Realistic and Efficient Modeling of Radiotracer Heterogeneity in Monte Carlo Simulations of PET Images With Tumors

Simon Stute, Sébastien Vauclin, Hatem Necib, Nicolas Grotus, Perrine Tylski, Niklas S. Rehfeld, Sébastien Hapdey, and Irène Buvat

Abstract—**Monte Carlo simulations are extensively used in PET to evaluate the accuracy with which PET images can yield reliable estimates of parameters of interest. For such applications, the simulated images should be as realistic as possible so that conclusions can be extrapolated to clinical PET images. In this work, we describe a method for introducing realistic modeling of radiotracer heterogeneity into Monte Carlo simulations of patient PET scans. The modeling of the complex physiological activity distribution in healthy regions is directly based on real patient PET/CT images, and realistic tumor shapes can be included into these regions. This method represents a competitive alternative to the use of complex anthropomorphic phantoms such as the XCAT, that require a fixed activity per structure. The method is extended to the simulation of serial PET scans with tumor changes, as acquired in the context of therapy monitoring, and this extension is validated using a patient study. Using the proposed method, very realistic patient PET images can be produced for evaluation purposes. In addition, a strategy to efficiently simulate many sets of pathological cases, based on a unique background physiological activity distribution, is described and carefully assessed using a numerical phantom. The background activity is simulated only once, while tumors are simulated separately. The data are then recombined in a specific way so that the final image has the same properties as images produced by simulating pathological and tumor activities at the same time.**

Index Terms—**Monte Carlo (MC) simulation, patient monitoring, PET imaging, realistic simulation.**

I. INTRODUCTION

MONTE CARLO (MC) simulations of positron emission tomography (PET) and single photon emission computed tomography (SPECT) have been proven useful to evaluate corrections and quantification methods. For instance, they have been used for the assessment of tumor segmentation [1], partial volume effect correction [2], scatter correction [3], [4],

lesion detection [5], [6], and motion correction [7], [8]. An advantage of using MC simulated data compared to clinical data is the availability of the ground truth. The PET scanner response can be modeled very accurately in MC codes. Therefore, the realism of the simulated data now mostly depends on the model of activity distribution.

In the context of PET tumor imaging, tumors are most often simulated as spheres with uniform activity. Tumors have also been more recently modeled by deriving their shapes from real tumors observed in patients [1], [9], [10]. The activity distribution within the tumor is either modeled as uniform (with or without a necrotic region) [9], or as multiple isolevels of activity [10].

Tumors are located in more or less sophisticated numerical phantoms. The most sophisticated phantoms are based on real clinical data, such as the Zubal phantom [11], or the XCAT phantom [12]. The XCAT phantom makes it possible to model cardiac and respiratory motions and can be customized to describe patient specific features. A recent review of computational anthropomorphic anatomical and physiological models can be found in [13].

All these phantoms can be viewed as piecewise phantoms, in which different organs or structures can be assigned different types of tissues ("materials") and different activity concentrations. When using such phantoms, the activity concentration to be assigned to each region has to be determined. Le Maitre *et al.* in [10] modeled the activity distribution measured in a specific patient by locally measuring activity concentration in different regions of the patient PET scan. However, this procedure requires many manual operations and thus makes the production of a large database impractical. Tomei *et al.* in [6] measured many local activity concentrations in a large group of clinical PET scans. They deduced a mean activity concentration and an associated standard deviation per region and used them to produce activity distributions inserted in the Zubal phantom [11]. Using this approach, large databases can be produced but nothing ensures that the combination of activity values set in the different organs is realistic (for instance, it might not be realistic if there is some systematic correlation between the uptake in different organs). Finally, accurate modeling of heterogeneous physiological uptake (like a gradually changing activity distribution within the lungs) remains difficult with such phantoms. Even if the latest XCAT versions include the ability to model airways and blood vessels inside the lungs, it is extremely difficult to know which activity concentration should be set in such

TABLE I
LIST OF ACRONYMS

Acronym	Meaning
HP	"Healthy part" of the simulation
TP	"Tumor part" of the simulation
GS	"Gold-standard" method
PM	"Proposed merging" method
MO	"Merging only" method

All non-standard acronyms used in the text are listed in this table in the order of appearance.

small regions as they cannot be easily measured with current PET scans. In any case, using such activity distributions may lead to simulated images that can be distinguished from clinical images, and using them for assessing performance in detection or quantification tasks might yield misleading conclusions.

In SPECT, Pereira *et al.* [14] used patient data as the background physiological activity and then added simulated small tumors (spheres) in the projection data. The same group later used the same approach but simulating the background physiological activity using the NCAT phantom [15], [16]. However, they gave no detail on the methodology. They use the simulated data for assessing lesion detection.

In this paper, we propose a method to efficiently simulate PET images of patients with tumor(s), including realistic modeling of heterogeneous physiological uptake in non-tumor regions based on patient PET acquisitions. We also show how the method can be extended to simulate serial PET scans for assessing quantification methods used in patient monitoring.

II. METHODS

Non-standard acronyms used in this paper are listed in Table I.

A. Efficient Simulation of a Single PET Scan

In this work, we will focus on simulating realistic PET scans including tumors. We first introduce a general method that was used to accelerate the simulations of many patient studies. This method does not directly speed-up a specific simulation, but reduces the computational cost needed to perform a large series of simulations, where minor changes, e.g. insertion of a tumor, are made to the object. Considering the total physiological activity C in non-tumor regions and the activity C' inside tumors, C is much larger than C'. Given that the execution time of MC simulations linearly depends on the total simulated activity, the idea is to separate the simulation of C and C', and then to merge the results. The simulation of C (called the "healthy" part: HP) is performed only once while many different cases of C' (called the "tumor" part: TP) can be simulated and added to C. In this way, the required CPU execution time can be reduced.

1) Simulation Process: MC simulations require the definition of an attenuation map to model the propagation medium, an activity map to model the activity distribution, and a scanner model and associated acquisition parameters.

In this work, the GATE software (Geant4 Application for Emission Tomography) [17] was used, together with the fictitious cross-section approach [18] for increased computational efficiency. The Philips GEMINI GXL PET scanner was modeled. The coincidence sorter of GATE was used and coincidence events were stored in a list-mode format. The emission position was recorded for each non-random detected coincidence. A delayed coincidence window was also used so that the simulated coincidences could be corrected for randoms as performed in a real scanner. When a multiple coincidence occurred (i.e., more than two photons detected within the coincidence time window), only the coincidence involving the two photons with the highest detected energies was kept (the others were discarded).

In our approach, the attenuation map of the HP is defined first. The attenuation map of the TP is a modified HP attenuation map, in which the current linear attenuation coefficient μ is replaced by μ_{tumor} in the added tumor regions. The activity map of the HP is defined as desired. The activity map of the TP is derived based on the attenuation map of the TP, by setting activity concentration in non-tumor regions to 0, and in tumor regions to a_{tumor} (that remains uniform for a given tumor, but that can differ from one tumor to another). The simulations of the HP and the TP are performed separately.

2) Pre-Processing: All detected coincidences of the HP simulation whose emission position is within a tumor region of the TP are removed. Prompt sinograms P and P' and delayed sinograms D and D' are derived from the list-mode files of the HP and TP simulations respectively. Prompt coincidences are the sum of true unscattered, true scattered and random coincidences. Sinograms of scatter coincidences S and S' are then estimated for HP and TP from the sinograms equal to the delayed sinogram subtracted from the prompt sinogram. In this paper, the convolution-subtraction scatter correction method [19] was used, but other methods could be used instead. The sinograms A and A' of attenuation coefficient factors are calculated by forward-projection of the attenuation maps at 511 keV of the HP and the TP.

3) Reconstruction: With linear reconstruction algorithms such as FBP, the data coming from HP and TP could be reconstructed separately. The resulting images could then be directly added to yield an image that would be identical to that obtained from a simulation of the HP including tumors (apart from statistical variations). For non-linear iterative algorithms such as MLEM [20], the data from the HP and the TP have to be pre-processed and combined to proceed to a single reconstruction. Here, we used the OP-OSEM algorithm [21]:

$$\lambda_j^{k+1} = \frac{\lambda_j^k}{\sum_i \frac{M_{i,j}}{\alpha_i}} \sum_i \frac{y_i}{\sum_l M_{i,l}\lambda_l^k + \beta_i} \qquad (1)$$

where λ_j^k is the value of the reconstructed image voxel j at iteration k, $M_{i,j}$ is the system matrix element for voxel j and sinogram element i (including only the geometric component here), and y_i, α_i and β_i are respectively the measured data, the attenuation correction factor and the estimated accidental coincidences (corrected for attenuation) in sinogram element i. For the attenuation correction factors α_i, we chose to use the TP

ones (i.e. A') including the tumor tissues. The data were then combined using:

$$y_i = P_i \frac{A_i}{A'_i} + P'_i \qquad (2)$$

The prompt sinogram P from the HP simulation is corrected for the attenuation affecting HP (i.e. A) and then assigned the TP attenuation (i.e. A'). The resulting data y are thus affected by the attenuation factor A', so that A' can be used as the attenuation correction factors in α (1). Finally, the accidental term β was estimated using:

$$\beta_i = [A_i(D_i + S_i) + A'_i(D'_i + S'_i)]_+ \qquad (3)$$

where $[\ldots]_+$ means that a negative value in the bracket term was set to zero. Negative values might indeed be found in the scatter sinograms S and S' as they result from a convolution of sinograms $(P - D)$ and $(P' - D')$ that can contain negative values. No normalization was introduced in the reconstruction for two reasons. First, no crystal efficiency correction was needed since all crystals had exactly the same characteristics in our simulations. Second, no geometric normalization (arc correction) was needed as the direct geometric raw LORs were used instead of interpolated sinograms [22].

4) Summary: The HP simulation is done once. If one wants to simulate realistic PET scans, the HP can be derived directly from the PET image of a real patient without lesions. Such an image will mimic a real physiological activity distribution containing physiological heterogeneity of tracer uptake in tissues (although the physiological heterogeneity as seen in the patient PET image might not perfectly reflect the true heterogeneity, due to noise, partial volume effect, or motion for instance). Different sets of tumors can be added in the real patient and simulated separately. The simulated data are finally combined to produce different realistic images of a specific patient with various sets of tumors, in an efficient manner.

B. Simulation of Serial PET Scans

In this section, we introduce a method for including tumors in serial PET simulations. The aim is to produce simulated data in the context of therapy monitoring, with accurate modeling of the tumor changes and of the evolution of physiological nontumoral activity.

For such simulations, N successive scans of a "healthy" subject are needed (attenuation and activity maps), in association with predefined tumor contours. For each scan n of the "healthy" subject, a region R of the attenuation map covering the region of interest (ROI) is used (Fig. 1). To minimize the registration error resulting from the assumption of rigid transformation, the region R is split into S sets $R'_1, R'_2, \ldots, R'_S$ of less slices than region R with an overlap between adjacent sets (Fig. 1). For each set R'_s of the nth scan, a rigid transformation T^n_s consisting of a rotation matrix and a translation matrix describing the transformation between set R'_s of the nth scan and set R'_s of 1st scan is calculated by maximizing a correlation criterion [23] using an ROI inside set R'_s (Fig. 1). Registering any scan n with respect to the 1st scan avoids any propagation of registration error that could occur if scan n was registered
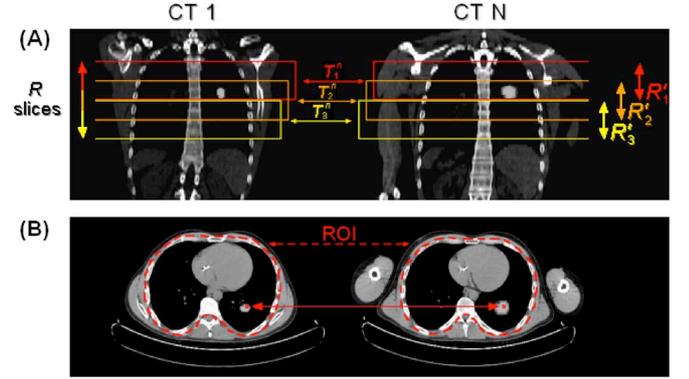
Fig. 1. Illustration of the method to simulate serial scans. Real CT data from a patient with no lesion in the lungs were used as the "healthy" subject to compute the registration parameters and to derive the attenuation maps. Images show coronal (A) and axial (B) views of the ROIs used for registration. In this example, $R = 40$ slices encompassed the lungs. $S = 3$ overlapping sets of 20 slices (R'_1, R'_2 and R'_3) covering these R slices were used to estimate 3 rigid transformations T^n_1, T^n_2 and T^n_3 (consisting in a translation and a rotation), one for each of the 3 regions between the 1st and the nth scans. The ROIs used to calculate the transformations in this example excluded the volume outside of the rib cage. The set of slices in the 1st scan in which the tumor is best centered axially is identified (R'_2 in the example). The tumor is then transferred from one scan to another by calculating the displacement of each tumor voxel using the corresponding registration parameters (here T^n_2).

with scan $n - 1$. Then an arbitrary tumor location is identified in the first scan of the "healthy" subject. First, the N tumor regions (corresponding to the same tumor in the N acquisitions) are "virtually" placed at this location in the first scan, from the biggest to the smallest. These placements obey two constraints: 1. a tumor region must overlap as much as possible all bigger tumor regions, 2. if several placements fulfill this first constraint, one of these placements is chosen at random. The "virtual" tumor region corresponding to acquisition n is then moved to the nth scan of the "healthy" subject by calculating the displacement of each voxel belonging to the region using the transformation matrix between the 1st and n^{th} scans (from T^n_s), and then linearly interpolating its new position in the voxels of this scan (Fig. 1). The same approach is repeated to locate any tumor in the $N - 1$ scans based on the first scan used for their "virtual" placement. Using this approach, any tumor region in the $n + 1^{th}$ scan is always realistically located with respect to the corresponding tumor region present in the nth scan. This tumor repositioning strategy is used to locate the tumors in the attenuation and activity maps, and thus simulate serial PET scans. The whole procedure is illustrated in Fig. 1.

III. ASSESSMENT OF THE PROPOSED METHODS

A. Simulation of a Single PET Scan

1) Numerical Phantom Study: Recombining HP and TP data simulated separately might yield different results compared to simulating directly a patient with tumors, due to possible different random coincidences rates, multiple coincidences rates, pile-up and dead-time effects (the latter was not taken into account in this work). When data are simulated in different attenuation media, the attenuation effect and the scatter distribution also differ compared to those observed when performing a single simulation. With this numerical phantom study, we assessed the impact of recombining the simulated data as

TABLE II
NUMERICAL PHANTOM PROPERTIES

Region	Shape	Radius along X (mm)	Radius along Y (mm)	Activity ratio with respect to R_1	Attenuation material	Attenuation coefficient μ (mm^{-1})
R_1	Elliptical	200	120	1	Soft Tissue	1.018×10^{-2}
R_2	Elliptical	150	100	0.6	Lung	2.496×10^{-3}
R_3	Sphere	10, 20, 40	---	0.6, 1, 2, 3, 4, 6, 8, 16	Soft Tissue	1.018×10^{-2}

Dimensions, activity ratios, attenuation materials and attenuation coefficient at 511 keV for the three regions of the numerical phantom.
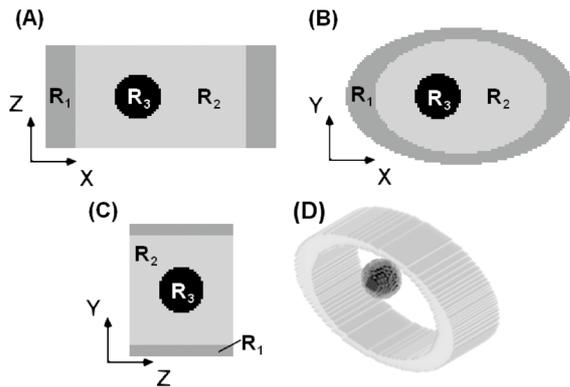
Fig. 2. (A) Coronal, (B) axial and (C) sagittal views of the numerical phantom including the sphere. The phantom consisted of 3 regions, R_1, R_2 and R_3. (D) 3D representation of the phantom.

described in Section II-A on the properties of the reconstructed images (as opposed to sinograms).

We designed a simple numerical phantom (Fig. 2). The phantom had an elliptical section and included a smaller compartment R_2 with an elliptical section too. These two compartments were invariant along the axial Z axis. A sphere with varying radius was introduced in compartment R_2 to mimic a lesion. The R_1 and R_2 volumes were centered in the field-of-view (FOV) but the sphere was 40 mm off-centered along the X axis.

The numerical phantom was used to define the attenuation and activity maps. For the HP simulation, the numerical phantom without the sphere was used, while for the TP, the phantom included the sphere. Twenty-four simulations were performed varying the sphere radius and the R_3/R_1 activity ratio. Dimensions of regions R_1 and R_2, R_2/R_1 activity ratio, and R_1, R_2 and R_3 attenuation coefficients were kept identical for all simulations. The dimensions and activity ratios used in the simulations are listed in Table II. The materials used for each region and their associated attenuation coefficients at 511 keV are also shown in this table and were obtained from a database of the National Institute of Standards and Technology [24].

A sampling of $124 \times 124 \times 45$ voxels of $4 \times 4 \times 4$ mm^3 was used for both attenuation and activity maps. The phantom could be scanned using a single bed position. In the HP simulation, activity in regions R_1 and R_2 only was simulated: the sphere corresponding to R_3 was removed and replaced by the "Lung" material and activity concentration of R_2. In the TP simulation, only activity in R_3 was simulated using the whole attenuation medium including the sphere. To mimic clinical conditions, the

activity in region R_1 in this study was set to 219 Bq per voxel (i.e. 3423 Bq/ml) and the acquisition time was set to 2 min. All prompt coincidences produced for each simulation were stored in a list-mode format, with a tag identifying them as unscattered trues, scattered trues or randoms.

To assess the accuracy of the proposed merging method (PM), the simulated data for each configuration (i.e. for each sphere radius and contrast) was compared to a gold-standard (GS). The GS consisted in simulating the whole phantom containing activity in R_1, R_2 and R_3, and for the same acquisition time (i.e. 2 min). The GS simulations were also compared to simulations following our methodology but in which the HP was simulated including "Soft Tissue" attenuation material in region R_3. These simulations are called "merging only" (MO). In this way the attenuation sinograms A and A' were the same and we therefore studied only the consequences of merging data that were simulated separately in the same medium. To thoroughly investigate the consequences of the merging operation, the reconstructions were first performed using only unscattered trues and then using prompts. The reconstructions were performed using 10 iterations and 10 subsets. No post-smoothing was applied to the reconstructed images. Several replicates ($k = 6$) were obtained for each simulation by running it k times, changing only the seed of the pseudo-random numbers generator. For all simulations, mean activity concentrations were measured within ROIs drawn inside R_1, R_2 and R_3. These measurements were repeated using all k replicates giving a mean and a standard deviation associated to the activity concentration in each region R_i and for each method: mA_{PMi} and σA_{PMi} for the proposed method, mA_{MOi} and σA_{MOi} for the "merging only" simulations, and mA_{GSi} and σA_{GSi} for the gold-standard. Mean relative errors $mE_{MO/GSi}$, $mE_{PM/MOi}$ and $mE_{PM/GSi}$ comparing the different methods were calculated for each region R_i using the following equation:

$$mE_{X/Yi} = \frac{(mA_{Xi} - mA_{Yi})}{mA_{Yi}} \times 100 \qquad (4)$$

Associated uncertainties $\sigma E_{MO/GSi}$, $\sigma E_{PM/MOi}$ and $\sigma E_{PM/GSi}$ on these errors were computed using (5):

$$\sigma E_{X/Yi} = \left[\frac{mA_{Xi}}{mA_{Yi}^2} \sigma A_{Yi} + \frac{1}{mA_{Yi}} \sigma A_{Xi} \right] \times 100 \qquad (5)$$

In (4) and (5), X and Y denote the methods that are compared.

2) Realistic Example: To simulate highly realistic PET scans of the thoracic region, data of a so-called "healthy" patient without any lesion in the thorax was obtained. The CT image of this patient was segmented into a finite number of
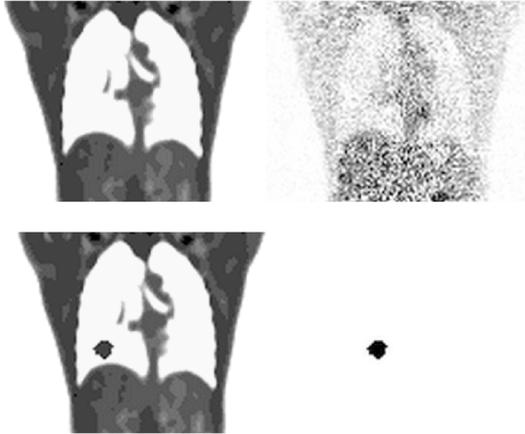
Fig. 3. (Upper row) Coronal slice of the attenuation (left) and activity maps (right) derived from the PET/CT acquisition of the "healthy" subject and used to simulate the HP. (Lower row) Coronal slice of the attenuation (left) and activity (right) maps used to simulate the TP.

different attenuation coefficient ranges. Each of these ranges was associated with a material defined in GATE. The first range, corresponding to the lowest CT numbers, was associated with material "Air". The second range was associated with material "Lung". Then 20 ranges were associated with materials having the same chemical composition (that of "Soft Tissue" material) but varying in density (in a linear way) between the lung density and the soft-tissues density. A last range was defined to represent bones (with an attenuation coefficient of 1.78×10^{-2} mm^{-1} at 511 keV). In this way the attenuation map of the HP, based on material definitions, was derived. The reconstructed PET image, expressed in Standardized Uptake Value (SUV) units, was directly used to derive the activity map of the HP, by assigning an amount of activity for each voxel proportional to the voxel value (assuming an injection of 5 MBq/kg). By doing so, the physiological uptake in non-tumor regions was simulated in a realistic attenuation medium (see Fig. 3). Input maps were sampled in $124 \times 124 \times 45$ voxels of $4 \times 4 \times 4$ mm^3 for each bed position. Three successive bed positions were used to cover the entire thoracic region (corresponding to a total of 87 slices for a bed translation of 21 slices).

For the TP simulation, tumor contours were used to define lesions that were inserted in the attenuation map of the HP (Fig. 3). These contours were obtained from the manual delineation of real lung tumors in PET/CT images by an experienced nuclear medicine physician. The attenuation coefficient corresponding to soft tissue was assigned to the added tumors. In this simulation, we inserted a tumor with uniform activity in each lung (SUV of 5 in the right lung and 3 in the left lung).

In a first study, the prompt coincidences from the HP and the TP were included in the reconstruction, and all corrections were performed based on (1), (2) and (3). In a second study, the prompt coincidences from the TP but only the unscattered trues from the HP were included. In this case sinograms D and S were set to 0 and sinogram P contained only unscattered trues. Monte Carlo scatter correction [3], [4] was applied to the original clinical PET data and to the simulated data. In order to obtain "clinical-like" PET images, 2 iterations and 16 subsets were used in the OSEM reconstruction, followed by a 3-D Gaussian post-smoothing using a 6 mm full-width at half maximum (FWHM) parameter. However, to produce the PET image used to design the HP simulation, 100 iterations and 16 subsets were used, without post-smoothing. This ensured the highest resolution possible in the PET image used as an input of the simulation.

The strategy described in this subsection, to simulate a realistic PET scan including tumors, is totally independent of the one described in Section II-A. In fact, the latter can be used to speed-up the simulations of multiple tumor sets of a given patient. This explains why we used it. However, the simulation of a realistic PET scan can be done all at once, i.e. simulating the tumoral and non-tumoral physiological activity at the same time. In this case, the attenuation map of the TP is used, and the activity map is defined by the one of the HP in which the tumor regions are included.

B. Simulation of Serial PET Scans in the Context of Patient Monitoring

Data of a "healthy" patient without any lesion in the thorax was used. Five successive PET/CT scans of this patient were available, acquired at t_0 (pre-therapy), $t_0 + 6$ months, $t_0 + 12$ months, $t_0 + 16$ months and $t_0 + 20$ months (Fig. 4). The PET images were reconstructed on the scanner console using the clinical reconstruction protocol based on 3D-RAMLA using blobs [25], with 2 iterations, 2.5 mm blob radius and 0.005 relaxation parameter. The CT data (1 mm pixel size and 5 mm slice thickness) were used to calculate the registration parameters between each scan, in ROIs that excluded the volume outside of the rib cage (as illustrated in Fig. 1). Real contours of lung tumors obtained from the manual delineation of real lung tumors on PET/CT images by an experienced nuclear physician were used. The delineation was performed for a patient that underwent 5 scans before and during chemotherapy. The contours were used to insert tumors in the 5 scans of the "healthy" subject using the method described in Section II-B. In this way, the tumor contours in a scan were always different from the tumor contours in the previous scan, and reflected the changes in tumor contours seen in the real patient between each scan. A total of six tumors were included in this patient. To model the activity concentration changes in the tumors from one scan to another, the changes seen in the real patient study were modeled in our simulated study. The tumor activity in each scan and each tumor was calculated as the mean activity in the tumor region delineated in the real PET scans and was assumed to be homogeneous within the tumor. The simulations were performed as described in Section II-A. Simulated data were corrected, reconstructed and post-processed as described in Section III-A-2.

The validity of the tumor relocation process was qualitatively investigated as follows: each simulated scan was registered on the first scan using a rigid transformation calculated in a different way from that described in Section II-B. For this registration, the entire region R was used (without splitting in sets of slices) and the algorithm minimized the squared difference between the 2 image volumes, and trilinear interpolation was used.
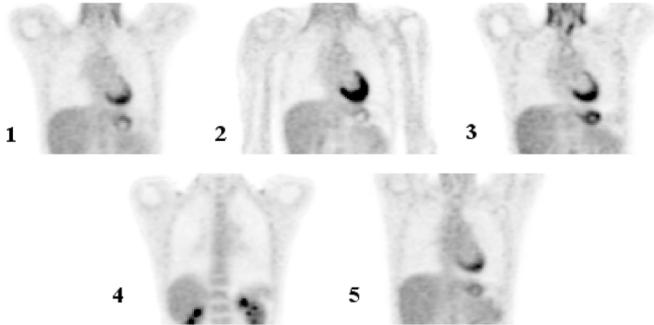
Fig. 4. 5 successive PET scans of the "healthy" patient (without any tumor in the lungs) scanned at t_0 (pre-therapy), $t_0 + 6$ months, $t_0 + 12$ months, $t_0 + 16$ months and $t_0 + 20$ months. All these coronal slices correspond to a same image location. The grey scale is the same for all images. There is non-negligible patient physiological uptake variability between scans along therapy.
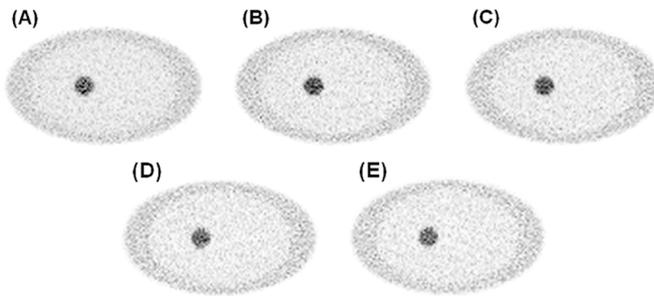
Fig. 5. Reconstructed images obtained using the 3 methods for a lesion radius of 20 mm and for a contrast R_3/R_1 of 4. (A) Unscattered trues from the GS simulation. (B) Unscattered trues from the MO simulation. (C) Unscattered trues from the PM simulation. (D) Prompts from the GS simulation. (E) Prompt from the PM simulation. No difference can be seen between A, B and C, nor between D and E.

Tumors were then segmented using a 40% threshold (T40%) with respect to the maximum SUV in the tumor [26]. Contours obtained using the T40% were superimposed on the tumor as seen on the 5 scans (Fig. 9).

IV. RESULTS

A. Simulation of a Single PET Scan

1) Numerical Phantom Study: Images resulting from the PM, MO and GS methods had no visual differences, and this was true when comparing images reconstructed using only the unscattered trues, or using all coincidences. Example of images obtained for a lesion radius of 20 mm and an activity ratio R_3/R_1 of 4 are shown in Fig. 5. Using only unscattered trues, the mean errors and associated variations over all radii and contrasts for R_1 were $0.02 \pm 0.32\%$ and $0.06 \pm 0.29\%$ for $mE_{MO/GS1}$ and $mE_{PM/GS1}$, respectively. For R_2 they were $0.04 \pm 0.21\%$ and $0.08 \pm 0.20\%$ for $mE_{MO/GS2}$ and $mE_{PM/GS2}$, respectively. Using all prompts and all corrections, errors were $-0.62 \pm 0.32\%$ and $-0.40 \pm 0.20\%$ for $mE_{PM/GS1}$ and $mE_{PM/GS2}$ respectively. These relative errors in regions R_1 and R_2 depended on neither lesion contrast nor lesion size. Errors obtained in region R_3 (the lesion) are shown in Fig. 6.

Graphs 6A shows the relative errors between the MO and GS simulations when using only the unscattered trues. For lesion radii of 20 mm and 40 mm, the errors were almost constant ($\sim 1.2 \pm 1.2\%$) and did not depend on the R_3/R_1 activity ratio.
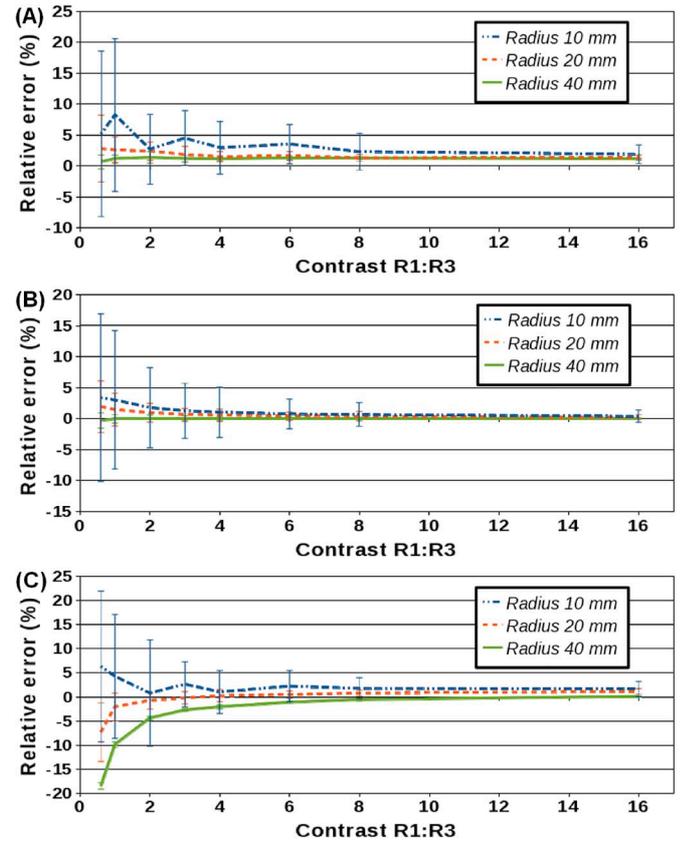
Fig. 6. Relative errors obtained in region R_3. (A) Relative errors $mE_{MO/GS3}$ between the MO and GS simulations using only the unscattered trues. (B) Relative errors $mE_{PM/MO3}$ between the PM and MO simulations using only the unscattered trues. (C) Relative errors $mE_{PM/GS3}$ between the PM and GS simulations using all prompt coincidences.

However, for the smallest lesion (i.e. 10 mm) the error reached $\sim 5\%$ or more for the low contrasts, but decreased when contrast increased and tended to $\sim 2\%$. For this 10 mm sphere and for low activity ratios, the variability of the relative error was about 10.5% while it decreased to 2.5% for activity ratios greater than 3. This variability can be explained by the very low statistics of these TP simulations ($\sim 6176 \pm 38$ unscattered trues for a 2 min scan and the lowest activity ratio), by the fact that only 6 replicates were available, and also because the ROI consisted of 56 voxels only.

Graph 6B shows the relative errors between the PM and MO simulations when using only the unscattered trues. It thus only reflects the errors due to recombining datasets simulated with different attenuation media. For lesion radii of 20 mm and 40 mm, the errors were almost zero ($\sim 0.3 \pm 1.1\%$). For the 10 mm radius, the error reached 4% for the lowest contrast, with a variability of 12.8%, for the same reason as that explaining the variability in graph 6A.

Graph 6C shows that when using all prompt coincidences and all corrections, the bigger the lesion radius, the higher the errors, when comparing PM to GS. However for high contrasts, errors converge to 0.5, 1 and 1.5% for lesion radii of 10, 20 and 40 mm, respectively.

All these results were obtained using the strategy to deal with multiple coincidences consisting in keeping only the coincidence involving the photons with the highest detected energy.
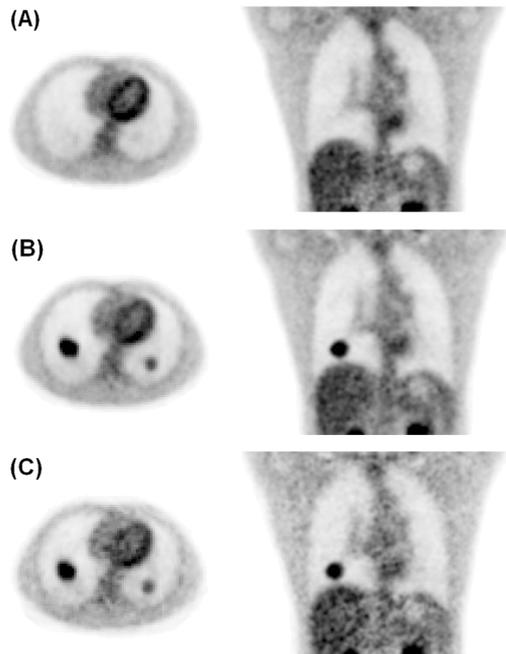
Fig. 7. Axial and coronal slices of PET images from (A) the original clinical data, (B) the simulated unscattered trues from the HP and prompts from the TP, and (C) the simulated prompts from the HP and the TP. The gray scale is the same for all images. Tumors shown in (B) and (C) have an SUV of 5 (in the right lung) and 3 (in the left lung).

Fig. 8. Coronal slices of a simulated patient study. From left to right, 5 successive scans of the "healthy" subject with inserted tumors based on real lung tumor contours.

Fig. 9. Reconstructed images showing the tumor changes along the simulated serial scans. From left to right, the same tumor from the 1st to the 5th scan (each scan registered on the 1st), cropped exactly in the same image location. The gray scale is the same in all images. Contours obtained using the T40% method are shown in colored lines, and superimposed in the last image to show the consistency of the tumor placement from one scan to another.

However, for the case of a tumor radius of 40 mm and activity ratio R_3/R_1 of 4, we ran the simulation with three different strategies for dealing with multiple coincidences: keep all, keep only the one with the highest energy (i.e. the default one), discard all. For each strategy, we calculated the relative error between the total number of unscattered trues coming from the tumor in the GS simulation, and in the TP simulation. The errors were 0.05%, 1.22% and 2.95% respectively for the three strategies listed above. The error of 1.22% corresponds to the point on graph 6A for the same tumor characteristics. These error magnitudes were consistent with the fact that the proportion of unscattered trues coming from multiple coincidences with respect to the total of unscattered trues was 2.80% and 0.16% respectively for the GS and TP simulations, due to the large difference in the amount of activity in the FOV.

2) Realistic Example: To illustrate the proposed method in a realistic context, Fig. 7 shows axial and coronal images resulting from the patient study described in Section III-A-2. All images were reconstructed using the clinical reconstruction parameters. Images 7A were obtained using the original clinical data, images 7B were obtained using simulated unscattered trues from the HP and prompts from the TP, while images 7C were obtained using simulated prompts from the HP and the TP. Images 7A and 7B look similar (except for the added tumors in 7B and a small loss of contrast) while in images 7C, some noise increase and contrast loss can be clearly observed compared to images 7A.

About 433 CPU hours ($\sim$18 days) were needed for the HP simulation, and about 3 CPU hours for the TP (on an Intel Xeon 2.66 GHz). This means that each time this patient is simulated with a new set of tumors, 433 CPU hours of computational time are saved.

B. Simulation of Serial PET Scans in the Context of Patient Monitoring

Fig. 8 shows coronal views of the reconstructed images obtained for the patient simulation described in Section III-B. The coronal slices shown for this study correspond to the coronal slices of the "healthy" patient shown in Fig. 4. The total CPU time for the 5 HP simulations was about 1719 hours ($\sim$72 days), while 61 CPU hours were needed for the TP simulations (Intel Xeon 2.66 GHz).

Using these simulated serial PET scans, we chose a tumor that could be seen on the 5 scans. Cropped images of this tumor are shown in Fig. 9 where the contours obtained with the T40% method are also shown. These cropped images were all taken exactly in the same image location (after image registration). Finally a superimposition of the 5 contours is also shown to demonstrate the changes in the tumor. The superimposed tumor contours show the consistency of the tumor placement in the serial scans.

V. DISCUSSION

A. Strategy for Efficient Data Simulations

Given the large computational time needed to perform realistic simulations of clinical PET scans, we introduced a strategy

in which the background activity distribution needs to be simulated only once, while various sets of tumors can be simulated separately. The simulated data corresponding to the background activity and to each tumor set are then recombined and reconstructed. Using this approach, the generation of a new dataset of the same patient but with a different set of tumors is more than 25 times faster (in these examples) than the complete simulation of the patient with tumors. This speed up factor depends on the total activity in the simulated set of tumors: the smaller it is, the higher the speed up factor.

The idea of inserting tumors into "healthy" datasets was proposed earlier [14]–[16]. However, no detail was given on how issues raised by this technique were dealt with. We investigated in details the approach, in the context of PET.

Given that a number of steps involved in PET imaging and reconstruction are non-linear, it was essential to first assess the biases that are potentially introduced in the simulated images by this recombination approach. For this evaluation, we used a simple numerical phantom. We used either true unscattered or prompt coincidences, simulated or not in different attenuation media, to be able to evaluate the different sources of errors.

The small and constant positive bias (between 1 and 2%) observed when comparing MO to GS with only trues (Fig. 6(a)) can be explained by the way multiple coincidences were processed. In our study, when a multiple coincidence occurred, only a coincidence involving the two photons with the highest energy was kept. There are statistically more multiple coincidences in the GS simulation than in the PM or MO simulations, given that the activity in the FOV is greater. In total, more true unscattered coincidences would thus be implied in the production of multiple coincidences. For a tumor radius of 40 mm and an activity ratio R_3/R_1 of 4, 2.80% of true unscattered coincidences coming from the tumor in the GS were implied in multiple coincidences (0.16% in the TP). Therefore, if multiple coincidences were all discarded, more true unscattered coincidences in the GS simulation would be lost than in the PM or MO simulations. This would result in a positive bias when comparing the MO (or PM) to the GS. Conversely, if all multiple coincidences were recorded, no true unscattered coincidences would be lost in any case, and we would not expect any bias when merging HP and TP simulated data. This was actually checked for the case of a tumor radius of 40 mm and activity ratio R_3/R_1 of 4 by running it with three different strategies for dealing with multiple coincidences (cf. end of Section IV-A-1). This also explains why the errors increased with smaller tumors (and/or smaller contrast), as the proportion of multiple coincidences for the TP simulations almost becomes 0, and this error was confined to smaller volumes.

As the HP and TP simulations are affected by different attenuation media, we first corrected the prompt sinogram P of the HP by multiplying it by the corresponding attenuation coefficient sinogram A and we then virtually attenuated the resulting sinogram by the attenuation coefficient sinogram A' of the TP (2). The data y in (2) were thus all affected by the same attenuation A', so that this sinogram A' could be used to weight for attenuation in the OP-OSEM reconstruction. This would obviously not be needed for tumors located in regions with an attenuation coefficient equal to that of the tumor: the A and A'

sinograms would then be identical (as in the MO simulations). The small error (Fig. 6(b)) due to this change in attenuation coefficient in the tumor region would therefore disappear.

In our approach consisting of merging HP and TP simulated data, some random coincidences cannot be simulated, namely those that would occur between a photon coming from a tumor region and a photon coming from the "healthy" part. In addition, some photons coming from the "healthy" part may have scattered in the tumor region. These photons cannot exist when separately simulating the HP and TP data and then merging them. These random and scattered coincidences can thus be regarded as "perfectly" corrected. When using all prompt events during the reconstruction, this phenomenon introduces a bias between the merged (MO and PM) and non-merged data (GS), given that these events are actually not perfectly corrected for in the GS simulations. However, this bias is negligible because the proportion of these coincidences is small compared to the total number of coincidences, and the current corrections for scatter and randoms have good accuracy.

When using the convolution based scatter correction method [19], a fixed scatter fraction must be set a priori to scale the scatter sinogram to the measured prompt sinogram. If a tumor region has a higher attenuation coefficient than the region in which it is located (as for lung tumors), the scatter fraction in the GS would be larger than in the HP plus TP. Moreover, the bigger the tumor, the larger this difference in scatter fraction. Therefore, when using a fixed scatter fraction as in the convolution based scatter correction, a negative bias is expected when comparing PM to GS. The bigger the tumor region, the larger the bias, as actually observed in Fig. 6(c). This problem might be reduced by using alternative scatter corrections, such as Single Scatter Simulation [27] technique or Monte Carlo based scatter simulation [3], [4].

In this study dead times were not modeled. With dead times, the number of lost events would be larger for the HP simulation than for the TP, as the FOV activity in larger in the HP simulation. Thus HP and TP simulated coincidences would need to be separately corrected for dead time effects.

B. Strategy for Realistic Data Simulations

The simulation of highly realistic datasets including tumors requires accurate modeling of heterogeneous background activity distributions. Although this can be partly achieved using complex phantoms consisting of many organs such as the XCAT phantom, these phantoms still assume that the activity in a given organ or structure is uniform. Even if there is an increasingly number of small structures, it is still hard to know how to set the activity concentration inside these structures.

To overcome this limitation, we proposed modeling background physiological activity distributions based on real activity distributions observed in patients' reconstructed images, for use as an input to the MC simulations. This technique avoids the arbitrary choice of activity concentrations in different regions, but ensures a realistic heterogeneous activity distribution. However, it also has limitations, namely the finite spatial resolution of the input activity map, the associated noise and possible artifacts. If no artifact is present in the images used as the input

data to the simulations, the trade-off between noise and spatial resolution can be well controlled, as shown in another study [28]. In [28], we show that the spatial resolution of the input activity map should be as high as possible, regardless of the noise level, as the noise in the input image does not propagate into the simulated data. We therefore used an activity map of the HP corresponding to a highly iterated (100 iterations and 16 subsets) image. However, for the simulation of serial PET scans, we could only access the clinical images reconstructed on the scanner console, but not the original sinograms of the "healthy" patient. As the resolution of these clinical images was poorer than in highly iterated images, a loss of contrast can be seen between original (Fig. 4) and simulated PET images (Fig. 8). A detailed investigation of the impact of the noise and resolution in the activity map used as an input for the simulation, for both PET and SPECT, can be found in [28].

Using clinical PET/CT scans as an input to the simulation can be a relevant alternative to using complex anthropomorphic phantoms such as the XCAT or Zubal phantoms. The simulated images can be used to get an assessment of quantification method performance in conditions close to those encountered in real patient scans. We previously showed that the performance of a quantification method was poorer when assessed using realistic simulated data than when using simpler phantom experiments [9]. The performance of quantification methods as measured on very realistic simulated data can be more readily extrapolated to real clinical conditions than when using simple phantoms or simulations.

In Fig. 7, images obtained using the prompts of the TP and either the prompts or only unscattered trues of the HP are both shown. Image 7B is visually similar to image 7A, suggesting that images reconstructed using only the unscattered trues from the HP might be appropriate for assessing detection, segmentation and quantification tasks. This is because the activity maps of the HP used for the simulation were already affected by imperfect scatter and random corrections. When prompt events of the HP are used, the reconstructed image (Fig. 7(c)) is more easily distinguished from the original image (Fig. 7(a)). This is because the activity distribution initially biased by imperfect corrections is again reconstructed with imperfect corrections (and associated noise increase due to non-smoothed random correction). This is also why we used a Monte Carlo scatter correction for the realistic example shown in Fig. 7, which is more accurate than the convolution-based method used for the simple phantom experiment.

The proposed method was extended to simulate serial PET scans. The only additional challenge was to properly relocate the tumor from one scan to another. By using appropriate constraints, our relocation approach produces realistic results (Figs. 8 and 9). Our simulation approach thus makes it possible to simulate serial PET scans appropriate for assessing methods that detect and characterize tumor changes between scans [29]. A similar approach could also be used to simulate respiratory gated scans when 4D PET/CT data are available, either using a model describing the tumor deformation through the respiratory cycle, or using tumor contours drawn from real respiratory gated PET images.

As mentioned previously, the strategy to simulate highly realistic PET images with tumors, whether in the context of therapy monitoring or not, is completely independent of the strategy to speed-up the simulations. A single PET scan of a patient with tumors can be made all at once, i.e. simulating all non-pathological and tumoral activity at the same time. In this way, however, the simulations of many cases are more time consuming, but ensure totally unbiased simulated data.

It is also important to state that healthy patient data were obtained from patients in which no visually detectable tumors were located inside the field of view used in the simulation (for instance, a patient with a colorectal cancer might be used to simulate thorax PET scans). Patients with a PET scan showing a complete response to therapy can thus also be used as the HP for simulations.

The methods described in this paper have already been used in the context of PET Monte Carlo simulations of lung cancer patients, using the GATE simulation software [1], [30]. They could be used with any simulation, and could be extended to any type of cancer, and heterogeneous uptake in tumors could be modeled to improve the realism of the simulations [9], [10], [31]. They can also be used in SPECT, as was already done in lesion detectability studies [14]–[16], using another Monte Carlo software.

VI. Conclusion

In this work, we described a method to simulate realistic PET images of patients including tumors and heterogeneous physiological uptake. This method is based on the use of real patient PET/CT data as the input of the simulation. We extended the method to the simulation of serial PET scans in the context of therapy monitoring, including realistic tumor placement and evolution.

We also described and assessed an efficient method to simulate databases of pathological cases for which the non-tumoral physiological uptake stays unchanged. The simulations of the physiological and tumoral uptakes are performed separately and then recombined, with the former being performed only once. We demonstrated that this procedure introduced small errors (in the order of 1–2%) compared to a single simulation in which all uptakes are simulated simultaneously.

References

[1] P. Tylski, S. Stute, N. Grotus, K. Doyeux, S. Hapdey, and I. Gardin *et al.*, "Comparative assessment of methods for estimating tumor volume and standardized uptake value in 18F-FDG PET," *J. Nucl. Med.*, vol. 51, pp. 268–276, 2010.

[2] M. Soret, P. Koulibaly, J. Darcourt, and I. Buvat, "Partial volume effect correction in SPECT for striatal uptake measurements in patients with neurodegenerative diseases: Impact upon patient classification," *Eur. J. Nucl. Med. Mol. Imaging*, vol. 33, pp. 1062–1072, 2006.

[3] C. H. Holdsworth, C. S. Levin, M. Janecek, M. Dahlbom, and E. J. Hoffman, "Performance analysis of an improved 3D PET Monte Carlo simulation and scatter correction," *IEEE Trans. Nucl. Sci.*, vol. 49, no. 1, pp. 83–89, Feb. 2002.

[4] O. Barret, T. A. Carpenter, J. C. Clark, R. E. Ansorge, and T. D. Fryer, "Monte Carlo simulation and scatter correction of the GE advance PET scanner with SimSET and Geant4," *Phys. Med. Biol.*, vol. 50, pp. 4823–4848, 2005.

[5] I. Castiglioni, G. Rizzo, M. Gilardi, V. Bettinardi, A. Savi, and F. Fazio, "Lesion detectability and quantification in PET/CT oncological studies by Monte Carlo simulations," *IEEE Trans. Nucl. Sci.*, vol. 52, no. 1, pp. 136–142, Feb. 2005.

[6] S. Tomei, A. Reilhac, D. Visvikis, N. Boussion, C. Odet, and F. Giammarile *et al.*, "OncoPET_DB: A freely distributed database of realistic simulated whole-body 18F-FDG PET images in oncology," *IEEE Trans. Nucl. Sci.*, vol. 57, no. 1, pp. 246–255, Feb. 2010.

[7] N. Grotus, A. J. Reader, S. Stute, J. C. Rosenwald, P. Giraud, and I. Buvat, "Fully 4D list-mode reconstruction applied to respiratory-gated PET scans," *Phys. Med. Biol.*, vol. 54, pp. 1705–1721, 2009.

[8] M. Jin, Y. Yang, X. Niu, T. Marin, J. G. Brankov, and B. Feng *et al.*, "A quantitative evaluation study of four-dimensional gated cardiac SPECT reconstruction," *Phys. Med. Biol.*, vol. 54, pp. 5643–5659, 2009.

[9] S. Stute, P. Tylski, N. Grotus, and I. Buvat, "LuCaS: Efficient Monte Carlo simulations of highly realistic PET tumor images," *IEEE Nucl. Sci. Symp. Conf. Rec.*, pp. 4010–4012, 2008.

[10] A. Le Maitre, W. P. Segars, S. Marache, A. Reilhac, M. Hatt, and S. Tomei *et al.*, "Incorporating patient-specific variability in the simulation of realistic whole-body 18F-FDG distributions for oncology applications," *Proc. IEEE*, vol. 97, no. 12, pp. 2026–2038, Dec. 2009.

[11] G. Zubal, G. Gindi, M. Lee, C. Harrell, and E. Smith, "High resolution anthropomorphic phantom for Monte Carlo analysis of internal radiation sources," *IEEE Comp. Based Med. Syst. Symp. Proc.*, pp. 540–547, 1990.

[12] W. P. Segars and B. M. W. Tsui, "MCAT to XCAT: The evolution of 4-D computerized phantoms for imaging research," *Proc. IEEE*, vol. 97, no. 12, pp. 1954–1968, Dec. 2009.

[13] H. Zaidi and B. M. W. Tsui, "Review of computational anthropomorphic anatomical and physiological models," *Proc. IEEE*, vol. 97, no. 12, pp. 1938–1953, Dec. 2009.

[14] N. F. Pereira, H. C. Gifford, T. H. Farncombe, M. Smyczynski, and M. A. King, "An investigation of reconstruction strategies for mediastinal lesion detection using hybrid Ga-67 SPECT images," in *IEEE Nucl. Sci. Symp. Conf. Rec.*, 2005, vol. 3, pp. 1691–1694.

[15] M. Smyczynski, H. C. Gifford, A. Lehovich, J. E. McNamara, and W. P. Segars *et al.*, "Impact of the respiratory motion on the detection of small pulmonary nodules in SPECT imaging," in *IEEE Nucl. Sci. Symp. Conf. Rec.*, 2007, vol. 5, pp. 3241–3245.

[16] M. Smyczynski, H. C. Gifford, A. Lehovich, J. E. McNamara, and W. P. Segars *et al.*, "Effectiveness of three alternative strategies in reducing the degrading impact of respiratory motion on the detection of small pulmonary nodules in SPECT imaging," in *IEEE Nucl. Sci. Symp. Conf. Rec.*, 2008, pp. 3663–3667.

[17] S. Jan, G. Santin, D. Strul, S. Staelens, K. Assié, and D. Autret *et al.*, "GATE: A simulation toolkit for PET and SPECT," *Phys. Med. Biol.*, vol. 49, pp. 4543–4561, 2004.

[18] N. S. Rehfeld, S. Stute, J. Apostolakis, M. Soret, and I. Buvat, "Introducing improved voxel navigation and fictitious interaction tracking in GATE for enhanced efficiency," *Phys. Med. Biol.*, vol. 54, pp. 2163–2178, 2009.

[19] D. L. Bailey and S. R. Meikle, "A convolution-subtraction scatter correction method for 3D PET," *Phys. Med. Biol.*, vol. 39, pp. 411–424, 1994.

[20] L. A. Shepp and Y. Vardi, "Maximum likelihood reconstruction for emission tomography," *IEEE Trans. Med. Imag.*, vol. MI-1, no. 2, pp. 113–122, Oct. 1982.

[21] D. G. Politte and D. L. Snyder, "Corrections for accidental coincidences and attenuation in maximum-likelihood image reconstruction for positron-emission tomography," *IEEE Trans. Med. Imag.*, vol. 10, no. 1, pp. 82–89, Mar. 1991.

[22] D. J. Kadrmas, "LOR-OSEM: Statistical PET reconstruction from raw line-of-response histograms," *Phys. Med. Biol.*, vol. 49, pp. 4731–4744, 2004.

[23] S. L. Bacharach, M. A. Douglas, R. E. Carson, P. J. Kalkowski, N. M. T. Freedman, and P. Perrone-Filardi *et al.*, "Three-dimensional registration of cardiac positron emission tomography attenuation scans," *J. Nucl. Med.*, vol. 34, pp. 311–321, 1993.

[24] J. H. Hubbel and S. M. Seltzer, Tables of X-Ray Mass Attenuation Coefficients and Mass Energy-Absorption Coefficients, *NIST website* [Online]. Available: nist.gov/physlab/data/xraycoef, 1996.

[25] J. A. Browne and A. R. Pierro, "A row-action alternative to the EM algorithm for maximizing likelihoods in emission tomography," *IEEE Trans. Med. Imag.*, vol. 15, no. 5, pp. 687–699, Oct. 1996.

[26] Y. E. Erdi, O. Mawlawi, S. M. Larson, M. Imbriaco, H. Yeung, and R. Finn *et al.*, "Segmentation of lung lesion volume by adaptative positron emission tomography image thresholding," *Cancer*, vol. 80, pp. 2505–2509, 1997.

[27] C. C. Watson, "New, faster, image-based scatter correction for 3D PET," *IEEE Trans. Nucl. Sci.*, vol. 47, no. 4, pp. 1587–1594, Aug. 2000.

[28] S. Stute, T. Carlier, K. Cristina, C. Noblet, and A. Martineau *et al.*, "Monte Carlo simulations of clinical PET and SPECT scans: Impact of the input data on the simulated images," *Phys. Med. Biol.*, vol. 56, pp. 6441–6457, 2011.

[29] H. Necib, C. Garcia, A. Wagner, B. Vanderlinden, P. Emonts, and A. Hendlisz *et al.*, "Detection and characterization of tumor changes in 18F-FDG PET patient monitoring using parametric imaging," *J. Nucl. Med.*, vol. 52, pp. 354–361, 2011.

[30] A.-S. Dewalle-Vignion, N. Betrouni, R. Lopes, D. Huglo, S. Stute, and M. Vermandel, "A new method for volume segmentation of PET images, based on possibility theory," *IEEE Trans. Med. Imag.*, vol. 30, no. 2, pp. 409–423, Feb. 2011.

[31] M. Hatt, C. Cheze Le Rest, P. Descourt, A. Dekker, D. De Ruysscher, and M. Oellers *et al.*, "Accurate automatic delineation of heterogeneous functional volumes in positron emission tomography for oncology applications," *Int. J. Radiat. Oncol. Biol. Phys.*, vol. 77, pp. 301–308, 2010.