Monte-Carlo simulations of clinically realistic respiratory gated ¹⁸F-FDG PET: Application to lesion detectability and volume measurements

S. Vauclin[a,b,1], C. Michel[c], I. Buvat[d], K. Doyeux[a,e], A. Edet-Sanson[a,f,g], P. Vera[a,f,g], I. Gardin[a,f,g], S. Hapdey[a,f,g,*]

[a] QuantIF-Litis, EA4108 – FR CNRS 3638, Rouen University, Rouen, France
[b] Siemens Medical, Saint-Denis, France
[c] Siemens Medical, Knoxville, TN, USA
[d] IMNC, UMR 8165 CNRS, Universités Paris 7 & 11, Orsay, France
[e] Radiotherapy Department, Henri Becquerel Center, Rouen, France
[f] Nuclear Medicine Department, Henri Becquerel Center, Rouen, France
[g] Rouen University Hospital, Rouen, France

ARTICLE INFO

ABSTRACT

In PET/CT thoracic imaging, respiratory motion reduces image quality. A solution consists in performing respiratory gated PET acquisitions. The aim of this study was to generate clinically realistic Monte-Carlo respiratory PET data, obtained using the 4D-NCAT numerical phantom and the GATE simulation tool, to assess the impact of respiratory motion and respiratory-motion compensation in PET on lesion detection and volume measurement. To obtain reconstructed images as close as possible to those obtained in clinical conditions, a particular attention was paid to apply to the simulated data the same correction and reconstruction processes as those applied to real clinical data. The simulations required 140,000 h (CPU) generating 1.5 To of data (98 respiratory gated and 49 ungated scans). Calibration phantom and patient reconstructed images from the simulated data were visually and quantitatively very similar to those obtained in clinical studies. The lesion detectability was higher when the better trade-off between lesion movement limitation (compared to ungated acquisitions) and image statistic preservation is considered (respiratory cycle sampling in 3 frames). We then compared the lesion volumes measured on conventional PET acquisitions versus respiratory gated acquisitions, using an automatic segmentation method and a 40%-threshold approach. A time consuming initial manual exclusion of noisy structures needed with the 40%-threshold was not necessary when the automatic method was used. The lesion detectability along with the accuracy of tumor volume estimates was largely improved with the gated compared to ungated PET images.

© 2014 Elsevier Ireland Ltd. All rights reserved.

1. Introduction

PET using ^{18}F-FluoroDeoxyGlucose (^{18}F-FDG PET) is a useful adjunct to Computed Tomography (CT) for radiotherapy planning. It yields a better delineation and discrimination of tumor tissues especially in patients with Non-Small Cell Lung Cancer (NSCLC). When combined with CT planning images, it often leads radiation oncologists to modify the definition of the Gross Tumor Volume (GTV) [1]. Using ^{18}F-FDG PET also decreases inter and intra observer variability in tumor delineation [2]. However, in PET/CT thoracic imaging, respiratory motion has been reported as a limiting factor reducing image quality and biasing lesion volume measurement. It introduces blurring, reduces contrast and leads to the underestimation of activity concentrations and overestimation of lesion volumes in PET data [3]. The commercially available solution consists in performing respiratory gated PET acquisitions, so that the respiratory cycle is segmented in multiple phases for which movement is limited [4]. Validation of such approach requires a perfect knowledge of the image parameters (such as organ motions, lesion contrast and volume, activity concentration). To that end, the Monte-Carlo simulations are the ideal tool since it allows a complete description of the data not accessible in the case of patient studies [5].

The first aim of this study was to generate clinically realistic respiratory gated ^{18}F-FDG-PET images. We focused our work on the data post-processing (correction and reconstruction) to produce realistic images. We used these simulated data to evaluate the contribution of respiratory gating on tumor volume and GTV definition knowing all image parameters. We also investigate the impact of respiratory gated images on tumor uptake estimation regarding the lesion location and displacement.

Other authors have already proposed the simulation of PET respiratory gating datasets. However, none of them have proposed to create a whole set incorporating multiple lesion positions, size and contrast or different respiratory bin sampling [6–8]. Since one single case, will not be sufficient to address the problem of respiratory gating, we had to use a scenario allowing the creation of multiple dataset based on the creation of a lesion-free patient and of multiple lesion varying in location and size.

First, the simulations were validated using a static cylindrical phantom mimicking the commercially available Jaszczak

Deluxe® phantom and comparing simulated and real, phantom and patient data. Second, the tumor volume (TV) and the displacement volume (DV), corresponding to the volume covered by the lesion during breathing, (i.e. to the volume to be considered for non-gated radiotherapy treatment planning), were compared depending on whether they were derived from ungated or respiratory gated PET acquisitions. Segmentation of these volumes (TV and DV) was performed using an automatic iterative and adaptive method [9] and a fixed threshold of 40% of the maximum intensity within the lesion.

2. Materials and methods

2.1. Simulation system and PET scanner

Geant4 Application for Emission Tomography (GATE) [10] was used to perform the simulations. The SIEMENS Biograph Sensation 16 Hi-Rez PET scanner corresponding to that used in the Department of Nuclear Medicine of the Henri Becquerel Center and Rouen University Hospital was modeled [11–13]. Its physical characteristics are given in Table 1.

The acquisition bed was also modeled. To mimic the limited energy and temporal resolution of a real scanner, the energy and detection time associated with each particle were blurred using Gaussian functions (FWHM$_{energy}$ = 15% at 511 keV and FWHM$_{time}$ = 1.15 ns). We also implemented a 500 ns delayed coincidence window (4.5 ns) used for random correction, applying a correction process similar to that used in the real scanner. The scanner dead-time was not considered. Sinograms were modified to account for the light sharing in the detector block, by spreading the detected events over the surrounding detectors, considering a Gaussian probability distribution. The Gaussian full-width at half maximum (FWHM) was derived by comparing simulated (FWHMs = 1.43 pixels) and experimental (FWHMe = 1.89 pixels) spatial resolutions in the sinogram of an ^{18}F-FDG PET point source.

Acquisitions were simulated in 3D mode and run on the CRIHAN supercomputer (http://www.crihan.fr). Prompts and delayed coincidences were recorded, as well as singles detected inside the energy window.

2.2. Phantoms

Two types of numerical phantoms were used. First, a cylindrical phantom containing 21 spheres with volumes ranging from 0.20 to 200 mL (i.e. 0.20, 0.43, 0.71, 0.99, 1.5, 2.1, 2.9, 3.8, 6.4, 9.0, 11.6, 15.4, 19.3, 23.6, 27.9, 43.0, 58.1, 77.7, 97.3, 148.6 and 200.0 mL) was defined. The simulations were performed considering 9 source-to-background activity ratios (S/B) ranging from 2.0 to 20.6 (i.e. 2.0, 2.7, 3.4, 4.9, 6.3, 7.7, 12.0, 16.3 and 20.6) and for 4 acquisition durations (1, 2, 3 and 4 min), corresponding to different noise levels, for each S/B. These durations yielded for example 17, 33, 50 and 66 million prompts for the lowest S/B (2.0). These simulations were used to calibrate the iterative segmentation algorithm described in Section 2.5.

Realistic ^{18}F-FDG PET thoracic scans were also simulated. The voxelized thoracic 4D NCAT phantom [14] of 128 × 128 × 86 isotropic voxels (size 4.68 mm) was used to simulate both cardiac and respiratory motions (see Fig. 2). The phantom covered 402 mm, that is approximately 2.5 times the axial FOV and was

Table 1 – Characteristics of the SIEMENS Biograph Sensation 16 Hi-Rez PET scanner.

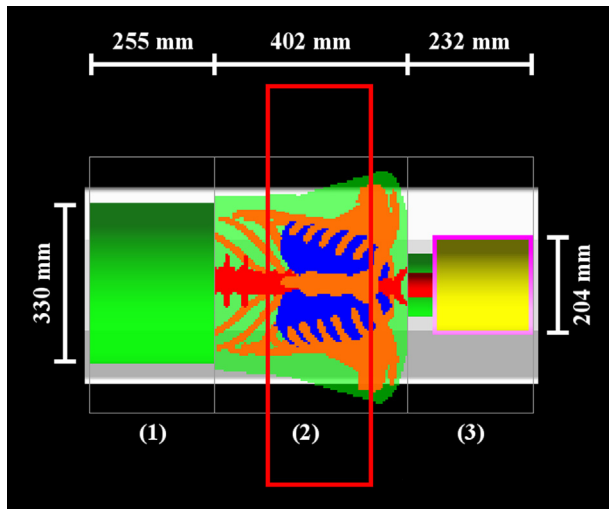| Characteristic	Value
# Rings	3
# Detector Blocks	48
Block geometry	13 × 13 crystal array
Crystal type	LSO
Crystal size (mm)	4.0 × 4.0 × 20.0
Ring diameter (mm)	842
Axial FOV (mm)	162
Energy Window (keV)	[425–650]
Coincidence Window (ns)	4.5
Delayed Coincidence Window (ns)	500

Fig. 1 – Patient model in GATE designed using both analytical ((1) and (3)) and voxelized geometries ((2) 4D NCAT phantom). The red box represents the scanner axial field of view. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

Table 2 – Median activity concentration of the different organs/structures simulated in gate.

Organ/anatomical structure VOI	Median activity concentration (Bq/mL)
Brain	21,857
Heart:blood compartments	5498
Myocardium	15,425
Blood	5498
Spine	5203
Ribs	2672
Stomach	6354
Liver	7983
Lungs	1405
Spleen	5760
Kidney	9444
Soft tissues	2104
Urethra	41,540
Bladder	133,590

centered on the upper border of the diaphragm. The arms were numerically removed to mimic clinical situations (patient with arms above the head). To take into account the contributions of randoms and multiple events coming from regions outside the FOV, head and neck, part of the abdomen (232 mm in length and 204 mm in diameter) and pelvis (255 mm in length and 330 mm in diameter) were also modeled. These body parts were modeled analytically (instead of using a voxelized description) to reduce computation time, since GATE allows the combination of analytical (using parametric functions) and voxelized description of the simulated objects. The whole simulated model is shown in Fig. 1. The geometric parameters of these analytical volumes of interest (VOIs) (volume, thickness, etc.) were derived from the NCAT phantom and/or from measurements (considering the median) performed in 10 patients with morphology close to that used for the design of the NCAT phantom (i.e. a 90 kg man). Fourteen different functional organs and structures were defined: bladder, blood pool (left and right atria/ventricle chambers, arteries and veins), brain, heart myocardium, kidney, liver, lungs, ribs, soft tissues/muscle, spine bone, spleen, stomach and urethra. ^{18}F-FDG activity concentrations were obtained from median measurements performed in the same set of 10 patients. These data are summarized in Table 2.

For tissue attenuation, 5 different media were considered: air, lungs, spine bone, ribs and soft tissues. A breathing cycle of 4 s with a motion amplitude of 2 cm for the diaphragm in the axial direction and of 1.2 cm in the anterior–posterior direction was defined. This model was first used to initially simulate tumor-free patient data, consisting of a set of 8 emission and transmission frames throughout the respiratory cycle with acquisition duration of 1 min for each frame (i.e. 13 million prompts). Then, spherical lesions (1, 2, 3, 4 cm in diameter) were simulated independently in 5 different locations (see

Fig. 2): lower part of the middle and inferior lobes of the right lung, upper and lower part of the inferior lobe of the left lung, upper part of the right lobe of the liver, considering 4 S/B (2, 4, 6, 8). These S/B were defined according to the background activity surrounding the lesion (i.e. lungs and liver tissue respectively).

In these simulations, most of the computation time is due to the simulation of the photons coming from healthy organs, compared to those coming from the tumors. To reduce the simulation time and reuse the simulated data in different configurations, we generated respiratory-gated ^{18}F-FDG PET thoracic scans of the patients (P) by simulating independently a healthy subject (HS) and only a tumor (TO). A particular attention was paid to the corrections and reconstruction of the recombined data as explained in Sections 2.3 and 2.4.

Finally, the simulated data were recombined as a function of the prompts time tag, to obtain 2 additional respiratory cycle sampling (RCS) in 5 (mimicking our clinical protocol) and 3 frames. Four RCS were therefore created: ungated, gated in 3, 5 and 8 frames. Transmission maps corresponding to each frame of the considered RCS were generated for attenuation correction purpose only.

2.3. Data correction

To obtain reconstructed images as close as possible to those obtained using the Biograph in clinical conditions, a particular attention was paid to apply the same type of reconstruction and correction processes on the simulated data as on clinical data. First, the normalization factors appropriate for the simulated tomograph (accounting for geometric effects, crystal efficiencies, dead time, etc.) were derived as one would do for real data.

Crystal efficiencies were assumed to be constant for all detectors and dead time was not considered because it was not simulated. Therefore, only the geometric effects (gap filling, cylindrical and spherical scanner geometry) were considered and directly obtained from the scanner normalization file. Although the random and scatter contributions could be determined using Gate, their corrections were actually performed using the same type of methods used in the real scanner. Random correction was performed by subtracting
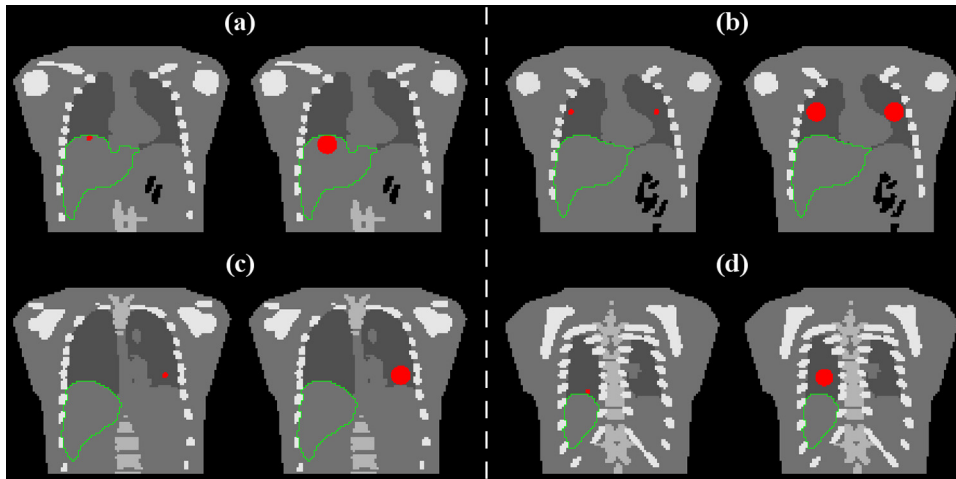
Fig. 2 – Illustration of the 5 lesion locations within: the liver (a), middle lungs (b), inferior left lung (c) and inferior right lung (d). In each case, the 1 and 4 cm lesion are represented, together with the liver contour.

an estimate of the randoms from the prompt. A variance reduced estimation of the random events recorded in the delay time window was obtained using the exact single-plane Casey method [15] expanded to 3D mode. Scatter was corrected following the algorithm proposed by Watson [16]. For attenuation correction, the NCAT transmission frames were smoothed in 3D to remove tissue transition artifacts, using a point spread function corresponding to the simulated Biograph PET image resolution. The FHWM parameters of the Gaussian function were 5.8 and 4.0 mm in the transaxial and axial direction respectively. Whatever the respiratory cycle sampling that was considered, each emission frame was corrected for attenuation with its corresponding transmission map to avoid biasing the attenuation correction. For each RCS, the attenuation map was generated, weighting the attenuation map with the time ratio between the considered RCS and the original simulated RCS (8 frames of 60 s). For example, when considering an RCS of 5 frames of 96 s (8 × 60/5), the 5 following attenuation maps were generated:

$$^5Att_1 = 60/96 \times {}^8Att_1 + 36/96 \times {}^8Att_2$$
$$^5Att_2 = 24/96 \times {}^8Att_2 + 60/96 \times {}^8Att_3 + 12/96 \times {}^8Att_4$$
$$^5Att_3 = 48/96 \times {}^8Att_4 + 48/96 \times {}^8Att_5$$
$$^5Att_4 = 12/96 \times {}^8Att_5 + 60/96 \times {}^8Att_6 + 24/96 \times {}^8Att_7$$
$$^5Att_5 = 36/96 \times {}^8Att_7 + 60/96 \times {}^8Att_8$$

As a result, both CT and PET acquisitions were considered to be gated.

2.4. Data recombination and images reconstruction

A data recombination approach similar to the one published recently [8] was used to generate the final simulated ^{18}F-FDG respiratory gated PET thoracic scans of the patient. The recombination consists in creating the patient prompt sinogram (P) by merging that of the healthy subject (HS) with that of the tumor only (TO). In lung lesions, the attenuating tissue at the location of the simulated lesion was not the same in the HS and TO data (lungs and soft tissue respectively). This difference has to be taken into account during the recombination and correction process of the summed prompt sinograms. Actually, neither the HS nor the TO attenuation maps can be directly used to accurately correct the merged data. To solve this issue, a method, consisting in 4 steps, was used [17,18] (Fig. 3). This approach was applied for each respiratory frame.

(1) In the HS sinogram, all prompts coming from the future tumor location were identified and removed. The resulting HS sinogram thus had no activity coming from the lung lesion location.
(2) HS and TO prompt sinograms were corrected for normalization, randoms, scatter and attenuation independently. The contributions of scatter and attenuation were estimated using respectively the HS and TO attenuation maps.
(3) Both previously corrected sinograms were added to generate that of the patient including the lesion.
(4) Finally, images were reconstructed into 81 slices of 168 × 168 voxels (4.06 mm × 4.06 mm × 2 mm) using Fourier rebinning (FORE) and attenuation weighted ordered subsets expectation maximization (AW-OSEM) iterative algorithms with the TO attenuation map (equivalent to the patient one). As performed for real clinical data, AW-OSEM was used to reduce the bias introduced by pre-correcting the sinograms for attenuation before reconstruction. Reconstruction parameters of 4 iterations, 8 subsets and a 3D post-reconstruction filtering of FWMH = 5 mm were chosen to match the clinical reconstruction protocol. To check the consistency of the whole procedure, images obtained from the combination of a 4 cm lung tumor with S/B = 1 and the HS were compared to those of the HS reconstructed directly with the standard clinical procedure.

2.5. Segmentation algorithms

Using our simulated data, we investigated the accuracy of tumor segmentation methods. The segmentation was applied on each frame of the gated data. Two methods were
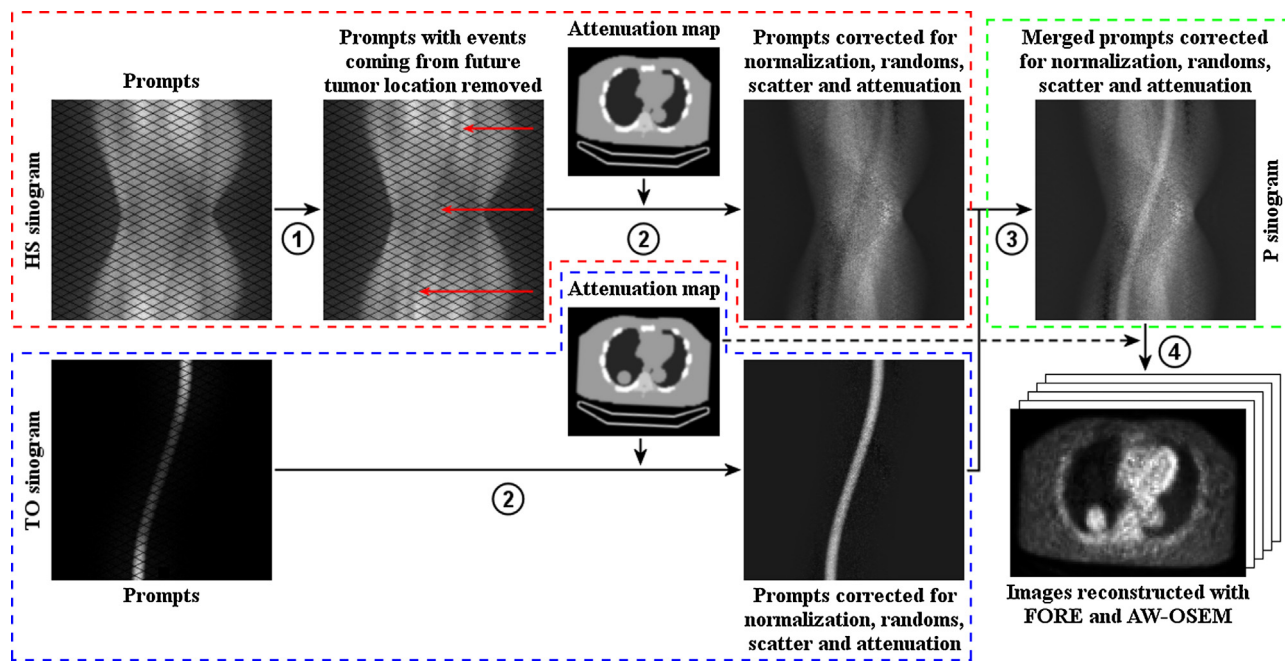
Fig. 3 – Data recombination and image reconstruction methodology used to generate the patient data by merging HS and TO data.

considered to segment the tumors. The first used a constant threshold of 40% of the maximum intensity within the lesion (M40%), as described in [19]. An automatic adaptive and iterative thresholding approach derived from a method previously proposed by Daisne et al. [20] and modified to be applicable to clinical data [9] was also used (MAut). It is known that the optimal threshold $ThOpt$ to accurately measure the true tumor volume from PET images depends on the local lesion contrast ($ContLocMeas$). In MAut, $ThOpt$ and $ContLocMeas$ were iteratively updated until convergence using: $ThOpt = A1 + A2/ContLocMeas$ where $A1$ and $A2$ are 2 constants previously adjusted from the cylindrical phantom data described in Section 2.2.

$ThOpt$ was defined according to an averaged maximum intensity voxel value within the lesion ($MaxAvg$, computed within a pseudo-spherical region of 0.5 mL, large enough to reduce the influence of noise and small enough to be used on small lesions) to limit the influence of noise. $ContLocMeas$ corresponds to the contrast between $MaxAvg$ and the local background of the lesion, estimated inside a shell surrounding the segmented volume. The method is decomposed in 4 steps. Once the user has selected one voxel within the lesion, the first step creates a volume of interest (VOI_{init}) including the lesion and computes both Max_{Avg} within the lesion and the initial threshold $Th_{i=0}$ of the iterative process. From this threshold value, the second step extracts the largest 3D volume, $VOI_{i=0}$, of connected voxels in VOI_{init}. The third step calculates $Cont_{LocMeas}$ between VOI_i and surrounding tissues. In the last step, the threshold value Th_i is optimized in an iteration loop to get Th_{Opt}. In addition, morphological mathematical operations were used before and during the segmentation process to avoid the integration of noisy or neighboring structures in the segmented volume and ensure its connexity, which is particularly important for small and/or low contrast lesions.

A more comprehensive description of MAut can be found in [9,21].

2.6. Data analysis

The impact of respiratory gating on GTV definition was assessed by measuring the DV, corresponding to the volume covered by the lesion during breathing, using the different datasets. As illustrates in Fig. 4, for a given gated acquisition, DV was computed by taking the union of the different segmented volumes obtained in each frame, using one of the approaches described in Section 2.5. Four estimates of the DV were performed depending on RCS (5 frames and ungated data) and threshold method considered (TM: MAut, M40%): DV_5MAut, DV_UnMAut, DV_5M40% and DV_UnM40%.
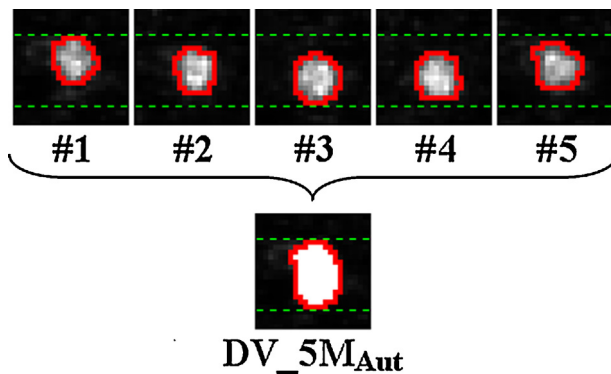

Fig. 4 – Illustration of DV computation when RCS = 5. The union of the 5 different segmented volumes (with M_{Aut} in this example) obtained from each frame of the gated acquisition was calculated to produce DV_5 M_{Aut}.
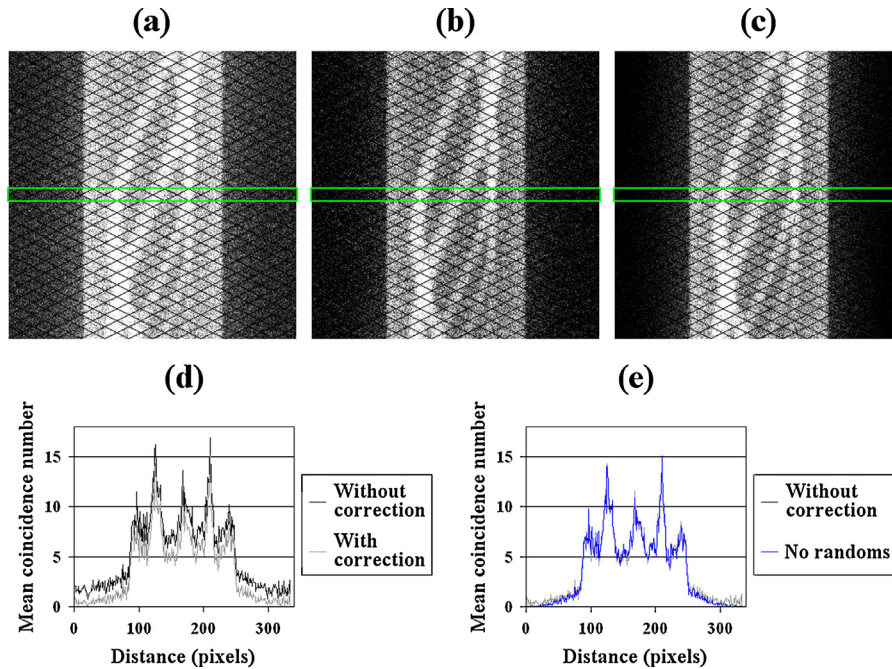
Fig. 5 – Random coincidence correction, assessed using the cylindrical phantom (S/B contrast = 20.6, 4 min acquisition). (a) and (b) Show the sinogram without and with correction. (c) Sinogram without random (known for simulations). (d) and (e) Represent the horizontal profiles of the mean detected events within the highlighted rectangle as a function of the radial position for sinograms a, b and c.

When MAut was used, A1 and A2 parameters were defined using the calibration data corresponding to the acquisition duration closest to that of each frame of the RCS. Knowing the simulated lesion size and displacement, we were able to determine the true DV (DVTr) for each lesion and displacement. The accuracy of DV estimate was then evaluated by computing the average error between each DV_RCSTM and its true associated DV: ErrDV = DV_RCSTM − DVTr.

3. Results

3.1. Simulations and data correction

The whole set of simulations (cylindrical phantom and patients) spread over 44 calculation processors of an IBM supercomputer, required a mono CPU equivalent computation time of 140,000 h and generating 1.5 T of data [18]. This includes the simulations of 36 calibration phantoms S/B and 4 noise levels, 192 patients and 4 HS scans. For each RCS considered (8, 5, 3 and ungated), 48 patients (all from the same corresponding HS) were generated with 1 (liver) or 2 tumors each (lungs) where the tumor varied in location (3: liver, left and right lungs), diameter (4) and S/B (4).

To produce realistic images, we implemented the random correction technique used in the real scanners, based on the single-plane Casey method expanded to 3D mode [15,22]. The performance of the correction is illustrated in Fig. 5. These Monte-Carlo simulated data show that the random correction is correctly performed within the phantom limit.

3.2. Data recombination and images reconstruction

No difference was observed at the tumor location when comparing the images of the HS alone, with the images obtained from the combination of a lung tumor with S/B = 1 and the HS. This indicates that we did not introduce any artifacts with our methodology of data recombination. A variation of the mean count value within the tumor location was ≪1%. As a result, the recombination and reconstruction approach allow generating patient images by merging independently the HS and TO simulated data.

3.3. Data comparison

Calibration phantom and patient reconstructed images from the simulated data were visually very similar to those obtained in clinical studies (Fig. 6). Fig. 7 presents the calibration curves of the MAut segmentation method obtained from the experimental and simulated data. This figure illustrates that the contrasts computed from the simulated images accurately match those computed from the physical experiment.

The calibration of the segmentation method using the simulated data was first compared to that obtained from the experimental PET/CT acquisition of a real phantom considering exactly the same sphere volumes (9), S/B (4) and acquisition time (n = 35 visible spheres out of 36 simulated ones). Results were close (see Fig. 7) with A1 and A2 respectively equal to 27.5 and 106.1 for the simulation compared to 26.4 and 102.4 using the experimental data. Correlation coefficients between experimental data and fit were slightly higher
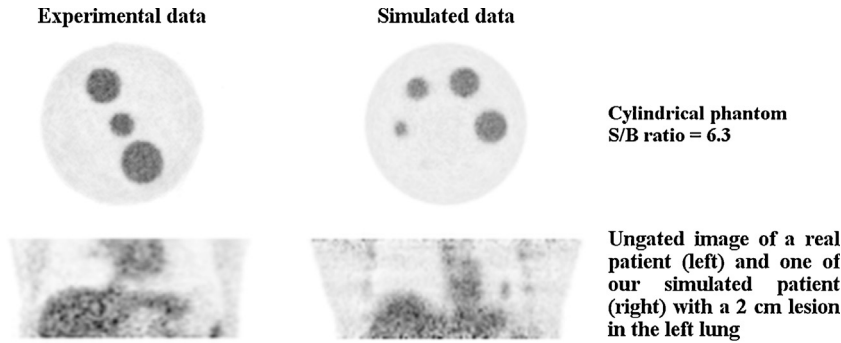
Experimental data　　**Simulated data**

Cylindrical phantom
S/B ratio = 6.3

Ungated image of a real patient (left) and one of our simulated patient (right) with a 2 cm lesion in the left lung

Fig. 6 – Examples of simulated and experimental images of the cylindrical calibration phantom and of clinical scans.

using the simulated data with $R^2 = 0.883$ vs. 0.825 for the experimental data. It underlines that simulated and experimental images were very close since $R^2 \approx 1$ between the 2 calibration curves for a comparable dispersion of *ThOpt/ContLocMeas* couples of points. This result indicates that the simulation and correction processes used in this work accurately mimic the clinical ones.

3.4. Detectability

Table 3 synthesizes the results of lesion visual detectability regarding the RCS (ungated, 3, 5 and 8 frames). The results are presented as a function of the lesion location and volume. We did not distinguish the lesion as a function of their contrast, since the not visible lesions correspond to the lowest contrast. This table shows that the RCS of 3 frames is the most adapted for lesion visual detectability, compare to ungated or higher RCS. This can be explained by the better trade-off between lesion movement limitation (unlike ungated acquisitions) and image statistic preservation. Indeed, with 5 or 8 frames, the same amount of coincidences is spread over more frames, inducing a statistic lost for each frame.
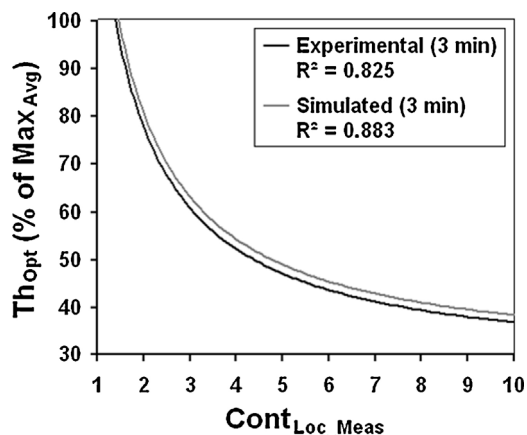

Fig. 7 – Comparison of the calibration results of the segmentation method using experimental and simulated data. The same sphere volumes, S/B and acquisition durations were considered.

3.5. Segmentation

The absolute errors in DV measurements are summarized in Table 4 for DV_8Aut, DV_5Aut, DV_3Aut, DV_UnAut and DV_Un40%. The segmentation of gated frames using the 40% fixed threshold was difficult due to the low signal in the images and the low lesion-to-background ratios. It needs a time consuming initial manual exclusion of noisy structures or organs close to the lesion and so could not be considered as an automatic method. Consequently, results were not incorporated to avoid bias introduced by the manual delineation. For the ungated data, the 40% fixed threshold was possible but also required few manual adjustments for about 50% of the lesions.

This exclusion of surrounding structures was not necessary when the MAut automatic iterative approach was used, thanks to the 3D connexity criterion included in the method. Moreover, accuracy of DV measurements was improved with the automatic approach when using gated compared to ungated images. For all contrasts and lesion locations, the median, 1st and 3rd quartile errors were −9.7, −15.6 and −3.9 mL and −21.5, −33.8 and −7.6 mL for DV_5Aut and DV_UnAut respectively yielding larger DV for gated data as illustrated in Fig. 8. Whatever the lesion, DV_5Aut was always underestimated. The best results were obtained when the RCS in 8 frames was used (−6.0, −12.0 and 0.6 mL for the median, 1st and 3rd quartile errors).

4. Discussion

4.1. Simulations

In this paper, our first goal was to obtain images as close as possible to our clinical data. To that end it was necessary to use Monte-Carlo simulation instead of analytical simulations, in order to produce simulated PET raw data as close as possible of the real PET data. We were limited in the choice of the MC simulation platform to the GATE platform which was the only one allowing the modelisation of any specific imaging device and the simulation of qualitatively realistic data (access to the raw data, allowing the random correction with the delayed window). Moreover, the use of GATE instead of other simulator package, will make it possible to reuse the simulated data in order to implement any software improvement such as the time-of-flight (TOF) correction.

Table 3 – Lesion detectability as a function of the simulated RCS (ungated, 3, 5 or 8 frames) and the lesion location and diameter.

		Number of detected lessions								Total
	Location	Lung				Liver				
	Lesion diameter (cm)	1	2	3	4	1	2	3	4	
	Number (n) simulated	16	16	16	16	4	4	4	4	80
RCS	Ungated	0	4	8	8	1	3	3	3	30
	3	0	12	12	12	2	3	3	3	47
	5	0	8	8	12	1	3	3	3	38
	8	0	8	8	8	1	3	3	3	34

Table 4 – Errors on the estimation of the displacement volume for DV_8MAut, DV_5MAut, DV_3MAut, DV_UnMAut and DV_8M40%.

Errors (mL)	Mean	1st Quartile	Median	3rd Quartile
DV_8M$_{Aut}$	−5.9	−12.0	−6.0	0.6
DV_5M$_{Aut}$	−11.0	−15.6	−9.7	−3.9
DV_3M$_{Aut}$	−15.7	−20.0	−15.4	−9.4
DV_UnM$_{Aut}$	−19.9	−33.8	−21.5	−7.6
DV_UnM$_{40\%}$	−17.2	−32.2	−19.6	−7.8

Simulation time is a crucial issue in term of simulation process. Unfortunately, the GATE version we used was not optimized in term of computation time, which was a limiting factor to integrate patient uptake variability in our dataset as it was proposed by other [7,8]. We focussed our work on the simulation of a large number of lesions (with different size and contrast) considering different gating protocols (3–8 time bins) instead of a limiting number of lesions with variable environment in a unique acquisition protocol (ungated or one respiratory gating protocol). It was not possible to simulate organ uptake heterogeneity as in real patients, neither to propose patient uptake variability due to the prohibitive calculation time necessary to produce respiratory gated Monte-Carlo simulation. Identically, it was not possible to generate more complex lesion movements without simulating a new complete set of HS and TO data.

For these simulations, we decided to truncate the NCAT phantom to remove the arms, usually positioned above the patient head. This truncation was used both for the emission and the attenuation map, so that with no impact in the attenuation or scatter corrections.

The system dead-time was not simulated in this study. Our goal was only to simulate clinically realistic data in term of image quality, and not to simulate the exact physical processes like it would be necessary for detector or scanner geometry design comparison. The simulation of dead time would dramatically increase the simulation computation time without any major impact on image quality.

We included light sharing in our model of the detector response, unlike other groups who ignored the effect or modeled it using a Gaussian blurring [23]. This second solution is probably correct for high counting rates but in the case of a patient realistic simulation, counting rates are rather low. We modeled light sharing by spreading the detected events over the surrounding detector, considering a Gaussian probability distribution, to keep integers in the lines of response, unlike what a convolution by a Gaussian blur would do. The same approach was integrated in GATE V5.0 and following.

4.2. Data recombination and images reconstruction

As previously described, simulating independently the HS and TO significantly reduced the processing time and make it possible to re-use the simulated data for other studies. Assuming N1 minutes were required to simulate one of the 8 initial emission frames of the HS (acquisition time of 1 min) whereas simulation of the associated tumor with the highest activity (liver, diameter = 4 cm and S/B = 8) required N2 minutes, the generation of each new patient was thus at least accelerated by a factor of N3 = ((N1 + N2))/N2.

The main drawback of this methodology is that the potential scattering of some HS event occurring inside the final patient lesion cannot be taken into account. Similarly, the modeling of randoms generated from the detection of singles coming from the HS and TO simulations was not possible. Random recombination could be perfectly corrected using the
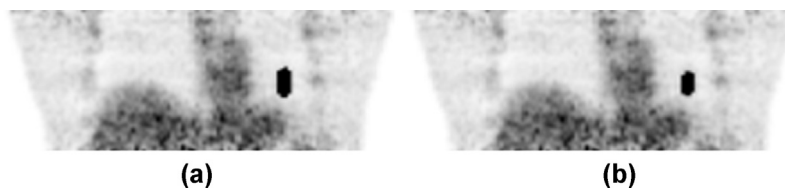

(a) **(b)**

Fig. 8 – Example of the DV segmentations, corresponding to the volume covered by the lesion during breathing, obtained using the automatic iterative method (M$_{Aut}$) for a lesion located in the lower lobe of the left lung (diameter = 2 cm, contrast = 8), estimated on the gated (a) and on the ungated (b) scans. The gated acquisition (a) allows to better restore the lesion displacement and so the volume covered by the lesion during the respiratory cycle.

high energies single data but at the expense of a more complicated post-processing.

The validity of our data recombination process was checked in one single case: a 4 cm lung lesion, with a low image statistic of 60 s (one of the 8 frames). A more extensive validation was published recently, based on other set of Monte-Carlo simulated data [8].

4.3. Detectability

This study emphasizes the limitation of visual detectability of moving lesion regarding their local contrast and volume. Indeed, in the best configuration (RCS in 3 frames), only 47 among the 80 simulated lesions are visible. None of the smallest (1 cm diameter) or the lowest contrasted pulmonary lesion (S/B ratio = 2) was detected. It can be also noticed that the liver lesions were more detectable thanks to a higher count rate in this organ compare to the lung. Indeed the S/B ratios were the same, only the background activity surrounding the lesion was different (i.e. lungs and liver tissue respectively). Finally, the clinical extrapolation of these results is to take with caution, since physiologically, the pulmonary lesion are often more contrasted than the lesions located in the liver.

4.4. Motion compensation technique and segmentation

The PET gating method subdivides the PET raw data in phases according to the respiratory cycle frequency [4]. The PET acquisition time is typically multiplied by the number of phases to maintain a sufficient statistic count in each phase. To avoid increasing the duration of PET acquisitions (due to the loss of statistics related to compensation techniques), more complicated techniques to correct for respiratory motion were proposed. Motion correction is usually performed by introducing a deformation field before, during, or after the data reconstruction. In this paper, we considered the lesion from the radiotherapist point-of-view, that is, we wanted to better estimate the volume covered by the lesion during the respiratory cycle. In this context, multigate compensation techniques are the only methods capable of describing the lesion's path during the respiratory cycle. Dawood et al. [24] demonstrated in a single case study that a 6-phase decomposition was optimal. On physical phantom data, Bettinardi et al. [25] determined that the lower the number of phases, the lower the SUVmax quantification. Our results suggest that the best trade-off between lesion detectability, volume estimation and noise is obtained for 5 gating bins when using multi-gate compensation techniques. This result was already found by Park et al. [26] based on images of physical phantom acquisitions, although the scan equivalent time used per frame was higher than ours (4 min/frame vs 3 min/frame).

Using more complicated motion-correction techniques (i.e. correction integrated during the reconstruction), the optimal number of phases would have probably been different, since with movement correction approaches, the higher the number of phases, better the blurring removal. Furthermore, the degree of motion compensation differs from one phase to another, due to differences in tissue velocity in the different phases of the respiratory cycle; some phases correspond to a very steep part of the respiratory cycle, where the tissue velocity is high. Moreover, respiratory amplitude varies significantly from one individual to another. Bettinardi et al. [25] emphasized that the number of gates needed is related to a lesion's size and displacement.

As expected, accuracy of DV measurement was improved with the automatic approach when using gated compared to ungated images. However, whatever the lesion, DV_MAut was always underestimated. This might be due to the breathing cycle undersampling which produces a reduction in the local lesion-to-background contrast due to signal spreading. In addition, the number of counts in the images was divided by the number of gated frames. *ContMeas* was not accurate, leading to an overestimation of ThOpt and therefore to DV underestimation.

It can be noticed that our analysis used only a limited amount of our simulated data to illustrate the usefulness of such realistic simulated data. Those data will further be used to investigate the influence of the RCS on lesion volume measurement (using more complicated lesion geometry and heterogeneity), SUV estimates or to assess the accuracy of partial volume corrections.

5. Conclusion

We generate clinically realistic respiratory gated ^{18}F-FDG-PET images and assess the impact of respiratory motion and respiratory-motion compensation in PET on lesion detection and volume measurement. Different physical phenomenon can affect the final image quality. To obtain reconstructed images as close as possible to those obtained in clinical conditions, a particular attention was paid to apply to the simulated data the same correction and reconstruction processes as those applied to real clinical data. We then showed that respiratory-gating of ^{18}F-FDG PET images improves the lesion detectability and the accuracy of DV estimates using an automatic iterative thresholding approach, compared to using non-gated PET images. This improvement might have a significant impact when patient treatment is performed using external radiotherapy. Further investigations will focus on the segmentation of more complex lesions.

REFERENCES

[1] C.B. Caldwell, K. Mah, M. Skinner, C.E. Danjoux, Can PET provide the 3D extent of tumor motion for individualized internal target volumes? A phantom study of the limitations of CT and the promise of PET, Int. J. Radiat. Oncol. Biol. Phys. 55 (2003) 1381–1393.

[2] C.B. Caldwell, K. Mah, Y.C. Ung, et al., Observer variation in contouring gross tumor volume in patients with poorly defined non-small-cell lung tumors on CT: the impact of

18FDG-hybrid PET fusion, Int. J. Radiat. Oncol. Biol. Phys. 51 (2001) 923–931.

[3] S.A. Nehmeh, Y.E. Erdi, C.C. Ling, et al., Effect of respiratory gating on quantifying PET images of lung cancer, J. Nucl. Med. 43 (2002) 876–881.

[4] A. Pepin, J. Daouk, P. Bailly, S. Hapdey, M.E. Meyer, Management of respiratory motion in PET/computed tomography: the state of the art, Nucl. Med. Commun. 35 (2014) 113–122.

[5] I. Buvat, I. Castiglioni, Monte Carlo simulations in SPET and PET, Q. J. Nucl. Med. 46 (2002) 48–61.

[6] F. Lamare, T. Cresson, J. Savean, R.C. Cheze Le, A.J. Reader, D. Visvikis, Respiratory motion correction for PET oncology applications using affine transformation of list mode data, Phys. Med. Biol. 52 (2007) 121–140.

[7] A. Le Maitre, W. Segars, S. Marache, et al., Incorporating patient specific variability in the simulation of realistic whole body ^{18}F-FDG distributions for oncology applications Nuclear Science Symposium Conference Record, 97, IEEE, 2009, pp. 2026–2038.

[8] S. Stute, S. Vauclin, H. Necib, et al., Realistic and efficient modeling of radiotracer heterogeneity in Monte Carlo simulations of PET images with tumors, IEEE Trans. Nucl. Sci. 59 (2012) 113–122.

[9] S. Vauclin, K. Doyeux, S. Hapdey, A. Edet-Sanson, P. Vera, I. Gardin, Development of a generic thresholding algorithm for the delineation of ^{18}FDG-PET-positive tissue: application to the comparison of three thresholding models, Phys. Med. Biol. 54 (2009) 6901–6916.

[10] S. Jan, G. Santin, D. Strul, et al., GATE: a simulation toolkit for PET and SPECT, Phys. Med. Biol. 49 (2004) 4543–4561.

[11] D. Lazaro, C. Michel, B. Bendriem, I. Buvat, Monte Carlo simulation of the Hi-Rez using GATE [abstract], J. Nucl. Med. 46 (2005) 487P.

[12] N. Karakatsanis, N. Sakellios, N.X. Tsantilas, et al., Comparison evaluation of two commercial PET scanners, ECAT HR+ and Biograph 2, using GATE, Nucl. Instrum. Meth. Phys. Res. A 569 (2006) 368–372.

[13] P. Gonias, N. Bertsekas, N. Karakatsanis, et al., Nucl. Instrum. Meth. Phys. Res. A 571 (2007) 263–266.

[14] W. Segars, Development and Application of the New Dynamic NURBS-based Cardiac-Torso (NCAT) Phantom (Thesis/Dissertation), 2001.

[15] M.E. Casey, E.J. Hoffman, Quantitation in positron emission computed tomography: 7. A technique to reduce noise in accidental coincidence measurements and coincidence

efficiency calibration, J. Comput. Assist. Tomogr. 10 (1986) 845–850.

[16] C.C. Watson, New, faster, image-based scatter correction for 3D PET, IEEE Trans. Nucl. Sci. 47 (2000) 1587–1594.

[17] P. Tylski, S. Stute, N. Grotus, et al., Comparative assessment of methods for estimating tumor volume and standardized uptake value in (18)F-FDG PET, J. Nucl. Med. 51 (2010) 268–276.

[18] S. Vauclin, S. Hapdey, C. Michel, et al., Monte Carlo simulations of respiratory gated ^{18}F-FDG PET for assessment of volume measurements methods, in: Nuclear Science Symposium Conference Record NSS'08, IEEE, 2008, pp. 4050–4053.

[19] Y.E. Erdi, O. Mawlawi, S.M. Larson, et al., Segmentation of lung lesion volume by adaptative positron emission tomography image thresholding, Cancer Res. 80 (1997) 2505–2509.

[20] J.F. Daisne, M. Sibomana, A. Bol, T. Doumont, M. Lonneux, V. Gregoire, Tri-dimensional automatic segmentation of PET volumes based on measured source-to-background ratios: influence of reconstruction algorithms, Radiother. Oncol. 69 (2003) 247–250.

[21] K. Doyeux, S. Vauclin, S. Hapdey, et al., Reproducibility of the adaptive thresholding calibration procedure for the delineation of ^{18}F-FDG-PET-positive lesions, Nucl. Med. Commun. 34 (2013) 432–438.

[22] R.D. Badawi, M.P. Miller, D.L. Bailey, P.K. Marsden, Randoms variance reduction in 3D PET, Phys. Med. Biol. 44 (1999) 941–954.

[23] F. Lamare, A. Turzo, Y. Bizais, C.C. Le Rest, D. Visvikis, Validation of a Monte Carlo simulation of the Philips Allegro/GEMINI PET systems using GATE, Phys. Med. Biol. 51 (2006) 943–962.

[24] M. Dawood, F. Buther, N. Lang, O. Schober, K.P. Schafers, Respiratory gating in positron emission tomography: a quantitative comparison of different gating schemes, Med. Phys. 34 (2007) 3067–3076.

[25] V. Bettinardi, E. Rapisarda, M.C. Gilardi, Number of partitions (gates) needed to obtain motion-free images in a respiratory gated 4D-PET/CT study as a function of the lesion size and motion displacement, Med. Phys. 36 (2009) 5547–5558.

[26] S.J. Park, D. Ionascu, J. Killoran, et al., Evaluation of the combined effects of target size, respiratory motion and background activity on 3D and 4D PET/CT images, Phys. Med. Biol. 53 (2008) 3661–3679.