

Optimal gating compared to 3D and 4D PET reconstruction for characterization of lung tumours

Wouter van Elmpt · James Hamill · Judson Jones ·
Dirk De Ruyscher · Philippe Lambin · Michel Öllers

Received: 2 September 2010 / Accepted: 16 December 2010 / Published online: 11 January 2011
© The Author(s) 2011. This article is published with open access at Springerlink.com

Abstract

Purpose We investigated the added value of a new respiratory amplitude-based PET reconstruction method called optimal gating (OG) with the aim of providing accurate image quantification in lung cancer.

Methods FDG-PET imaging was performed in 26 lung cancer patients during free breathing using a 24-min list-mode acquisition on a PET/CT scanner. The data were reconstructed using three methods: standard 3D PET, respiratory-correlated 4D PET using a phase-binning algorithm, and OG. These datasets were compared in terms of the maximum SUV (SUV_{max}) in the primary tumour (main endpoint), noise characteristics, and volumes using thresholded regions of SUV 2.5 and 40% of the SUV_{max}. **Results** SUV_{max} values from the 4D method (13.7±5.6) and the OG method (14.1±6.5) were higher (4.9±4.8%, $p < 0.001$ and 6.9±8.8%, $p < 0.001$, respectively) than that from the 3D method (13.1±5.4). SUV_{max} did not differ between the 4D and OG methods (2.0±8.4%, $p = \text{NS}$). Absolute and relative threshold volumes did not differ between methods, except for the 40% SUV_{max} volume in which the value from the 3D method was lower than that from the 4D method (−5.3±7.1%, $p = 0.007$). The OG method exhibited less noise than the 4D method. Variations in volumes and SUV_{max} of up to 40% and 27%, respectively, of the individual gates of the 4D method were also observed.

Conclusion The maximum SUVs from the OG and 4D methods were comparable and significantly higher than that from the 3D method, yet the OG method was visibly less noisy than the 4D method. Based on the better quantification of the maximum and the less noisy appearance, we conclude that OG PET is a better alternative to both 3D PET, which suffers from breathing averaging, and the noisy images of a 4D PET.

Keywords PET · Lung cancer · Respiratory motion · Respiratory gating · SUV

Introduction

Accurate delineation of tumour lesions in lung cancer based on PET/CT is required in radiotherapy treatment planning. Moreover, as the amount of ¹⁸F-fluorodeoxyglucose (FDG) uptake is prognostic for survival, precise measurement of the standardized uptake value (SUV) may be of clinical value [1–3]. However, delineation of lung tumours is fraught with error and suffers from significant inter- and intraobserver variability [4, 5]. Tumour definition based on FDG-PET/CT is the current standard, but the method is far from ideal [6–9]. In particular, the FDG uptake pattern in 3D PET imaging suffers from breathing averaging, which prevents accurate SUV and volume measurements.

Respiratory-correlated PET imaging (4D PET) is an attractive alternative [10–15]. In 4D PET, a surrogate measurement of the patient's breathing is made during the scan, and based on this information the PET data are sorted according to respiratory phase or amplitude. Hybrid methods have also been proposed [16]. Phantom studies have shown that 4D PET predicts lesion volume recovery more accurately than 3D PET and yields more accurate SUV values

W. van Elmpt (✉) · D. De Ruyscher · P. Lambin · M. Öllers
Department of Radiation Oncology (MAASTRO), GROW –
School for Oncology and Developmental Biology, Maastricht
University Medical Centre,
Dr. Tanslaan 12,
NL-6229 ET, Maastricht, The Netherlands
e-mail: wouter.vanelmpt@maastro.nl

J. Hamill · J. Jones
Siemens Medical Solutions,
Knoxville, TN 37932, USA

[17–22]. In spite of its promise, 4D PET has had limited clinical acceptance. Perhaps the most significant limitation is the technical challenge of making the respiratory measurement in a clinical setting, handling the large amount of 4D data, and applying the respiratory motion information in treatment planning. Moreover, 4D PET images have relatively more statistical noise than 3D PET images, since only a fraction of the acquired data is used in each image. There is concern about inconsistent attenuation correction in 4D PET due to mismatches between PET and CT [23]. Even if 4D CT is used for attenuation correction of the corresponding 4D PET, errors due to the mismatch of CT and corresponding PET gates may be introduced due to differences in the binning methods used for reconstructing 4D CT and 4D PET [24]. Despite its drawbacks, 3D PET is still used for tumour delineation and characterization of tumours and lymph nodes.

In this study evaluated a new method, optimal gating (OG), for applying respiratory-correlated measurements to PET. OG finds a single amplitude interval which minimizes the blurring due to respiratory motion while maximizing the number of coincidence events. We compared 3D, 4D and OG PET in terms of SUV, tumour volume and noise.

Materials and methods

Patient overview and PET acquisition

An FDG-PET scan was performed in 26 patients with lung cancer treated in the period between August 2008 and March 2009. The scan was performed during free breathing using a Siemens Biograph 40 PET/CT scanner (Siemens Medical Solutions, Knoxville, TN, USA) for radiotherapy treatment planning purposes. Patients were positioned supine with a dedicated arm support in the radiotherapy position. Our standard clinical protocol included a 4D CT scan and a 24-min list-mode PET acquisition of a single bed position (16.2 cm axial extent) centred on the primary tumour. The signal was measured using a respiratory monitor system (RMON) with a pressure sensor in a belt strapped around the patient's chest (AZ-733 V; Anzai Medical Corporation, Tokyo, Japan) [25, 26]. The injected activity (megabecquerel) of FDG depended on the weight (kilograms) of the patient and for the first nine patients in the study was four times the body weight plus 20 MBq. For the second set of 17 patients, the injected activity was 2.5 times the body weight following the recommendations of Boellaard et al. [27].

3D, 4D and OG PET image reconstruction

The list-mode data were reconstructed using three methods: (1) standard 3D PET reconstruction, resulting in a free-

breathing PET image for a 24-min acquisition; (2) phase-based 4D PET reconstruction, with eight gates; and (3) reconstruction using the OG algorithm.

The OG PET reconstruction was applied after processing the respiratory signals, $s(t)$, acquired during the scan. Prototype software, described in Appendix A was used to create a modified PET data list with the motion largely removed, and this modified list was processed normally on the clinical PET/CT system. The first step in the OG method is to form a histogram of the $s(t)$ values. Second, the histogram is converted to a cumulative distribution function, $cdf(s)$, representing the probability of observing a signal of amplitude s or greater. Third, an algorithm considers all possible combinations of lower (L) and upper (U) levels, subject to a constraint on U that forces the total sensitivity to equal a specific percentage (e.g. 35%) of the acquired breathing signals. The specific percentage is a parameter that is called the optimal gating yield parameter, OG_{yield} . The constraint can be written as:

$$cdf(U) - cdf(L) = OG_{yield} \quad (1)$$

To emphasize that this is a parameter, we write:

$$OG_{yield} = 0.35. \quad (2)$$

A rationale for choosing the specific value of 0.35 is presented in Appendix B. The OG method determines the narrowest such interval by selecting the L value which makes the difference $U - L$ as small as possible. In most patients, the result is a narrow range of breathing amplitudes near the end-expiration phase. Figure 1 shows a 30-s plot of breathing amplitudes in a representative patient study, along with the histogram from the entire 24-min acquisition. The optimal amplitude interval is represented by a shaded region representing a range of breathing amplitudes, or alternatively a portion of the histogram.

For all PET image reconstructions, attenuation correction was performed using the maximum exhalation phase of the 4D CT scan. In most patients, this is the most representative phase. For some patients 4D CT images were not available at the maximum exhalation phase. In these cases we used the CT phase closest to maximum exhalation. Reconstructions were performed using a Fourier-rebinning-based OSEM 2D algorithm with four iterations and eight subsets, with a final image size of 168×168 pixels with a typical pixel size of approximately 4 mm and a 3-mm slice thickness.

Phantom validation

To validate the OG PET reconstruction we performed a phantom study. We used a sphere with a diameter of approximately 3.2 cm and volume of 17 cm^3 filled with 10 MBq of FDG that was mounted on a motorized platform

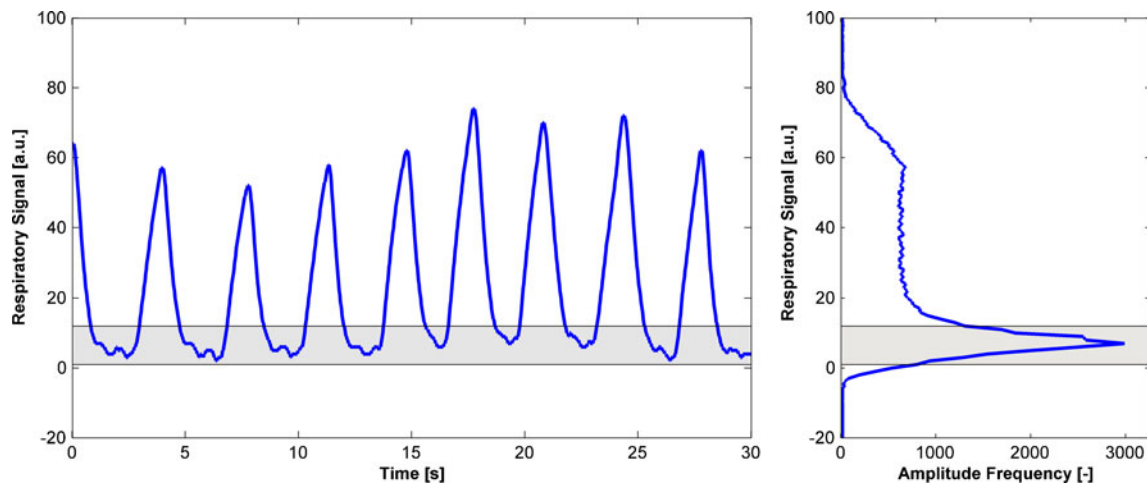


Fig. 1 Example breathing pattern (*left*) over 30 s showing the optimal gating window. The histogram (*right*) shows the amount of breathing amplitudes for the entire 24-min list-mode acquisition. The OG

method selects the narrowest bandwidth (*shaded area*) containing 35% of the respiratory signal

(Respiratory Phantom; Anzai Medical Corporation, Tokyo, Japan). The RMON pressure sensor was attached to the platform to monitor the motion. The sphere moved at a rate of ten cycles per minute on a trajectory with an amplitude of 1.7 cm. The period and the amplitude of the motion correspond to the movement of lung tumours *in vivo*. A 10-min acquisition time was used for these experiments. Signal quality was very high because the source was strong and there was little attenuating medium around the phantom. A 4D CT scan was also acquired. In a similar manner to the PET reconstructions for the patients, 3D, 4D and OG PET reconstructions were performed. The average over all eight gates of the 4D PET measurements was used as the reference value to which the other reconstructions were compared.

Analysis of PET images

PET images of the patients were analysed using TrueD software (version VC60; Siemens, Erlangen, Germany). For all ten datasets, i.e. the 3D PET dataset, the 4D PET dataset with eight gates and the OG PET dataset, the volume of interest (VOI) was selected manually in the transverse, sagittal and coronal anatomical planes around the primary tumour. This VOI was carefully selected so as not to include any high FDG uptake regions of nonprimary tumour tissue such as the heart or the involved mediastinal lymph nodes. In the VOI the threshold for volume calculation was based on two autocontouring criteria: first, a fixed SUV threshold of 2.5 [28], and second, a relative threshold of 40% of the maximum SUV (SUV_{max}) in the primary tumour [8, 29]. We recorded the volume of the region to which the threshold was applied, and the minimum, average, maximum and standard deviation of SUV values inside the region.

To provide a measure of image noise, we created an additional VOI in the contralateral lung, selecting a region that appeared to be homogeneous. This VOI was copied to all datasets, and the standard deviation of the SUV values was computed and recorded as a fraction of the mean value in the VOI.

In addition, the displacement of the centre of mass of the SUV volumes in each of the corresponding gates of the 4D PET was automatically calculated in three dimensions by the TrueD software. A 3D motion vector, defined as the square root of the quadratic sum of motion amplitudes in the transverse, sagittal and coronal directions, was determined for each patient.

Evaluation strategy

Our hypothesis was that OG would produce less blur than 3D reconstruction and less noise than 4D reconstruction, and yet be quantitatively accurate. The most important objective in the analysis was to retrieve quantitative values for the maximum SUV inside the primary tumour. Because a gold standard was missing, we chose the average of the maximum SUV in the 4D PET as the reference. As a secondary objective, we quantified differences in the volume as determined by different reconstruction methods and different delineation criteria. Finally, noise in the images was evaluated to give an objective surrogate for image quality.

Statistical analysis

Paired variables were compared using a Wilcoxon Signed Rank's test (Matlab R2009a; The Mathworks, Natick, MA). Because a Bonferroni correction was applied to

correct for multiple testing (i.e. three comparisons are performed in this analysis), p values obtained from statistical testing were multiplied by 3. P values smaller than 0.05 were assumed to be statistically significant. Results are expressed as means \pm standard deviation, unless mentioned otherwise.

Results

Phantom experiment

Figure 2 shows coronal sections through 4D, 3D, and OG images from the phantom experiment. The 3D PET image was visibly blurred, whereas the OG image appeared to be as sharp as the 4D images. Quantitative differences in maximum pixel value were similar in all measurements. Compared to the average over all phases of the 4D PET dataset, the 3D image was 1.2% lower and the OG image was 1.2% higher. The variation in SUVmax within the 4D acquisition was 1.5%. The volumes were compared by computing autothresholded volumes using the 40% of the maximum SUV criterion. The volume in 4D PET was $15.8 \pm 0.3 \text{ cm}^3$, averaged over phases. The volume in 3D PET was 13.8 cm^3 . The volume in OG PET was 15.1 cm^3 . Average SUVs within autothresholded volumes varied by 4.7% across gates in the 4D PET acquisition, and in OG PET the average SUV was 3.4% higher. In 3D PET, the average SUV was 10.2% lower.

Patient characteristics

Table 1 shows the patient characteristics. The average length of the motion vector in 3D of the primary lung tumour of the 26 patients was $3.8 \pm 2.7 \text{ mm}$, with a range between 1.3 and 11.3 mm. Examples of the 3D, 4D and OG PET images are shown in Fig. 3.

SUVmax values

Table 2 compares average SUVmax values in the primary tumours. The maximum values for the 3D, 4D and OG PET reconstruction methods were 13.1 ± 5.4 , 13.7 ± 5.6 and 14.1 ± 6.5 , respectively. The large standard deviations were caused by the large interpatient variability and SUVmax ranges from approximately 5 to 35. The SUVmax values for the 3D PET method were significantly lower than the values for the 4D and OG methods. The differences were

$4.9 \pm 4.8\%$ ($p < 0.001$) lower than the 4D values, and $6.9 \pm 8.8\%$ lower than the OG values ($p < 0.001$; Table 3). The average SUVmax values for the OG and 4D methods were similar, the differences being $2.0 \pm 8.4\%$ ($p = \text{NS}$). Individual SUVmax values are shown in Fig. 4.

Differences were observed between the SUVmax of the different gates of the 4D PET images. In several cases, these variations exceeded 10%; this is shown in Fig. 5. The standard deviation of the SUVmax values within the individual patients, expressed as ratios in relation to the mean 4D PET value, was $4.0 \pm 2.1\%$ for the entire patient population, ranging from 1.9% (patient 19) to 12.7% (patient 17). In the case of patient 17, the SUVmax ranged from 6.8 to 10.0 with a median value of 7.7 for a small tumour of 1.7 cm^3 .

Fixed SUV threshold

The volumes computed with autocontours at a fixed threshold of 2.5 SUV were not significantly different between the three methods (the differences were of the order of 1 cm^3): $74.2 \pm 76.8 \text{ cm}^3$, $75.1 \pm 77.9 \text{ cm}^3$ and $74.9 \pm 76.2 \text{ cm}^3$ for the 3D, 4D and OG methods, respectively. The average SUV values in the volume within the contour were 5.4 ± 1.2 , 5.4 ± 1.1 and 5.6 ± 1.3 , and the values from both the 3D and 4D methods differed slightly, but significantly ($p < 0.001$), from the average OG PET values.

Relative SUV threshold

Volumes computed by the 40% thresholding criterion did show significant differences between the 3D and 4D methods, but otherwise no significant differences were seen. This is shown in Table 2. A significant reduction of $5.3 \pm 7.1\%$ ($p = 0.007$) in the average 4D volume compared to the 3D volume was found, although the average absolute numbers were all within 2 cm^3 . Figure 5 shows the variation in autothresholded volumes for the different gates of the 4D reconstruction. As an example, considering all gates for patient 17, the volumes determined by the 40% SUVmax method in this dataset varied between 61% and 139% of the mean volume. The average SUV values within the 40% SUVmax contour were 7.8 ± 3.1 , 8.1 ± 3.2 and 8.4 ± 3.8 , for 3D, 4D and OG methods, respectively. Again, average values for the 3D method differed significantly from those for the OG and 4D methods, whereas there was no significant difference between the values for the OG and 4D methods.

Fig. 2 Phantom experiment showing 4D PET reconstructed images together with the motion blurred (static) 3D PET and the OG PET images

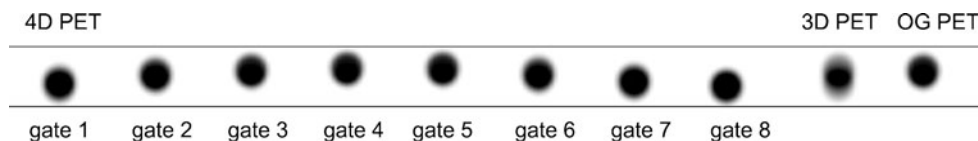


Table 1 Patient characteristics

Characteristic	Value
Gender	
Male	16
Female	10
Age (years)	
Median	67
Range	55–83
Cancer type	
NSCLC	14
SCLC	2
NSCLC stage	
I/II	4
IIIA	8
IIIB	8
IV	4
Location of primary tumour	
Left upper lobe	8
Left lower lobe	3
Right upper lobe	6
Right lower lobe	7
Central	2

Noise quantification

For the evaluation of image noise based on the contralateral lung, VOIs with an average size of $37.8 \pm 18.5 \text{ cm}^3$ were created. The standard deviations of the SUV values of the VOI inside the contralateral lung are shown for all patients in Fig. 6. Compared to the 3D PET method, the noise was higher for the 4D PET method ($89 \pm 52\%$, $p < 0.001$) and the OG PET method ($31 \pm 21\%$, $p < 0.001$). Noise for the OG PET method was $44 \pm 30\%$ ($p < 0.001$) lower than that for the 4D PET method.

Discussion

Accurate quantification of FDG uptake is important in oncology. PET imaging is used for staging and diagnosis of lung tumours, for quantifying the SUV, and for delineating the actual border of the tumour. For conventional radiotherapy schemes, a better delineation of the primary tumour will in general lead to smaller safety margins used in radiotherapy to define the treated volume. These smaller irradiated volumes will encompass less normal lung tissue

Fig. 3 Axial images from an example patient (patient 18) comparing OG, 3D and 4D PET reconstruction methods: *top row* CT images, *middle row* PET images, *bottom row* fused PET/CT images; *left column* OG reconstruction, *middle column* 3D reconstruction, *right column* 4D reconstruction

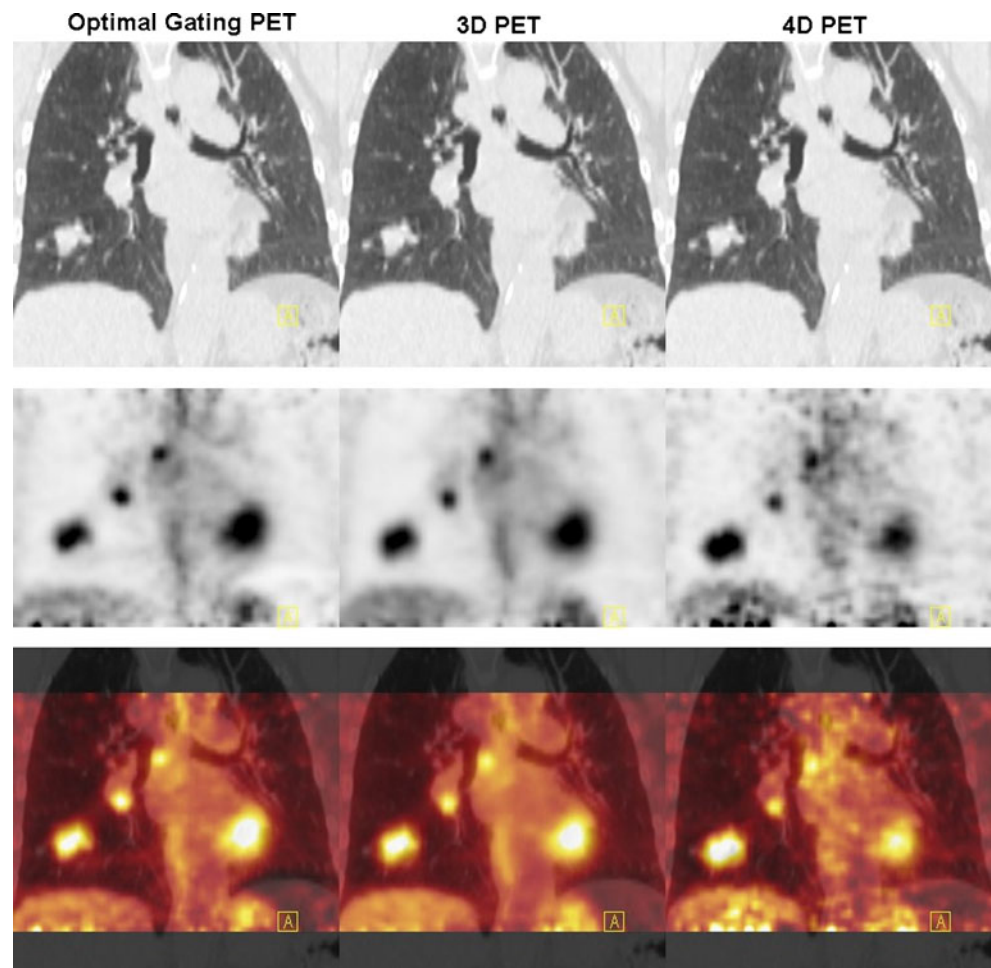


Table 2 Results of the 3D, 4D and OG PET reconstruction: maximum SUV, and volumes within the SUV 2.5 and 40% of the SUVmax cut-off (values are means±SD)

VOI	Parameter	3D	4D ^a	OG
n.a.	Maximum SUV	13.1±5.4	13.7±5.6	14.1±6.5
40% of SUVmax	Average SUV	7.8±3.1	8.1±3.2	8.4±3.8
	Volume (cm ³)	32.2±32.9	30.1±29.6	31.4±32.2
SUV>2.5	Average SUV	5.4±1.2	5.4±1.1	5.6±1.3
	Volume (cm ³)	74.2±76.8	75.1±77.9	74.9±76.2

^aThe 4D values are the averages over all eight gates of the dataset.

to be irradiated and hence lower (lung) toxicity, or if an isotoxic dose-escalation strategy is used this will lead to a higher tumour dose and hence higher overall survival in lung cancer patients [30]. Delineating subvolumes of the tumour are currently also a focus of research, as areas with high SUV are presumably more therapy-resistant [2, 31, 32]. Especially in the context of more advanced radiotherapy techniques that rely on dose painting or subvolume boosting strategies, accurate quantification of the PET uptake is essential to define these regions. Furthermore, accurate quantification becomes more important in current trials and clinical practice. In a large Dutch randomized trial (NVALT 8), patients were randomized between two treatment arms based on a SUVmax cut-off value of 7. Applying this to the cohort of patients analysed in this study would mean an increase of approximately 10% in the number of patients in the SUVmax >7 group if 4D or OG PET was used for quantification.

Validation of automatic segmentation methods and delineations in patients with lung cancer is difficult. Results from pathological studies are limited and hampered by technological difficulties in specimen extraction from thoracic surgery [4, 5]. Hence, a gold standard for tumour delineation is difficult to obtain [33]. Validation of new PET reconstruction methods for accurate SUV quantification techniques in vivo is also not directly possible and is limited to phantom evaluations and validation. We therefore performed a phantom validation, demonstrating correct

technical behaviour of the optimal gating method providing confidence that the method is applicable in a clinical setting. The phantom movement was similar, even slightly smaller, compared to the size. Hence, no large differences in SUVmax were expected for these reconstructions. The volume recovery of the 3D PET reconstruction differed from that of the 4D PET reconstruction, whereas that of the OG PET method was similar. However, the images in Fig. 2 clearly show that either the 4D or OG PET reconstruction methods was necessary to reduce blurring artefacts from respiratory motion. Figure 2 also shows the correct processing of the list-mode files and OG PET reconstruction technique.

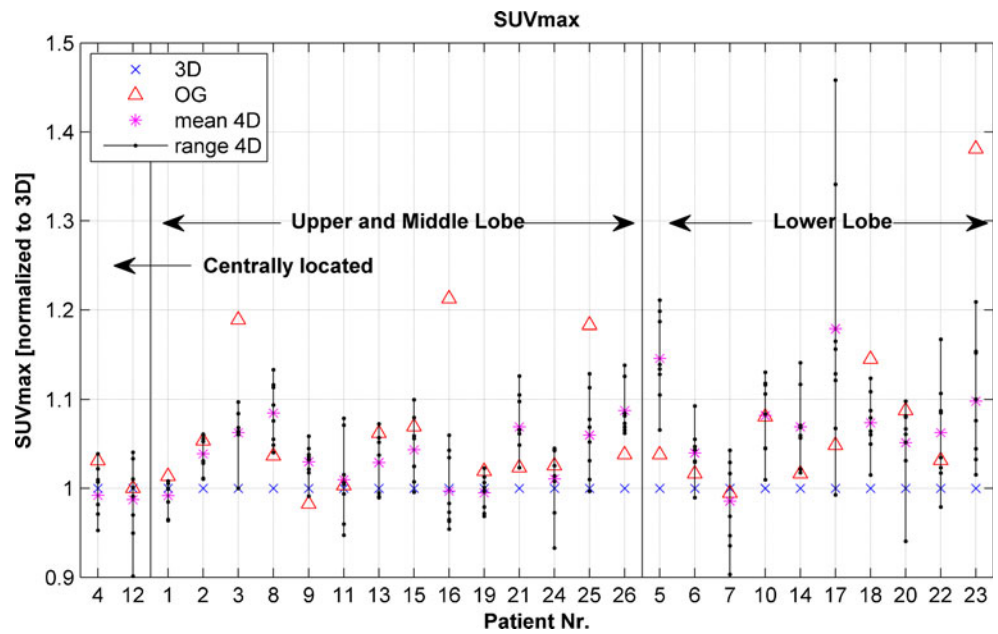
The prototype software was expected to be quantitative in the case of regular breathing, as described in Appendix A. The prototype software treated the OG sinogram as if it came from a static acquisition, and therefore could not accurately correct for effects that are correctly modelled in 4D PET, for example changes in activity distribution during the scan, detector dead time, and decay of the radioisotope. In spite of this limitation, in most scans the quantitative values (SUVmax) were similar for the OG and the average of the 4D method. Figure 4 demonstrates the close agreement and supports our hypothesis that motion blurring would be reduced by OG and that it is possible to extract accurate quantitative values from the OG method. In 3D PET, the SUVmax was significantly reduced due to respiratory motion blurring.

Table 3 Differences between maximum SUV, volumes within the SUV 2.5 and 40% of the SUVmax cut-off for the 3D, 4D and OG PET reconstruction

VOI	Parameter	Difference 3D vs. OG PET (%)		Difference 4D vs. OG PET (%)		Difference 3D vs. 4D PET (%)	
		Mean±SD	<i>p</i> value	Mean±SD	<i>p</i> value	Mean±SD	<i>p</i> value
n.a.	Maximum SUV	6.9±8.8	<0.001*	2.0±8.4	1.271	4.9±4.8	<0.001*
40% of SUVmax	Average SUV	6.4±9.7	<0.001*	2.7±8.3	0.547	3.5±4.1	0.001*
	Volume (cm ³)	-0.8±9.0	0.102	5.3±12.3	0.220	-5.3±7.1	0.007*
SUV>2.5	Average SUV	3.4±4.8	<0.001*	2.8±3.8	0.001*	0.6±1.8	1.271
	Volume (cm ³)	3.6±7.1	0.355	2.5±6.3	2.195	1.0±3.4	0.175

**p*<0.05.

Fig. 4 SUVmax values for the OG, 3D and 4D PET reconstruction methods classified according to tumour location inside the lung



Residual motion is also likely to play an effect in determining the SUVmax for the 4D phase-based PET reconstruction we used. Choosing an amplitude-based binning approach might reduce the variation to some extent by reducing the residual motion component, but the counts in each gate would vary and lead to different statistics for the various gates. For the volumes delineated using 40% of the SUVmax, the values from the 4D method were smaller. There could be several reasons for this. Accurate determination of SUVmax of the individual

phases was hampered due to increased noise leading to a somewhat larger SUVmax which resulted in smaller delineated volumes. Another possibility could be that the volume of the tumours in this study were quite large and motion was limited for these tumours, making volume definition in 3D already a good estimate of the actual volume of the lesion. For smaller tumours and in the case of more respiratory motion this might be different. However, our dataset was too limited to draw definite statistical conclusions.

Fig. 5 Top row: box plot of the SUVmax values normalized to the mean value for the different gates of the 4D PET method. Bottom row: box plot of the volumes defined by 40% of the SUVmax normalized to the mean volume for the individual patients. In both plots, the edges of the boxes represent the 25th and 75th percentiles, and the whiskers represent the extreme data points. Outliers are plotted individually (circles)

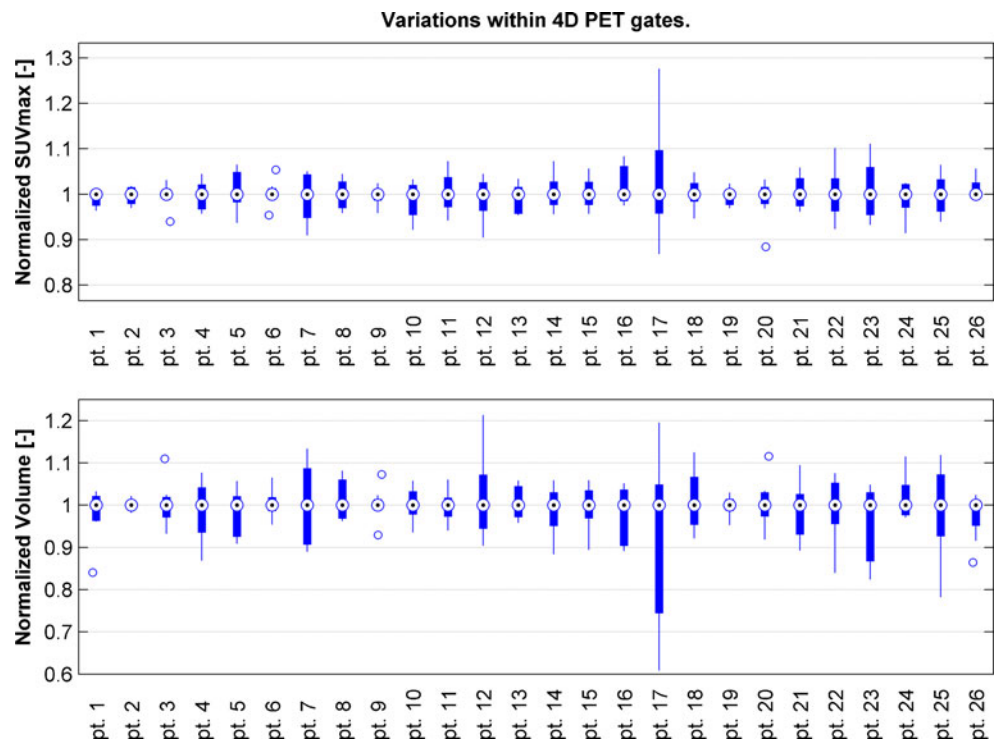
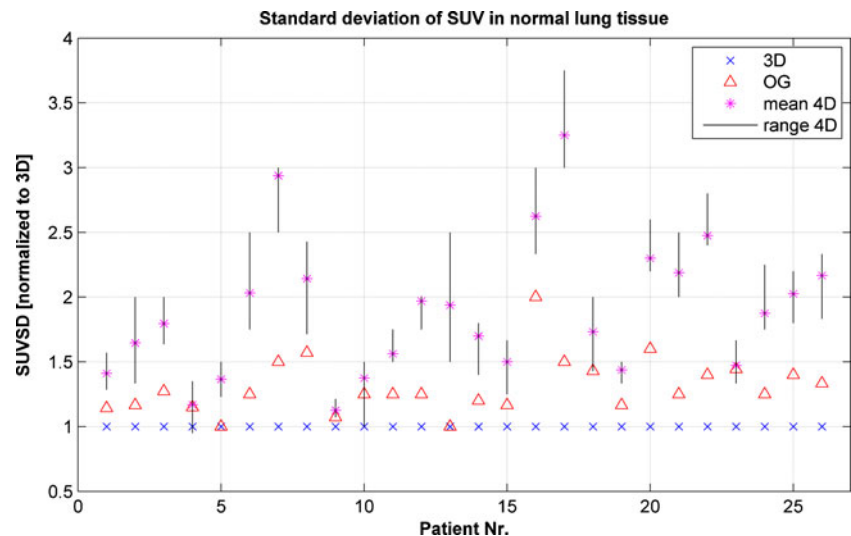


Fig. 6 Standard deviations of the SUV values inside the VOI in the contralateral lung tissue for all patients relative to the value of the 3D PET reconstruction



OG PET images were sharper than conventionally gated 4D PET images, and not as noisy. Since image quality is difficult to quantify directly in absolute numbers, we used a surrogate of noise defined as the standard deviation of the SUV in a homogeneous part of the lung, and this clearly showed reduced noise for OG PET compared to 4D PET. Also, OG PET images were visually sharper at the border of the primary tumour (Fig. 3) compared to the 3D PET images which had intrinsic motion blurring. The average motion of the primary tumour was about 4 to 5 mm which was comparable to the resolution and voxel size of the PET scanner. Some deterioration of the edges of the primary tumour was visible in the static 3D PET images.

Noise in the 4D PET images was also reflected in the variations between the various gates of the 4D reconstruction. Volume differences up to 40%, and differences as large as 27% in maximum SUV were shown in this study. These results were similar to variations reported by Erdi et al. [23]. These effects might be partly caused by a wrong attenuation correction due to a mismatch between the CT and PET scan [22]. Using the end-expiration phase of the 4D CT would be expected to minimize the errors caused by the attenuation correction. In patients who breathe normally, end-expiration is the most stable breathing phase. In our experience, the OG algorithm typically found an amplitude interval range near to that of the end-expiration phase. One could prospectively gate the CT scanner to acquire attenuation data at the respiratory amplitude determined by the OG calculation, and the patient would receive very little additional radiation exposure from the prospectively gated CT scan. This was not possible in the retrospective study described here. For patients with an irregular or nonphysiological breathing pattern, choosing a different CT gate for the attenuation correction might circumvent possible errors in attenuation correction. However, for the

patient group described in this study all OG amplitude intervals were close to the end-exhale position.

Others have tried breath-hold techniques for motion freezing, mainly for the CT acquisition [34, 35]. For PET imaging, breath hold techniques might not be feasible because the patient needs to breath-hold repeatedly, because acquisition times are typically in the order of minutes, whereas the CT scan can generally be performed in a few seconds. Reproducibility of the breath-hold phase for multiple breath holds is difficult, and may require specially designed equipment. The approach described in this paper is simple, and comfortable for the patient as free breathing is allowed during the scan. If 4D images are not required for evaluation of the tumour motion, and OG images are to be used instead, the duration of the study can be significantly reduced, further increasing patient comfort.

The simplicity of the OG method is expected to make it comparatively robust. This simple method may not be able to recover all of the resolution possible with a PET scanner. It is our view that better resolution imaging might one day be achievable with accurate modelling of the complex motions that occur during human breathing, perhaps using methods such as optical flow-based methods of Dawood et al. [36]. In Appendix B we show that OG reduces the blurring from 1 to 2 cm, in extreme cases, to the range 6 to 7 mm that is normal in PET imaging today. Indeed, the OG methodology should work well in the imaging of small tumours, for example metastatic lesions. In such cases, the motion trajectory may exceed the size of the tumour. Another improvement of the OG method compared to 3D and 4D PET would be the detection of involved lymph nodes in the mediastinum. The OG PET might reveal these nodes whereas for the 3D PET these might be blurred due to motion and the 4D PET contains too much noise to distinguish them from the surrounding uptake. However, this has to be validated in clinical practice. Similarly, the

OG method is also applicable in cardiac imaging, where respiratory motion is in many cases larger in magnitude than the heart's contractile motions. Indeed, Frey et al. have recently reported examples in which the OG method significantly sharpened the images in tumour imaging and cardiac PET/CT, with better image quality than 4D PET [37].

However, motion amplitude of the primary tumour is difficult to predict and selection of 3D vs. 4D PET acquisition is frequently not possible prior to the imaging session. The OG reconstruction might be the intermediate step suitable for all patients providing good signal-to-noise ratios for quantitative analysis of the SUV parameters and sharp (non-motion-blurred) images for visual assessment.

Conclusion

The SUVmax for the OG and 4D PET reconstruction methods were comparable and were both significantly higher than that for the 3D PET reconstruction method. Optimal gating accurately determined the SUVmax values, while reducing image noise and providing accurate volume determination, comparable to that from 3D and 4D PET. Based on the better quantification of the maximum value and the less noisy appearance, we conclude that OG PET is a better alternative both to 3D PET, which suffers from breathing averaging, and to 4D PET, which produces noisy images.

Conflicts of interest James Hamill and Judson Jones are employees of Siemens Medical Solutions. MAASTRO has received technical and financial support from Siemens Medical Solutions for performing this study.

Open Access This article is distributed under the terms of the Creative Commons Attribution Noncommercial License which permits any noncommercial use, distribution, and reproduction in any medium, provided the original author(s) and source are credited.

Appendix A. Formation of respiratory gated images

The data flow for ordinary 4D PET is illustrated in Fig. 7a. A respiratory monitor system (RMON) analysed the respiratory amplitude waveform $s(t)$ from the patient in near-real time and provided a logic pulse, or trigger, whenever the patient's breathing reached a specified point in the breathing cycle. Typically, we set the system to provide triggers at 100% inspiration. The triggers were communicated to the PET scanner, which recorded these as gates in the list-mode file, along with other information needed for quantitative PET. A reconstruction computer converted the list to static (3D) images and 4D images.

Amplitude-based reconstructions were performed by modifying this system in a prototype configuration of computers and software, illustrated in Fig. 7b. The RMON included a computer that recorded the waveform, the triggers, and

timing information relative to the start of the respiratory acquisition (clock 2) in a separate respiratory data list file. These data and the PET data list were communicated to an auxiliary computer which converted them to a new optimally gated (pruned) data list, which was sent to the reconstruction computer to create the optimally gated image.

Clock 1 in the PET list and clock 2 in the respiratory data list ran at the same rate but were offset from one another by a time offset Δt . In each patient study, we determined Δt by digital calculation. Since the RMON was configured with a 0.025-s time resolution and the PET scanner had a 0.001-s time resolution, we defined two arrays in computer memory representing 0.025-s samples of the trigger signals in the RMON (the T array) and gate signals in the PET scanner (the G array). A value of 1 represented intervals in which the trigger or gate was on, and 0 represented intervals in which the trigger or gate was off. The length of each array represented a time interval at least twice as long as the acquisition time. Array indices represented the time from the beginning of the recorded information. Figure 7c shows the triggers and gates and their relation to the waveform. Alignment of the two clocks was based on aligning triggers with gates. This was analogous to matching fingerprints and used a technique that, like fingerprint matching, was robust in the case of missing data points. The correlation of T and G was:

$$c(\Delta t) = \langle T(t) \times G(t - \Delta t) \rangle, \quad (\text{A.1})$$

where $\langle \dots \rangle$ represents averaging. We calculated this with an approach based on the Fourier shift theorem [38]:

$$c(\Delta t) = F^{-1} \left(F(T) \times F(G)^* \right), \quad (\text{A.2})$$

where F is the discrete Fourier transform, F^{-1} is its inverse, and the asterisk represents complex conjugation. The correlation function for this patient study is shown in Fig. 7d. A sharp, prominent maximum defined the actual time offset. This can be expressed by the equation:

$$\Delta t_{\text{actual offset}} = \arg \max \{ c(\Delta t) \} \quad (\text{A.3})$$

The plot also shows broad peaks representing harmonics. When the gates were temporally shifted by $\Delta t_{\text{actual offset}}$, they were perfectly aligned with the triggers, as shown in Fig. 7e.

With the temporal relation between RMON and PET determined in this way for each patient study, we created the new optimally-gated PET data list using a program in the auxiliary computer. Event words from the PET scanner were included in the new list whenever the amplitude fell between the L and U thresholds defined in [Materials and Methods](#). Other events were discarded. The new list contained the optimally gated part of the original list, about 35%, and represented the conditions of a static study in

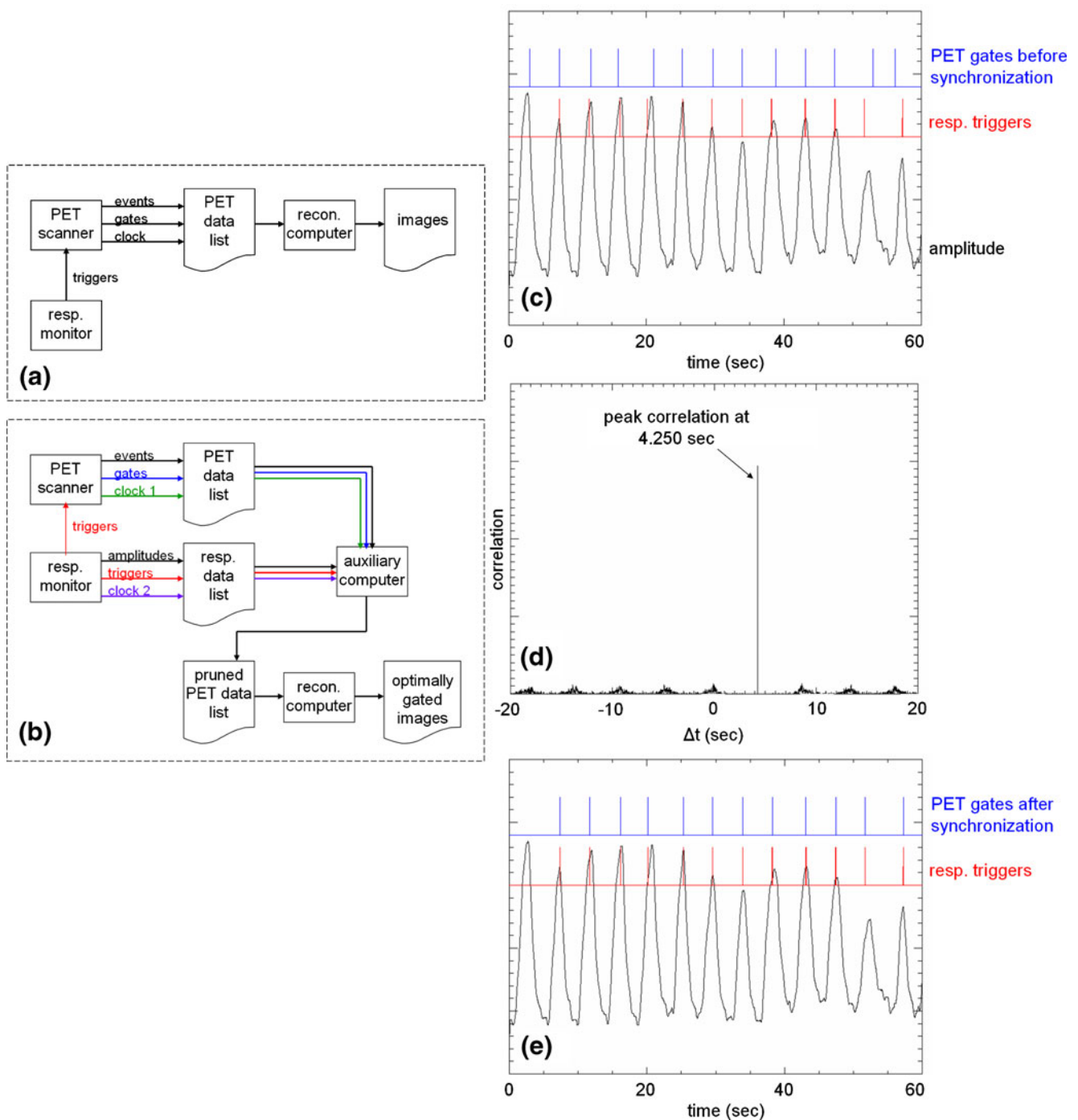


Fig. 7 Processing for respiratory gating. **a** Data flow in ordinary 4D PET. **b** Data flow with amplitude-based respiratory gating. **c** A 60-s trace showing gates before synchronization, triggers, and respiratory

amplitudes. **d** Time correlation between triggers and gates. **e** Gates and triggers after synchronization

which the data rate was 100% when $s(t)$ was in range, and 0% when it was out of range. The program recorded the total time, $T_{in\ range}$ when the signal was in range, which could be compared to the total time of the acquisition, T_{acq} . The reconstruction computer created a nonquantitative static image. The images were multiplied by the ratio T_{acq}

$/T_{in\ range}$ to make them more nearly quantitative. This simple scaling approach would be accurate in the case of a constant breathing pattern during the PET acquisition, with the optimal gate opening and closing regularly during the scan, resulting in a simple data loss of $T_{acq} / T_{in\ range}$. An irregular breathing pattern would result in more data being

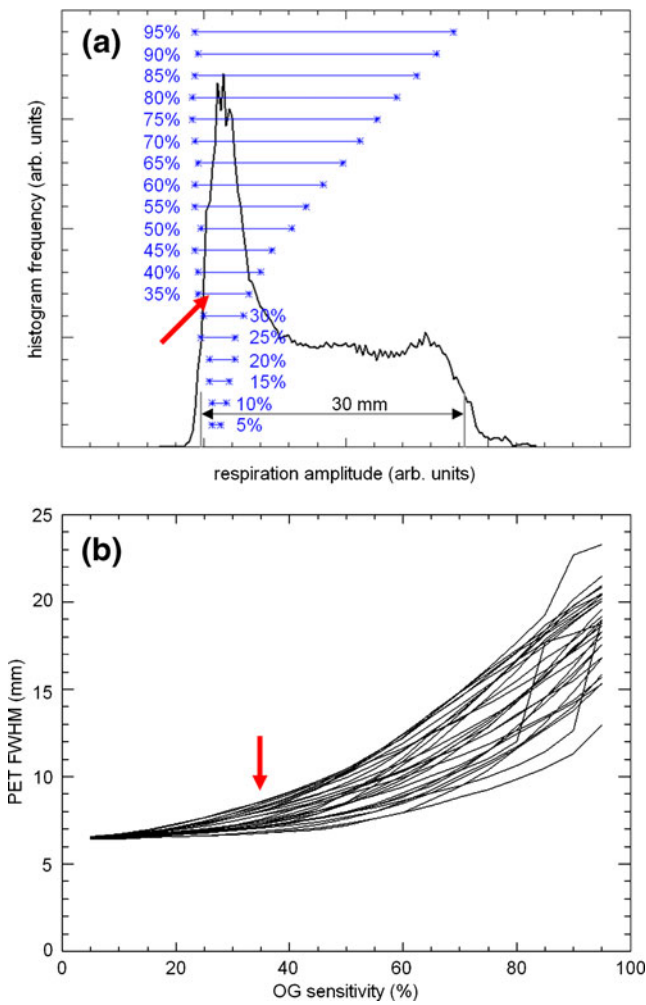


Fig. 8 **a** Histogram of a patient’s respiration amplitudes. The amplitude ranges found by the OG algorithm for sensitivities from 5% to 95% are shown in *blue*. It was assumed that tumour excursion is 30 mm (*black arrow*). The 35% sensitivity value (*red arrow*) was used in this research. **b** For each of the 26 patients in our study, the PET FWHM resolution modelled by Eq. B.6 is plotted as a function of OG sensitivity (*red arrow* 35% sensitivity)

acquired at the beginning or end of the scan, leading to quantitative errors.

Appendix B. Rationale for choosing the optimal gating yield parameter

The optimal gating algorithm has only one parameter, the yield parameter, which was set to $OG_{yield}=0.35$, or 35%. We determined several years ago that this value is appropriate because most patients pause between breaths, representing about one-third of the breathing cycle. We believed that in most cases an analysis of a patient’s breathing pattern should find the quiescent part of the breathing cycle. The shaded band in Fig. 1 is a typical example.

In addition, a physics-based argument demonstrates that 35% is an appropriate choice. We assume that the tumour moves in a definite direction, say z , and that the position relative to a reference point is proportional to $s(t)$ described in *Materials and methods*. The position at time t is:

$$z(t) = z_{ref} + k \times s(t), \tag{B.1}$$

where k is a proportionality constant. Since one does not know the actual motion at any point in the chest, we assume a nearly worst-case scenario. We assume that the tumour excursion is 30 mm except for the lowest 2% and the highest 2% of the amplitudes seen in $s(t)$. To make this concrete, we refer to the histogram $h(s)$ and the cumulative distribution $cdf(s)$ described in *Materials and methods*. Let a_2 and a_{98} be the amplitudes corresponding to cdf values of 0.02 and 0.98. In our model, the coefficient k is:

$$k = \frac{30mm}{a_{98} - a_2}. \tag{B.2}$$

In an optimally gated measurement, data are processed when s is between the limits L and U . Relative to the lowest observed s position, the average s value is:

$$s_{avg} = \frac{\int_L^U ds \times s \times h(s)}{\int_L^U ds \times h(s)}, \tag{B.3}$$

and the variance is:

$$\sigma_s^2 = \frac{\int_L^U ds \times (s - s_{avg})^2 \times h(s)}{\int_L^U ds \times h(s)}. \tag{B.4}$$

During the OG measurement period, the variance in z positions is:

$$\sigma_{respiratory}^2 = k^2 \sigma_s^2, \tag{B.5}$$

This can be used to estimate PET image resolution when OG is used. In the absence of motion, the PET image is characterized by a full-width at half-maximum (FWHM) resolution parameter due to system resolution and post-reconstruction filtering. With system resolution of about 4 mm and post-reconstruction filtering of 5 mm used in our clinical work, that parameter is 6.4 mm. It is reasonable to invoke the central limit theorem and convert variances to FWHM in a gaussian model:

$$PETFWHM = 2.355 \times \sqrt{\sigma_{recon}^2 + \sigma_{respiration}^2}. \tag{B.6}$$

Since this resolution figure of merit depends on the patient’s breathing pattern and on the OG sensitivity parameter, we applied (Eq. B.6) to the respiration curves of all patients in our study, and varied the sensitivity from 5% to 95%. Figure 8a indicates the lower and upper edges

of the optimal windows for each sensitivity parameter. Figure 8b shows the resolution curves for all patients.

When the OG sensitivity is 35%, the resolution of the OG method (7.6 mm average over all patients) is only a little larger than the resolution without OG (6.4 mm). This is true for all patients, and it is true even though we assumed an unusually large tumour motion (30 mm). If we assumed a more typical tumour motion of 15 mm, the resolution would be 6.7 mm.

In this sense, our method is optimal. One can use sensitivity values much smaller than 35%, but this would be at the cost of lower statistics with little advantage in resolution. One can use larger sensitivity values, making the image less noisy, but the resolution may be significantly worse, up to 20 mm in some patients. The 35% setting strikes a balance between resolution and noise.

References

1. Eschmann SM, Friedel G, Paulsen F, Reimold M, Hehr T, Budach W, et al. Is standardised (18)F-FDG uptake value an outcome predictor in patients with stage III non-small cell lung cancer. *Eur J Nucl Med Mol Imaging*. 2006;33:263–9.
2. Aerts HJ, van Baardwijk AA, Petit SF, Offermann C, Loon J, Houben R, et al. Identification of residual metabolic-active areas within individual NSCLC tumours using a pre-radiotherapy (18) Fluorodeoxyglucose-PET-CT scan. *Radiother Oncol*. 2009;91:386–92.
3. van Baardwijk A, Bosmans G, Dekker A, van Kroonenburgh M, Boersma L, Wanders S, et al. Time trends in the maximal uptake of FDG on PET scan during thoracic radiotherapy. A prospective study in locally advanced non-small cell lung cancer (NSCLC) patients. *Radiother Oncol*. 2007;82:145–52.
4. Stroom J, Blaauwgeers H, van Baardwijk A, Boersma L, Lebesque J, Theuvs J, et al. Feasibility of pathology-correlated lung imaging for accurate target definition of lung tumors. *Int J Radiat Oncol Biol Phys*. 2007;69:267–75.
5. van Baardwijk A, Bosmans G, Boersma L, Buijsen J, Wanders S, Hochstenbag M, et al. PET-CT-based auto-contouring in non-small-cell lung cancer correlates with pathology and reduces interobserver variability in the delineation of the primary tumor and involved nodal volumes. *Int J Radiat Oncol Biol Phys*. 2007;68:771–8.
6. Biehl KJ, Kong FM, Dehdashti F, Jin JY, Mutic S, El Naqa I, et al. 18F-FDG PET definition of gross tumor volume for radiotherapy of non-small cell lung cancer: is a single standardized uptake value threshold approach appropriate. *J Nucl Med*. 2006;47:1808–12.
7. Daisne JF, Sibomana M, Bol A, Doumont T, Lonneux M, Gregoire V. Tri-dimensional automatic segmentation of PET volumes based on measured source-to-background ratios: influence of reconstruction algorithms. *Radiother Oncol*. 2003;69:247–50.
8. Nestle U, Kremp S, Schaefer-Schuler A, Sebastian-Welsch C, Hellwig D, Rube C, et al. Comparison of different methods for delineation of 18F-FDG PET-positive tissue for target volume definition in radiotherapy of patients with non-small cell lung cancer. *J Nucl Med*. 2005;46:1342–8.
9. Ollers M, Bosmans G, van Baardwijk A, Dekker A, Lambin P, Teule J, et al. The integration of PET-CT scans from different hospitals into radiotherapy treatment planning. *Radiother Oncol*. 2008;87:142–6.
10. Nehmeh SA, Erdi YE, Pan T, Pevsner A, Rosenzweig KE, Yorke E, et al. Four-dimensional (4D) PET/CT imaging of the thorax. *Med Phys*. 2004;31:3179–86.
11. Boucher L, Rodrigue S, Lecomte R, Benard F. Respiratory gating for 3-dimensional PET of the thorax: feasibility and initial results. *J Nucl Med*. 2004;45:214–9.
12. Vedam SS, Keall PJ, Kini VR, Mostafavi H, Shukla HP, Mohan R. Acquiring a four-dimensional computed tomography dataset using an external respiratory signal. *Phys Med Biol*. 2003;48:45–62.
13. Nehmeh SA, Erdi YE, Ling CC, Rosenzweig KE, Schoder H, Larson SM, et al. Effect of respiratory gating on quantifying PET images of lung cancer. *J Nucl Med*. 2002;43:876–81.
14. Nehmeh SA, Erdi YE, Pan T, Yorke E, Mageras GS, Rosenzweig KE, et al. Quantitation of respiratory motion during 4D-PET/CT acquisition. *Med Phys*. 2004;31:1333–8.
15. Wolthaus JW, van Herk M, Muller SH, Belderbos JS, Lebesque JV, de Bois JA, et al. Fusion of respiration-correlated PET and CT scans: correlated lung tumour motion in anatomical and functional scans. *Phys Med Biol*. 2005;50:1569–83.
16. Dawood M, Buther F, Lang N, Schober O, Schafers KP. Respiratory gating in positron emission tomography: a quantitative comparison of different gating schemes. *Med Phys*. 2007;34:3067–76.
17. Hamill JJ, Bosmans G, Dekker A. Respiratory-gated CT as a tool for the simulation of breathing artifacts in PET and PET/CT. *Med Phys*. 2008;35:576–85.
18. Nagel CC, Bosmans G, Dekker AL, Ollers MC, De Ruyscher DK, Lambin P, et al. Phased attenuation correction in respiration correlated computed tomography/positron emitted tomography. *Med Phys*. 2006;33:1840–7.
19. Park SJ, Ionascu D, Killoran J, Mamede M, Gerbaudo VH, Chin L, 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*. 2008;53:3661–79.
20. Ponisch F, Richter C, Just U, Enghardt W. Attenuation correction of four dimensional (4D) PET using phase-correlated 4D-computed tomography. *Phys Med Biol*. 2008;53:N259–68.
21. Vines DC, Keller H, Hoisak JD, Breen SL. Quantitative PET comparing gated with nongated acquisitions using a NEMA phantom with respiratory-simulated motion. *J Nucl Med Technol*. 2007;35:246–51.
22. Rosario T, Ollers MC, Bosmans G, De Ruyscher D, Lambin P, Dekker A. Phased versus midventilation attenuation-corrected respiration-correlated PET for patients with non-small cell lung cancer. *J Nucl Med Technol*. 2009;37:208–14.
23. Erdi YE, Nehmeh SA, Pan T, Pevsner A, Rosenzweig KE, Mageras G, et al. The CT motion quantitation of lung lesions and its impact on PET-measured SUVs. *J Nucl Med*. 2004;45:1287–92.
24. Wink N, Panknin C, Solberg TD. Phase versus amplitude sorting of 4D-CT data. *J Appl Clin Med Phys*. 2006;7:77–85.
25. Li XA, Stepaniak C, Gore E. Technical and dosimetric aspects of respiratory gating using a pressure-sensor motion monitoring system. *Med Phys*. 2006;33:145–54.
26. Ionascu D, Jiang SB, Nishioka S, Shirato H, Berbeco RI. Internal-external correlation investigations of respiratory induced motion of lung tumors. *Med Phys*. 2007;34:3893–903.
27. Boellaard R, Oyen WJ, Hoekstra CJ, Hoekstra OS, Visser EP, Willemsen AT, et al. The Netherlands protocol for standardisation and quantification of FDG whole body PET studies in multi-centre trials. *Eur J Nucl Med Mol Imaging*. 2008;35:2320–33.
28. Paulino AC, Johnstone PA. FDG-PET in radiotherapy treatment planning: Pandora's box? *Int J Radiat Oncol Biol Phys*. 2004;59:4–5.
29. Erdi YE, Rosenzweig K, Erdi AK, Macapinlac HA, Hu YC, Braban LE, et al. Radiotherapy treatment planning for patients with non-small cell lung cancer using positron emission tomography (PET). *Radiother Oncol*. 2002;62:51–60.

30. van Baardwijk A, Wanders S, Boersma L, Borger J, Ollers M, Dingemans AM, et al. Mature results of an individualized radiation dose prescription study based on normal tissue constraints in stages I to III non-small-cell lung cancer. *J Clin Oncol*. 2010;28:1380–6.
31. Lambin P, Petit SF, Aerts HJ, van Elmpt WJ, Oberije CJ, Starmans MH, et al. The ESTRO Breur Lecture 2009. From population to voxel-based radiotherapy: exploiting intra-tumour and intra-organ heterogeneity for advanced treatment of non-small cell lung cancer. *Radiother Oncol*. 2009;96:145–52.
32. Petit SF, Van Elmpt W, Dehing C, Vegt E, Dingemans AM, Lambin P, et al. [18F]Fluorodeoxyglucose uptake patterns in lung before radiotherapy identify areas more susceptible to radiation-induced lung toxicity in non-small-cell lung cancer patients. *Int J Radiat Oncol Biol Phys*. 2010;in press. doi:10.1016/j.ijrobp.2010.06.016
33. Siedschlag C, van Loon J, van Baardwijk A, Rossi MM, van Pel R, Blaauwgeers JL, et al. Analysis of the relative deformation of lung lobes before and after surgery in patients with NSCLC. *Phys Med Biol*. 2009;54:5483–92.
34. Fin L, Daouk J, Morvan J, Bailly P, El Esper I, Saidi L, et al. Initial clinical results for breath-hold CT-based processing of respiratory-gated PET acquisitions. *Eur J Nucl Med Mol Imaging*. 2008;35:1971–80.
35. Torizuka T, Tanizaki Y, Kanno T, Futatsubashi M, Yoshikawa E, Okada H, et al. Single 20-second acquisition of deep-inspiration breath-hold PET/CT: clinical feasibility for lung cancer. *J Nucl Med*. 2009;50:1579–84.
36. Dawood M, Lang N, Jiang X, Schafers KP. Lung motion correction on respiratory gated 3-D PET/CT images. *IEEE Trans Med Imaging*. 2006;25:476–85.
37. Frey K, Hamill JJ, Jones J, Koeppe R. Respiratory motion gated PET with end-expiratory, amplitude-based gating. *J Nucl Med*. 2009 (Suppl 2);50:348.
38. Bracewell RN. *The Fourier transform and its applications*. New York: McGraw Hill; 1986.