Published in final edited form as:

Phys Med Biol. 2009 December 21; 54(24): 7345–7362. doi:10.1088/0031-9155/54/24/007.

The impact of respiratory motion on tumor quantification and delineation in static PET/CT imaging

Chi Liu¹, Larry A Pierce II, Adam M Alessio, and Paul E Kinahan

Department of Radiology, University of Washington, Seattle, WA, USA

Abstract

Our aim is to investigate the impact of respiratory motion on tumor quantification and delineation in static PET/CT imaging using a population of patient respiratory traces. A total of 1295 respiratory traces acquired during whole body PET/CT imaging were classified into three types according to the qualitative shape of their signal histograms. Each trace was scaled to three diaphragm motion amplitudes (6 mm, 11 mm and 16 mm) to drive a whole body PET/CT computer simulation that was validated with a physical phantom experiment. Three lung lesions and one liver lesion were simulated with diameters of 1 cm and 2 cm. PET data were reconstructed using the OS-EM algorithm with attenuation correction using CT images at the end-expiration phase and respiratory-averaged CT. The errors of the lesion maximum standardized uptake values (SUV_{max}) and lesion volumes between motion-free and motion-blurred PET/CT images were measured and analyzed. For respiration with 11 mm diaphragm motion and larger quiescent period fraction, respiratory motion can cause a mean lesion SUV_{max} underestimation of 28% and a mean lesion volume overestimation of 130% in PET/CT images with 1 cm lesions. The errors of lesion SUV_{max} and volume are larger for patient traces with larger motion amplitudes. Smaller lesions are more sensitive to respiratory motion than larger lesions for the same motion amplitude. Patient respiratory traces with relatively larger quiescent period fraction yield results less subject to respiratory motion than traces with long-term amplitude variability. Mismatched attenuation correction due to respiratory motion can cause SUV_{max} overestimation for lesions in the lower lung region close to the liver dome. Using respiratory-averaged CT for attenuation correction yields smaller mismatch errors than those using end-expiration CT. Respiratory motion can have a significant impact on static oncological PET/CT imaging where SUV and/or volume measurements are important. The impact is highly dependent upon motion amplitude, lesion location and size, attenuation map and respiratory pattern. To overcome the motion effect, motion compensation techniques may be necessary in clinical practice to improve the tumor quantification for determining the response to therapy or for radiation treatment planning.

1. Introduction

Whole body F-18 fluoro-2-deoxyglucose (^{18}F -FDG) positron emission tomography/computed tomography (PET/CT) has become an important method for detecting tumors, planning radiation treatment and evaluating response to therapy. However, PET/CT imaging of the lung and abdomen region is generally affected by patient respiratory motion, which can lead to underestimation of SUV_{max} of a region of interest, overestimation of tumor volume and mismatched PET and CT images that yield attenuation correction (AC) errors, registration errors and tumor mislocalization (Nehmeh and Erdi 2008).

¹Author to whom any correspondence should be addressed: cliu2@u.washington.edu.

Different methods have been developed to manage the respiratory motion problem. These methods include respiratory-gated PET/CT (Nehmeh *et al* 2004a, 2004b, Pan *et al* 2004, 2007, Rietzel *et al* 2005, Wolthaus *et al* 2005, Abdelnour *et al* 2007, Dawood *et al* 2007, Guckenberger *et al* 2007, Lu *et al* 2006, Wink *et al* 2006), deep-inspiration-breathhold (DIBH) PET/CT (Kawano *et al* 2008, Meirelles *et al* 2007, Nehmeh *et al* 2007), postprocessing methods (Dawood *et al* 2006, Yamazaki *et al* 2006) and motion-corrected PET reconstruction (Lamare *et al* 2007, Li *et al* 2006, Qiao *et al* 2006, 2007).

Although these methods have been shown to reduce respiratory artifacts, the effect of respiratory motion in routine clinical practice (where correction methods are not used) is not well understood. It has been reported that various respiratory motion correction methods can cause the tumor SUV_{max} to increase as much as 159% and tumor volume to decrease as much as 43% compared to those of the conventional measurement in patient studies (Kawano *et al* 2008, Meirelles *et al* 2007, Nehmeh *et al* 2002a, 2002b, 2007). In particular, a study on 108 patients showed that the SUV_{max} increased as much as 51.8% on average from free-breathing PET to DIBH PET for lesions in the lower lung region (Kawano *et al* 2008). Yet, the true SUV_{max} and tumor volume in motion-free PET/CT, which is hypothetically obtained from a patient who does not breath, are unknown in these studies, and instead, the tumor SUV_{max} and volume obtained from either respiratory-gated PET/CT or DIBH PET/CT were used as a surrogate for the true values. Other studies using the phantom experiment with known true tumor SUV_{max} in motion-free images showed that the SUV_{max} can be underestimated by as much as 75% and the recovery coefficient can be as low as 0.2 in motion-blurred images, depending on different lesion size and motion amplitude (Park *et al* 2008, Pevsner *et al* 2005). However, these studies used a simple oscillatory motion to drive the phantom movement, which differs greatly from realistic patient motion traces and does not capture the variation in respiratory motion patterns.

The purpose of this work is to investigate the impact of respiratory motion on tumor quantification in terms of SUV_{max} and delineation in terms of tumor volume in static PET/CT imaging, with a single helical CT for attenuation correction, using both phantom experiments and computer simulations driven by a large population of 1295 actual patient respiratory traces. In other words, with current routine clinical protocols (single helical CT for attenuation correction, SUV_{max} for tumor quantification and simple threshold-based segmentation algorithms for tumor boundary delineation in PET images), the magnitude of error with respect to the true values in the motionless patient is investigated. Our study with known truth of tumor uptake and volume in the context of no respiratory and actual patient respiratory motion leads to a more complete understanding of the impact of respiratory motion in PET/CT imaging.

With a large number of respiratory traces, we are able to study the effect of respiratory motion on patients with certain breathing patterns. We classified the 1295 traces into different types according to their breathing patterns, and the simulation results were further stratified according to the different trace types. We envision that this classification could lead to more knowledge of personalized motion correction methods for individual patients in the future, as patients with different breathing patterns would benefit differently from different motion correction methods.

2. Methods

2.1. Respiratory motion traces

We collected 3803 anonymized patient respiratory traces using the real-time position management (RPM) respiratory gating system (Varian Medial Systems, Palo Alto, CA) during routine PET/CT studies at University of Washington Medical Center. The RPM system uses an infrared camera mounted on the PET/CT bed to image a reflective block marker placed on

the chest of the patient (Vedam *et al* 2003). The respiratory trace acquired by the RPM system is the vertical displacement of the marker as a function of time. This trace is considered to be a surrogate for internal organ motion. Due to the limitations in Varian software, only eight consecutive minutes of respiratory motion can be recorded at a time. The subjects were neither trained nor coached in this study; thus, the acquired RPM traces represent natural free breathing patterns during a PET/CT study. About 1300 traces had missing data for a portion of the scan, which exhibited a flat trace, mainly due to operator mistakes and were omitted. About 900 traces contained significant baseline shift clearly due to bed deflection and/or other motions and were omitted. Any trace less than 5 min in total length was omitted. In the end, 1295 viable patient traces were selected for this study. According to the record of patients scheduled for the PET/CT study at University of Washington Medical Center over a long period, about 70% of patients are males and about 30% are females. We expect this is close to the gender ratio of RPM traces in this study.

2.2. Classification of traces

By inspecting all 1295 traces, we classified each trace into one of three types according to its displacement histogram. The classification used a semi-automated method with a first pass using histogram pattern recognition and a final visual confirmation by the observer. As shown in figure 1(A), the type-1 trace has a prominent peak at the low end of the displacement histogram, which we call the quiescent peak. This peak is caused by the fact that the patient's breathing cycle consistently returns to a similar location at end-expiration and tends to spend more time near end-expiration. This type accounts for ~60% of the studied patient respiratory traces. As shown in figure 1(B), the histogram of type-2 trace shows a Gaussian or Poisson-like distribution of displacement locations, which could be caused by numerous respiratory patterns including variable end-expiration locations or similar inspiratory and expiratory phase lengths. This type accounts for ~20% of the traces studied. As shown in figure 1(C), the histogram of type-3 trace exhibits no recognizable shape and is simply spread out over the histogram bins. This is usually caused by the long-term variability of the respiratory trace and accounts for the remaining ~20% of the traces studied. According to our visual observation, the RPM traces become progressively irregular moving from type-1 to type-3.

2.3. Determination of patient diaphragm motion

To determine the diaphragm movement range in patient studies, we acquired cine CT data for 24 patients with normal tidal breathing during PET/CT exam. For each cine CT data, three-dimensional images of the temporal maximum intensity projection and minimum intensity projection were generated. The Department of Radiation Oncology at University of Washington routinely uses these images to determine the boundaries of tumor motion during radiation treatment planning for lung cancer patients. The highest location of diaphragm was measured from the maximum intensity projection with the smallest lung volume. The lowest location of diaphragm was measured from the minimum intensity projection with the largest lung volume. The diaphragm motion amplitude is determined as the difference between the highest and lowest diaphragm locations for each patient. This measurement was separated for left and right diaphragms. The mean, standard deviation, maximum and minimum of diaphragm amplitude derived from the 24 patient data were used to guide the trace scaling to drive the computer simulation in this paper.

2.4. Processing of traces

We used patient RPM traces to drive the diaphragm motion directly in our phantom experiment and computer simulation. Since the RPM traces acquired by the Varian system are measurements of the anterior-posterior motion, a transformation is needed to convert the acquired RPM traces to diaphragm displacement traces. To determine the amplitude in the

original RPM traces, we used an adaptive triggering method to determine the peaks and valleys in each cycle. The mean amplitude of the whole trace was determined by averaging the amplitudes of each individual cycle. To transform the RPM traces to diaphragm traces, we first shifted each original RPM trace to have zero median displacement and scaled it to an amplitude based on the results in section 2.3, regardless of the amplitude of the original RPM traces. As several studies demonstrated, there is a good correlation between the internal tissue motion and external marker motion in general, with different amplitudes (Beddar *et al* 2007, Gierga *et al* 2005, Hoisak *et al* 2004, Ionascu *et al* 2007). Therefore, it is reasonable to use the scaled RPM trace to drive the internal diaphragm motion in this study.

2.5. Physical phantom experiment

For a phantom validation study, the Data Spectrum[®] Anthropomorphic cardiac-torso phantom including heart, lungs, liver and spine inserts shown in figure 2(A) was used to acquire experimental data. A total activity of 148 MBq ¹⁸F-FDG at the start time of acquisition was injected into the phantom. The activity ratios for heart:liver:background:lung were 18:15:8:5. Two spherical ‘lesions’ were inserted into the lung and liver, respectively. The lung lesion has an inner diameter of ~1.1 cm and the liver lesion has an inner diameter of ~1.4 cm. Both lesions had a lesion-to-background ratio of 8:1. The phantom was positioned on the QUASAR[®] programmable respiratory motion platform from Modus Medical Devices Inc. (London, Ontario, Canada).

The phantom data were acquired on a General Electric (GE) DSTE PET/CT scanner operated in 2D mode with one bed position. We first scanned the stationary phantom without motion and then scanned the moving phantom driven by the QUASAR[®] platform. We scaled a canonical 8 min, type-1 trace to an 11 mm motion and used it to drive the QUASAR[®] platform to translate the phantom during scanning. The total acquisition time for both scans was 7 min. During the stationary scan, the phantom was placed at the center of the motion range and a CT image was acquired for attenuation correction with a CT technique used in our typical clinical PET/CT protocol (120 kVp, 120 mAs, pitch 1, 2.5 mm slice thickness). The PET sinogram data were reconstructed into $128 \times 128 \times 47$ image volumes with a voxel size of $5.47 \times 5.47 \times 3.27$ mm using the ordered-subset expectation-maximization (OS-EM) reconstruction algorithm with 4 iterations and 20 subsets (Hudson and Larkin 1994). Following the routine parameters used in one of the PET suites at University of Washington, a Gaussian post-filter with 7 mm full-width-at-half-maximum (FWHM) and a heavy *z*-direction filter were applied to the reconstructed images. Attenuation correction was applied in both the stationary and motion-blurred images with the attenuation map converted from the CT image acquired during the stationary scan. We measured weight-based SUV in this study as

$$\text{SUV} = \text{FDG activity concentration} / \text{Injected dose} / \text{Phantom weight.} \quad (1)$$

The $\text{SUV}_{\max}$ of both lesions in the stationary and motion-blurred PET images were measured for evaluation.

2.6. Computer simulation and validation

We used the NURBS-based cardiac torso (NCAT) phantom (Segars 2001, Segars *et al* 2001) with the same FDG organ activity distribution as that of the phantom experiment. The organ surfaces of the NCAT phantom were delineated using the non-uniform rational B-splines (NURBS). The NCAT phantom realistically models the respiratory motion in the body based on high-resolution respiratory-gated CT datasets. The attenuation maps generated by the

NCAT phantom at 511 keV were used to incorporate the attenuation effect in data simulation and correct for attenuation in image reconstruction.

To simulate whole-body PET acquisition of a free breathing patient, we first created a four-dimensional (4D) NCAT phantom dataset by generating 200 three-dimensional (3D) NCAT phantoms with diaphragm displacements equally sampled over a 4 cm amplitude. Each 3D NCAT phantom had a matrix size of $512 \times 512 \times 282$ and was forward projected with the matched attenuation map to generate 200 sinograms, which had a matrix size of $293 \times 280 \times 141$, using a distance-driven forward projector (De Man and Basu 2004). To form a motion-blurred sinogram according to a given scaled RPM trace, a displacement histogram with 200 bins was first generated; then the 200 simulated sinograms spanning the 4 cm motion range were weighted according to the histogram value, which characterizes the dwell time, of each displacement. The weighted sinograms were finally summed and normalized to form a motion-blurred sinogram. To simulate the PET acquisition of a stationary patient for ground truth, we selected one of the 200 sinograms with a displacement at the center of the diaphragm motion range. All the sinograms were convolved with a spatially variant point-spread function (PSF) measured from the GE DSTE scanner to simulate the detection blurring effect caused by inter-crystal penetration and scatter (Alessio and Kinahan 2008). In an effort to assess the influence of only respiratory motion, no noise from limited photon counts, scatter events or random events were included in the simulation.

To compare our computer simulation with the physical phantom experiment measurement, we generated two spherical lesions located in the lung and liver with the same size and contrast as those used in the phantom experiment. The same respiratory trace data acquisition, and reconstruction used in the phantom experiment were used in the computer simulation. The NCAT-generated attenuation map corresponding to the center of motion range is chosen for attenuation correction. For both the phantom experiment and NCAT simulation, the fraction errors of the SUV_{max} from motion-free to motion-blurred images were compared. This validation step was intended to ensure that the computer simulation is adequately modeling the influence of respiratory motion and the resolution degrading components of the PET acquisition.

2.7. Computer simulation using all RPM traces

We simulated the motion-blurred sinograms for all 1295 patient RPM traces with different amplitudes using the NCAT phantom. Two sets of spherical lesions with 1 cm and 2 cm diameters, respectively, both with 8:1 contrast, were generated to study the effect of motion for different tumor sizes. Lesions of each diameter set were inserted into the phantom with locations in upper-right lung, middle-left lung, lower-right lung and central liver regions. The same simulation and reconstruction processes described in the previous section, except that the attenuation maps corresponding to the end-expiration location of each individual trace were chosen for attenuation correction, were used to generate the final reconstructed motion-blurred images.

Each of the 1295 RPM traces was scaled to three diaphragm motion amplitudes of 6 mm, 11 mm and 16 mm based on the cine CT patient results in sections 2.3 and 3.1, to study the impact of different motion amplitudes. Thus, the total number of traces used in this simulation study was $1295 \times 3 = 3885$, which represents a large variety of the patient population. For lesion quantification, the fractional errors of the lesion SUV_{max} and volume change from motion-free to motion-blurred images were calculated for each patient trace and the mean, standard deviation and quartiles of these errors were used to characterize the impact of respiratory motion. To investigate the tumor volume change in motion-blurred PET images, we used two methods consistent with methods from published studies. One method determines the lesion volume using thresholds at a percentage (we used 50% in this study) of the lesion SUV_{max}

(Biehl *et al* 2006); the other method, which is essentially the same as the methods utilizing an absolute SUV value as a threshold (Nestle *et al* 2005), determines the lesion volume by using a fixed threshold at two times the local background. The results of this analysis were further stratified based on different respiratory histogram types to investigate the respiratory motion impact in each patient population sub-category.

To study the effect of the choice of the attenuation map on the motion-blurred images reconstructed with attenuation correction, we reconstructed the images simulated from the type-1 traces using attenuation maps derived from respiratory-averaged CT (ACT) (Pan *et al* 2005) and compared the results with those reconstructed using attenuation maps at the end-expiration phase.

3. Results

3.1. Patient diaphragm motion range

Table 1 shows the mean, standard deviation and range of diaphragm motion amplitude measured from 24 cine CT studies. There are only 22 measurements for the left diaphragm, because the left diaphragms of two patient data merged with the chest wall and were not measurable. The mean motion amplitude is 10.9 mm for the left diaphragm and 10.1 mm for the right diaphragm. The maximum and minimum amplitudes for the left and right diaphragms are also shown. These findings are also consistent with other published studies (Segars *et al* 2007, Seppenwoolde *et al* 2002).

3.2. Phantom experiment and NCAT simulation validation

Figure 2(B) shows sample reconstructed images from the phantom experiment. The images show that lesions in both the lung and liver regions were elongated in the axial direction due to respiratory motion in the motion-blurred image. The SUV_{max} for both lesions were reduced in the motion-blurred image compared to those in the no-motion image.

Table 2 compares the fractional error of SUV_{max} for the phantom experiment and computer simulation with the same trace and amplitude. The phantom experiment resulted in fractional errors of 18% and 19% for lung and liver lesions, respectively. The computer simulation resulted in fractional errors of 19% and 23%, which are in reasonable agreement with those of the phantom experiment.

3.3. Computer simulation using all patient traces

Figure 3 shows sample reconstructed NCAT images with 1 cm lesions without motion and with different diaphragm motion amplitudes. For each diaphragm motion amplitude, the lesions in the lower lung region are more blurred than those in the middle and upper lung regions, and upper lung lesions were the least affected by respiratory motion in general. With larger diaphragm motion amplitudes, lesions are more blurred at all locations. Artifacts caused by the mismatched attenuation correction, such as the appearance of bone structure, increased voxel values in the liver dome and around the intestines, are also visible in the motion-blurred images but not in the motion-free image.

Figure 4 shows box plots of the SUV_{max} errors of 1 cm lesions for different lesion locations and trace pattern types. For each trace type, the errors were generally larger for the lesions closer to the diaphragm. This was not the case for the lower lung lesion, where the mismatched attenuation correction artifactually boosted the lesion voxel value around the liver dome as seen in figure 3.

Figure 5 shows the fractional errors of lesion SUV_{max} for different trace types with 1 cm lesions for different diaphragm motion amplitudes and lesion locations. For lesions in the upper lung, middle lung and liver regions, the errors are larger with larger motion amplitudes; for the same amplitude, type-1 traces gave the smallest errors and type-3 traces gave the largest errors. For the lower lung lesion, there is a trend that larger motion amplitudes lead to smaller errors, mainly as a result of artifacts due to mismatched attenuation correction at larger motion amplitudes.

Table 3 provides the errors of volume measurements for different lesion locations, diaphragm motion amplitudes and trace pattern types. As shown in figure 3, the lower lung lesions with larger motion amplitudes merged with the liver due to motion and mismatched attenuation correction, which prevented meaningful volume measurements based on a threshold. Therefore, we did not include the lower lung lesions in volume measurements in this study. For the other three lesion locations shown in table 3, the errors of lesion volume with thresholds at 50% SUV_{max} increased with the same trend as that of SUV_{max} , with larger errors for larger motion amplitudes. Again, the type-3 traces gave the largest volume error, while the type-1 traces gave the smallest error. However, the errors of lesion volume with a fixed threshold at two times the local background showed an opposite trend for lesions in middle lung and liver, with smaller errors for larger motion amplitudes, and the negative signs of the errors indicate that the volume measurement from motion-blurred images are even smaller than those from motion-free images, which is not expected.

Figure 6 shows the box plot of the fractional errors of lesion SUV_{max} for the type-1 traces with 11 mm diaphragm motion amplitude for 1 cm and 2 cm lesion diameters. For the lesions in the middle lung, lower lung and liver regions, the fractional errors of 2 cm lesions are apparently smaller than those of 1 cm lesions. The upper lung lesions do not show a significant difference between 1 cm and 2 cm lesions.

Figure 7 shows the box plot of the fractional errors of lesion SUV_{max} for the type-1 traces with 11 mm diaphragm motion amplitude with a 1 cm lesion reconstructed using different attenuation maps. For the upper lung, middle lung and liver lesions, the errors with respiratory-averaged CT attenuation map are slightly smaller than those reconstructed using attenuation maps of the end-expiration phase. For the lower lung lesion, the trend is the opposite, showing larger errors with respiratory-averaged CT attenuation maps.

Table 4 presents a comparison of the mean and standard deviation of the lesion SUV_{max} and lesion volume between 1 cm and 2 cm lesion diameters and those using different attenuation maps. Similar to the results in figures 6 and 7, the errors of 1 cm lesions are larger than those of the 2 cm lesions, for both quantification and volume measurement, and respiratory-averaged attenuation maps gave slightly smaller error for lesions at locations other than lower lung and larger errors for lower lung lesions.

4. Discussion

In this study, we investigated the impact of respiratory motion on tumor tracer uptake quantification and delineation in static PET/CT using 1295 patient respiratory traces and computer simulations validated by a phantom experiment. For both lung and liver lesions, the fractional errors of lesion SUV_{max} are in reasonably good agreement between the simulation and phantom study. The focus of this paper is to understand the range of error caused by respiratory motion; potential motion management solutions are beyond the scope of this paper. The results show that respiratory motion can have a significant impact on measures derived from PET/CT imaging, depending on motion amplitude, lesion location and size, choice of attenuation map and respiratory pattern, and can cause, on average, 28% underestimation of

lesion SUV_{max} and 130% overestimation of lesion volume with a threshold at 50% SUV_{max} for a 1 cm lesion with type-1 trace and 11 mm diaphragm motion. The largest mean SUV_{max} error was found to be ~50% for liver lesions with 16 mm diaphragm motion and type-3 traces. Therefore, respiratory motion can be a major factor that degrades PET/CT image quality. Special attention should be paid to the impact of respiratory motion when evaluating response to therapy using quantitative PET images. The various tumor quantification errors in phantom and patient studies reported in the literature (mentioned in the introduction) all fall into the range of errors in this simulation study. With a large number of patient traces in this study, we are able to characterize a more complete range of SUV and volume quantification errors for different breathing patterns, lesion sizes and attenuation corrections.

The results in table 3 unexpectedly show that the lesion volume measurements with a threshold at two times the local background decreased with larger motion amplitudes for middle lung and liver lesions for all the three trace types, and the measured volumes are smaller than those measured from motion-free images. This may be caused by the decrease of lesion activity due to motion blurring, consequently resulting in a decrease in lesion volume above the threshold for larger motion amplitudes, as shown in figure 8. These results indicate that the volume measurements with thresholds at 50% SUV_{max} , which increased with larger motion amplitudes as shown in figure 8, may be more appropriate than the method using a SUV threshold in radiation treatment planning for size estimation from PET images.

For reconstructing the motion-blurred data with attenuation correction, an attenuation map at the end-expiration phase was chosen, as using attenuation maps other than the end-expiration phase results in greater SUV_{max} underestimation and motion artifacts (Beyer *et al* 2003, Erdi *et al* 2004). Since the CT acquisition images a snapshot scan and the PET acquisition images an averaged scan during several minutes, the motion-blurred PET image was reconstructed with a mismatched attenuation map as a consequence. The liver lesion in this study has a smaller effect caused by the mismatched attenuation map because the attenuation factors are uniform throughout the liver. In contrast, the lung lesions are expected to have greater errors from mismatched attenuation correction because the lesion translated through adjacent lung space.

Our results in figure 7 and table 4 indicate that using respiratory-averaged CT for attenuation correction yields slightly smaller errors in quantification and volume measurement for lesions with locations other than lower lung. However, for lower lung lesions, using respiratory-averaged CT leads to larger errors. As demonstrated by sample images in figure 9, this increased error can be explained by the elimination of mismatched attenuation artifact of lower lung lesions and liver dome with respiratory-averaged CT for attenuation correction. Therefore, one potential solution to address artifacts around the diaphragm interface is using respiratory-averaged CT for attenuation correction.

Generally, larger motion amplitude leads to greater errors of SUV_{max} and volume in motion-blurred images as shown in figures 4 and 5 and table 3. For the same motion amplitude as shown in figure 6 and table 4, lesions with 2 cm diameter lead to smaller errors in SUV_{max} and volume compared to those of the 1 cm lesions, indicating that larger lesions are less subject to respiratory motion than smaller lesions. As tissues near the diaphragm generally have larger motion amplitudes than those further away from the diaphragm, lesions in the upper lung region have smaller errors in SUV_{max} and volume than those in the middle and lower lung regions. However, as shown in figure 3, lesions in the lower lung region near the liver dome enter the liver space in the attenuation map with larger motion amplitudes due to motion and mismatched attenuation correction. The voxel values of the incorrectly attenuation-corrected part of lower lung lesions usually get boosted and result in a smaller SUV_{max} error than that of the middle lung lesion, as shown in figures 4 and 5. For the lower lung lesions in figures 4, 6 and 7, there are some outliers with positive fractional errors, indicating that respiratory motion can also

result in SUV_{max} overestimation in some cases, probably due to the mismatched attenuation correction.

In motion-blurred PET/CT images, the impact of respiratory motion on tumor quantification and delineation is conceptually determined by the tumor motion amplitude and the duration the tumor spends at each displacement. Therefore, after exploring several mathematical classification methods, we believe that the classification based on the displacement histogram of each RPM trace is most relevant to the PET/CT image generation. The results for different trace types shown in figures 4 and 5 and table 3 indicate larger errors for type-3 traces and smaller errors for type-1 traces. Patients with type-1 traces usually breathe to a similar expiration location and spend a longer amount time near that location, while patients with type-3 traces usually have long-term amplitude changes. Therefore, patients spending more time at end-expiration are generally less subject to respiratory motion than patients with long-term motion amplitude variability. As shown in figure 5(A) and table 3, for type-1 traces, the mean errors for upper lung lesions, regardless of motion amplitude, are less than 5% for both SUV_{max} and volume measurements with thresholds at 50% SUV_{max} . Therefore, patients with an upper lung lesion and a type-1 breathing pattern may not need further motion compensation for a relatively accurate quantification and delineation. However, patients with lesions at other locations and with other breathing patterns may need motion compensation techniques to ensure quantification and delineation accuracy.

All the computer simulations are noise free, while the SUV_{max} measurement in clinical practice is affected by image noise. As shown by Barrett *et al* (1994), the noiseless ML-EM reconstruction results in the same image as the average of reconstructed images from multiple noise realizations of the same object. We anticipate ML-EM (OS-EM) behaves similarly for reconstructing noiseless and noisy data with clinical counts level for the same object. Also, the noise will affect the SUV_{max} and volume measurement similarly for both motion-free and motion-blurred images. Therefore, we do not expect that the results of error analysis presented in this paper will change significantly for noisy data.

In our simulation, we assume that the tumor moved along the same path in the superior–inferior direction during inspiration and expiration. Studies showed that patient tumor motion is primarily superior–inferior and motion in other directions is significantly smaller and hard to predict, due to various factors such as hysteresis (Seppenwoolde *et al* 2002). Therefore, our study design may lead to a slight underestimation in errors for patients with irregular tumor movement orbits. The detector point-spread function model used in this study only incorporated radial blurring for the major source of error from parallax and did not account for axial blurring (Alessio and Kinahan 2008). In the validation study, the slightly larger errors of computer simulation than that of the phantom experiment may be due to these factors.

5. Conclusion

We have investigated the impact of respiratory motion on tumor quantification and delineation in static PET/CT imaging in terms of the estimation errors of lesion SUV_{max} and volume. Depending on patient breathing pattern, motion amplitude, lesion size and location, and the choice of attenuation correction, respiratory motion can result in significant underestimation in the SUV_{max} and overestimation in volume measurements. Mismatched attenuation correction, which can be partly compensated by using respiratory-averaged CT as the attenuation map, further complicates tumor quantification and delineation. This study suggests that it can be critical to correct for respiratory motion in clinical practice to improve quantification and treatment planning in oncological PET/CT imaging.

Acknowledgments

We thank William Paul Segars for help with the NCAT phantom, Jenine Yager for proof reading, and Alexander Ganin and Scott Wollenweber for helpful discussions. Thanks also go to Joshua Busch and Steve Kohlmyer for collecting the cine CT patient data. This work is funded by NIH grant R01-CA115870 and a research contract from GE Healthcare.

References

- Abdelnour AF, et al. Phase and amplitude binning for 4D-CT imaging. *Phys Med Biol* 2007;52:3515–29. [PubMed: 17664557]
- Alessio, AM.; Kinahan, PE. Application of a spatially variant system model for 3-D whole-body PET image reconstruction. *ISBI 2008. 5th IEEE Int. Symp. on Biomedical Imaging: From Nano to Macro, 2008; 2008. p. 1315-8.*
- Barrett HH, Wilson DW, Tsui BM. Noise properties of the EM algorithm: I. Theory. *Phys Med Biol* 1994;39:833–46. [PubMed: 15552088]
- Beddar AS, Kainz K, Briere TM, Tsunashima Y, Pan T, Prado K, Mohan R, Gillin M, Krishnan S. Correlation between internal fiducial tumor motion and external marker motion for liver tumors imaged with 4D-CT. *Int J Radiat Oncol Biol Phys* 2007;67:630–38. [PubMed: 17236980]
- Beyer T, Antoch G, Blodgett T, Freudenberg LF, Akhurst T, Mueller S. Dual-modality PET/CT imaging: the effect of respiratory motion on combined image quality in clinical oncology. *Eur J Nucl Med Mol Imaging* 2003;30:588–96. [PubMed: 12582813]
- Biehle KJ, Kong FM, Dehdashti F, Jin JY, Mutic S, El Naqa I, Siegel BA, Bradley JD. 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. [PubMed: 17079814]
- 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. [PubMed: 17822014]
- 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. [PubMed: 16608062]
- De Man B, Basu S. Distance-driven projection and backprojection in three dimensions. *Phys Med Biol* 2004;49:2463–75. [PubMed: 15248590]
- Erdi YE, et al. The CT motion quantitation of lung lesions and its impact on PET-measured SUVs. *J Nucl Med* 2004;45:1287–92. [PubMed: 15299050]
- Gierga DP, Brewer J, Sharp GC, Betke M, Willett CG, Chen GT. The correlation between internal and external markers for abdominal tumors: implications for respiratory gating. *Int J Radiat Oncol Biol Phys* 2005;61:1551–8. [PubMed: 15817361]
- Guckenberger M, Weininger M, Wilbert J, Richter A, Baier K, Krieger T, Polat B, Flentje M. Influence of retrospective sorting on image quality in respiratory correlated computed tomography. *Radiother Oncol* 2007;85:223–31. [PubMed: 17854931]
- Hoisak JD, Sixel KE, Tirone R, Cheung PC, Pignol JP. Correlation of lung tumor motion with external surrogate indicators of respiration. *Int J Radiat Oncol Biol Phys* 2004;60:1298–306. [PubMed: 15519803]
- Hudson HM, Larkin RS. Accelerated image reconstruction using ordered subsets of projection data. *IEEE Trans Med Imaging* 1994;13:601–9. [PubMed: 18218538]
- 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. [PubMed: 17985635]
- Kawano T, Ohtake E, Inoue T. Deep-inspiration breath-hold PET/CT of lung cancer: maximum standardized uptake value analysis of 108 patients. *J Nucl Med* 2008;49:1223–31. [PubMed: 18632812]
- Lamare F, Ledesma Carbayo MJ, Cresson T, Kontaxakis G, Santos A, Le Rest CC, Reader AJ, Visvikis D. List-mode-based reconstruction for respiratory motion correction in PET using non-rigid body transformations. *Phys Med Biol* 2007;52:5187–204. [PubMed: 17762080]
- Li T, Thorndyke B, Schreiber E, Yang Y, Xing L. Model-based image reconstruction for four-dimensional PET. *Med Phys* 2006;33:1288–98. [PubMed: 16752564]

- Lu W, Parikh PJ, Hubenschmidt JP, Bradley JD, Low DA. A comparison between amplitude sorting and phase-angle sorting using external respiratory measurement for 4D CT. *Med Phys* 2006;33:2964–74. [PubMed: 16964875]
- Meirelles GSP, Erdi YE, Nehmeh SA, Squire OD, Larson SM, Humm JL, Schoder H. Deep-inspiration breath-hold PET/CT: clinical findings with a new technique for detection and characterization of thoracic lesions. *J Nucl Med* 2007;48:712–9. [PubMed: 17475958]
- Nehmeh SA, Erdi YE. Respiratory motion in positron emission tomography/computed tomography: a review. *Semin Nucl Med* 2008;38:167–76. [PubMed: 18396177]
- Nehmeh SA, Erdi YE, Ling CC, Rosenzweig KE, Schoder H, Larson SM, Macapinlac HA, Squire OD, Humm JL. Effect of respiratory gating on quantifying PET images of lung cancer. *J Nucl Med* 2002a; 43:876–81. [PubMed: 12097456]
- Nehmeh SA, Erdi YE, Meirelles GSP, Squire O, Larson SM, Humm JL, Schoder H. Deep-inspiration breath-hold PET/CT of the thorax. *J Nucl Med* 2007;48:22–6. [PubMed: 17204695]
- Nehmeh SA, et al. Effect of respiratory gating on reducing lung motion artifacts in PET imaging of lung cancer. *Med Phys* 2002b;29:366–71. [PubMed: 11929020]
- Nehmeh SA, et al. Four-dimensional (4D) PET/CT imaging of the thorax. *Med Phys* 2004a;31:3179–86. [PubMed: 15651600]
- Nehmeh SA, et al. Quantitation of respiratory motion during 4D-PET/CT acquisition. *Med Phys* 2004b; 31:1333–8. [PubMed: 15259636]
- Nestle U, Kremp S, Schaefer-Schuler A, Sebastian-Welsch C, Hellwig D, Rube C, Kirsch CM. Comparison of different methods for delineation of ^{18}F -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. [PubMed: 16085592]
- Pan T, Lee TY, Rietzel E, Chen GT. 4D-CT imaging of a volume influenced by respiratory motion on multi-slice CT. *Med Phys* 2004;31:333–40. [PubMed: 15000619]
- Pan T, Sun X, Luo D. Improvement of the cine-CT based 4D-CT imaging. *Med Phys* 2007;34:4499–503. [PubMed: 18072515]
- Pan TS, Mawlawi C, Nehmeh SA, Erdi YE, Luo DS, Liu HH, Castillo R, Mohan R, Liao ZX, Macapinlac HA. Attenuation correction of PET images with respiration-averaged CT images in PET/CT. *J Nucl Med* 2005;46:1481–7. [PubMed: 16157531]
- Park SJ, Ionascu D, Killoran J, Mamede M, Gerbaudo VH, Chin L, Berbeco R. 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. [PubMed: 18562782]
- Pevsner A, Nehmeh SA, Humm JL, Mageras GS, Erdi YE. Effect of motion on tracer activity determination in CT attenuation corrected PET images: a lung phantom study. *Med Phys* 2005;32:2358–62. [PubMed: 16121593]
- Qiao F, Pan T, Clark JWW, Mawlawi OR. A motion-incorporated reconstruction method for gated PET studies. *Phys Med Biol* 2006;51:3769–83. [PubMed: 16861780]
- Qiao F, Pan T, Clark JWW, Mawlawi O. Joint model of motion and anatomy for PET image reconstruction. *Med Phys* 2007;34:4626–39. [PubMed: 18196790]
- Rietzel E, Pan T, Chen GT. Four-dimensional computed tomography: image formation and clinical protocol. *Med Phys* 2005;32:874–89. [PubMed: 15895570]
- Segars, WP. PhD Thesis. University of North Carolina; 2001. Development of a new dynamic NURBS-based cardiac-torso (NCAT) phantom.
- Segars WP, Lalush DS, Tsui BMW. Modeling respiratory mechanics in the MCAT and spline-based MCAT phantoms. *IEEE Trans Nucl Sci* 2001;48:89–97.
- Segars, WP.; Mori, S.; Chen, GTY.; Tsui, BMW. Modeling respiratory motion variations in the 4D NCAT phantom. 2007 IEEE Nuclear Science Symp. Conf. Record; 2007. p. 4531–33.
- Seppenwoolde Y, Shirato H, Kitamura K, Shimizu S, van Herk M, Lebesque JV, Miyasaka K. Precise and real-time measurement of 3D tumor motion in lung due to breathing and heartbeat, measured during radiotherapy. *Int J Radiat Oncol Biol Phys* 2002;53:822–34. [PubMed: 12095547]
- Vedam SS, Kini VR, Keall PJ, Ramakrishnan V, Mostafavi H, Mohan R. Quantifying the predictability of diaphragm motion during respiration with a noninvasive external marker. *Med Phys* 2003;30:505–13. [PubMed: 12722802]

- Wink N, Panknin C, Solberg TD. Phase versus amplitude sorting of 4D-CT data. *J Appl Clin Med Phys* 2006;7:77–85. [PubMed: 16518319]
- Wolthaus JW, van Herk M, Muller SH, Belderbos JS, Lebesque JV, de Bois JA, Rossi MM, Damen EM. Fusion of respiration-correlated PET and CT scans: correlated lung tumour motion in anatomical and functional scans. *Phys Med Biol* 2005;50:1569–83. [PubMed: 15798344]
- Yamazaki, T.; Ue, H.; Haneishi, H.; Hirayama, A.; Sato, T.; Nawano, S. An attenuation correction method for respiratory-gated PET/CT image. *IEEE Nuclear Science Symp. Conf. Record*; 2006.

Figure 1.
Sample patient motion traces and corresponding displacement histogram for the three types:
(A) type-1, (B) type-2 and (C) type-3.

Figure 2.

(A) Experiment setup of the phantom study. (B) Sample reconstructed coronal slices of the phantom experiment without and with motion.

Figure 3.
Sample reconstructed coronal slices of computer simulation driven by a type-1 trace, without motion and with different motion amplitudes.

Figure 4.

Box plot of the fraction errors of SUV_{max} of 1 cm diameter lesions at different locations with 11 mm diaphragm motion for trace pattern types: (A) type-1, (B) type-2 and (C) type-3. The box stretches from the lower 25% quartile to the upper 75% quartile with a line across the box representing the median. Whiskers extend from each end of the box to the most extreme values in the data within 1.5 times the interquartile range, which is defined as the difference between the upper quartile and the lower quartile, from each end of the box. Outliers with values beyond the ends of the whiskers are displayed with a + sign.

Figure 5.

Means and standard deviations of the fraction errors of SUV_{max} with different diaphragm motion amplitudes, trace types and lesion locations at (A) upper lung, (B) middle lung, (C) lower lung and (D) liver. The data of different trace types are slightly shifted avoid error bars overlap.

Figure 6.
Box plot of the fraction errors of SUV_{max} for type-1 traces and 11 mm diaphragm motion amplitude with different lesion locations (UL: upper lung; ML: middle lung; LL: lower lung).

Figure 7.

Box plot of the fraction errors of SUV_{max} for type-1 traces and 11 mm diaphragm motion amplitude with different lesion locations and attenuation maps for attenuation correction (UL: upper lung; ML: middle lung; LL: lower lung; EE: attenuation map at end-expiration; ACT: attenuation map derived from respiratory-averaged CT).

Figure 8.

One-dimensional superior-inferior profiles through motionless and motion-blurred tumor driven by a type-1 trace with an 11 mm amplitude. As shown, by using the 50% SUV_{max} method, the volume (length in this 1D case) of the motion-blurred lesion is larger than the motionless profile. However, by using a threshold of two times the local background, the volume of the motion-blurred lesion is smaller than the motionless one, in this case.

Figure 9. Sample reconstructed motion-blurred images driven by an 11 mm type-1 trace with end-expiration CT and respiratory-averaged CT for attenuation correction (AC).

Table 1

Diaphragm motion amplitudes measured from cine CT acquisitions of 24 patients with normal tidal breathing during PET/CT exam.

Diaphragm	Number of patient data	Mean $\pm$ standard deviation of amplitude (mm)	Maximum amplitude (mm)	Minimum amplitude (mm)
Left	22	10.9 $\pm$ 4.3	20	5
Right	24	10.1 $\pm$ 3.5	15	4

Table 2

Comparison of SUV_{max} error from motion-free to motion-blurred images in the phantom experiment and computer simulation.

Lesion location	Experiment	SUV_{max} error
Lung lesion	Quasar experiment	-18%
	NCAT simulation	-19%
Liver lesion	Quasar experiment	-19%
	NCAT simulation	-23%

Lesion volume errors from motion-free to motion-blurred images in computer simulation for different motion amplitudes, lesion locations and trace types.

Table 3

Lesion location	Diaphragm motion amplitude	Percent error (mean ± standard deviation)					
		Volume with thresholds at 50% SUV _{max}			Volume with fixed threshold at two times the local background		
		Type-1	Type-2	Type-3	Type-1	Type-2	Type-3
Upper lung	6 mm	-3% ± 5%	-7% ± 4%	-5% ± 9%	14% ± 4%	19% ± 4%	20% ± 5%
	11 mm	-3% ± 7%	0% ± 9%	5% ± 18%	19% ± 4%	25% ± 4%	26% ± 4%
	16 mm	1% ± 10%	11% ± 13%	20% ± 21%	22% ± 5%	29% ± 4%	29% ± 3%
Middle lung	6 mm	22% ± 40%	60% ± 64%	96% ± 125%	-21% ± 2%	-17% ± 4%	-16% ± 6%
	11 mm	92% ± 88%	229% ± 149%	277% ± 180%	-17% ± 4%	-9% ± 7%	-7% ± 10%
	16 mm	191% ± 141%	410% ± 148%	419% ± 153%	-13% ± 6%	-2% ± 9%	-1% ± 10%
Liver	6 mm	40% ± 30%	48% ± 32%	128% ± 390%	-25% ± 8%	-27% ± 8%	-29% ± 7%
	11 mm	129% ± 249%	194% ± 308%	547% ± 1042%	-15% ± 8%	-20% ± 7%	-21% ± 7%
	16 mm	291% ± 469%	619% ± 905%	1420% ± 1683%	-9% ± 7%	-15% ± 6%	-16% ± 6%

Table 4

Comparison of lesion SUV_{max} and volume errors from motion-free to motion-blurred images in computer simulation for different lesion sizes and attenuation map. Data are shown for type-1 traces and 11 mm motion amplitude.

Lesion location	Lesion diameter and attenuation map	SUV_{max}	Percent error of lesion volume with thresholds at 50%
			SUV_{max}
Upper lung	2 cm EE	$-2.4\% \pm 1.0\%$	$0\% \pm 1\%$
	1 cm EE	$-2.2\% \pm 2\%$	$-3\% \pm 7\%$
	1 cm ACT	$-1.9\% \pm 1.7\%$	$0\% \pm 9\%$
Middle lung	2 cm EE	$-7.0\% \pm 5.0\%$	$4\% \pm 14\%$
	1 cm EE	$-22.2\% \pm 6.7\%$	$92\% \pm 88\%$
	1 cm ACT	$-20.1\% \pm 5.6\%$	$66\% \pm 66\%$
Lower lung	2 cm EE	$-11.8\% \pm 4.8\%$	N/A
	1 cm EE	$-20.5\% \pm 5.3\%$	N/A
	1 cm ACT	$-25.3\% \pm 5.3\%$	N/A
Liver	2 cm EE	$-15.5\% \pm 5.8\%$	$17\% \pm 15\%$
	1 cm EE	$-28.1\% \pm 6.3\%$	$129\% \pm 249\%$
	1 cm ACT	$-27.2\% \pm 6.2\%$	$131\% \pm 262\%$

^aEE: end-expiration; ACT: respiratory-averaged CT.

^bData are shown as mean $\pm$ standard deviation.