

## Partial volume effect estimation and correction in the aortic vascular wall in PET imaging

This content has been downloaded from IOPscience. Please scroll down to see the full text.

2013 Phys. Med. Biol. 58 7527

(<http://iopscience.iop.org/0031-9155/58/21/7527>)

View [the table of contents for this issue](#), or go to the [journal homepage](#) for more

Download details:

IP Address: 132.166.112.252

This content was downloaded on 11/07/2014 at 15:23

Please note that [terms and conditions apply](#).

## Partial volume effect estimation and correction in the aortic vascular wall in PET imaging

S Burg<sup>1,7</sup>, A Dupas<sup>2,7</sup>, S Stute<sup>3</sup>, A Dieudonné<sup>4</sup>, P Huet<sup>5</sup>,  
D Le Guludec<sup>1,2,6</sup> and I Buvat<sup>5</sup>

<sup>1</sup> APHP-Service de médecine nucléaire, Hôpital Bichat-Claude-Bernard, F-75018 Paris, France

<sup>2</sup> INSERM U698, F-75018 Paris, France

<sup>3</sup> CEA-Service Hospitalier Frédéric Joliot, F-91400 Orsay, France

<sup>4</sup> APHP-Service de médecine nucléaire, Hôpital Beaujon, F-92110 Clichy, France

<sup>5</sup> IMNC, UMR 8165 CNRS, F-91400 Orsay, France

<sup>6</sup> Université Paris VII, F-75013 Paris, France

E-mail: [samuel\\_burg@yahoo.fr](mailto:samuel_burg@yahoo.fr) and [buvat@imnc.in2p3.fr](mailto:buvat@imnc.in2p3.fr)

Received 4 April 2013, in final form 27 August 2013

Published 8 October 2013

Online at [stacks.iop.org/PMB/58/7527](http://stacks.iop.org/PMB/58/7527)

### Abstract

We evaluated the impact of partial volume effect (PVE) in the assessment of arterial diseases with <sup>18</sup>FDG PET. An anthropomorphic digital phantom enabling the modeling of aorta related diseases like atherosclerosis and arteritis was used. Based on this phantom, we performed GATE Monte Carlo simulations to produce realistic PET images with a known organ segmentation and ground truth activity values. Images corresponding to 15 different activity-concentration ratios between the aortic wall and the blood and to 7 different wall thicknesses were generated. Using the PET images, we compared the theoretical wall-to-blood activity-concentration ratios (WBRs) with the measured WBRs obtained with five measurement methods: (1) measurement made by a physician (Expert), (2) automated measurement supposed to mimic the physician measurements (Max), (3) simple correction based on a recovery coefficient (Max-RC), (4) measurement based on an ideal VOI segmentation (Mean-VOI) and (5) measurement corrected for PVE using an ideal geometric transfer matrix (GTM) method. We found that Mean-VOI WBRs values were strongly affected by PVE. WBRs obtained by the physician measurement, by the Max method and by the Max-RC method were more accurate than WBRs obtained with the Mean-VOI approach. However Expert, Max and Max-RC WBRs strongly depended on the wall thickness. Only the GTM corrected WBRs did not depend on the wall thickness. Using the GTM method, we obtained more reproducible ratio values that could be compared across wall thickness. Yet, the feasibility of the implementation of a GTM-like method on real data remains to be studied.

<sup>7</sup> Both authors have contributed equally to this work.

## 1. Introduction

Inflammatory vascular pathologies (vasculitis) and atherosclerosis play a major role in cardiovascular events. Vasculitis most striking consequence is the acute obstruction of major arteries, typically the central retinal artery in the Horton disease or other arteries as in Takayasu vasculitis leading to brain stroke, upper limb ischemia or myocardial ischemia. Atherosclerosis main consequence is the development of plaques including a lipidic core in the vascular wall of arteries. The rupture of such atherosclerotic plaques, which leads to the formation of thrombi, is known to be a major cause of morbidity and mortality in cardiovascular diseases (Lendon *et al* 1992). Forecasting the vasculitis evolution or the rupture of plaques is therefore important to assess the cardiovascular risk and propose appropriate treatments. Non-invasive imaging techniques proved to be extremely valuable to this diagnostic effort (Vallabhajosula and Fuster 1997). One of the leading functional imaging techniques in clinical setting is positron emission tomography (PET). Among the numerous PET radio-tracers,  $^{18}\text{F}$ FDG targets cellular metabolism. Although  $^{18}\text{F}$ FDG is currently mostly used in cancer detection (Hawkins and Hoh 1994),  $^{18}\text{F}$ FDG can also assess hypermetabolism associated with the inflammation of tissues. Inflammation is the basis of vasculitis and retroperitoneal fibrosis and plays a key role in the genesis of the plaque, its growth and rupture. Several studies have already shown the correlation between  $^{18}\text{F}$ FDG concentration in PET images and the risk of plaque rupture (Rudd *et al* 2002, Ogawa *et al* 2004) or the link between  $^{18}\text{F}$ FDG uptake and disease activity in vasculitis (Arnaud *et al* 2009). Clinicians usually characterize the uptake of the tracer as the ratio between the maximum of the wall activity concentration and the mean blood activity concentration (WBR: wall-to-blood ratio, also named TBR—target to background ratio) (Rudd *et al* 2008). The higher this WBR, the higher the aggressivity of vasculitis or the higher the risk of plaque rupture for the patient. However, the small size of arteries makes accurate and precise quantitation extremely challenging given the spatial resolution and the voxel size in the PET images. For instance, the aorta has an average diameter of 20 mm (lumen plus wall) with a vascular wall of about 2 mm thickness (for a healthy wall). Vasculitis increases the wall thickness of the arteries in a wide range of values from millimeters to centimeters. The presence of a plaque increases the thickness of the wall up to about 4 mm depending on the state and development of the disease. With a voxel sampling between 1 and 4 mm and a spatial resolution between 5 and 7 mm, accurate uptake measurement in the vascular wall is difficult, mainly due to partial volume effect (PVE) and also noise. Atherosclerosis plaques being a very localized phenomenon of a diffuse pathology, we first focused on PVE in vasculitis, a more segmental phenomenon.

PVE commonly refers to two distinct phenomena that alter the intensity values in images (Soret *et al* 2007, Erlandsson *et al* 2012). The first phenomenon is due to the finite spatial resolution of the imaging system and appears as a 3D blurring of the image. Spatial resolution depends on the physics of PET imaging (positron range, acollinearity, scatter), on the detector design and on the reconstruction method. The second effect is caused by image sampling. The reconstruction process creates a sampled image of the radio-tracer distribution over a voxel grid. As tissue contours do not match voxel borders, a voxel uptake value represents the mean uptake of the tissues present in the voxel. This phenomenon is also called the tissue fraction effect.

PVE strongly affects small anatomical structure such as the aortic wall: the uptake of the arterial wall spills over adjacent structures, like the lumen or other neighboring organs, while the uptake in these regions also spills over the arterial wall. In addition, with a spatial resolution twice as large as the average aortic wall width, and a voxel sampling of about the same size, delineating the border of the aortic wall on PET images is extremely challenging. In

current hybrid PET/CT systems, PET images are registered with higher resolution computed tomography (CT) images, which facilitates the delineation of the anatomical structures. But even with new cardio-respiratory gating systems, accurate spatio-temporal registration of small structures remains difficult.

In this paper, our goal was to characterize errors in WBR estimates in PET images as a function of the measurement approach, without and with PVE correction, focusing on the aorta. The aorta is indeed the largest artery in the human body, with the thickest wall, making it more appropriate for accurate quantification from PET images. Our study is based on simulated images, with known ratios between the aorta wall and blood activity concentration. Working with real patient images to achieve our goal was indeed impossible since the exact WBRs cannot be obtained.

Delso *et al* showed that lesions smaller than the voxel size could be seen when their activity was greater than  $\approx 250$  Bq with an activity ratio with respect to background over 50–100 (Bazañez-Borgert *et al* 2008, Delso *et al* 2011). The two major factors of signal loss were PVE ( $\approx 90\%$  signal loss in phantom acquisitions) and cardio-respiratory motions ( $\approx 75\text{--}90\%$  signal loss in GATE simulations with the NCAT phantom).

Recently, Reeps *et al* (Bazañez-Borgert *et al* 2008, Reeps *et al* 2013) built a customized aortic phantom based on radioactive wax to study abdominal aortic aneurysm. Yet, the aortic wall thickness ranged from 5 to 20 mm, which is not suitable to study PVE on a realistic aortic wall thickness (from 2 to 5 mm). Thus, to our knowledge, there is no physical phantom suitable to study WBRs in the aorta.

In recent years, numerical models have proven useful in medical imaging research. When combined with accurate modeling of the imaging system, they can yield highly realistic images for which the ground truth is known. In this work, we present how these tools can be used to produce a model of arterial wall pathologies and corresponding PET images for PVE estimation and for the assessment of PVE corrections.

The aim of our study was to assess the magnitude of PVE in aortic vascular wall in  $^{18}\text{F}$ -PET imaging, and to determine whether WBRs could be accurately measured assuming an ideal segmentation of the wall.

In section 2.1, we present our model of the aortic arterial wall in an anthropomorphic digital phantom. Section 2.2 describes the simulation and reconstruction mechanisms that produced realistic PET images from the numerical phantom. Section 2.3 presents four methods for estimating the WBR. Statistical methods are introduced in section 2.4, results are given in section 3 and discussed in section 4.

## 2. Material and methods

### 2.1. Numerical phantom of the aortic vascular wall

We used the NCAT anthropomorphic phantom (Segars and Tsui 2009, 2010), which is a non-uniform rational basis splines (NURBS)-based Cardiac-Torso phantom, where NURBS surfaces are used to describe the organ shapes. The main blood vessels are included in the NCAT phantom but small anatomical structures like vascular walls are missing. To use the NCAT phantom in our study, we therefore addressed two issues: properly sampling the 3D labeled image (each label representing an organ) to have an accurate yet compact representation of the volume of interest, and inserting the aortic arterial wall.

To obtain a fine sampling of the arterial wall (healthy aorta walls are about 2 mm thick), the phantom was sampled with 0.25 mm cubic voxels. To simulate realistic images, the volume of interest is a complete thoracoabdominal image, included in a box of  $600 \times 500 \times 300$  mm<sup>3</sup>.

The size of the complete image encoded using 16 bits integers as labels was about 11 GB. As the labeled image has large areas of uniform values, we compressed it using run length encoding (Golomb 1966) hence reducing its size to about 150 MB.

In the NCAT phantom, the maximum diameter of the aorta is 23 mm near the heart, and the diameter is lower in the descending aorta. Comparing this diameter to real patient aorta diameters (averaging around 20 mm), the NCAT aorta volume is more reflective of the lumen than of the aorta and associated wall. In the labeled image, all arteries have the same label. Since the only artery with a diameter greater than 10 mm is the aorta, we extracted the aorta using morphological operations and assigned a specific label to voxels within the aorta. To add a wall around the aorta, we dilated the lumen using a spherical structuring element with a radius equal to the desired thickness of the wall. The resulting combination of aortic lumen and wall had a diameter similar to those observed in real patients: the maximal lumen diameter was 23 mm at the origin of the ascending thoracic aorta, then 19 mm in the arch of aorta, 17 to 15 mm in the thoracic descending aorta and 15 to 11 mm in the abdominal aorta. To study the impact of the aortic wall thickness on the measured WBR, we created seven phantoms with a wall thickness varying from 2 to 5 mm by step of 0.5 mm. The final phantom contained 33 labels representing 33 anatomical structures and organs from superior mediastinum down to the aortic bifurcation. No cardiac nor respiratory motion were modeled.

## 2.2. Simulations

The  $^{18}\text{F}$ FDG uptakes assigned to each anatomical structure of our phantom were derived from measurements made in real patients in our department. In our clinical acquisition protocol, the patients are administered an activity between 3 and 4 MBq kg<sup>-1</sup> and the PET acquisition is performed 60 min post-injection. We studied images of three healthy patients (two males and one female, non-diabetic, non-alcoholic, BMI  $\approx$ 23–25) initially scheduled for cancer screening with  $^{18}\text{F}$ FDG PET. The average activity concentration in each organ was measured relative to the liver activity concentration. We set the total activity to simulate an injection of 300 MBq one hour before acquisition. The virtual patient weight was set to 75 kg and we simulated a PET acquisition with 3 min per bed position.

Monte Carlo simulations of the acquisitions were performed using the GATE tool (Jan *et al* 2011, 2004). No variance reduction technique was used. A precise model of the Philips Gemini GXL PET scanner with 4 mm  $\times$  6 mm  $\times$  30 mm GSO-crystals without time of flight was modeled (Stute *et al* 2012). Random events were subtracted from the detected events based on a delayed window. Scattered events were removed from the sinogram, mimicking a perfect scatter correction. The resulting coincidences were reconstructed using a 3D OSEM algorithm with ten subsets and three iterations, including attenuation correction (AC) and geometric normalization. A Gaussian filter (FWHM = 5 mm) was applied after each iteration to reduce noise propagation. The reconstruction of a bed position produced an image volume of 256  $\times$  256  $\times$  64 voxels with a voxel size of 2.25 mm  $\times$  2.25 mm  $\times$  2.8125 mm. Three bed positions were used, with an overlap of 45 mm, so that the resulting volume was 256  $\times$  256  $\times$  160 voxels.

For each wall thickness, images were simulated with 15 different wall to blood activity-concentration ratios: 1/1, 1.5/1, 2/1, 2.5/1, 3/1, 3.5/1, 4/1, 5/1, 6/1, 7/1, 8/1, 10/1, 12/1, 14/1, 16/1. Combined with the 7 wall thicknesses, we ended up with a data set of 105 image volumes.

## 2.3. WBR measurement methods

Four WBR measurement methods were studied. For each method, in each single simulated PET volume, the WBR was measured four times from four different segments of the aorta

that roughly corresponded to the ascending aorta (AscAo), the arch of aorta (ArcAo), the descending thoracic aorta (DesAo) and the abdominal aorta (AbdAo) as in Rudd *et al* (2002). The theoretical WBR in these four segments was the same in a given dataset but the four estimated WBRs were expected to be a little different because of varying local cross-contamination, noise, and possibly reconstruction artifacts.

**2.3.1. Manual measurement.** The first measurement was performed by an expert, taking the ratio of the maximum uptake in the wall to the mean blood activity concentration measured in a sphere located inside the right atrium, so that the sphere was not too close to the atrium wall. In real patients, venous blood from superior vena cava or inferior vena cava could also be measured, but these two veins are not modeled in the NCAT phantom. A region in the left ventricle is not a good choice for blood activity-concentration measurement as the myocardial uptake is high and introduces substantial spill-in. To avoid some major measurement errors caused by spill-in, physicians manually limit the measurement areas of the arterial wall using only reduced parts of the four aorta segments. For instance, to estimate the activity concentration of the ascending aorta section, physicians exclude parts near the heart because of the high spill-in resulting from the hypermetabolism of the cardiac muscle. As it was a long and repetitive procedure, the physician only measured WBR in 30 PET images (all 15 true WBR times 2 thicknesses, namely 2.5 mm and 4 mm) on all 4 segments of the aorta, leading to 120 WBR. In the following, we refer to the resulting values as the *Expert WBR* obtained by the *Expert method*.

**2.3.2. Mean values in ideal VOI.** As a second measurement method, we computed the average intensity value for each of the eight regions of the aorta (vascular wall and lumen for each of the four segments) using the labels in the input data of the simulations to identify the regions. These regions were thus ideal and would not be achievable in practice. For these measurements, the PET images were oversampled without interpolation to match the resolution of the phantom image (0.25 mm voxel sampling). The mean values in the aortic wall and in the blood of the aorta lumen yielded the estimated WBR. In the following, we refer to this method as the *Mean-VOI method* with the corresponding *Mean-VOI WBR*.

**2.3.3. Max method.** In this method, box regions were drawn around each of the four aorta segments, excluding hypermetabolic extra vascular regions surrounding these segments to avoid including high activity voxels not belonging to the aorta. The wall activity concentration was obtained using the maximum intensity in each of the four reduced segments of the aorta (wall plus lumen). The average blood activity concentration was obtained from the blood in the right atrium, using a sphere located inside the right atrium, so that the sphere was not too close to the atrium wall. This method is referred to as *the Max method* with the corresponding *Max WBR*.

**2.3.4. Max corrected with a recovery coefficient.** This method included a PVE correction based on the Max value and the vessel wall thickness estimated by the user. The NCAT exact wall thickness was used as the user input so that the vessel wall thickness estimate was not biased. We modeled the PVE in two dimensions using an annulus of 17 mm inner-diameter, a wall thickness varying from 2.0 to 5.0 mm (steps of 0.5 mm) and a Gaussian point spread function (PSF) of FWHM = 6.5 mm (see next section for details about the FWHM estimation). This allowed us to compute a lookup table of recovery coefficients (observed maximum in the

wall/real value in the wall) that we then applied on the Max values. In the following, we refer to this method as the *Max-RC method* with the corresponding *Max-RC WBR*.

**2.3.5. GTM corrected measurement.** Last, we assessed a measurement method involving PVE correction using the geometric transfer matrix (GTM) method. This method, initially developed to correct PVE in PET brain imaging (Rousset *et al* 1998), uses anatomical data (from a registered MRI in brain PET) to define different functional compartments. Each compartment is supposed to have a uniform uptake. The method first produces a theoretical blurred image of each compartment  $i$  by blurring the true compartment by the PSF in the reconstructed image. The fraction of the signal emitted by compartment  $i$  and detected in compartment  $j$  can then be determined from this blurred image, for each pair of compartments  $i$  and  $j$ . These fractions  $W_{ij}$  give the entries of the GTM. Based on the GTM model, the estimated uptake in a compartment is a linear combination of the true uptake in all compartments. The small set of linear equations (as many equations and unknown as the number of compartments) can be easily solved to estimate the true uptake in each compartment. In our application, we used the *Boost* (Karlsson 2005) software to solve the system. The GTM was computed from a 1 mm voxel sampling of the original labeled volume representing the organs. There were 41 compartments in the image, one for each organ (33), and two for each segment on the aorta (one for the wall, one for the blood). Those aorta regions were exactly the same as for the Mean-VOI method.

In the resulting  $41 \times 41$  GTM matrix, we set to zero values lower than  $10^{-6}$  to avoid instabilities, resulting in a matrix with 34% of non zero values. The PSF in the reconstructed images was estimated as a Gaussian with a FWHM of 6.5 mm in the three directions. The procedure used to assess the PSF FWHM consisted in averaging the FWHM value over 20 cm in the middle of the FOV using a digital 20 cm tilted cube, as described in Fujita *et al* (1992). This measurement method based on a GTM correction is referred to as *the GTM method* with the corresponding *GTM WBR*.

#### 2.4. Statistical analysis method

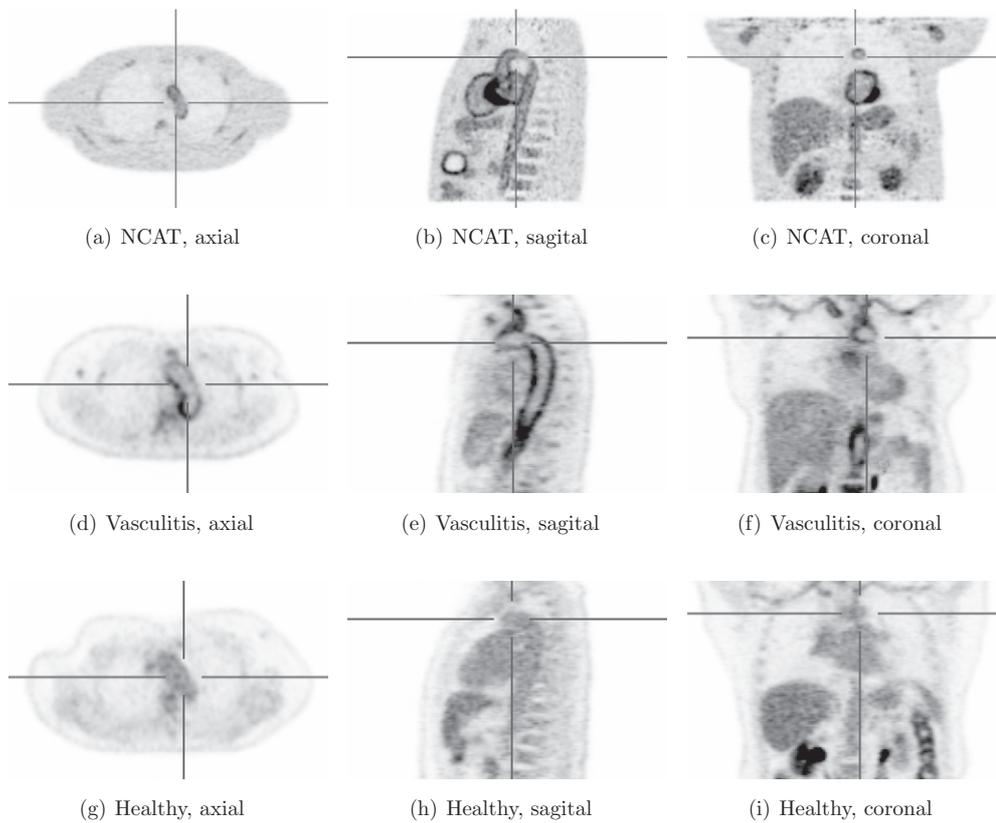
For all statistical tests, we used *The R Project for Statistical Computing* software (R Core Team 2012).

All statistical tests assumed that values were Gaussian distributed in each samples. We therefore systematically checked the validity of this assumption using the Shapiro–Wilk test.

The five measurement methods were tested against each other with analysis of variance (ANOVA) in a randomized block design (Kirk 1982).

Two way analysis of variance (2W-ANOVA) was also used to check whether there was a cross effect between the true WBR and the wall thickness for each measurement method. Then, WBRs obtained by each method were analyzed separately against the WBRs and wall thicknesses.

ANOVA tests were applied on log-scaled data, as usually performed to work on normally distributed data. Therefore, the resulting coefficients are not directly meaningful. We thus applied a linear regression on the not log-scaled data and reported  $\hat{\beta}$  as beta estimate (aka ‘the slope’) and  $R^2$  as the coefficient of determination. If the measured WBRs were linear with respect to the underlying true WBR, then  $\hat{\beta}$  would give an indication about how much the observable changes when the true WBR changes by 1 (although the intercept also has to be considered).



**Figure 1.** Reconstructed slices from the simulated NCAT phantom compared with clinical images. The NCAT phantom images shown here corresponded to an arterial wall thickness of 2.5 mm and a WBR of 4:1 between the aortic wall activity concentration and the blood activity concentration: (a) axial, (b) sagittal, (c) coronal slices. Real patient with a proved aortitis: (d) axial, (e) sagittal, (f) coronal slices. Real patient without any known vascular abnormalities: (g) axial, (h) sagittal, (i) coronal slices.

Post-hoc Tukey-Honestly-Significant-Difference (Tukey-HSD) tests followed each statistically significant ANOVA test. We also computed the mean Tukey-HSD threshold (MTT) to determine the mean minimal factor variation (wall thickness or true WBR) needed to observe a significant difference in resulting WBR.

For all tests, type I error cut-off was set to 0.05.

### 3. Results

#### 3.1. Simulated images

We simulated 105 PET volumes with activities in each organ as listed in table 1. Figure 1 shows the coronal, sagittal, and transverse simulated slices along with slices from real patient images. In these simulated images, the true WBR was set to 4 and the phantom had a 2.5 mm thick aorta wall.

Activities reconstructed in each section of the aorta were different despite the fact they were simulated with homogeneous values within the whole aorta. As an example, WBRs measured with the Mean-VOI method compared across aorta regions in a randomized block

**Table 1.** Simulated and reconstructed activities in the NCAT phantom for each organ. Activity concentrations were selected and adjusted to simulate a clinical  $^{18}\text{F}$ FDG-PET acquisition 1 h after intra-vascular injection of 300 MBq in a 75 kg patient. Reconstructed activity concentrations were measured for the perfectly delineated compartments as defined in the NCAT phantom.

NCAT organ	Simulated activity 1 h post-IV $\text{kBq ml}^{-1}$	Simulated activity normalized to the liver	Reconstructed activity normalized to the liver
Body, all what is not described below	2.3	0.33	0.37
Thorax			
Airway tree	0.0	0.00	0.23
Lung	1.4	0.20	0.24
Blood pool in heart	4.6	0.67	1.08
Pericardium	2.3	0.33	0.69
Heart, LV	36.9	5.33	4.36
Heart, LA	18.4	2.67	1.40
Heart, RV	18.4	2.67	1.48
Heart, RA	18.4	2.67	1.21
Vascular system			
Lymph	2.3	0.33	0.56
Blood, vein	4.6	0.67	0.67
Blood, artery	4.6	0.67	0.57–3.70
Artery wall	4.6–73.7	0.67–10.67	0.49–6.31
Abdomen			
Ascending large intestine, content	6.0	0.87	0.99
Transverse large intestine, content	6.0	0.87	0.65
Descending large intestine, content	4.6	0.67	0.73
Small intestine, content	6.0	0.87	0.98
Stomach, wall	7.4	1.07	0.99
Ascending large intestine, wall	10.4	1.50	1.06
Transcending large intestine, wall	8.8	1.27	1.43
Descending large intestine, wall	5.5	0.80	0.71
Small intestine, wall	8.3	1.20	0.93
Ureter	18.4	2.67	1.20
Spleen	6.9	1.00	0.93
Kidney	10.4	1.50	1.38
Gallbladder	1.4	0.20	0.48
Liver	6.9	1.00	1.00
Skeleton			
Spine head	0.0	0.00	0.38
Spine process	0.0	0.00	0.34
Bone marrow	7.4	1.07	0.72
Rib	0.0	0.00	0.46
Cartilage	0.0	0.00	0.26

design showed systematic higher signal in the ascending aorta, systematic lower signal in the arch of aorta and non-significant differences between the descending thoracic aorta and the abdominal aorta table 2.

### 3.2. Statistical comparison of the methods

We first tested the distribution of each WBRs sample, i.e. each subset of estimated WBRs corresponding either to a given true WBR or to a given wall thickness. The Shapiro–Wilk test suggested that our samples had a distribution not significantly different from a normal law, allowing us to proceed with the ANOVA.

**Table 2.** Mean-VOI WBRs comparison in the four aortic regions. The ANOVA test on the WBRs was processed in a randomized block design with the randomized factor being the four aorta regions and the blocking factor being the combination of the two factors true WBR and wall thickness (all levels for the two factors). The global ANOVA test was significant with  $p < 0.001$ . Second line of the table gives the average WBR Recovery coefficient (RR)  $\pm$  standard deviation of the method for each aortic region. Other entries give the difference between RR with the post-hoc Tukey-HSD sub-test results (NS: not significant; \*:  $p < 0.001$ ). Difference for each test is given as a percentage: RR in region from the first line minus RR in region from the left column divided by RR in region from the first line.

RR	AscAo	ArcAo	DesAo	AbdAo
	$0.34 \pm 0.22$	$0.26 \pm 0.16$	$0.29 \pm 0.21$	$0.30 \pm 0.22$
AscAo	0	-28% *	-16% *	-13% *
ArcAo	+22% *	0	+9% *	+12% *
DesAo	+14% *	-10% *	0	+3% NS
AbdAo	+11% *	-13% *	-3% NS	0

For wall thickness of 2.5 and 4.0 mm, the mean WBR Recovery coefficient (RR = ratio of measured WBR to true WBR) and associated standard deviation for each of the five methods was  $0.87 \pm 0.40$  for the Expert method,  $0.73 \pm 0.26$  for the Max method,  $1.67 \pm 0.62$  for the Max-RC method,  $0.29 \pm 0.21$  for the Mean-VOI method and  $0.58 \pm 0.14$  for the GTM method. ANOVA in a randomized block design suggested that the Max method reproduced the expert measurement well enough: Max WBRs were not significantly different from Expert WBRs ( $p = 0.235$ ). This same test showed significant differences between the WBRs of all other methods ( $p < 0.05$  for Max versus GTM comparison,  $p < 0.001$  for all other comparisons). This ANOVA test was performed in a randomized block design with the dependent variable being the measured WBRs. The randomized factor was the set of the five measurement methods. The blocking factor was the combination of true WBR, wall thickness and aortic region. We use all levels for the true WBR and aortic region factors, but only levels 2.5 and 4.0 mm of the wall thicknesses factor because only these two levels were available for the Expert method.

As Max WBRs were not significantly different from Expert WBRs, in the following, we therefore use the Max method as an expert reference to be compared to the Max-RC, Mean-VOI and GTM methods.

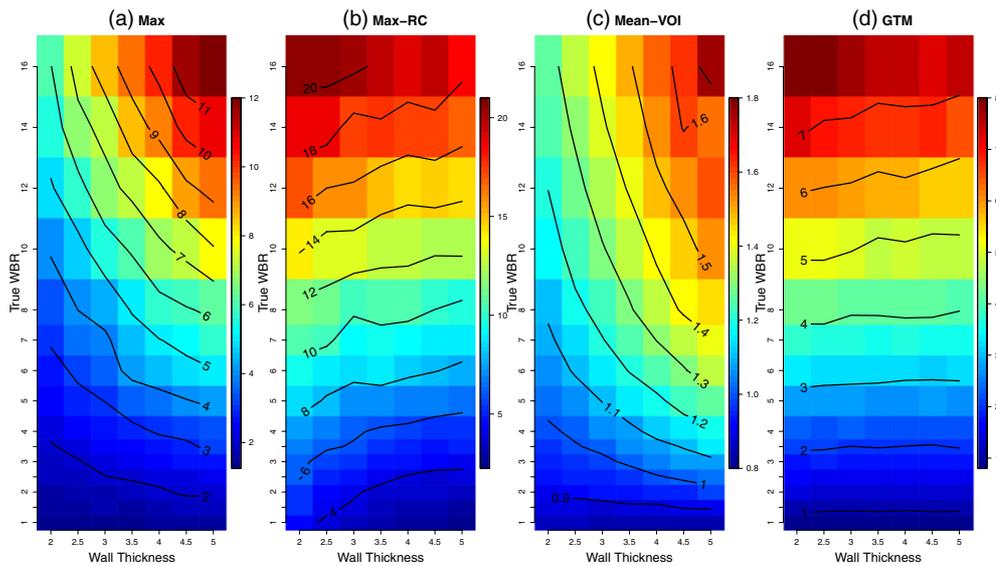
Considering all wall thicknesses, the RRs for the Max, Max-RC, Mean-VOI and GTM methods were respectively  $0.76 \pm 0.27$ ,  $1.65 \pm 0.63$ ,  $0.30 \pm 0.21$ ,  $0.58 \pm 0.14$ . Using ANOVA in a randomized block design as described in the previous paragraph, but including all seven wall thickness levels from 2.0 to 5.0 mm, we found that each method produced results significantly different from the other three ( $p < 0.001$ ).

A summary of our results is represented in table 3 while figure 2 represents the results as iso-level graphs. This figure shows that the same estimated WBR represented by a fixed color can be obtained for different combinations of wall thickness and true WBR. These graphs also allow for a qualitative comparison of the methods: the flatter the black lines, the less estimated WBR depends on wall thickness.

### 3.3. Performance of the methods

**3.3.1. Expert method.** WBRs measured by the expert for wall thickness of 2.5 mm and of 4 mm (for true WBRs from 1 to 16) are presented in figures 3(a) and 4(a).

WBRs estimated by the expert are correlated with the true WBR but they are biased. For a given wall thickness, the larger the Expert WBR, the larger the true ratio (2W-ANOVA:



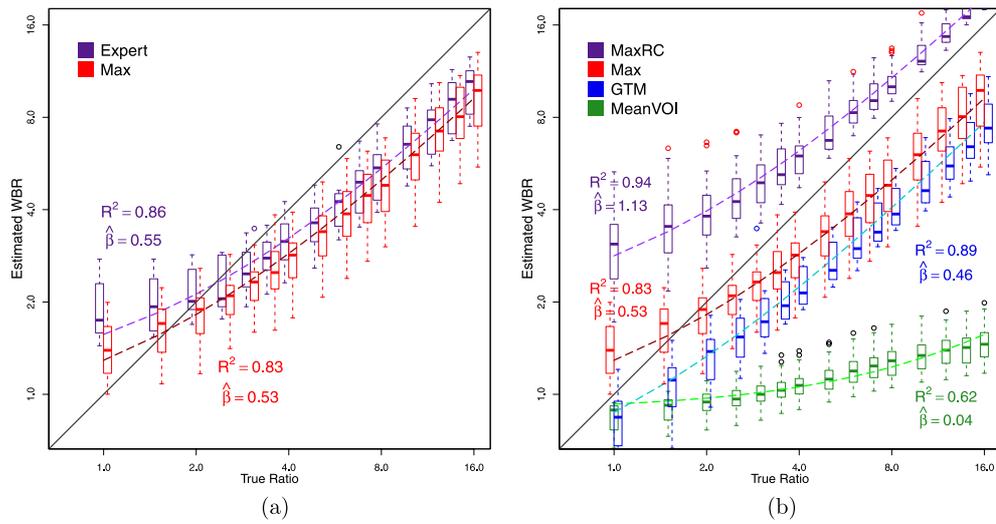
**Figure 2.** Measured WBRs against true WBRs shown as iso-level maps. (a) Max method. (b) Max-RC method. (c) Mean-VOI method. (d) GTM method. Horizontal axis: wall thickness. Vertical axis: true WBR. Each color level represents a measured WBR value. Black lines represent iso-level measured WBR values.

**Table 3.** Summary of the performance of the five methods. The table shows the mean RR  $\pm$  standard deviation of the five methods for each aortic region (columns two to five). The last column gives the average performance of the five methods.

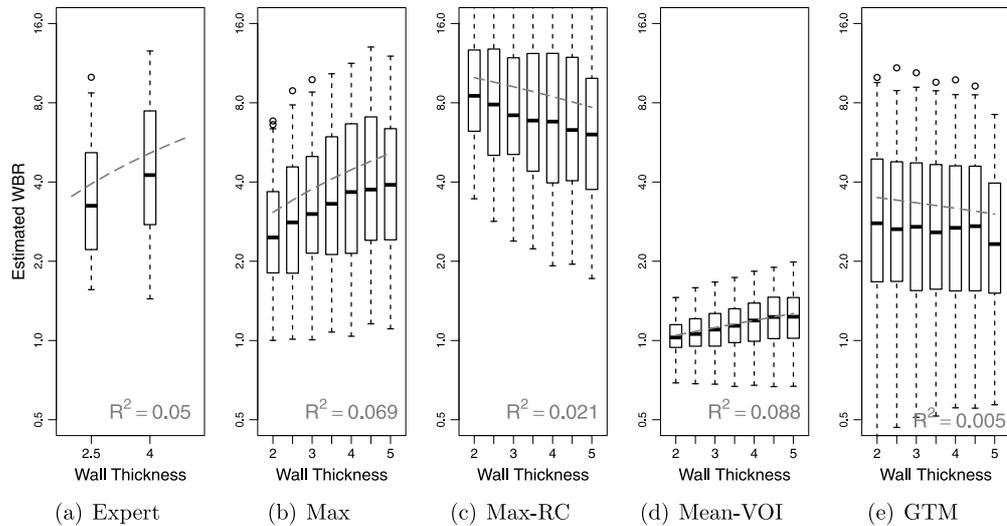
Method	AscAo	ArcAo	DesAo	AbdAo	Average
Expert	0.78 $\pm$ 0.34	0.75 $\pm$ 0.26	1.01 $\pm$ 0.57	0.95 $\pm$ 0.30	0.87 $\pm$ 0.16
Max	0.65 $\pm$ 0.17	0.70 $\pm$ 0.22	0.80 $\pm$ 0.32	0.88 $\pm$ 0.29	0.76 $\pm$ 0.27
Max-RC	1.41 $\pm$ 0.38	1.52 $\pm$ 0.52	1.76 $\pm$ 0.79	1.92 $\pm$ 0.65	1.65 $\pm$ 0.63
Mean-VOI	0.34 $\pm$ 0.22	0.26 $\pm$ 0.16	0.29 $\pm$ 0.21	0.30 $\pm$ 0.22	0.30 $\pm$ 0.21
GTM	0.70 $\pm$ 0.06	0.46 $\pm$ 0.05	0.52 $\pm$ 0.12	0.63 $\pm$ 0.17	0.58 $\pm$ 0.14

$p = 0.316$ ; estimated WBR versus true ratio:  $p < 0.001$ ;  $\hat{\beta} = 0.555$ ,  $R^2 = 0.864$ ). Yet, for the same true WBR but different wall thicknesses, the expert measured significantly different WBRs (estimated WBR versus wall thickness:  $p < 0.001$ ;  $R^2 = 0.050$ ) demonstrating that due to PVE, the expert measurement method is sensitive to wall thickness.

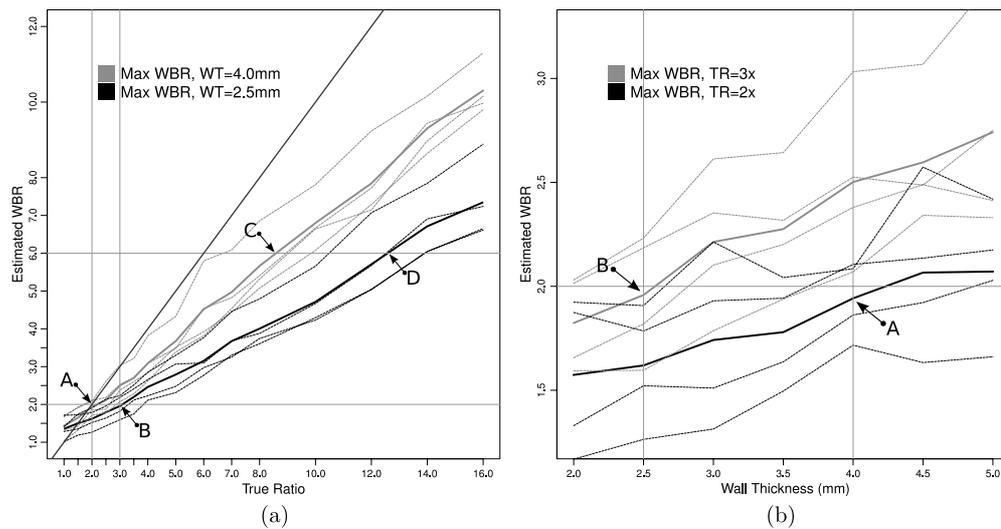
**3.3.2. Max method.** As an example, figure 5(a) shows some WBRs estimated with the Max method as a function of the true ratio. For two different wall thicknesses (2.5 and 4.0 mm), we plotted the average value for the four aorta regions. In figure 5(b), we plotted the Max WBR as a function of the wall thickness (average values for the four aorta regions for two different true ratios (2 and 3)). These graphs show that the Max method yields the same WBR value for different true WBRs, because the Max WBR not only depends on the true WBR but also depends on the wall thickness. For instance, in figures 5(a) and (b), we identified two points (A and B) where the Max WBR was about the same (1.96 and 1.94) but the real ratio differed by 50% (2 and 3). The same was observed for points C and D (measured WBR of 5.67 and 5.71 for actual WBRs of 8 and 12).



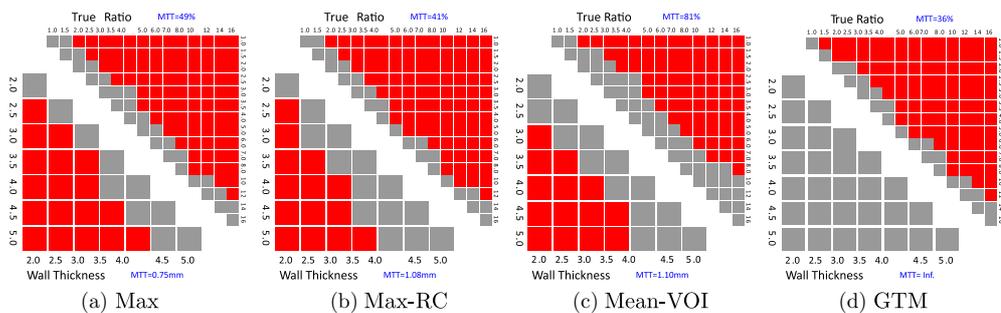
**Figure 3.** Estimated WBRs for the five measurement methods versus simulated WBRs. Boxes represent the repartition of the WBRs for all wall thicknesses and all aortic regions: central thick mark is the median, box extends from first to third quartile, whiskers corresponds to a 95% confidence interval on the median, small circles are outliers. Color dashed lines are results obtained when fitting a linear model to each data set, with  $\hat{\beta}$  the estimated 'slope' and  $R^2$  the correlation coefficient. (a) Expert and Max WBRs versus true ratio (b) Max, Max-RC, Mean-VOI and GTM WBRs versus true ratio.



**Figure 4.** Estimated WBRs for the five measurement methods versus simulated wall thickness. (a) Expert WBRs, (b) Max WBRs, (c) Max-RC WBRs, (d) Mean-VOI WBRs, (e) GTM WBRs. Boxes, color dash lines and legends are as explained in figure 3. Boxes represent the repartition of the WBRs for all true WBRs and aortic regions.



**Figure 5.** Samples of WBR measured with the Max method as a function of the true ratio and the wall thickness. Thick lines are average values of the single values shown as thin dashed lines. (a) Max WBR versus true ratio (b) Max WBR versus wall thickness.



**Figure 6.** Discrimination power of the measurement methods. Tukey-HSD tests determined whether a change in true WBR (top triangle) or in wall thickness (bottom triangle) leads a significantly different WBR for each WBR estimation methods. Each square represents the result of the Tukey-HSD test for a change, with red color for significant result ( $p < 0.05$ ) and gray color for non-significant result ( $p \geq 0.05$ ). Ideally, changes in true WBR should all be detected (top triangle including only red squares) while the estimated WBR should not change with wall thickness (bottom triangle including only gray squares). Mean Tukey-HSD threshold (MTT): mean minimal factor variation (wall thickness or true ratio) needed to observe a significant difference in WBR. (a) Max (b) Max-RC (c) Mean-VOI (d) GTM.

All the Max WBRs are shown in figures 2(a), 3(a) and (b) and 4(b) for wall thickness from 2 to 5 mm and for true WBRs from 1 to 16.

Estimates were close to those obtained by the expert. The measured Max WBRs strongly depended on the true WBRs (2W-ANOVA:  $p = 0.787$ ; estimated WBR versus true ratio:  $p < 0.001$ ;  $\hat{\beta} = 0.529$ ,  $R^2 = 0.835$ ) but also depended on the aorta wall thickness (estimated WBR versus wall thickness:  $p < 0.001$ ;  $R^2 = 0.069$ ). The larger the wall, the greater the measured ratio.

Post-hoc Tukey-HSD test results computed after the aforementioned 2W-ANOVA are shown in figure 6(a). This figure suggests that the Max method is not appropriate to compare

aortic wall uptake between patients when aortic wall thickness varied by more than 0.75 mm (MTT) and could detect statistically significant changes in WBR only for true WBR change  $\geq 49\%$  (MTT).

**3.3.3. Max-RC method.** The Max-RC WBRs are shown in figures 2(b), 3(b) and 4(c) for wall thickness from 2 to 5 mm and for true WBRs from 1 to 16.

The measured Max-RC WBRs strongly depended on the true WBRs (2W-ANOVA:  $p = 0.787$ ; estimated WBR versus true ratio:  $p < 0.001$ ;  $\hat{\beta} = 1.132$ ,  $R^2 = 0.939$ ) but also significantly depended on the aorta wall thickness (estimated WBR versus wall thickness:  $p < 0.001$ ;  $R^2 = 0.021$ ). The larger the wall, the smaller the measured WBR.

Post-hoc Tukey-HSD test results computed after the 2W-ANOVA are shown in figure 6(b). This figure suggests that the Max-RC method can detect statistically significant changes in WBR for true WBR change  $\geq 41\%$  (MTT) and is less sensitive to wall thickness than the Max method, but it still cannot compare WBR between patients when aortic wall thickness varies by more than 1.08 mm (MTT). As an example, figure 2(b) shows that a Max-RC of 6 (second black line from the bottom) can correspond to a true WBR from 2.5 to 4.5 depending on the wall thickness.

**3.3.4. Mean-VOI method.** Figures 2(c), 3(b) and 4(d) show WBRs obtained using the Mean-VOI method on images with wall thickness varying from 2 to 5 mm and with true WBR varying from 1 to 16.

The Mean-VOI method led to largely biased WBR estimates, with a mean WBR recovery coefficient of  $30\% \pm 21\%$ . The magnitude of WBR change between true WBR of 1.0 and true WBR of 16.0 was rather small, with an increase by a factor of 3 only (from 0.67 to 1.98) instead of 16. The measured WBR depended on the true WBR (2W-ANOVA:  $p = 0.999$ ; estimated WBR versus true WBR:  $p < 0.001$ ) but with a weaker correlation ( $\hat{\beta} = 0.043$ ,  $R^2 = 0.617$ ) than for the three previous methods. Because of PVE, wall thickness also affected the measured WBR (estimated WBR versus wall thickness:  $p < 0.001$ ) with a higher correlation coefficient ( $R^2 = 0.088$ ) than for the three previous methods.

Tukey-HSD test results are shown in figure 6(c). The Mean-VOI method does not enable an accurate nor precise comparison of aortic wall uptake when the aortic wall thickness varies by more than 1.10 mm (MTT) and could detect changes in WBR for true WBR variation  $\geq 81\%$  (MTT). This method is therefore less dependent on wall thickness than the Max method, but does not distinguish between different true WBRs as well as the Max or Max-RC method.

**3.3.5. GTM corrected method.** Figures 2(d), 3(b) and 4(e) summarizes all WBR values obtained with the GTM method for wall thickness ranging from 2 to 5 mm and for true ratio ranging from 1 to 16.

Estimates are similar to the ones obtained with the Max method and by the expert. The measured WBRs still depend on the theoretical ratio (2W-ANOVA:  $p = 0.999$ ; estimated WBR versus true ratio:  $p < 0.001$ ,  $\hat{\beta} = 0.456$ ,  $R^2 = 0.889$ ). In addition, the GTM WBRs do not correlate anymore with the wall thickness (estimated WBR versus wall thickness:  $p = 0.913$ ,  $R^2 = 0.005$ ) suggesting that a change in GTM WBR is more likely to describe a real change in true WBR, instead of just a change in wall thickness.

Tukey-HSD test results are shown in figure 6(d). MTT for wall thickness was not computed as the WBRs measured with this GTM method did no longer depend on the wall thickness. This GTM method is sensitive to the true WBR variation when the MTT is  $\geq 36\%$ .

#### 4. Discussion

The simulated images used in this study are consistent with images observed in patients with vascular pathologies. Our simulated WBRs between 8 to 16, leading to Expert WBRs and Max WBRs in the range of 4 to 8 (figure 3(a)), are consistent with the expert observations in the clinics for large vessel vasculitis. They also agree with other author results, e.g. the 1.69–3.09 activity-concentration ratio between arteries and liver reported in Arnaud *et al* (2009). Yet, atherosclerosis leads to less intense WBRs in clinical PET scans, with WBRs about 1.5–2.5 on active atherosclerotic plaques (Rudd *et al* 2002, 2008). The geometrical distribution of  $^{18}\text{F}$ FDG uptake also differs as a function of the pathology, with focal abnormalities, very focal inflammation, and thus very focal  $^{18}\text{F}$ FDG uptake in plaques compared to vasculitis.

Although each simulated configuration corresponded to a homogeneous activity concentration along the entire aortic wall and lumen, different WBR values were measured in the different segments with all measurement methods (table 2). Indeed, the ascending aorta is near the heart whose activity spills in the neighboring regions. Conversely, the arch of aorta is, in the digital phantom, completely surrounded by lung tissue with a very low uptake, yielding less spill-in. Also, aortic arch main axis is in the  $X$ – $Y$  plan, when the other three sections of the aorta are mainly oriented along the  $Z$  axis. Thus Mean-VOI WBRs are lower in the aortic arch but we cannot explain if this is mainly due to the vessel axis or to the surrounding tissues. The descending thoracic aorta and the abdominal aorta have a more homogeneous neighborhood that produces an average contamination yielding similar WBR estimates for the aortic arterial wall in both segments.

The reconstruction algorithm and the effect of the sampling of the AC map might also explain some differences in WBR estimates in different segments. We indeed observed some small artifacts near the interfaces between tissues with a highly different density, like the lung and the mediastinum or the digestive wall and the air inside the digestive track. These artifacts were actually also observed in real patient images, leading physicians to read both AC images and non-AC images depending on patients pathologies, typically for small lesions in the mediastinum (lung hilar lymph nodes metastasis) and on the skin (melanoma).

The Max, Expert, and Mean-VOI methods did not correctly estimate the true WBR in our experiments. The Mean-VOI method was the least reliable method: a difference of  $\geq 81\%$  was needed to detect some significant difference in the estimated WBR. Investigating this Mean-VOI method was useful to assess how much the signal was altered by PVE. The Max method detected smaller differences in WBR, but still yielded WBR values depending on the wall thickness (figure 6).

The Max-RC method largely decreased the WBR's correlation coefficient against the wall thickness ( $R^2 = 0.021$ ) compared to the Max and Mean-VOI methods but led to a large WBR over-estimation ( $RR = 165\% \pm 63\%$ ) because, among other reasons, it did not take the spill-in into account. We also assumed that the vessel wall thickness was not biased, which is a true challenge in practice. Last, it also suffers from increased variance.

GTM PVE correction improved the precision of the estimates compared to the Max method and the accuracy and precision of the estimates compared to the Mean-VOI method. Still, the estimated GTM WBRs were not accurate, because of noise and approximation in the GTM due to PSF estimation error (Sattarivand *et al* 2012). In our application, PVE is extremely severe and more severe than in brain applications for which the GTM method has been widely assessed and used so far (Rousset *et al* 1998, Frouin *et al* 2002). The regions in which we aim at accurately estimating the uptake have dimensions below the FWHM of system PSF, so it is expected that the correction will not be perfect. Still, the improvement in WBR estimates brought by the GTM approach was significant compared to no PVE correction. In

addition, for clinical applications, it is preferable to have a biased method with high precision (i.e. always the same bias), rather than an accurate method with high variability. A caveat of our study is that the GTM method as implemented in this study is unrealistic, as it assumed perfect segmentation of the wall and lumen and of all local and distant tissues. Local PVE corrections as described in Moore *et al* (2012) and Southeikal *et al* (2012) would avoid the need for careful segmentation of many organs. Yet, the main issue that still need to be addressed in a practical setting is how to properly segment the vascular wall: even if the wall and lumen segmentations could benefit from contrast-enhanced CT images, real data will likely be affected by respiratory and cardiac motions and segmentation accuracy will be limited. Still, it was important to characterize how well the GTM would perform for a perfect segmentation of the wall and lumen. These results set the limit of accuracy and precision that could be achieved using the GTM method, when assuming a stationary PSF in the reconstructed images (at least in the region where the measurements are performed). The stationarity of the PSF in the reconstructed images is also known to be an approximation, especially when the detector response is not modeled within the reconstruction process as in our study. To check the impact of our PSF FWHM value, we reanalyzed all data by varying the FWHM of the assumed PSF from 4.0 to 8.0 mm (by steps of 0.5 mm). There was a strong effect on the results, with clear outlier values for under- and over-estimation of the PSF FWHM by 1 mm. The best results were obtained with our initial 6.5 mm PSF FWHM.

Last, we assumed that scatter correction was both unbiased and noise free, to assess the quantification accuracy independently from other confounding factors. Again, this means that our results set the optimal results than could be expected using the quantification approaches we studied.

The next steps for understanding the reliability of quantitative characterization of vascular wall uptake will include assessing voxel-based methods to reduce partial volume correction (Teo *et al* 2007, Tohka and Reilhac 2008, Erlandsson *et al* 2012) and evaluating the robustness of region-based PVE corrections with respect to segmentation and registration errors.

## 5. Conclusion

In this work, we used realistic simulated PET images of a human torso to model vascular wall disease with known ground truth values and assessed the accuracy and precision with which the WBR could be estimated, as a function of the measurement method.

Measuring the WBR based on average activity concentrations in the exact wall and blood compartments yielded WBR estimates strongly affected by PVE. When using methods mimicking clinical practice (Expert method or Max method), WBR estimates were found to be highly dependent not only on the actual WBR but also on the wall thickness, making it difficult to assign differences in measured WBRs to changes in actual uptake or in the wall thickness. A practical correction method based on a lookup table of recovery coefficients without spill-in modeling (Max-RC method) led to an overcorrection and to WBRs that were still correlated with the wall thickness. When the GTM method was implemented in an ideal situation (no segmentation nor registration errors), it significantly improved the WBR estimates and had potential to better distinguish between changes in wall thickness and change in true WBRs, although the estimates remained biased due to the very low dimensions of the vascular wall.

## Acknowledgments

This work was performed as part of the DRCD-DIRC–St Louis–APHP interface project and was supported by the MEDICEN IMOVA project.

## References

- Arnaud L *et al* 2009 Is  $^{18}\text{F}$ -fluorodeoxyglucose positron emission tomography scanning a reliable way to assess disease activity in Takayasu arteritis? *Arthritis Rheum.* **60** 1193–200
- Bazañez-Borgert M, Bundschuh R A, Herz M, Martínez M J, Schwaiger M and Ziegler S I 2008 Radioactive spheres without inactive wall for lesion simulation in PET *Z. Med. Phys.* **18** 37–42
- Delso G, Martínez-Möller A, Bundschuh R A, Nekolla S G, Ziegler S I and Schwaiger M 2011 Preliminary study of the detectability of coronary plaque with PET *Phys. Med. Biol.* **56** 2145–60
- Erlandsson K, Buvat I, Pretorius P H, Thomas B A and Hutton B F 2012 A review of partial volume correction techniques for emission tomography and their applications in neurology, cardiology and oncology *Phys. Med. Biol.* **57** R119–59
- Frouin V, Comtat C, Reilhac A and Grégoire M C 2002 Correction of partial-volume effect for PET striatal imaging: fast implementation and study of robustness *J. Nucl. Med.* **43** 1715–26
- Fujita H, Tsai D Y, Itoh T, Doi K, Morishita J, Ueda K and Ohtsuka A 1992 A simple method for determining the modulation transfer function in digital radiography *IEEE Trans. Med. Imaging* **11** 34–39
- Golomb S W 1966 Run-length encodings *IEEE Trans. Inform. Theory* **12** 399–401
- Hawkins R A and Hoh C K 1994 PET FDG studies in oncology *Nucl. Med. Biol.* **21** 739–47
- Jan S *et al* 2004 GATE: a simulation toolkit for PET and SPECT *Phys. Med. Biol.* **49** 4543–61
- Jan S *et al* 2011 GATE V6: a major enhancement of the GATE simulation platform enabling modelling of CT and radiotherapy *Phys. Med. Biol.* **56** 881–901
- Karlsson B 2005 *Beyond the C++ Standard Library: an Introduction to Boost* (Reading, MA: Addison-Wesley) ([www.boost.org](http://www.boost.org))
- Kirk R E 1982 Experimental design *Handbook of Psychology* (New York: Wiley) ISBN 9780471264385
- Lendon C, Born G V, Davies M J and Richardson P D 1992 Plaque fissure: the link between atherosclerosis and thrombosis *Nouv. Rev. Fr. Hematol.* **34** 27–29
- Moore S C, Southekal S, Park M A, McQuaid S J, Kijewski M F and Muller S P 2012 Improved regional activity quantitation in nuclear medicine using a new approach to correct for tissue partial volume and spillover effects *IEEE Trans. Med. Imaging* **31** 405–16
- Ogawa M *et al* 2004  $^{18}\text{F}$ -FDG accumulation in atherosclerotic plaques: immunohistochemical and PET imaging study *J. Nucl. Med.* **45** 1245–50
- R Core Team 2012 *R: a Language and Environment for Statistical Computing* (Vienna: R Foundation for Statistical Computing) ISBN 3-900051-07-0 ([www.r-project.org](http://www.r-project.org))
- Reeps C, Bundschuh R A, Pellisek J, Herz M, van Marwick S, Schwaiger M, Eckstein H H, Nekolla S G and Essler M 2013 Quantitative assessment of glucose metabolism in the vessel wall of abdominal aortic aneurysms: correlation with histology and role of partial volume correction *Int. J. Cardiovasc. Imaging* **29** 505–12
- Rousset O G, Ma Y and Evans A C 1998 Correction for partial volume effects in PET: principle and validation *J. Nucl. Med.* **39** 904–11
- Rudd J H F *et al* 2002 Imaging atherosclerotic plaque inflammation with  $^{18}\text{F}$ -fluorodeoxyglucose positron emission tomography *Circulation* **105** 2708–11
- Rudd J H F *et al* 2008 Atherosclerosis inflammation imaging with  $^{18}\text{F}$ -FDG PET: carotid, iliac, and femoral uptake reproducibility, quantification methods, and recommendations *J. Nucl. Med.* **49** 871–8
- Sattarivand M, Kusano M, Poon I and Caldwell C 2012 Symmetric geometric transfer matrix partial volume correction for PET imaging: principle, validation and robustness *Phys. Med. Biol.* **57** 7101–16
- Segars W P and Tsui B M W 2009 MCAT to XCAT: the evolution of 4-D computerized phantoms for imaging research *Proc. IEEE* **97** 1954–68
- Segars W P and Tsui B M W 2010 *Handbook of Anatomical Models for Radiation Dosimetry* (London: Taylor and Francis) pp 105–33 chapter 5
- Soret M, Bacharach S L and Buvat I 2007 Partial-volume effect in PET tumor imaging *J. Nucl. Med.* **48** 932–45
- Southekal S, McQuaid S J, Kijewski M F and Moore S C 2012 Evaluation of a method for projection-based tissue-activity estimation within small volumes of interest *Phys. Med. Biol.* **57** 685–701
- Stute S, Vauclin S, Necib H, Grotus N, Tylski P, Rehfeld N S, Hapdey S and Buvat I 2012 Realistic and efficient modeling of radiotracer heterogeneity in Monte Carlo simulations of PET images with tumors *IEEE Trans. Nucl. Sci.* **59** 113–22
- Teo B K, Seo Y, Bacharach S L, Carrasquillo J A, Libutti S K, Shukla H, Hasegawa B H, Hawkins R A and Franc B L 2007 Partial-volume correction in PET: validation of an iterative postreconstruction method with phantom and patient data *J. Nucl. Med.* **48** 802–10
- Tohka J and Reilhac A 2008 Deconvolution-based partial volume correction in Raclopride-PET and Monte Carlo comparison to MR-based method *NeuroImage* **39** 1570–84
- Vallabhajosula S and Fuster V 1997 Atherosclerosis: imaging techniques and the evolving role of nuclear medicine *J. Nucl. Med.* **38** 1788–96