SUV}_{\text{max}}$ of the muscle for ^{18}F -FLT. The methods using an SUV of 1.4 and the method using $1.5 \times \text{SUV}_{\text{max}}$ of the aorta could be used for the 2 tracers (^{18}F -FMISO and ^{18}F -FLT) with relatively good sensitivity and specificity. It is worth noting that, for a slightly larger value ($1.6 \times \text{SUV}_{\text{max}}$), the sensitivity declines from 0.71 to 0.14. The threshold change ($\sim 7\%$) is within the range of the statistical fluctuations of the SUV_{max} , indicating that even the best method for ^{18}F -FMISO is unstable.

An optimal fraction for all reference values (max/mean aorta/lung/muscle) can be estimated empirically in a given study sample. It is likely that a different fraction would be optimal if the sample changed (e.g., different PET scanner, reconstruction, scan time). The issue is the degree of stability of these methods if there

are small variations. The method stability is especially important in the context of multicenter trials. The above results indicate that the stability is different for ^{18}F -FMISO and ^{18}F -FLT. Therefore, an absolute threshold with an SUV of 1.4 is deemed preferable in view of its simplicity.

For these images with low contrast, statistical methods, such as random walk as defined by Onoma et al. during the International Symposium on Biomedical Imaging (ISBI) Congress 2012 or fuzzy locally adaptive Bayesian (FLAB) methods (38), might generate some interest if tested in future work.

CONCLUSION

This study showed excellent reproducibility of the analysis of ^{18}F -FMISO and ^{18}F -FLT PET/CT images when the images were analyzed using binary scales. This reproducibility allows for the use of these images in multicenter studies. In contrast, the poor reproducibility of the delineation of hypoxia and proliferative volumes requires great caution about their use for the management of patients and for therapeutic decision-making. The best methods found were $1.5 \times \text{SUV}_{\text{max}}$ of the aorta for ^{18}F -FMISO and $1.3 \times \text{SUV}_{\text{max}}$ of the muscle for ^{18}F -FLT. The method using an SUV of 1.4 and the method using $1.5 \times \text{SUV}_{\text{max}}$ of the aorta could be used for the 2 tracers (^{18}F -FMISO and ^{18}F -FLT). With respect to the low tumor-on-background ratio of these tracers, we suggest the use of a fixed threshold (SUV, 1.4).

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

REFERENCES

1. Bradley JD, Moughan J, Graham MV, et al. A phase I/II radiation dose escalation study with concurrent chemotherapy for patients with inoperable stages I to III non-small-cell lung cancer: phase I results of RTOG 0117. *Int J Radiat Oncol Biol Phys.* 2010;77:367-372.
2. De Ruyscher D, Faivre-Finn C, Nestle U, et al. European Organisation for Research and Treatment of Cancer recommendations for planning and delivery of high-dose, high-precision radiotherapy for lung cancer. *J Clin Oncol.* 2010; 28:5301-5310.
3. Ung Y, Gu C, Cline K, et al. An Ontario Clinical Oncology Group (OCOG) randomized trial (PET START) of FDG PET/CT in patients with stage III non-small cell lung cancer (NSCLC): predictors of overall survival [abstract]. *J Clin Oncol.* 2011;29 (suppl):7018
4. Fleckenstein J, Hellwig D, Kremp S, et al. F-18-FDG-PET confined radiotherapy of locally advanced NSCLC with concomitant chemotherapy: results of the PET-PLAN pilot trial. *Int J Radiat Oncol Biol Phys.* 2011;81:e283-e289.
5. Nestle U, Kremp S, Schaefer-Schuler A, et al. Comparison of different methods for delineation of ^{18}F -FDG PET-positive tissue for target volume definition in radiotherapy of patients with non-small cell lung cancer. *J Nucl Med.* 2005;46:1342-1348.
6. Edet-Sanson A, Dubray B, Doyeux K, et al. Serial assessment of FDG-PET FDG uptake and functional volume during radiotherapy (RT) in patients with non-small cell lung cancer (NSCLC). *Radiother Oncol.* 2012;102:251-257.
7. Nordmark M, Bentzen SM, Rudat V, et al. Prognostic value of tumor oxygenation in 397 head and neck tumors after primary radiation therapy: an international multi-center study. *Radiother Oncol.* 2005;77:18-24.
8. Choi W, Lee SW, Park SH, et al. Planning study for available dose of hypoxic tumor volume using fluorine-18-labeled fluoromisonidazole positron emission tomography for treatment of the head and neck cancer. *Radiother Oncol.* 2010; 97:176-182.
9. Hendrickson K, Phillips M, Smith W, Peterson L, Krohn K, Rajendran J. Hypoxia imaging with [F-18] FMISO-PET in head and neck cancer: potential for

- guiding intensity modulated radiation therapy in overcoming hypoxia-induced treatment resistance. *Radiother Oncol.* 2011;101:369–375.
10. Han D, Yu J, Yu Y, et al. Comparison of ¹⁸F-fluorothymidine and ¹⁸F-fluoro-deoxyglucose PET/CT in delineating gross tumor volume by optimal threshold in patients with squamous cell carcinoma of thoracic esophagus. *Int J Radiat Oncol Biol Phys.* 2010;76:1235–1241.
 11. Roels S, Slagmolen P, Nuyts J, et al. Biological image-guided radiotherapy in rectal cancer: is there a role for FMISO or FLT, next to FDG? *Acta Oncol.* 2008;47:1237–1248.
 12. Lee NY, Mechalakos JG, Nehmeh S, et al. Fluorine-18-labeled fluoromisonidazole positron emission and computed tomography-guided intensity-modulated radiotherapy for head and neck cancer: a feasibility study. *Int J Radiat Oncol Biol Phys.* 2008;70:2–13.
 13. Vera P, Bohn P, Edet-Sanson A, et al. Simultaneous positron emission tomography (PET) assessment of metabolism with F-fluoro-2-deoxy-d-glucose (FDG), proliferation with F-fluoro-thymidine (FLT), and hypoxia with fluoro-misonidazole (F-miso) before and during radiotherapy in patients with non-small-cell lung cancer (NSCLC): a pilot study. *Radiother Oncol.* 2011;98:109–116.
 14. Hicks RJ, Rischin D, Fisher R, et al. Utility of FMISO PET in advanced head and neck cancer treated with chemoradiation incorporating a hypoxia-targeting chemotherapy agent. *Eur J Nucl Med Mol Imaging.* 2005;32:1384–1391.
 15. Rischin D, Hicks RJ, Fisher R, et al. Trans-Tasman Radiation Oncology Group Study 98.02. Prognostic significance of [18F]-misonidazole positron emission tomography-detected tumor hypoxia in patients with advanced head and neck cancer randomly assigned to chemoradiation with or without tirapazamine: a sub-study of Trans-Tasman Radiation Oncology Group Study 98.02. *J Clin Oncol.* 2006;24:2098–2104.
 16. Cohen J. A coefficient of agreement for nominal scales. *Educ Psychol Meas.* 1960;20:37–46.
 17. Fleiss JL, Lewin B, Paik MC. *Statistical Methods for Rates and Proportions.* Hoboken, NJ: John Wiley and Sons; 2003.
 18. Byrt T, Bishop J, Carlin JB. Bias, prevalence and kappa. *J Clin Epidemiol.* 1993;46:423–429.
 19. Feinstein AR, Cicchetti DV. High agreement but low kappa: I. The problems of Two Paradoxes. *J Clin Epidemiol.* 1990;43:543–549.
 20. Palie O, Michel P, Ménard JF, et al. The predictive value of treatment response using FDG PET performed on day 21 of chemoradiotherapy in patients with oesophageal squamous cell carcinoma: a prospective, multicentre study (RTEP3). *Eur J Nucl Med Mol Imaging.* May 29, 2013 [Epub ahead of print].
 21. Caldwell CB, Mah K, Ung YC, et al. Observer variation in contouring gross tumor volume in patients with poorly defined non-small-cell lung tumors on CT: the impact of ¹⁸FDG-hybrid PET fusion. *Int J Radiat Oncol Biol Phys.* 2001;51:923–931.
 22. Steenbakkers RJ, Duppen JC, Fitton I, et al. Reduction of observer variation using matched CT-PET for lung cancer delineation: a three-dimensional analysis. *Int J Radiat Oncol Biol Phys.* 2006;64:435–448.
 23. Aerts HJ, Bussink J, Oyen WJ, et al. Identification of residual metabolic-active areas within NSCLC tumours using a pre-radiotherapy FDG-PET-CT scan: a prospective validation. *Lung Cancer.* 2012;75:73–76.
 24. MacManus M, Nestle U, Rosenzweig KE, et al. Use of PET and PET/CT for radiation therapy planning: IAEA expert report 2006–2007. *Radiother Oncol.* 2009;91:85–94.
 25. Vauclin S, Doyeux K, Hapdey S, Edet-Sanson A, Vera P, Gardin I. Development of a generic thresholding algorithm for the delineation of ¹⁸FDG-PET-positive tissue: application to the comparison of three thresholding models. *Phys Med Biol.* 2009;54:6901–6916.
 26. Hatt M, Cheze-le Rest C, van Baardwijk A, Lambin P, Pradier O, Visvikis D. Impact of tumor size and tracer uptake heterogeneity in ¹⁸F-FDG PET and CT non-small cell lung cancer tumor delineation. *J Nucl Med.* 2011;52:1690–1697.
 27. Devic S, Tomic N, Faria S, Menard S, Lisbona R, Lehnert S. Defining radiotherapy target volumes using ¹⁸F-fluoro-deoxy-glucose positron emission tomography/computed tomography: still a Pandora's box? *Int J Radiat Oncol Biol Phys.* 2010;78:1555–1562.
 28. Hatt M, Visvikis D. Defining radiotherapy target volumes using ¹⁸F-fluoro-deoxy-glucose positron emission tomography/computed tomography: still a Pandora's box?: in regard to Devic et al. (Int J Radiat Oncol Biol Phys 2010). *Int J Radiat Oncol Biol Phys.* 2010;78:1605.
 29. Daisne JF, Duprez T, Weynand B, et al. Tumor volume in pharyngolaryngeal squamous cell carcinoma: comparison at CT, MR imaging, and FDG PET and validation with surgical specimen. *Radiology.* 2004;233:93–100.
 30. Nehmeh SA, Erdi YE, Ling CC, et al. Effect of respiratory gating on quantifying PET images of lung cancer. *J Nucl Med.* 2002;43:876–881.
 31. Nehmeh SA, Erdi YE, Pan T, et al. Quantitation of respiratory motion during 4D-PET/CT acquisition. *Med Phys.* 2004;31:1333–1338.
 32. Carlin S, Humm JL. PET of hypoxia: current and future perspectives. *J Nucl Med.* 2012;53:1171–1174.
 33. Nehmeh SA, Lee NY, Schröder H, et al. Reproducibility of intratumor distribution of ¹⁸F-fluoromisonidazole in head and neck cancer. *Int J Radiat Oncol Biol Phys.* 2008;70:235–242.
 34. Grosu AL, Souvatzoglou M, Röper B, et al. Hypoxia imaging with FAZA-PET and theoretical considerations with regard to dose painting for individualization of radiotherapy in patients with head and neck cancer. *Int J Radiat Oncol Biol Phys.* 2007;69:541–551.
 35. Tian J, Yang X, Yu L, et al. A multicenter clinical trial on the diagnostic value of dual-tracer PET/CT in pulmonary lesions using 3'-deoxy-3'-¹⁸F-fluorothymidine and ¹⁸F-FDG. *J Nucl Med.* 2008;49:186–194.
 36. Xu B, Guan Z, Liu C, et al. Can multimodality imaging using ¹⁸F-FDG/¹⁸F-FLT PET/CT benefit the diagnosis and management of patients with pulmonary lesions? *Eur J Nucl Med Mol Imaging.* 2011;38:285–292.
 37. Gagel B, Piroth M, Pinkawa M, et al. pO polarography, contrast enhanced color duplex sonography (CDS), [18F] fluoromisonidazole and [18F] fluorodeoxyglucose positron emission tomography: validated methods for the evaluation of therapy-relevant tumor oxygenation or only bricks in the puzzle of tumor hypoxia? *BMC Cancer.* 2007;28:7:113.
 38. Hatt M, Cheze le Rest C, Turzo A, Roux C, Visvikis D. A fuzzy locally adaptive Bayesian segmentation approach for volume determination in PET [letter]. *IEEE Trans Med Imaging.* 2009;28:881–893.