

Variability and Uncertainty of ^{18}F -FDG PET Imaging Protocols for Assessing Inflammation in Atherosclerosis: Suggestions for Improvement

Pauline Huet^{1,2}, Samuel Burg³, Dominique Le Guludec³, Fabien Hyafil³, and Irène Buvat¹

¹U1023 Inserm/CEA/Paris Sud University–ERL 9218 CNRS, CEA-SHFJ, Orsay, France; ²IMNC UMR 8165 CNRS, Paris Sud University, Orsay, France; and ³Department of Nuclear Medicine, Bichat University Hospital, Assistance Publique—Hôpitaux de Paris, UMR 1148, Inserm and Paris Diderot-Paris 7 University, Département Hospitalo-Universitaire Fire, Paris, France

PET with ^{18}F -FDG shows promise for the evaluation of metabolic activities in atherosclerotic plaques. Although recommendations regarding the acquisition and measurement protocols to be used for ^{18}F -FDG PET imaging of atherosclerosis inflammation have been published, there is no consensus regarding the most appropriate protocols, and the image reconstruction approach has been especially overlooked. Given the small size of the targeted lesions, the reconstruction and measurement methods might strongly affect the results. We determined the differences in results due to the protocol variability and identified means of increasing the measurement reliability. **Methods:** An extensive literature search was performed to characterize the variability in atherosclerosis imaging and quantification protocols. Highly realistic simulations of atherosclerotic carotid lesions based on real patient data were designed to determine how the acquisition and processing protocol parameters affected the measured values. **Results:** In 49 articles, we identified 53 different acquisition protocols, 51 reconstruction protocols, and 46 quantification methods to characterize atherosclerotic lesions from ^{18}F -FDG PET images. The most important parameters affecting the measurement accuracy were the number of iterations used for reconstruction and the postfiltering applied to the reconstructed images, which could together make the measured standardized uptake values (SUVs) vary by a factor greater than 3. Image sampling, acquisition duration, and metrics used for the measurements also affected the results to a lesser extent (SUV varying by a factor of 1.3 at most). For an acceptable SUV variability, the lowest bias in SUV was observed using an 8-min acquisition per bed position; ordered-subset expectation maximization reconstruction with at least 120 maximum likelihood expectation maximization equivalent iterations, including a point spread function model using a 1 mm³ voxel size; and no postfiltering. Because of the partial-volume effect, measurement bias remained greater than 60%. The use and limitations of the target-to-blood activity ratio metrics are also presented and discussed. **Conclusion:** ^{18}F -FDG PET protocol harmonization is needed in atherosclerosis imaging. Optimized protocols can significantly reduce the measurement errors in wall activity estimates, but PET systems with higher spatial resolution and advanced partial-volume corrections will be required to accurately assess plaque inflammation from ^{18}F -FDG PET.

Key Words: atherosclerosis; ^{18}F -FDG PET; vascular lesions; partial-volume effect

J Nucl Med 2015; 56:552–559

DOI: 10.2967/jnumed.114.142596

Postmortem observations demonstrated that besides progressive stenosis, plaque rupture is the primary cause for myocardial infarction and stroke. Coronary thrombosis and occlusion occur in about two thirds of cases of sudden coronary death. Plaques prone to fissure have been defined as vulnerable plaques and identified by specific anatomic and biologic features. Accordingly, molecular imaging techniques such as PET/CT that visualize molecular targets using different ligands have been developed to identify these high-risk lesions. ^{18}F -FDG is the most evaluated tracer targeting the macrophage activity in the plaques and shows promise in this setting (1,2) but also in the evaluation of Takayasu arteritis (3) or aortic aneurysm (4). For these applications, reliable and reproducible quantification of the intensity of vascular inflammation would be a considerable asset for patient monitoring and assessment of therapeutic response. Yet, an accurate estimate of the tracer uptake in vascular lesions is extremely challenging given the small size of the lesions, compared with the spatial resolution of PET. In particular, uptake measurements in the vascular walls are strongly affected by partial-volume effect (PVE; Table 1) (5), which causes large activity underestimation in structures that are typically less than 3 times the spatial resolution in the reconstructed images. Assuming a constant uptake in a lesion, the bias in uptake measurements introduced by PVE depends on several parameters, including the volume of the lesion, its shape and contrast with respect to surrounding tissue, and the spatial resolution in the PET images. It also depends on how the uptake is locally measured (6,7). Because of the large number of parameters PVE depends on, it is extremely challenging to predict and compensate. An analysis of the literature shows that there is currently no consensus on the acquisition protocol and quantification procedures used to characterize vascular abnormalities from PET images. Yet, because of PVE, how images are acquired, reconstructed, and analyzed might actually strongly affect the reliability of the resulting measurements. This lack of consensus thus makes it difficult to objectively compare results reported in different studies and also prevents any meta-analysis of the literature. In that context, focusing on atherosclerotic

Received Oct. 30, 2014; revision accepted Jan. 13, 2015.

For correspondence or reprints contact: Irène Buvat, CEA-SHFJ, 4 Place du Général Leclerc, 91400 Orsay, France.

E-mail: irene.buvat@u-psud.fr

Published online Feb. 26, 2015.

COPYRIGHT © 2015 by the Society of Nuclear Medicine and Molecular Imaging, Inc.

TABLE 1
List of Abbreviations Used in Manuscript

Abbreviation	Meaning
BA	Blood activity
dX-vY protocol	Protocol involving an X-min duration acquisition with reconstructed voxel size of Y mm
dX-vY-f protocol	dX-vY protocol, followed by 3-dimensional gaussian postfiltering of 4 mm in FWHM in each direction
dX-vY-psf protocol	dX-vY protocol involving a gaussian kernel model of 4 mm in FWHM within the reconstruction
ENI	Effective number of iterations
I_{counts}	Index characterizing the number of counts in the raw data to be reconstructed
LBR	Lesion-to-blood activity ratio
OSEM	Ordered-subset expectation maximization
OWA	Observed wall activity
PSF	Point spread function
PVE	Partial-volume effect
RWA	Real wall activity
SUV	Standardized uptake value
SUV_{max}	Maximum SUV in the volume of interest
SUV_{mean}	Mean SUV in the volume of interest
$SUV_{\text{meanofmax}}$	Mean of the SUV_{max} measured in all transaxial slices encompassing the lesion
TBR	Target-to-blood activity ratio
TBR_{max}	TBR calculated using the maximum uptake in the target
TBR_{mean}	TBR calculated using the mean uptake in the target
$TBR_{\text{meanofmax}}$	TBR calculated by averaging the maximum uptake measured in all transaxial slices encompassing the target
VOI	Volume of interest

lesions only, the purpose of this study was 2-fold: first, to review the various image acquisition, reconstruction, and analysis protocols that are currently used to assess vascular lesions, and second, to determine the magnitude of the differences in results that can be due to these variations in protocols.

MATERIALS AND METHODS

We first performed an extensive literature search to characterize the variability in atherosclerosis imaging, reconstruction, and quantification protocols. We then performed highly realistic simulations of atherosclerotic lesions based on real patient data to determine the impact of various parameters involved in the acquisition and processing protocols on the values derived from the images.

Analysis of Literature

A literature search was performed using PubMed (<http://www.ncbi.nlm.nih.gov>) based on the regular expression “(fluorodeoxyglucose OR FDG) AND (athero*) & (Positron Emission Tomography OR PET),” covering all PubMed content until November 2013. A manual cross-referencing of articles cited in the papers resulting from that search was also performed to complete the review. Only studies involving patient PET images were included, with papers reporting preclinical data only discarded.

All articles resulting from this selection were analyzed to precisely identify the acquisition protocol, the reconstruction method and associated parameters, and the way quantitative analysis was performed. Table 2 summarizes all parameters that were systematically noted for each article.

On the basis of the injected activity, postinjection delay, and acquisition duration, an index I_{counts} characterizing the number of counts in the raw

data was systematically calculated. This index was defined as the mean injected activity per kilogram at acquisition time multiplied by the mean acquisition duration. The means of injected activity per kilogram and acquisition duration were used instead of the patient individual values because these values were never reported.

The reconstruction algorithm was noted, and when an iterative subsets was used, we recorded the effective number of iterations (ENIs), defined as the product between the number of subsets and the number of iterations. Resolution recovery based on the modeling of the point spread function (PSF) of the imaging system within the reconstruction procedure was noted when performed.

For quantitative analysis, the measured metric, the size of the volume of interest (VOI) in which the measurement was performed (single voxel value or region of interest), and the VOI delineation method were recorded. We also noted whether data were corrected for PVE. The mean and range (reported here as $[\text{max} - \text{min}]/2$) of the measured metric values observed over all patients were also recorded.

Assessment of Impact of Various Acquisition/Reconstruction/Analysis Procedures

To study the impact of variations in acquisition, reconstruction, and data analysis methods, we performed highly realistic simulations of atherosclerotic lesions based on real patient data. The advantage of using simulations is that unlike in real scans, the precise lesion features are fully known, and the bias and variability of the estimated quantity of interest can be objectively determined. We simulated 2 lesions with distinct features to illustrate how various parameters involved in the acquisition and processing protocols affect the values derived from the images in different situations.

TABLE 2
Parameters Systematically Recorded in All Articles as Part of Bibliographic Study

Image acquisition	Reconstruction	Quantitative analysis
Acquisition system	Voxel size*	Measured metric*
Injected activity*	Spatial resolution in reconstructed images*	VOI size
Postinjection time	Reconstruction algorithm	VOI delineation method
Acquisition duration*	Effective number of iterations*	PVE correction
Investigated arterial bed	Postfiltering*	Metric values
	Attenuation correction	
	Scatter correction	
	Resolution recovery*	

*We investigated influence of parameters noted with asterisk using simulated data (see "Assessment of Impact of Various Acquisition/Reconstruction/Analysis Procedures" section).

Simulations and Default Settings. A model of an atherosclerotic lesion was developed based on the XCAT numeric phantom (8). Because this phantom does not include any arterial wall (only the blood in the arteries and veins is modeled), an extended version of the XCAT was developed including walls with a thickness linearly related to the arterial diameter (9). An atherosclerotic lesion was modeled as a lipid core surrounded by an inflammatory region consisting of activated and inflammatory cells with enhanced ^{18}F -FDG uptake in the wall in which macrophages accumulate (10). In the numeric model, the lesion was defined as the intersection of the wall with a portion of an off-centered extruded cylinder with variable angular coverage and length (Fig. 1). Two lesions were considered. The first was a 10-mm-long, 60°-wide lesion (0.032-mL lesion) with an 8:1 lesion-to-blood activity ratio (LBR) corresponding to a true SUV of 14.1, in which we expected a strong PVE. The second lesion was a 30-mm-long, 300°-wide lesion (0.477-mL lesion) with an 8:1 LBR and true SUV of 14.1, in which PVE should be less pronounced because of the larger lesion size (Figs. 1A and 1B). Images were sampled with voxels of $1 \times 1 \times 1$ mm. To mimic a realistic PET scan, activity values in the surrounding organs were set as described in Table 3, assuming 4 MBq/kg injected 120 min before the PET acquisition. The values set in each compartment listed in Table 3 were derived from the mean values measured in the same compartments in 6 healthy subjects. An example of the resulting activity map is shown in Figure 1C. The elemental composition of the lesion and surrounding

tissues was taken from Woodard (11). On the basis of these activity and tissue composition distributions, PET scans were simulated using the GATE V6.1 Monte Carlo simulation software (12), without any variance reduction technique to properly reproduce the statistical properties of real data. They corresponded to 1 bed position (18-cm long) located on the neck acquired with a Gemini GXL PET/CT scanner (Philips) (13). The acquisition duration was set to 8 min. To assess the variability of the measurements, 10 replicates of each acquisition set-up were simulated for which only the initial seed of the Monte Carlo engine was changed.

The simulated data were reconstructed using an ordinary Poisson-weighted ordered-subset expectation maximization (OSEM) algorithm (14) with 10 subsets varying the number of OSEM iterations from 1 to 15 (i.e., 10 to 150 ENIs to reach a different trade-off between spatial resolution and noise). The reconstruction included attenuation and scatter corrections. The algorithm optionally included an image space model of the system PSF (15) that was turned off by default. The voxel size in the reconstructed images was $4 \times 4 \times 4$ mm (0.064 mL). No postfiltering was used by default.

Lesion Characterization. To characterize the lesions from the reconstructed PET images, VOIs corresponding to the true hypermetabolic atheromatous volume were used. Three metrics were systematically calculated: SUV_{max} , defined as the maximum voxel value in the VOI; SUV_{mean} , equal to the mean value in the VOI; and $\text{SUV}_{\text{meanofmax}}$, defined as the mean of the SUV_{max} measured in all transaxial slices encompassing the simulated lesion.

Impact of Changes Induced by Differences in Protocols. Using the simulated data, we studied how a change in the acquisition or data processing protocol affected the 3 metrics derived from the images.

To determine how the injected dose or acquisition duration affected the measurements, we compared the measured values obtained from the 8-min-duration PET scan with those obtained with 4-min PET scans obtained using half of the simulated events. This comparison was performed without any PSF modeling during the reconstruction and without postfiltering of the images reconstructed using a $4 \times 4 \times 4$ mm voxel size. All subsequent comparisons were performed using the data resulting from the 8-min-duration acquisition. Hereafter, we refer to these protocols as the d8-v4 and d4-v4 protocols.

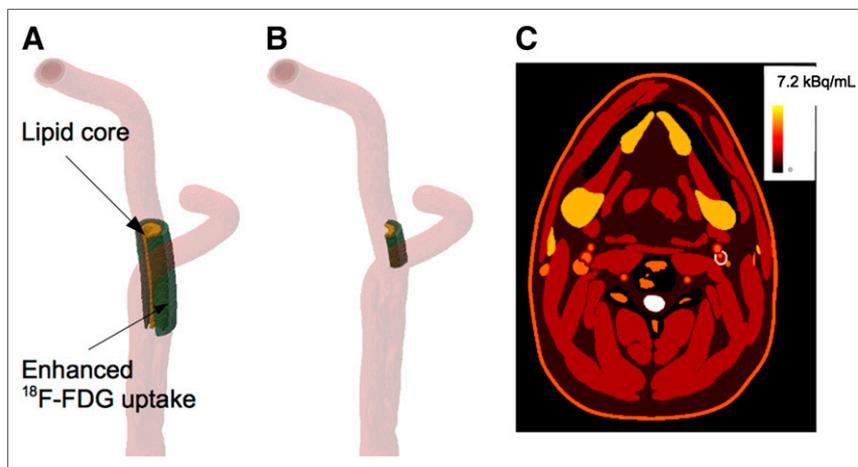


FIGURE 1. (A and B) Numeric model of 2 simulated lesions. (C) Activity map of modified XCAT phantom including atherosclerotic lesion.

TABLE 3
Activity Values Assigned to Different Compartments
of XCAT Phantom

Label	Activity (kBq/mL)
Blood	3.3
Healthy adventitia, intima, media	1.6
Inflammatory adventitia	LBR × 3.3
Atheroma	1.4
Esophagus	6.3
Salivary gland	4.0
Thyroid	3.7
Bone marrow	3.3
Parathyroid	2.5
Body	0.6
Muscle	1.6
Rib	6.3
Skin	3.2
Brain	19.7
Spinal cord	7.16
Laryngopharynx	0.60

The impact of the postreconstruction filter in the images reconstructed using a $4 \times 4 \times 4$ mm voxel size was investigated by comparing the metric values obtained without PSF modeling and without postfiltering (d8-v4 protocol) with those obtained without PSF modeling but with a 3-dimensional gaussian postfiltering of 4 mm in full width at half maximum (FWHM) in each direction (d8-v4-f protocol).

The change in metric measurements brought by an isotropic 4 mm in FWHM gaussian kernel PSF modeling within the reconstruction was assessed by comparing the metric values when no PSF was modeled and no postfiltering was applied with the values obtained with PSF modeling and without postfiltering. This comparison was performed for images reconstructed using $4 \times 4 \times 4$ mm voxels (d8-v4 and d8-v4-psf protocols) but also for images reconstructed using a voxel size of $1 \times 1 \times 1$ mm (d8-v1 and d8-v1-psf protocols).

The impact of the voxel size in the reconstructed images on the metric values was studied by comparing the measurements obtained for image sampling of $1 \times 1 \times 1$ mm (0.001 mL) and $4 \times 4 \times 4$ mm (0.064 mL) in the 2 lesions, with PSF modeling (d8-v4-psf and d8-v1-psf protocols) and without PSF modeling (d8-v4 and d8-v1 protocols) during the reconstruction and without postfiltering of the reconstructed images.

For all configurations, the 10 simulated replicates were reconstructed so that a sample of 10 values of the metric of interest (SUV_{max} , SUV_{mean} , or $SUV_{meanofmax}$) could be calculated. All results are reported as a curve of the mean error over the 10 replicates (expressed in percentage of the true $SUV = 14.1$) for which the percentage error is defined as:

$$\%Error = 100 \times (\text{estimatedSUV} - \text{trueSUV}) / \text{trueSUV} \quad \text{Eq. 1}$$

as a function of the SD of the SUV (in SUV units) over the 10 replicates, for which each point of the curve corresponds to a given number of iterations. We displayed only results corresponding to an acceptable level of noise in the reconstructed images—that is, an SUV SD equal to or less than 0.5 SUV units. This threshold has been chosen given that the measured SUVs are typically between 1 and 5, with 0.5 thus corresponding to between 10% and 50% of the measured SUV.

RESULTS

Analysis of Literature

The number of publications obtained using our PubMed search greatly increased over time, with 5 publications per year at most from 1997 to 2005, between 10 and 15 publications per year from 2006 to 2008, and more than 25 publications per year since 2009, to reach 50 publications per year in 2013. After disregarding articles reporting preclinical studies only and completing the list based on cross-referencing, we ended up with 49 articles (a list is provided in the supplemental materials available at <http://jnm.snmjournals.org>).

Acquisition Protocols. In the 49 articles, we listed 55 acquisition protocols (6 articles reported results obtained using 2 different protocols), among which 53 differed in the acquisition system or the injected dose or postinjection delay or acquisition duration. Forty-two of 55 (76%) protocols involved a hybrid PET/CT scanner sometimes in combination with MR data ($4/42 = 10\%$). Only 1 acquisition protocol involved a hybrid PET/MR scanner, 7 of 55 (13%) involved a PET scanner in combination with CT data, 2 of 55 (4%) involved a PET scanner plus MR data, and 3 of 55 (5%) involved PET data acquired on a standalone PET scanner combined with both CT and MR data. The investigated arterial beds were most often the carotid artery (in $48/55 = 87\%$ protocols) or the aorta ($24/55 = 44\%$), whereas the other protocols focused on the iliac ($9/55 = 16\%$), femoral ($8/55 = 15\%$), coronary ($3/55 = 5\%$), and subclavian ($1/55 = 2\%$) arteries.

The injected activity varied from 2.52 to 10.57 MBq/kg, and the time between injection and PET acquisition was between 45 and 190 min. The acquisition duration per bed position varied between 1 and 30 min. The mean I_{counts} was 35.21 ± 27.20 MBq/kg min. In 16 of 55 (29%) protocols, neither the injected dose nor the postinjection delay nor the acquisition duration was reported.

Reconstruction Procedures. We listed 51 different reconstruction procedures that varied in the reconstruction algorithm or associated corrections. In 9 of 51 (18%), image reconstruction was performed using filtered backprojection. OSEM-like algorithms were used in 23 of 51 (45%) and row-action maximum likelihood algorithm in 3 of 51 (6%), and 17 of 51 (33%) procedures did not mention the reconstruction algorithm used to produce the images. When using OSEM, 62 ± 30 ENIs. They were used (range, 24–126). ENI was not reported in 11 of 23 (48%) reconstructions involving OSEM.

Attenuation correction was explicitly mentioned in 41 of 51 (80%) reconstruction procedures and scatter correction in 12 of 51 (24%), and 2 articles (16,17) explicitly mentioned PSF modeling within the reconstruction algorithm. Ten of 51 (20%) procedures did not provide any information regarding the corrections that were used.

The spatial resolution in the reconstructed images was given for only 4 of 51 procedures (8%), whereas for 7 of 51 (15%), only the National Electrical Manufacturers Association spatial resolution was mentioned. No information regarding spatial resolution was provided in the other 41 of 51 (80%) procedures.

The voxel size in the reconstructed images varied from $0.7 \times 0.7 \times 2.03$ mm (0.001 mL) (16) to $4.3 \times 4.3 \times 4.25$ mm (0.079 mL) (18) and was not given for 36 of 51 (71%) procedures. A single article (16) explicitly mentioned a postreconstruction 2-mm filtering step.

Quantification Methods. Regarding image analysis, the 49 articles reported 97 quantification protocols, among which 36 differed either by the metric used or by the method used to calculate that metric. Most protocols (93/97) derived a metric from the PET images, and only 4 of 97 used visual grading (19–22). Regarding the quantitative metric, uptake was characterized using

SUVs in 37 of 94 (39%) protocols and using target-to-blood activity ratios (TBRs) in 52 of 94 (55%) cases. Other metrics were the ratio of the lesion uptake to the normal wall activity concentration (23) and the ratio of the lesion uptake to the lung uptake (22). Two articles used ^{18}F -FDG PET dynamic imaging to normalize the activity measured in the lesion by the integral of the input function (24) or to calculate the Ki influx constant using a Patlak analysis (25). To assess the overall severity of the disease by accounting for several lesions in a segment or in multiple vessels (at least carotid, aorta, iliac, or femorals), the number of lesions was measured in 3 of 94 (3%) quantification protocols (26–28). Atherosclerotic burden defined as the sum of SUV or TBR over targeted lesions was also calculated in 2 articles (27,29).

In addition to differences in the metric of interest, even a given metric did not always obey the same definition. In particular, SUVs were normalized using either the body weight (17/37 = 46%), the lean body mass (4/37 = 11%) (30), or the body surface area (3/37 = 8%) (25). In 15 of 37 (41%) articles using SUV, the normalization approach was not mentioned. A single article (31) corrected the SUV for the blood glucose level. SUV and TBR were often calculated from mean (36/94 = 38%) or maximum values (30/94 = 32%) over a region of interest, leading to SUV_{max} and TBR_{max} or SUV_{mean} and TBR_{mean} metrics. Some protocols (26/94 = 28%) averaged the maximum values over all transaxial slices encompassing the lesion, corresponding to what we call $\text{SUV}_{\text{meanofmax}}$ or $\text{TBR}_{\text{meanofmax}}$ in our paper (32). SUV_{mean} was calculated in a VOI that was manually drawn, either from the PET only or based on PET and anatomic information derived from the associated CT (29) or MR (25). $\text{SUV}_{\text{meanofmax}}$ was calculated by extracting the list of N maximum values for each of the N slices encompassing the lesion and either calculating the mean of these N maximum values or calculating the mean of the M maximum values in M consecutive slices ($M < N$) chosen so as to maximize that mean. This latter metric characterized what was called the most-diseased segment (33). TBR metrics varied as a function of how the lesion uptake was measured and how the blood uptake was measured, either in jugular veins or in the vena cava. The sizes of the VOI used to calculate SUV_{mean} or the numerator of TBR-like metrics were precisely reported in only 6 of 94 (6%) protocols. Depending on

the study, the SUV_{max} corrected for body weight averaged over all patients included in the study varied between 1.76 and 2.87, whereas the averaged patient SUV_{mean} was between 1.24 and 1.87. Similarly, the TBR_{max} averaged over all patients in a study varied between 1.46 and 2.68, whereas TBR_{mean} varied between 1.19 and 1.64.

Only 1 article corrected the uptake measurement from PVE (25) using the geometric transfer matrix approach (34) based on the segmentation of an MR image associated with the PET scan. In this article, the authors used PVE-corrected SUV, TBR, and Ki to characterize the lesion.

Impact of Acquisition and Processing Parameters on Measured Values

Simulated images closely mimicked patient images, in terms of both lesion appearance and surrounding activity distribution (Fig. 2). The same simulated lesion produced highly different images as a function of the acquisition and reconstruction protocols (Fig. 3).

The spatial resolution in the reconstructed images estimated from a point source off-centered by 10 cm in the transaxial plane of the field of view and reconstructed with 150 ENIs, 1-mm voxel size, and PSF modeling was 4.7 mm in FWHM.

For the large lesion (Fig. 4), the mean error in SUV_{max} estimates (y-axes in the graphs of Fig. 4) varied between –91% and –65% of the true 14.1 SUV ($\text{SUV}_{\text{max}} = 1.2$ –5.0 instead of 14.1) as a function of protocol (Fig. 4A), between –93% and –81% for SUV_{mean} ($\text{SUV}_{\text{mean}} = 1.0$ –2.6) (Fig. 4B), and between –92% and –53% for $\text{SUV}_{\text{meanofmax}}$ (1.1–6.6) (Fig. 4C). For the smallest lesion (Fig. 5), these mean errors varied between –94% and –83% as a function of the protocol with SUV_{max} (0.8–2.4) (Fig. 5A), between –95% and –89% with SUV_{mean} (0.8–1.5) (Fig. 5B), and between –95% and –87% with $\text{SUV}_{\text{meanofmax}}$ (0.8–1.8) (Fig. 5C).

When looking at the variability of the estimated SUV over the 10 replicates, for the large lesion (x-axes of Fig. 4), the SD of SUV_{max} was less than 0.5 SUV units for all protocols except d8-v1, suggesting that the bias in measurement was quite stable over replicates (Fig. 4A). For SUV_{mean} , it was even more stable because the SD was always less than 0.1 SUV units (Fig. 4B), whereas for $\text{SUV}_{\text{meanofmax}}$, it was less than 0.3 SUV units except for d8-v1 (Fig. 4C). The SD of SUV_{max} in the small lesion (Fig. 5) was less than 0.5 SUV units for all protocols except d8-v1 (Fig. 5A), whereas for SUV_{mean} (Fig. 5B) and $\text{SUV}_{\text{meanofmax}}$ (Fig. 5C), it remained less than 0.4 SUV units except for the d8-v1 protocol.

Overall, the biases in plaque SUV estimates are large (>60% whatever the lesion and the measurement method), but they can be significantly reduced by increasing the number of iterations—that is, moving toward the top of each graph without excessively increasing the variability of the measurement (Figs. 4 and 5). Indeed, the SD of the estimated SUV (x-axis of the graphs) was always less than 0.5 SUV units except for d8-v1 when measuring SUV_{max} (only the point corresponding to 10 ENIs has been shown in Fig. 4A; the point corresponding to 20 ENIs would correspond to an SD of 1.8 SUV units and a mean error of –46% and the points corresponding to 150 ENIs would correspond to an SD of 3.4 SUV units and a mean error of +56%).

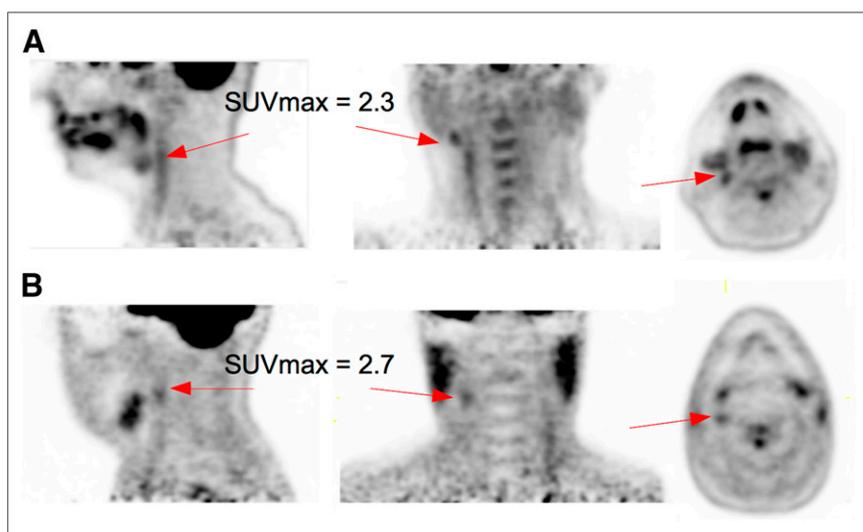


FIGURE 2. Patient (A) and simulated (B) images (OSEM algorithm, 1-mm voxel size, PSF modeling). Arrows point to lesions.

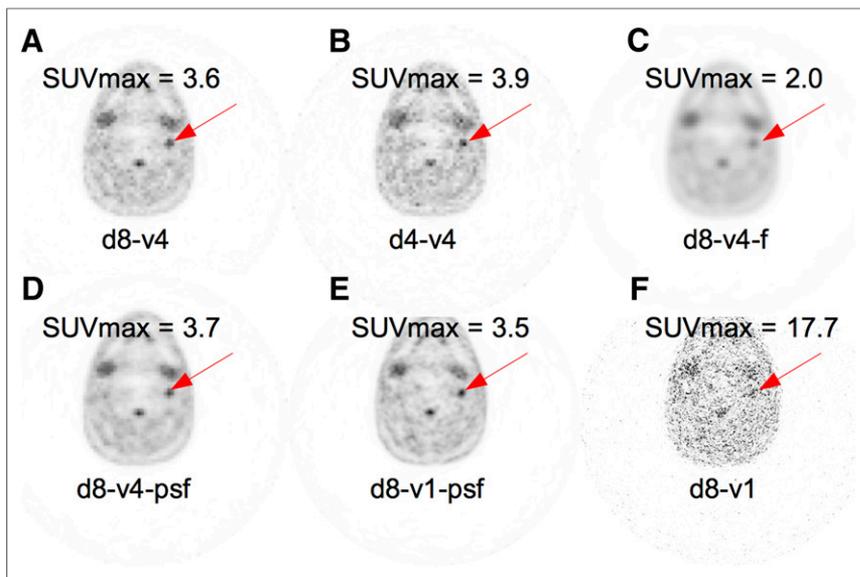


FIGURE 3. Sample images obtained for same lesion acquired and reconstructed using different protocols (50 ENIs). Images that look most pleasant visually due to low noise are not the ones in which SUV estimate is most accurate (true SUV = 14.1).

DISCUSSION

Our analysis of the literature showed the huge variability in the ^{18}F -FDG PET imaging protocols used to assess the inflammation associated with the plaques present in the arterial bed. The differences first concerned the acquisition protocol. This variability could be due to the fact that about one third of the studies (15/49) were retrospective studies using data that were initially acquired for cancer investigation and then revisited to investigate vascular inflammation. Yet, even when these 15 studies were removed, the variation in I_{counts} was still

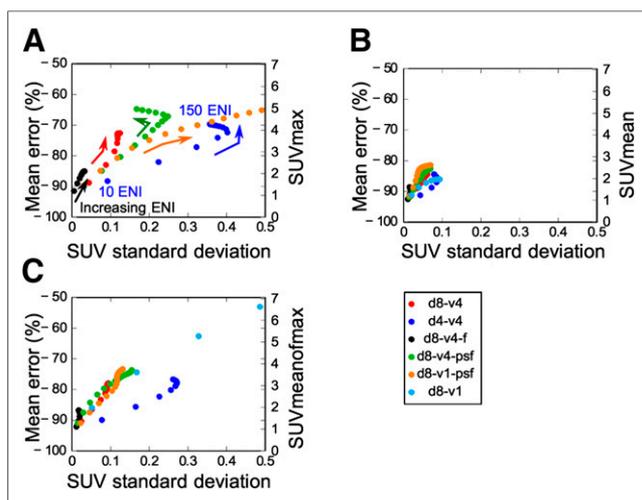


FIGURE 4. Large lesion. Mean error in SUV estimate (expressed as percentage of true SUV, scale on left of graph) as function of SD of that mean percentage error over 10 replicates. Results are shown for SUV_{max} (A), SUV_{mean} (B), and $\text{SUV}_{\text{meanofmax}}$ (C) and for 6 different protocols. dX = acquisition duration of X minutes; vY = voxel size = Y mm; f , with a 4-mm in FWHM postfiltering included; psf , with a 4-mm in FWHM PSF modeling included. Colored arrows indicate direction of increasing iterations (1 point every 10 ENIs), showing that increasing iteration number reduces average percentage error. Corresponding estimated SUVs are given by scale on right of each graph.

37 ± 24 MBq/kg min (against 35 ± 27 MBq/kg min for all 49 studies). Differences also related to the reconstruction protocols. Iterative reconstruction was used in half of the studies mentioning the reconstruction algorithm, but one third of the reports did not mention which algorithm was used to produce the images. Again, these differences might be due to the fact that part of the studies were initially designed for tumor imaging, so that images were not necessarily reconstructed using a protocol suitable for accurate quantification in sub-centimeter lesions. Key parameters such as the spatial resolution and the voxel size were not even mentioned in almost 70% of the articles. This lack of appropriate reporting prevented a sound comparison of the results given that spatial resolution and voxel size highly affect the severity of PVE and tissue fraction effect (5), which in turn strongly biases the measurements in small lesions such as plaques. Last, protocols highly varied in the metrics used to

assess the severity of the disease. This high variability reveals a lack of consensus regarding which parameter should be measured and how this should be done. Most studies (89%) measured either an SUV or a TBR but still differed in the way these metrics were estimated. The precise definition and size of the region used to make the measurement were reported in less than 10% of the articles. Only 6% of the studies used metrics that gave an overall assessment of the disease by combining measurements performed in different locations of single or different targeted segments.

These observations demonstrate the need for harmonization and systematic reporting of ^{18}F -FDG PET imaging procedures of vascular wall inflammation to facilitate the comparison of results between studies. Also, such a comparison should account for the fact that in retrospective studies involving oncology patients, there might be confounding radiation- and chemotherapy-related vascular inflammation and alterations in the metabolic milieu related to the underlying malignancy itself. Therefore, SUV arterial wall signals assessed in oncology subjects will remain difficult to compare with those measured in noncancer patients.

Using 2 simulated lesions, we observed that the bias in SUV estimates was large whatever the acquisition and reconstruction protocol. This large bias is due to the small thickness of the inflammatory part of the atherosclerotic lesions (typically ~ 1 mm) associated with a small longitudinal extent (typically 1–3 cm) with respect to the spatial resolution in the reconstructed images (~ 5 mm). In particular, the lesion thickness was well below 3 times the FWHM of the PSF characterizing the imaging system, making it impossible to properly recover the activity concentration without any sophisticated PVE correction.

Despite these high errors, we identified various approaches that minimized the biases. The parameter that most affected the error was the number of iterations used in reconstruction. The use of too-few iterations maximizes the bias, especially for SUV_{max} , whereas SUV_{mean} is the least sensitive to that number. The benefit of using an increased ENI was observed for all reconstruction schemes (Figs. 4 and 5), without excessive increase in the variability of SUV estimates. Using a high ENI is especially important when

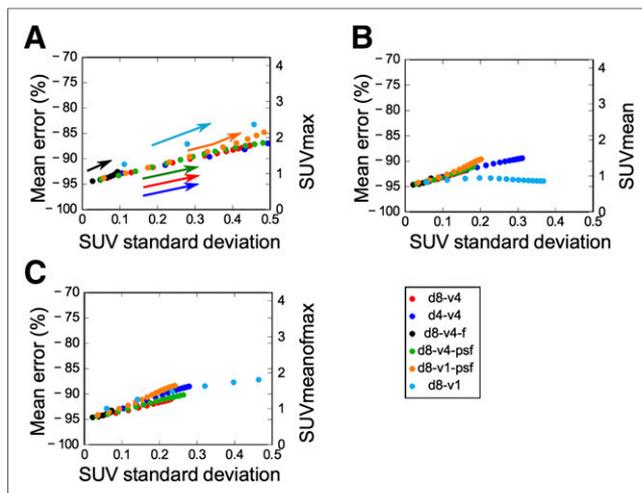


FIGURE 5. Small lesion. Mean error in SUV estimate (expressed as percentage of true SUV, scale on left of graph) as function of SD of that mean percentage error over 10 replicates. Results are shown for SUV_{max} (A), SUV_{mean} (B), and $SUV_{meanofmax}$ (C) and for 6 different protocols. dX = acquisition duration of X minutes; vY = voxel size = Y mm; f, with a 4-mm in FWHM postfiltering included; psf, with a 4-mm in FWHM PSF modeling included. Colored arrows indicate direction of increasing iterations (1 point every 10 ENIs), showing that increasing iteration number reduces average percentage error. Corresponding estimated SUVs are given by scale on right of each graph.

modeling the PSF in the reconstruction, because such a modeling reduces the speed of convergence. Yet, in most articles surveyed as part of our literature analysis, most reconstruction protocols used fewer than 60 ENIs. For the largest lesion reconstructed with a 1-mm voxel size and PSF modeling, varying the ENIs from 60 to 150 increased SUV_{max} from 3.8 to 5.1.

Postfiltering always had a detrimental effect on SUV recovery, as it further blurred images, hence amplifying PVE. Although using a low iteration number and postfiltering improved visual image quality (Fig. 3), this was at the expense of quantification accuracy, suggesting that the images that are used for visual interpretation should not be the same as the ones from which measurements are performed.

Using finely sampled reconstructed images associated with PSF modeling also systematically reduced the bias in uptake estimates, although the effect was not as high as that brought by increasing the number of iterations or removing postfiltering. When small voxels are used, noise regularization is absolutely needed to avoid excessive noise in the reconstructed images as shown by comparing Figures 3E and 3F. Such noise regularization can be introduced by modeling the PSF during the reconstruction process, as shown by comparing the orange and light blue curves in Figures 4 and 5. The impact of the voxel size was the highest for SUV_{max} and the lowest for SUV_{mean} . It was almost negligible for the large lesion.

The acquisition duration modified the measured values to a small extent: increasing the acquisition duration decreased the variability of the error for a given bias. This was especially true for SUV_{max} .

The metrics systematically affected the resulting error, with the smallest biases observed with SUV_{max} and the largest bias observed with SUV_{mean} .

Overall, for an SD in the measurement less than 0.5 SUV units, the lowest bias was always estimated using the protocol involving an 8-min acquisition and a reconstruction including a PSF model using a $1 \times 1 \times 1$ mm voxel size and no postfiltering, with at least 120

ENIs when using an OSEM approach. This was true whatever the metrics and the lesion. Yet, even with these optimized parameters, the mean error in SUV estimate remained high because of PVE. Several strategies can be considered to reduce these errors: improving the spatial resolution of the PET images based on hardware development and implementing explicit partial-volume correction (7). Also, in our study, we did not simulate motion, which introduces extra PVE, so motion should be compensated as well for increased accuracy (35). PVE correction is difficult to implement because the inflammatory part of the lesion cannot be seen in CT. PET/MR protocols might substantially ease the implementation of PVE by facilitating the delineation of plaque components (25). Improvements in these 3 directions will certainly enhance the value of plaque imaging using ^{18}F -FDG PET. Still, even before such new features become practical in the clinics, the systematic and precise reporting of the acquisition, processing, and measurement protocols would considerably facilitate the reviewing of the literature and the interpretation of the data and might help explain contradictory findings (31,36).

Given the high biases affecting the measurements, one can wonder how ^{18}F -FDG PET plaque imaging has been found useful to identify patients at high risk of plaque rupture (37). Although the bias is high, the variability of the bias is not so high, especially because the number of iterations is always rather low in published reports. Thus, for a given protocol, the magnitude of the error will be related to the lesion size and contrast. Even if the measurement is wrong, it bears information regarding the lesion features. The measured metrics describe a combination of activity concentration and lesion size (5,7). For a given lesion size, despite the bias, the measured value will reflect the activity concentration. For a given activity concentration, the measured value will be larger for larger lesions. When both the size and the activity concentration decrease, the measured metrics decrease as well. As the measured metrics reflect a combination of activity concentration and lesion size, they still convey useful information regarding the lesion.

In this article, we analyzed SUV only. The use of TBR has been justified by indicating that it was a blood-corrected uptake measurement (31,38). Yet, blood activity actually adds to wall activity due to the spatial blurring of the PET images (imperfect spatial resolution) and to the tissue fraction effect. Mathematically, we have:

$$OWA = a.RWA + b.BA, \quad \text{Eq. 2}$$

where OWA stands for observed wall activity, RWA stands for real wall activity, BA stands for blood activity, and a and b are 2 unknown weighting factors. The weight a is less than 1 because of the spill-out of wall activity in neighboring regions. Even if BA could be precisely measured, the resulting so-called TBR used in current articles would be:

$$TBR = OWA/BA = a.RWA/BA + b, \quad \text{Eq. 3}$$

which is an accurate estimate of RWA only if $b = 0$ and $a/BA = 1$, that is, $OWA = BA.RWA$, as if the blood contamination was a multiplicative process instead of an additive one. Therefore, there is no legitimate rationale for using TBR instead of SUV. In addition, the use of TBR increases the variability of the measurements because of the biologic and measurement variability of the blood uptake, as actually underlined or suggested in a few articles (37,39). The variance of TBR is indeed the sum of the variance in OWA and of the variance of the estimated BA, making TBR less reproducible than SUV.

Last, our study focused on ^{18}F -FDG as a marker of the inflammatory component of atherosclerosis. Other tracers targeting plaque features or some other components of an atherosclerotic lesion, such as the promising ^{18}F -fluoride (40), would also lead to focal uptake of spatial extent similar to ^{18}F -FDG, suggesting that the observations reported in this paper would have some relevance for those tracers.

CONCLUSION

Current ^{18}F -FDG PET protocols in atherosclerosis imaging present a huge variability, calling for some harmonization. Optimized protocols can significantly reduce the measurement errors in characterizing plaques although the remaining biases remain large. PET systems with higher spatial resolution and advanced partial-volume corrections will be required to accurately assess plaque inflammation from ^{18}F -FDG PET.

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. The work has been performed as part of the IMOVA/MEDICEN Paris Region project and is partly funded by the Conseil Général de l'Essonne, France. No other potential conflict of interest relevant to this article was reported.

ACKNOWLEDGMENTS

We deeply thank the reviewers for insightful suggestions.

REFERENCES

- Rudd JHF, Hyafil F, Fayad Z. Inflammation imaging in atherosclerosis. *Arterioscler Thromb Vasc Biol.* 2009;29:1009–1016.
- Orbay H, Hong H, Zhang Y, Cai W. Positron emission tomography imaging of atherosclerosis. *Theranostics.* 2013;3:894–902.
- Cheng Y, Ly N, Wang Z, Chen D, Dang A. ^{18}F -FDG-PET in assessing disease activity in Takayasu arteritis: a meta-analysis. *Clin Exp Rheumatol.* 2013;31:S22–S27.
- Hope MD, Hope TA. Functional and molecular imaging techniques in aortic aneurysm disease. *Curr Opin Cardiol.* 2013;28:609–618.
- Erlandsson K, Buvat I, Pretorius PH, Thomas B, Hutton BF. A review of partial volume correction techniques for emission tomography and their applications in neurology, cardiology and oncology. *Phys Med Biol.* 2012;57:R119–R159.
- Soret M, Bacharach SL, Buvat I. Partial-volume effect in PET tumor imaging. *J Nucl Med.* 2007;48:932–945.
- Burg S, Dupas A, Stute S, et al. Partial volume effect estimation and correction in the aortic vascular wall in PET imaging. *Phys Med Biol.* 2013;58:7527–7542.
- Segars WP, Sturgeon G, Mendonca S, Grimes J, Tsui BMW. 4D XCAT phantom for multimodality imaging research. *Med Phys.* 2010;37:4902–4915.
- Avolio AP. Multi-branched model of the human arterial system. *Med Biol Eng Comput.* 1980;18:709–718.
- Michel J-B, Thaumant O, Houard X, Meilhac O, Caligiuri G, Nicoletti A. Topological determinants and consequences of adventitial responses to arterial wall injury. *Arterioscler Thromb Vasc Biol.* 2007;27:1259–1268.
- Woodard HQ. The composition of body tissue. *Br J Radiol.* 1986;59:1209–1218.
- Jan S, Benoit D, Becheva E, et al. GATE V6: a major enhancement of the GATE simulation platform enabling modelling of CT and radiotherapy. *Phys Med Biol.* 2011;56:881–901.
- Stute S, Vauclin S, Necib H, et al. Realistic and efficient modeling of radiotracer heterogeneity in Monte Carlo simulations of PET images with tumors. *IEEE Trans Nucl Sci.* 2012;59:113–122.
- Michel C, Sibomani M, Boll A, et al. Preserving Poisson characteristics of PET data with weighted OSEM reconstruction. *IEEE Nucl Sci Symp Conf Rec.* 1998;2:1323–1329.
- Sureau FC, Reader AJ, Comtat C, et al. Impact of image-space resolution modeling for studies with the high-resolution research tomograph. *J Nucl Med.* 2008;49:1000–1008.
- Ripa RS, Knudsen A, Hag AMF, Lebeck A, Loft A, Sune H. Feasibility of simultaneous PET/MR of the carotid artery: first clinical experience and comparison to PET/CT. *Am J Nucl Med Mol Imaging.* 2013;3:361–371.
- Rogers IS, Nasir K, Figueroa AL, et al. Feasibility of FDG imaging of the coronary arteries: comparison between acute coronary syndrome and stable angina. *JACC Cardiovasc Imaging.* 2010;3:388–397.
- Menezes LJ, Kayani I, Ben-Haim S, Hutton B, Eli PJ, Groves AM. What is the natural history of ^{18}F -FDG uptake in arterial atheroma on PET/CT? Implications for imaging the vulnerable plaque. *Atherosclerosis.* 2010;211:136–140.
- Ben-Haim S, Kupzov E, Tamir A, Israel O. Evaluation of ^{18}F -FDG uptake and arterial wall calcifications using ^{18}F -FDG PET/CT. *J Nucl Med.* 2004;45:1816–1821.
- Ben-Haim S, Kupzov E, Tamir A, Frenkel A, Israel O. Changing patterns of abnormal vascular wall F-18 fluorodeoxyglucose uptake on follow-up PET/CT studies. *J Nucl Cardiol.* 2006;13:791–800.
- Tatsumi M, Cohade C, Nakamoto Y, Wahl RL. Fluorodeoxyglucose uptake in the aortic wall at PET/CT: possible findings for active atherosclerosis. *Radiology.* 2003;229:831–837.
- Paulmier B, Duet M, Khayat R, et al. Arterial wall uptake of fluorodeoxyglucose on PET imaging in stable cancer disease patients indicates higher risk for cardiovascular events. *J Nucl Cardiol.* 2008;15:209–217.
- Davies JR, Rudd JHF, Fryer TD, et al. Identification of culprit lesions after transient ischemic attack by combined ^{18}F fluorodeoxyglucose positron-emission tomography and high-resolution magnetic resonance imaging. *Stroke.* 2005;36:2642–2647.
- Rudd JHF, Warburton EA, Fryer TD, et al. Imaging atherosclerotic plaque inflammation with [^{18}F]-fluorodeoxyglucose positron emission tomography. *Circulation.* 2002;105:2708–2711.
- Izquierdo-Garcia D, Davies JR, Graves MJ, et al. Comparison of methods for magnetic resonance-guided [^{18}F]-fluorodeoxyglucose positron emission tomography in human carotid arteries: reproducibility, partial volume correction, and correlation between methods. *Stroke.* 2009;40:86–93.
- Tahara N, Kai H, Nakaura H, et al. The prevalence of inflammation in carotid atherosclerosis: analysis with fluorodeoxyglucose-positron emission tomography. *Eur Heart J.* 2007;28:2243–2248.
- Lee SJ, On YK, Lee EJ, Choi JY, Kim B-T, Lee K-H. Reversal of vascular ^{18}F -FDG uptake with plasma high-density lipoprotein elevation by atherogenic risk reduction. *J Nucl Med.* 2008;49:1277–1282.
- Bucerius J, Duivenvoorden R, Mani V, et al. Prevalence and risk factors of carotid vessel wall inflammation in coronary artery disease patients: FDG-PET and CT imaging study. *JACC Cardiovasc Imaging.* 2011;4:1195–1205.
- Bural GG, Torigian D, Chamroomrat W, et al. Quantitative assessment of the atherosclerotic burden of the aorta by combined FDG-PET and CT image analysis: a new concept. *Nucl Med Biol.* 2006;33:1037–1043.
- Tahara N, Kai H, Ishibashi M, et al. Simvastatin attenuates plaque inflammation: evaluation by fluorodeoxyglucose positron emission tomography. *J Am Coll Cardiol.* 2006;48:1825–1831.
- Bucerius J, Mani V, Moncrieff C, et al. Impact of noninsulin-dependent type 2 diabetes on carotid wall ^{18}F -fluorodeoxyglucose positron emission tomography uptake. *J Am Coll Cardiol.* 2012;59:2080–2088.
- Rudd JHF, Myers KS, Bansilal S, et al. Atherosclerosis inflammation imaging with ^{18}F -FDG PET: carotid, iliac, and femoral uptake reproducibility, quantification methods, and recommendations. *J Nucl Med.* 2008;49:871–878.
- Fayad ZA, Mani V, Woodward M, et al. Safety and efficacy of dalcetrapib on atherosclerotic disease using novel non-invasive multimodality imaging (dal-PLAQUE): a randomised clinical trial. *Lancet.* 2011;378:1547–1559.
- Rousset OG, Ma Y, Evans AC. Correction for partial volume effects in PET: principle and validation. *J Nucl Med.* 1998;39:904–911.
- Petibon Y, El Fakhri G, Nezafat R, Johnson N, Brady T, Ouyang J. Towards coronary plaque imaging using simultaneous PET-MR: a simulation study. *Phys Med Biol.* 2014;59:1203–1222.
- Calcagno C, Ramachandran S, Izquierdo-Garcia D, et al. The complementary roles of dynamic contrast-enhanced MRI and ^{18}F -fluorodeoxyglucose PET/CT for imaging of carotid atherosclerosis. *Eur J Nucl Med Mol Imaging.* 2013;40:1884–1893.
- Marnane M, Merwick A, Sheehan OC, et al. Carotid plaque inflammation on ^{18}F -fluorodeoxyglucose positron emission tomography predicts early stroke recurrence. *Ann Neurol.* 2012;71:709–718.
- Okane K, Ibaraki M, Toyoshima H, et al. ^{18}F -FDG accumulation in atherosclerosis: use of CT and MR co-registration of thoracic and carotid arteries. *Eur J Nucl Med Mol Imaging.* 2006;33:589–594.
- Menezes LJ, Kotze CW, Agu O, et al. Investigating vulnerable atheroma using combined ^{18}F -FDG PET/CT angiography of carotid plaque with immunohistochemical validation. *J Nucl Med.* 2011;52:1698–1703.
- Joshi NV, Vesey AT, Williams MC, et al. ^{18}F -fluoride positron emission tomography for identification of ruptured and high-risk coronary atherosclerotic plaques: a prospective clinical trial. *Lancet.* 2014;383:705–713.



The Journal of
NUCLEAR MEDICINE

Variability and Uncertainty of ^{18}F -FDG PET Imaging Protocols for Assessing Inflammation in Atherosclerosis: Suggestions for Improvement

Pauline Huet, Samuel Burg, Dominique Le Guludec, Fabien Hyafil and Irène Buvat

J Nucl Med. 2015;56:552-559.

Published online: February 26, 2015.

Doi: 10.2967/jnumed.114.142596

This article and updated information are available at:
<http://jnm.snmjournals.org/content/56/4/552>

Information about reproducing figures, tables, or other portions of this article can be found online at:
<http://jnm.snmjournals.org/site/misc/permission.xhtml>

Information about subscriptions to JNM can be found at:
<http://jnm.snmjournals.org/site/subscriptions/online.xhtml>

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

© Copyright 2015 SNMMI; all rights reserved.

The logo for the Society of Nuclear Medicine and Molecular Imaging (SNMMI) features the letters 'S', 'N', 'M', and 'I' in a white, sans-serif font, each contained within a red square. The squares are arranged in a 2x2 grid, with the 'S' and 'N' in the top row and the 'M' and 'I' in the bottom row.
SOCIETY OF
NUCLEAR MEDICINE
AND MOLECULAR IMAGING