Data-driven gating in PET: Influence of respiratory signal noise on motion resolution

Florian Büther[a)]
Department of Nuclear Medicine, University Hospital Münster, Albert-Schweitzer-Campus 1, Münster 48149, Germany

Iris Ernst
German CyberKnife Centre, Senator-Schwartz-Ring 8, Soest 59494, Germany

Lynn Johann Frohwein
European Institute for Molecular Imaging, University of Münster, Waldeyerstr. 15, Münster 48149, Germany

Joost Pouw
Department of Nuclear Medicine, University Hospital Münster, Albert-Schweitzer-Campus 1, Münster 48149, Germany
Magnetic Detection and Imaging Group, University of Twente, Drienerlolaan 5, Enschede 7522 NB, The Netherlands

Klaus Peter Schäfers
European Institute for Molecular Imaging, University of Münster, Waldeyerstr. 15, Münster 48149, Germany

Lars Stegger
Department of Nuclear Medicine, University Hospital Münster, Albert-Schweitzer-Campus 1, Münster 48149, Germany

Purpose: Data-driven gating (DDG) approaches for positron emission tomography (PET) are interesting alternatives to conventional hardware-based gating methods. In DDG, the measured PET data themselves are utilized to calculate a respiratory signal, that is, subsequently used for gating purposes. The success of gating is then highly dependent on the statistical quality of the PET data. In this study, we investigate how this quality determines signal noise and thus motion resolution in clinical PET scans using a center-of-mass-based (COM) DDG approach, specifically with regard to motion management of target structures in future radiotherapy planning applications.

Methods: PET list mode datasets acquired in one bed position of 19 different radiotherapy patients undergoing pretreatment $[^{18}F]$FDG PET/CT or $[^{18}F]$FDG PET/MRI were included into this retrospective study. All scans were performed over a region with organs (myocardium, kidneys) or tumor lesions of high tracer uptake and under free breathing. Aside from the original list mode data, datasets with progressively decreasing PET statistics were generated. From these, COM DDG signals were derived for subsequent amplitude-based gating of the original list mode file. The apparent respiratory shift d from end-expiration to end-inspiration was determined from the gated images and expressed as a function of signal-to-noise ratio SNR of the determined gating signals. This relation was tested against additional 25 $[^{18}F]$FDG PET/MRI list mode datasets where high-precision MR navigator-like respiratory signals were available as reference signal for respiratory gating of PET data, and data from a dedicated thorax phantom scan.

Results: All original 19 high-quality list mode datasets demonstrated the same behavior in terms of motion resolution when reducing the amount of list mode events for DDG signal generation. Ratios and directions of respiratory shifts between end-respiratory gates and the respective nongated image were constant over all statistic levels. Motion resolution d/d_{max} could be modeled as $d/d_{max} = 1 - e^{-1.52(SNR-1)^{0.52}}$, with d_{max} as the actual respiratory shift. Determining d_{max} from d and SNR in the 25 test datasets and the phantom scan demonstrated no significant differences to the MR navigator-derived shift values and the predefined shift, respectively.

Conclusions: The SNR can serve as a general metric to assess the success of COM-based DDG, even in different scanners and patients. The derived formula for motion resolution can be used to estimate the actual motion extent reasonably well in cases of limited PET raw data statistics. This may be of interest for individualized radiotherapy treatment planning procedures of target structures subjected to respiratory motion. © *2018 American Association of Physicists in Medicine* [https://doi.org/10.1002/mp.12987]

1. INTRODUCTION

Respiratory motion has been identified as a major source of image degradation in positron emission tomography (PET) imaging of thoracic and abdominal structures for several decades now.[1] It leads to blurred images, a loss of effective spatial resolution, reduced standardized uptake values (SUV) of lesions, and increased apparent lesion volumes.[2] Furthermore, artifacts caused by attenuation correction based on computer tomography (CT) or magnetic resonance imaging (MRI) data are often connected to the presence of respiration-related co-registration failures of emission and transmission data.

Several potential solutions have been developed to reduce the influence of respiratory motion on PET image quality. Clinically, the most wide-spread solution is the concept of "gating" the acquired PET data.[3,4] Here, respiratory information is acquired in parallel to the PET data, usually by dedicated hardware such as pressure-sensitive belts (e.g., the Respiratory Gating System AZ-733V, Anzai, Tokyo, Japan; www.anzai-med.co.jp) or video camera systems (e.g., the Real-time Position Management System, Varian, Palo Alto, USA; www.varian.com). The respiration signals can then easily be used to prospectively or retrospectively sort the PET data into one or more subsets ("gates") with highly reduced motion influence. Multi-gated datasets can additionally provide estimates of respiratory motion extent.

However, although demonstrated to result in better image quality, these clinically available solutions are only slowly starting to be accepted in clinical routine scans. This may be connected to several potential drawbacks of these methods. First, they can be time-consuming and difficult to install, and require trained personnel. Additionally, they may fail to record accurate motion information.[5] Furthermore, they usually acquire external motion information as surrogate parameters of respiratory motion, not necessarily well correlated to internal organ movements.[6]

Data-driven gating (DDG) approaches have been developed as alternatives to conventional hardware-based methods. The basic idea in DDG is to analyze the measured PET raw data in order to extract respiratory motion information. This potentially allows for user-independence, requires no additional setup time, and is a software-based rather than a hardware-based approach.[7] In contrast to hardware-based methods, motion information is collected mainly from PET-active structures within the body instead of relying on external surface motion effects which could be problematic if effects like hysteresis come into play.[6] Several DDG approaches have been explored in the last few years, relying on different strategies such as exploiting the inhomogeneous geometric sensitivity of 3D PET systems,[8,9] measuring the center-of-mass (COM) of PET-active structures in PET list mode data,[10–12] analyzing sinogram fluctuations[13,14] or spectra of sinogram bin time-activity curves,[15,16] or applying dimension reduction techniques like principal component analysis (PCA).[17,18]

Recently, the first studies have been published investigating the effect of DDG in PET scans.[19,20] The results demonstrate that DDG methods on average perform as well as conventional methods in typical clinical settings.

Similar DDG methods have been successfully developed for other medical imaging devices like CT.[21] The well-known navigators in MR imaging can also be considered as DDG-like methods. In this respect, DDG methods for medical imaging have already proven to be valuable in standard clinical environments.

However, several aspects concerning DDG particular to PET applications still need to be investigated. One of these is related to the fact that the success of DDG is directly dependent on the statistical quality of measured PET raw data. Therefore, we initiated a retrospective study based on acquired high-quality PET list mode data in order to investigate how the amount of noise in DDG gating signals influences the success of motion resolution in a DDG approach. Eventually, this work may be of great interest for radiotherapy applications where specific knowledge of target motion trajectories may contribute to smaller planning volumes with consecutive sparing of healthy tissues, and may prove to be a valuable alternative to conventional CT- or fiducial marker-based approaches to determine target motion caused by respiration.

2. MATERIALS AND METHODS

2.A. PET systems

PET systems used within this study were the Biograph mCT PET/CT and the Biograph mMR PET/MRI (Siemens Healthcare GmbH, Erlangen, Germany).

The PET/CT scanner covered an axial PET field of view (FOV) of 21.8 cm and had a patient bore diameter of 78 cm and a spatial resolution of 4 mm FWHM at the center of the FOV.[22] List mode data were acquired in 3D with axial compression (span 11, 9 segments). Time-resolved list mode data comprised prompt $p(r,\varphi,s,\tau,t)$ and delayed $d(r,\varphi,s,\tau,t)$ sinogram datasets with 400 radial bins r, 168 angular bins φ, 621 total slices s, with 109 direct and indirect planes and 512 oblique planes, and 13 time-of-flight bins τ of 312 ps each.

The PET/MRI scanner covered an axial PET FOV of 25.8 cm with a patient bore diameter of 60 cm and a spatial resolution of 4 mm FWHM at the center of the FOV.[23] List mode data were acquired in 3D without axial compression. Prompt $p(r,\varphi,s,t)$ and delayed $d(r,\varphi,s,t)$ sinogram datasets comprise 344 radial bins r, 252 angular bins φ, and 4084 total slices s. This scanner did not offer time-of-flight information for measured coincidence events.

2.B. Patient/phantom data and scanning procedures

PET list mode datasets that were retrospectively analyzed within this study were taken from a patient cohort undergoing PET/CT or PET/MR imaging prior to high-precision radiotherapy treatment of malignant lesions (primary tumors or metastases) located in the liver or in the lower lungs. This

study was performed in accordance with the ethical standards of the local ethics committee and with the 1964 Helsinki declaration and its later amendments. PET/CT and PET/MR data were acquired with clinical indication, hence no explicit patient consent was needed for this retrospective analysis as per local institutional guidelines.

The patients fasted overnight before the scan. An intravenous dose of approximately 4 MBq [^{18}F]FDG per kg body mass was injected 1 h prior to scanning. Clinical whole-body PET/CT or PET/MRI scans were then performed in sinogram mode. In order to get high-quality PET images for treatment planning purposes, single bed-position list mode PET scans were finally performed for 5 to 10 min under free breathing with the malignant lesions centered in the PET FOV.

As a first step, 19 list mode datasets (14 from PET/CT, 5 from PET/MRI) with organs (myocardium, kidneys) or tumor lesions of high tracer uptake (SUV_{max} > 10 in the clinical whole-body scans) were chosen to investigate DDG motion resolution with progressively decreasing data statistics as described below ("training datasets").

Furthermore 25 PET/MRI list mode datasets comprising lesions or organs of SUV_{max} < 10 were available for subsequent testing ("testing datasets"). These were acquired in parallel to the acquisition of repeated fast sagittal T1w 2D TurboFLASH MR slices through the liver dome that were measured during the whole list mode PET scan (TR/TE = 140 ms/1.4 ms, flip angle: 12°, matrix size: 128 × 160, voxel size: 2.5 × 2.5 mm, slice thickness: 7 mm). This resulted in around 2140 frames for a 5 min PET scan. Synchronization to the list mode data was achieved by inserting specific tags within the PET list mode stream whenever acquisition of MR data for a single slice began and ended, thus precisely identifying when MR information was measured. Measurements using a moving phantom demonstrated that the synchronization accuracy between these MR tags and the PET list mode stream is below the repetition time TR of the MR scan (140 ms). This was deemed sufficient for respiratory gating purposes.

Finally, an in-house developed PET/MRI-compatible anthropomorphic thorax phantom was scanned in the PET/MRI scanner.[24] It comprised realistic inflatable lungs, a diaphragm-like membrane, a heart insert, and small movable tubes simulating PET-active lesions. For this study, [^{18}F]FDG solution was filled into the background (7 kBq/mL at scan start), the heart (33 kBq/mL), and a lesion located at the diaphragm (180 kBq/mL, volume: 0.3 mL). The respiratory motion amplitude of the lesion was set to 20 mm in axial direction; the mean respiratory cycle duration amounted to 6.5 s. PET list mode data was acquired for 5 min, with the lesion located in the center of the FOV.

2.C. Gating and reconstruction

The list mode datasets were processed for respiratory motion signals using the segmented COM gating approach.[11] This DDG method was utilized as it gives local motion signals from a specified volume-of-interest (VOI) and since it is expected to be the most direct measurement of (cranio-caudal) target motion within the list mode data. Briefly, this automatic method worked to determine the brightest regions in the reconstructed non-attenuation-corrected, nongated PET image. Then, a sufficiently large cylindrical volume-of-interest was determined around these, which was forward-projected into sinogram space. This led to a binary mask m with $m(r,\varphi,i) = 1$ for lines-of-response (LOR) crossing the VOI, and $m(r,\varphi,i) = 0$ for those missing the VOI, thus excluding the latter ones from further processing. A raw respiratory signal was calculated by determining the axial COM of all single slice-rebinned prompt events in 50 ms frames:

$$R_{raw}(t) = \frac{\sum_{r,\varphi,s} z \cdot p(r,\varphi,z,t) \cdot m(r,\varphi,z)}{\sum_{r,\varphi,s} p(r,\varphi,z,t) \cdot m(r,\varphi,z)} \tag{1}$$

where z denotes the axial coordinate of the prompt LOR after single slice rebinning. Time-of-flight information in the case of data originating from PET/CT was neglected. A smoothed gating signal $R(t)$ was determined from $R_{raw}(t)$ using a Fourier low-pass filter cutting out frequencies above 2/3 Hz, as these essentially represent noise.

For each of the 19 PET/CT and PET/MRI list mode files of the training datasets, this respiration signal was also calculated from modified raw data where only every 2nd, 4th, 8th, ..., 256th measured event was taken into account (corresponding to statistic levels x = 1/2, 1/4, 1/8, ..., 1/256), thereby effectively reducing the statistical quality of PET raw data for signal calculation. This resulted in nine respiratory gating signals for every bright region that was segmented from the non-attenuation-corrected, nongated PET image.

A signal-to-noise (SNR) metric that was already used to assess the quality of DDG methods in earlier publications is as follows:

$$SNR[R_{raw}(t)] = \frac{f_3 - f_2}{f_1 - f_0} \cdot \frac{\int_{f_0}^{f_1} |\mathrm{FT}[R_{raw}(t)]|\mathrm{d}f}{\int_{f_2}^{f_3} |\mathrm{FT}[R_{raw}(t)]|\mathrm{d}f} \tag{2}$$

with FT denoting the Fourier transform, $[f_0, f_1]$ = [1/12 Hz, 1/3 Hz] representing an interval of typical respiratory frequencies (corresponding to respiratory periods of 3 s to 12 s), and $[f_2, f_3]$ = [3 Hz, f_{max}] representing high frequencies caused by noise alone.[19,25] This concept is similar to the "quality factor" defined in Ref. [13]. Large SNR values reflect comparatively high amounts of respiratory signal within the raw signal, while values around 1.0 indicate practically white noise with no obvious respiratory features in the calculated signal. Typical values encountered for the employed DDG method in a similar study amounted to 1.8 ± 1.3.[25] This metric was calculated for every obtained raw signal.

Further respiration signals were derived from the sagittal MR frames for the 25 PET/MRI list mode files of the testing datasets and used for validation. Here, an axial profile was defined at the liver dome, and the axial coordinate of the liver — lung border was determined based on a threshold representing the mean value of averaged lung and liver values (MR navigator-like signal). With the help of the synchronization tags written into the list mode data, and after using a

spline interpolation to the DDG signal frames of 50 ms length, these coordinates (as a function of scanning time) were considered as ground truth respiratory signals for internal organ motion. These were thus used for additional gated PET reconstructions.

To evaluate motion resolution qualities for different respiratory signals, amplitude-based gated reconstructions of end-expiratory and end-inspiratory phases according to the signals were performed. More specifically, the amplitude interval from the maximum amplitude value to a lower threshold amplitude, defined so that 10% of the scanning time falls into this interval, was determined, and events measured during this interval were reconstructed (end-inspiration gate). Likewise, the interval from the minimum amplitude value to a higher one encompassing 10% of scanning time was determined and similarly reconstructed (end-expiration gate). These reconstructions were performed using an ordered subsets expectation maximization (OSEM) algorithm with 3 iterations and 21 subsets with normalization but without correcting for attenuation and scatter, as only geometric shifts of structures with high tracer uptake instead of quantified radioactivity units were considered in this study. This procedure avoided attenuation correction artifacts induced by mismatches between PET and CT/MRI images. All reconstructions were done using the STIR reconstruction software.[26] A 5 mm FWHM Gaussian filter was applied after reconstruction. Image dimensions for PET/CT reconstructions were $400 \times 400 \times 109$ voxels with volumes of $2.04 \times 2.04 \times 2.03$ mm^3, while PET/MRI-based images amounted to $400 \times 400 \times 127$ voxels with volumes of $2.08 \times 2.08 \times 2.03$ mm^3.

A COM DDG signal and its *SNR* value was similarly calculated from the thorax phantom list mode data. Gated reconstructions using DDG were performed just as in the patient scans.

2.D. Data analysis

Reconstructed images of end-expiration and end-inspiration phases were imported into the AMIDE visualization software (available at amide.sourceforge.net). For the 19 PET/CT and PET/MRI training datasets, the coordinates $(x_{insp}, y_{insp}, z_{insp})$, $(x_{exp}, y_{exp}, z_{exp})$ of the COM of the hottest structure were determined in the gated images, the *respiratory shift vector* $\vec{d} = (x_{insp} - x_{exp}, y_{insp} - y_{exp}, z_{insp} - z_{exp})$ was calculated, and its Euclidean length was regarded as the apparent *respiratory shift* $d = \sqrt{(x_{insp} - x_{exp})^2 + (y_{insp} - x_{exp})^2 + (z_{insp} - z_{exp})^2}$ of the structure when using the calculated DDG signal for gating. This parameter was used as a metric for the evaluation of motion resolution as it is robust against noise effects, as opposed to standard image intensity values like SUV$_{max}$ or derived values like lesion volumes. It has thus been used to characterize gating methods in the past.[10,12,19,20] Furthermore, an exact knowledge of the average shift vector between end-expiratory and end-inspiratory phases of target structures is

expected to be of benefit in radiotherapy treatment planning procedures to account for respiratory movements of target structures, and DDG PET-based motion delineation may have advantages over alternative methods.[27,28]

In order to investigate any general relationship between *SNR* and motion resolution, every plot of *d* against *SNR* was analyzed and fitted using the curve fitting tool of MATLAB (version R2016a; MathWorks, Natick, Mass, USA) to the following expression:

$$d(SNR) = d_{max}\left[1 - e^{-b(SNR-1)^c}\right] \qquad (3)$$

with d_{max} denoting the actual respiratory shift resolvable using end-expiratory and end-inspiratory gates as defined above. This is a general expression with the desired features $d(1) = 0$ and $\lim_{SNR \to \infty} d(SNR) = d_{max}$; furthermore, d/d_{max} can easily be interpreted as a metric for motion resolution. Optimal values for d_{max}, b, and c were individually determined for every dataset, and globally. The optimal values b and c found in the latter analysis served as a general model of motion resolution d/d_{max} vs *SNR* of the raw DDG signal.

To test this general model, the statistic level x which gave *SNR* values closest to the arbitrarily chosen value of 2.0 (which is on the order of typically encountered values[25]) was determined. For this level, the corresponding measured respiratory shift d was taken, and d_{max} was calculated according to the general model as

$$d_{max} = d \left/ \left[1 - e^{-b(SNR-1)^c}\right]\right. \qquad (4)$$

This was then compared to the individually determined d_{max} in order to demonstrate the internal consistency of the general model to the individual datasets. Additionally, for the same statistic level x resulting in *SNR* values around 2.0, the general model was used to estimate d at the original full data statistic level ($x = 1$), and compared to the actually measured value.

Furthermore, the COM of the considered structures in nongated reconstructions was determined. The Euclidean distances from this COM to the already determined COM in end-expiration and end-inspiration were calculated for all statistic levels and termed d_{exp} and d_{insp}, respectively. The ratios d_{insp}/d_{exp} were analyzed for changes with varying statistic levels x.

For the additional 25 PET/MRI testing datasets, reconstructions were done only for full list mode statistics ($x = 1$). Distance measurements between end-expiratory and end-inspiratory phases were performed for all organs or lesions demonstrating noticeable tracer uptake present in the PET images. The determined respiratory shifts d of the DDG-based images were additionally extrapolated to d_{max} according to the general model mentioned above, and the shift values d based on DDG, on the generalized model [d_{max} as calculated using Eq. (4), *SNR* and d], and on the MR navigator-like signal (d_{MR}), respectively, were compared. Additionally, the angles α between DDG-based and MR-based shift vectors $\vec{d}$ and $\vec{d}_{MR}$ were determined as:

$$\alpha = \cos^{-1} \frac{\vec{d} \cdot \vec{d}_{MR}}{d \cdot d_{MR}} \tag{5}$$

This was done in order to not only compare the lengths d and d_{MR} of the different shift vectors $\vec{d}$ and $\vec{d}_{MR}$, respectively, but also their relative orientation in space, since not only deviations in length but also in relative angles will lead to different motion trajectories and, when used in radiotherapy planning procedures, different radiation volumes.

Both the image-derived and extrapolated shift values from the phantom scan were compared to the predefined value of 20 mm.

Finally, a similarity analysis of DDG-based and MR-derived respiratory curves was performed by calculating Pearson correlation coefficients r and relating them to the SNR value of the DDG raw signals.

d and the extrapolated d_{max} calculated using the achieved SNR and the general model were also determined for the phantom scan data and compared to the predefined actual shift value of 20 mm.

All numerical data are given as mean $\pm$ standard deviation. Statistical testing was performed using paired Wilcoxon's signed rank tests with Holm–Bonferroni corrections in MATLAB. P values smaller than 0.05 were considered to show significant differences.

3. RESULTS

Within the training group, DDG signals were calculated using automatically determined myocardial VOI in 10 cases, and liver/lung lesion VOI in four cases, for the PET/CT datasets. Prompt and true event count rates amounted to 371 ± 61 kcps and 206 ± 36 kcps, respectively. For the PET/MRI datasets, automatically determined myocardial VOI were used in four cases, and liver lesion VOI were used in one case. Prompt and true event count rates amounted to 448 ± 103 kcps and 221 ± 37 kcps, respectively. Within the testing group, a myocardial VOI was used in nine cases for DDG signal generation, a kidney VOI was used in three cases, while a liver/lung lesion VOI was used in 13 cases for DDG determination. Prompt and true event count rates amounted to 360 ± 99 kcps and 188 ± 45 kcps, respectively.

SNR values of the original list mode datasets of the training group containing all acquired events amounted to 9.2 ± 2.8 (range: 6.0–16.0).

Decreasing the amount of events x from 1 to 1/256 used for DDG signal generation led to higher amounts of noise in the calculated DDG signals (Fig. 1) and a corresponding decrease in SNR values. This relation could be expressed as $SNR = 1 + a_1 x^{a_2}$ in all cases, both for PET/CT and PET/MRI datasets, with $a_1 \approx SNR(x = 1) - 1$ and a_2 being determined to be 0.59 ± 0.04 in the training group (Fig. 2).

The determined respiratory shifts when using the original list mode datasets at $x = 1$ amounted to 13.8 ± 5.7 mm (range: 5.9–24.4 mm) in the training group. These shifts generally decreased when excluding progressively more data
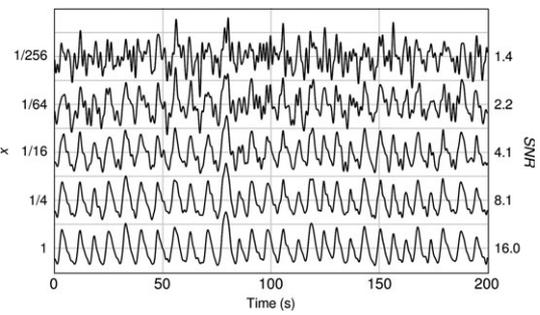
FIG. 1. Example of calculated DDG signals for different statistic levels x derived from a list mode dataset of the PET/CT subgroup used for training.
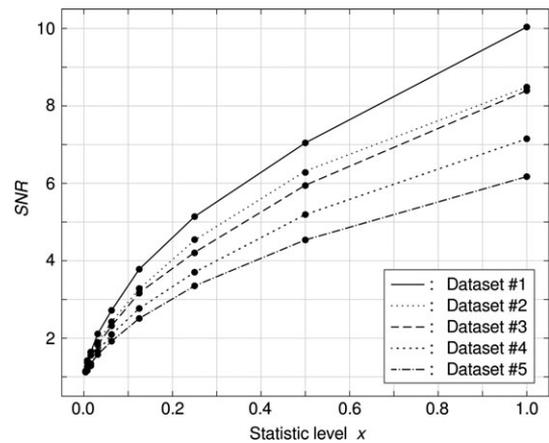
FIG. 2. SNR for different statistic levels x, shown here only for the PET/MRI subgroup of the training group for the sake of clarity.

from calculation of DDG signals. Qualitatively, for large SNR values, d values start on a plateau, but quickly decrease toward 0 when SNR reaches values of around 1. This behavior can be expressed by $d(SNR) = d_{max}\left[1 - e^{-b(SNR-1)^c}\right]$, where d_{max} denotes the plateau height, and b and c describe the specific appearance of the curve. Within the training group, average b and c were determined by individual fitting as 1.57 ± 0.24 (range: 1.22–2.17) and 0.52 ± 0.11 (range: 0.31–0.73), respectively.

Fitting of all individually determined d/d_{max} values of the training group to $1 - e^{-b(SNR-1)^c}$ led to optimal values of $b = 1.52$ and $c = 0.52$ (95% confidence interval: 1.47–1.57 and 0.49–0.55 for b and c, respectively; Fig. 3), resulting in the following general motion resolution model for COM DDG:

$$d/d_{max} = 1 - e^{-1.52(SNR-1)^{0.52}} \tag{6}$$

No apparent difference in this motion resolution behavior between the PET/CT and PET/MR subgroups were detected (Fig. 3).

Measured d values at those statistic levels with SNR values closest to 2.0 were 11.1 ± 4.8 mm, significantly smaller than the individually determined d_{max} values at 14.1 ± 5.8 mm ($P = 0.00013$), while there was no significant difference between the latter and the d_{max} values resulting from the
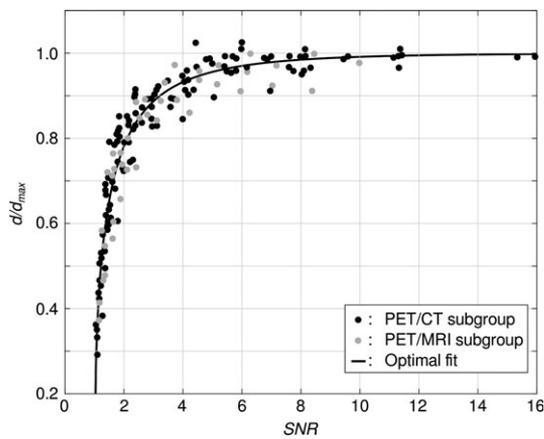
FIG. 3. Normalized respiratory shifts for different *SNR* in the PET/CT subgroup (black) and in the PET/MRI subgroup (grey). Black line denotes the optimal fit.
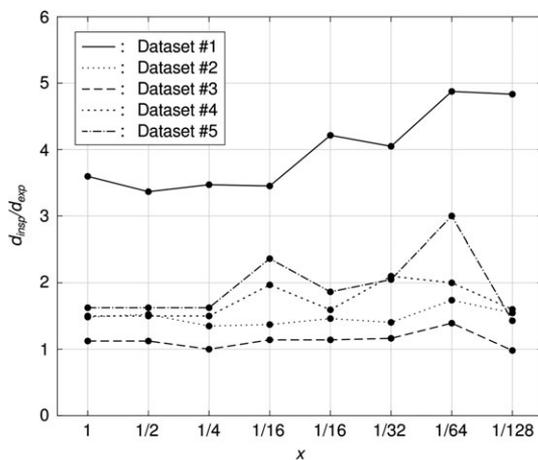
FIG. 4. Ratios d_{insp}/d_{exp} for different statistic levels in the PET/MRI subgroup of the training group.

general model [Eq. (6); 14.3 ± 6.1 mm; $P = 0.38$]. Furthermore, there was no significant difference between the general model-based shifts at statistic level $x = 1$ calculated using the values at $SNR \approx 2.0$ (14.1 ± 6.1 mm) and the actually measured values using all emission events for DDG signal calculation (13.8 ± 5.7 mm, $P = 0.17$).

The ratios d_{insp}/d_{exp} in the training group remained constant over all statistic levels; generally, no systematic drifts or shifts were detectable (Fig. 4). For $x = 1$, the ratios amounted to 2.1 ± 0.9 (range: 1.1–4.6), for $x = 1/8$, they were 2.2 ± 1.0 (range: 1.1–5.3), and for $x = 1/64$ they were 2.4 ± 1.1 (range: 1.4–5.4). Differences between d and $d_{insp} + d_{exp}$ were small and less than a deviation of 1.5 mm for all datasets and statistic levels (on average always less than 10% of d for all statistic levels), indicating that the angle between the inspiration vector and the expiration vector is always very close to 180°.

The respiratory shifts of 55 target structures in the 25 PET/MRI PET scans of the testing group were analyzed. These comprised 10 hearts, 6 kidneys, and 39 hepatic and lower lung lesions. On average, *SNR* values for the testing group amounted to 3.9 ± 1.6 (range: 1.4–8.1); the actually

measured respiratory shifts d using DDG were 11.4 ± 5.9 mm (range: 2.4–29.3 mm). d_{max} calculated from the general model [Eq. (6)] amounted to 12.7 ± 6.2 mm [range: 2.7–29.7 mm; Fig. 5(a)], while the shifts d_{MR} based on the MR-derived gating were 12.7 ± 6.2 mm [range: 3.0–30.2 mm; Fig. 5(b)]. The difference in motion resolution was highly significant between the original DDG shifts d and the MR-based d_{MR} ($P < 10^{-9}$) as expected, but no significant difference was found between the extrapolated shifts d_{max} and d_{MR} ($P = 0.63$). This was also apparent in the determined linear fits. These demonstrated major differences in slope and intercept to identity, that is, $f(x) = x$, when comparing d to either d_{max} or d_{MR} (Fig. 5).

On average, α amounted to 5.5 ± 5.1° (range: 0–24°); larger angles were only found when d or d_{MR} were comparatively small [Fig. 6(a)]. Mean Pearson correlation coefficients r between DDG- and MR-derived respiratory signals amounted to 0.87 ± 0.09 (range: 0.58–0.98); there was a clear correlation between these and *SNR* of the DDG raw signal which could be approximated by a similar expression as the general motion resolution model:

$$r(SNR) = 1 - e^{-1.40(SNR-1)^{0.46}} \qquad (7)$$

The measured shift value d for the simulated lesion in the phantom scan amounted to 17 mm, underestimating the shift by 3 mm. Usage of the general model and the achieved *SNR* value of 2.8 results in d_{max} of 19.5 mm, 0.5 mm smaller than the actual predefined shift.

4. DISCUSSION

DDG methods are promising alternatives to conventional hardware-based gating approaches. They have already proven to perform similarly well in first clinical trials.[19,20] On the other hand, there are still challenges that need to be investigated. One is the question of how the statistical quality of the acquired PET data influences the outcome of gating. The presented study contributes to this issue by progressively decreasing the amounts of PET raw data for COM-derived DDG signal calculation.

As expected, the ability to resolve respiratory motion decreases monotonically with decreased amounts of data, as this introduces statistical noise masking the true respiratory signal (Fig. 1). Surprisingly, this loss in motion resolution was found to be fairly similar in a wide range of investigated patient scans (and thus different patient geometries, lesion locations, actual respiratory shifts etc.) and different scanner systems, enabling a general model of motion resolution as a function of *SNR* calculated from the DDG raw signal [Fig. 3, Eq. (6)]. The *SNR* value thus seems to be a potentially robust metric not only for intra-, but also for inter-patient and -scanner comparisons (although this has to be proven in a wider range of scans). This allows judging how large the *SNR* has to be in order to resolve, for example, at least 90% of the actual motion shift d_{max}: $d/d_{max} > 0.9$, thus $SNR > 3.2$ according to the general model. Alternatively, if the motion should be
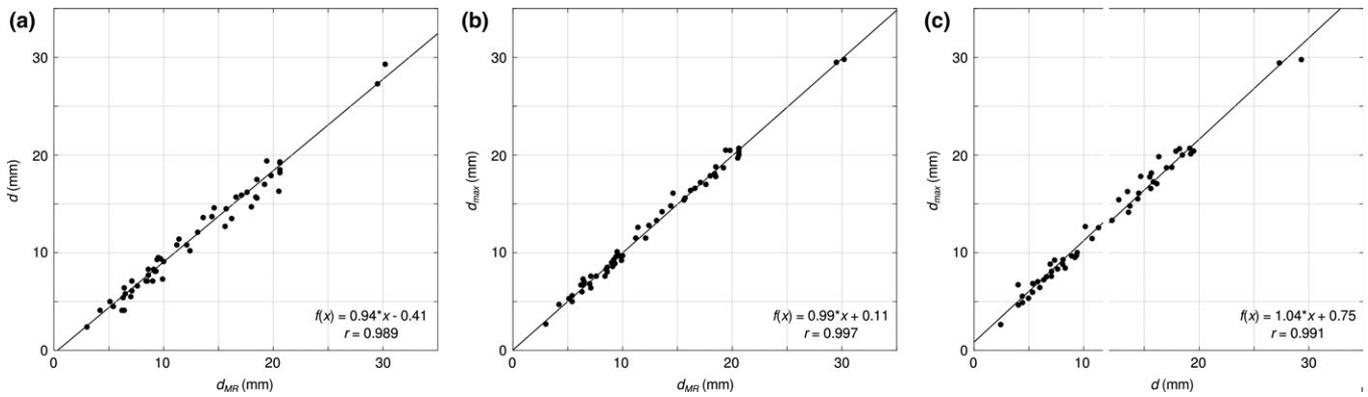
FIG. 5. Scatter plots of d (a) and d_{max} (b) against d_{MR}, and of d_{max} against d (c) of the testing group. Black lines denote linear fits.
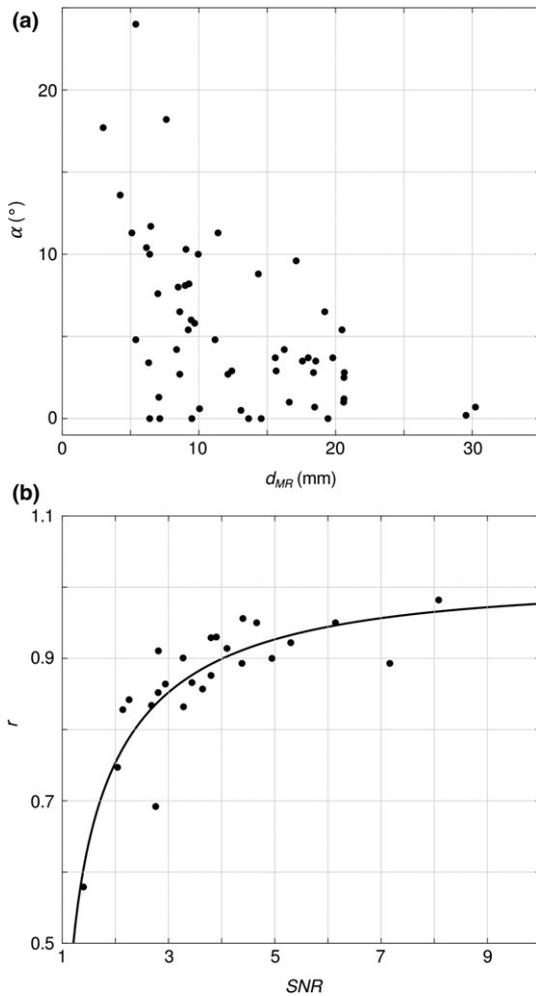
FIG. 6. Determined angles α against d_{MR} (a); correlation coefficients between DDG- and MR-based respiratory signals against SNR (b). Black line denotes the optimal fit.

resolved to below the spatial resolution of the scanner (4 mm FWHM) at a determined shift d of 15 mm, then $d/d_{max} > 15/(15 + 4)$, therefore, SNR should at least be 2.0.

Additionally, the found relation allows a reasonable extrapolation to d_{max} in cases where limited statistical qualities of measured PET data prohibit an exact measurement from the gated images. This was demonstrated using this relation to

estimate d_{max} in several PET/MRI scans where internal motion information from fast sagittal MR images was available as ground truth gating signals, and also in a scan using a dedicated human thorax phantom with realistic motion patterns. It was found that this extrapolation led to an excellent agreement between DDG and MR-based shifts. On the other hand, the native original DDG-based shifts underestimated the whole motion extent (Fig. 5). In the phantom scan, the general model almost exactly reproduced the predefined motion extent of 20 mm. Furthermore, the deviations in angle between DDG- and MR-based shift vectors are small, especially for larger shifts, demonstrating that the whole trajectory vector calculated using COM DDG and MR-based gating are very similar (Fig. 6). Finally, the ratio d_{insp}/d_{exp} was found to be fairly constant over all statistic levels (Fig. 4). Therefore, not only the absolute value of the actual respiratory shift d_{max} but also its location relative to the averaged lesion position can be approximated by the general model.

It should be noted that the found motion resolution model is only valid for the chosen 10% gated end-inspiration and end-expiration gates and for SNR as defined here. Any other definition will lead to different curves and fitting parameters. Specifically, the apparent respiratory shift of structures will be smaller when increasing the effective amount of used data per gate.[29]

While the results of this study may seem to be more of theoretical interest, it is interesting to note that SNR as defined here was found to be a reliable metric of gating success for a motion resolution-related parameter like the actual respiratory shift of PET-active structures. This is different from hardware-based methods where there is no way of judging whether the acquired signal really reflects internal motion. The results shown here are in line with previously published results, where a quality factor derived solely from a frequency analysis of the DDG signal was shown to predict the correlation to a hardware-driven respiratory signal.[13] The general model presented here quantifies this in terms of motion resolution and thus gating success.

Furthermore, it may be of special interest within the context of personalized radiotherapy treatment planning. A precise knowledge of respiratory trajectories of thoracic and abdominal target structures is needed in order to deliver high

doses to the target while at the same time irradiating as little healthy surrounding tissue as possible. DDG-based trajectories may be an attractive alternative to other methods of measuring and incorporating motion information into the treatment planning process like using breath-hold CT, 4D CT, or tracking with invasive fiducial markers. Due to prolonged PET scanning times, many respiratory cycles are measured, allowing statistically defined trajectories which on average potentially represent physiological motion better than CT-based approaches with their typically limited number of monitored respiratory cycles. Future work will have to show if DDG can really be used for this purpose, but first simulation studies seem to demonstrate promising results.[30] A potential disadvantage of DDG PET-based motion information may be the limited spatial resolution of current PET systems (around 4 mm FWHM for the scanners used in this study), which is still inferior to high-resolution modalities such as CT. However, especially for large structures and high target-to-background uptake ratios, differences in target COM positions between different respiratory phases (and hence averaged respiratory trajectories) can be measured with high precision if the quality of the PET raw data is high enough. This had already been demonstrated in phantom scans with predefined respiratory shifts[11] and is again confirmed in the phantom scan performed in this study.

It remains to be seen if other DDG methods yield comparable results. The COM method was chosen for this study as it gives the most direct measurement of actual internal motion of an isolated PET-active structure. Furthermore, based on our experience, it tends to give raw signals with less noise than sensitivity-based methods,[11] thus making it easier to find high-quality cases for analyzing effects of decreasing statistics. Additionally, the definition of *SNR* as given here is straight-forward for the COM method. For other methods like PCA-based ones, an accurate SNR metric will have to be defined in a different way and may lead to other results, potentially not giving simple relations like the one found here for a wide range of patients.

This study has some limitations. First, the amount of patient and phantom data analyzed is limited, and patients with extreme breathing patterns may not behave according to the determined relation between noise and motion resolution. Additionally, it remains to be seen how the findings of this study hold up in cases where no isolated high-uptake structures are present in the FOV of the scanner.

Furthermore, it remains to be proven that knowledge of a respiratory shift vector $\vec{d}$ (or, slightly more complex, two shift vectors $\vec{d}_{insp}$, $\vec{d}_{exp}$ pointing from an averaged, ungated position to the extreme respiration phases) is sufficient to adequately describe an averaged target trajectory for use in radiotherapy treatment planning. Considering the amount of respiration irregularities encountered in patients under free breathing, this is certainly a challenging task especially for small target volumes. Additionally, the basic question if target trajectories determined by DDG PET can be used for radiotherapy treatment which might take place on a different day in a different setting remains to be answered. However,

considering that some CT-based techniques with acquisition in end-inspiration and end-expiration are still used for radiotherapy planning, DDG PET-based approaches may prove to be a step forward in terms of accuracy in nongated personalized high-precision applications of radiotherapy.

Finally, it is not clear if the MR navigator-like respiratory signals, used here as ground truth gating signals, really reflect valid signals for structures other than the lung–liver border. There is evidence that hysteresis effects could lead to a decrease in motion correlation between different parts of the body.[6] On the other hand, the strong correlation between d_{max} and d_{MR} seems to demonstrate that this was not a large effect in the presented study. Nevertheless, this should be investigated in greater detail in future studies.

5. CONCLUSION

The influence of raw data statistics on motion resolution success in a COM-based DDG method for PET has been investigated. A simple general model connecting motion resolution and SNR of the DDG raw signal was found, giving both reasonable quality estimates of the gating success in COM-based DDG as well as accurate motion vector estimates. This was successfully tested against MR-based respiratory data serving as internal ground truth signals. The model may thus allow precise estimations of lesion movements due to respiration. This may be of potential interest in personalized radiotherapy treatment planning approaches, especially for those cases where PET imaging is part of pretreatment imaging strategies anyway. In these cases DDG PET may not only result in "sharp" motion-resolved images well suited for gaining detailed information about metabolic active parts of target structures, but also in their respiratory trajectories. Knowledge of these could help to deliver high radiation doses to the target while keeping the burden for healthy tissue as low as possible. The comparatively long PET acquisition duration guarantees statistically sound data compared to 4D CT applications which usually comprise only a limited number of respiratory cycles. However, future studies comprising larger patient cohorts will have to confirm whether DDG-based motion trajectories are really accurate enough for this purpose; additionally, the question of how to optimally incorporate this knowledge into the treatment planning or radiation procedure has to be answered. Finally, other DDG PET methods should be investigated for their statistical properties and abilities to resolve respiratory motion.

ACKNOWLEDGMENTS

CONFLICTS OF INTEREST

a)Author to whom correspondence should be addressed. Electronic mail: butherf@uni-muenster.de; Telephone: +49 251 8349336; Fax: +49 251 8347363.

REFERENCES

1. Ter-Pogossian MM, Bergmann SR, Sobel BE. Influence of cardiac and respiratory motion on tomographic reconstructions of the heart: implications for quantitative nuclear cardiology. *J Comput Assist Tomogr*. 1982;6:1148–1155.

2. Liu C, Pierce LA, Alessio AM, Kinahan PE. The impact of respiratory motion on tumor quantification and delineation in static PET/CT imaging. *Phys Med Biol*. 2009;54:7345–7362.

3. Nehmeh SA, Erdi YE, Ling EE, et al. Effect of respiratory gating on reducing lung motion artifacts in PET imaging of lung cancer. *Med Phys*. 2002;29:366–371.

4. Boucher L, Rodrigue S, Lecomte R, et al. Respiratory gating for 3-dimensional PET of the thorax: feasibility and initial results. *J Nucl Med*. 2004;45:214–219.

5. Didierlaurent D, Ribes S, Batatia H, et al. The retrospective binning method improves the consistency of phase binning in respiratory-gated PET/CT. *Phys Med Biol*. 2012;57:7829–7841.

6. Dasari PK, Shazeeb MS, Könik A, et al. Adaptation of the modified Bouc-Wen model to compensate for hysteresis in respiratory motion for the list-mode binning of cardiac SPECT and PET acquisitions: testing using MRI. *Med Phys*. 2014;41:112508.

7. Kesner AL, Schleyer PJ, Büther F, Walter MA, Schäfers KP, Koo PJ. On transcending the impasse of respiratory motion correction applications in routine clinical imaging – a consideration of a fully automated data driven motion control framework. *EJNMMI Phys*. 2014;1:8.

8. He J, O'Keefe GJ, Gong SJ, et al. A novel method for respiratory motion gating with geometric sensitivity of the scanner in 3D PET. *IEEE Trans Nucl Sci*. 2008;55:2557–2565.

9. He J, O'Keefe GJ, Geso M. Motion image compensation based on dynamic data in PET acquisition. *J Inf Comp Sci*. 2010;7:885–891.

10. Büther F, Dawood M, Stegger L, et al. List mode-driven cardiac and respiratory gating in PET. *J Nucl Med*. 2009;50:674–681.

11. Büther F, Ernst I, Dawood M, et al. Detection of respiratory tumour motion using intrinsic list mode-driven gating in positron emission tomography. *Eur J Nucl Med Mol Imaging*. 2010;37:2315–2327.

12. Heß M, Büther F, Schäfers KP. Data-driven methods for the determination of anterior-posterior motion in PET. *IEEE Trans Med Imaging*. 2017;36:422–432.

13. Kesner AL, Bundschuh RA, Detorie NC, et al. Respiratory gated PET derived in a fully automated manner from raw PET data. *IEEE Trans Nucl Sci*. 2009;56:677–686.

14. Kesner AL, Kuntner C. A new fast and fully automated software based algorithm for extracting respiratory signal from raw PET data and its comparison to other methods. *Med Phys*. 2010;37:5550–5559.

15. Schleyer PJ, O'Doherty MJ, Barrington SF, Marsden PK. Retrospective data-driven respiratory gating for PET/CT. *Phys Med Biol*. 2009;54:1935–1950.

16. Schleyer PJ, O'Doherty MJ, Marsden PK. Extension of a data-driven gating technique to 3D, whole body PET studies. *Phys Med Biol*. 2011;56:3953–3965.

17. Thielemans K, Rathore S, Engbrant F, Razifar P. Device-less gating for PET/CT using PCA. In: 2011 IEEE Nuclear Science Symposium and Medical Imaging Conference (NSS/MIC). IEEE; 2011.

18. Bertolli O, Arridge S, Wollenweber SD, Stearns CW, Hutton BF, Thielemans K. Sign determination methods for the respiratory signal in data-driven PET gating. *Phys Med Biol*. 2017;62:3204–3220.

19. Büther F, Vehren T, Schäfers KP, Schäfers M. Impact of data-driven respiratory gating in clinical PET. *Radiology*. 2016;281:229–238.

20. Kesner AL, Chung JH, Lind KE, et al. Validation of software gating: a practical technology for respiratory motion correction in PET. *Radiology*. 2016;281:239–248.

21. Pan T, Lee T, Rietzel E, Chen GTY. 4D-CT imaging of a volume influenced by respiratory motion on multi-slice CT. *Med Phys*. 2004;31:333–340.

22. Jakoby BW, Bercier Y, Conti M, Casey ME, Bendriem B, Townsend DW. Physical and clinical performance of the mCT time-of-flight PET/CT scanner. *Phys Med Biol*. 2011;56:2375–2389.

23. Delso G, Fürst S, Jakoby B, et al. Performance measurements of the Siemens mMR integrated whole-body PET/MR scanner. *J Nucl Med*. 2011;52:1–9.

24. Bolwin K, Czekalla B, Frohwein LJ, Buther F, Schäfers KP. Anthropomorphic thorax phantom for cardio-respiratory motion simulation in tomographic imaging. *Phys Med Biol*. 2018;63:035009.

25. Büther F, Ernst I, Hamill J, et al. External radioactive markers for PET data-driven respiratory gating in positron emission tomography. *Eur J Nucl Med Mol Imaging*. 2013;40:602–614.

26. Thielemans K, Tsoumpas C, Mustafovic S, et al. STIR: software for tomographic image reconstruction release 2. *Phys Med Biol*. 2012;57:867–883.

27. Ernst I, Buether F, Moustakis C, et al. Detection of target motion for liver SBRT by 4D list mode PET/CT: an analysis based on phantom measurements, interfractional cone beam CT scans, and 510 clinical data. *Int J Radiat Oncol Biol Phys*. 2014;90:S374.

28. Frood R, Prestwich R, Tsoumpas C, Murray P, Franks K, Scarsbrook A. Effectiveness of respiratory-gated positron emission tomography/computed tomography for radiotherapy planning in patients with lung carcinoma: a systematic review. *Clin Oncol*. 2018;30:225–232.

29. Dawood M, Büther F, Stegger L, et al. Optimal number of respiratory gates in positron emission tomography: a cardiac patient study. *Med Phys*. 2009;36:1775–1784.

30. Reinartz G, Haverkamp U, Wullenkord R, et al. 4D-Listmode-PET-CT and 4D-CT for optimizing PTV margins in gastric lymphoma: determination of intra- and interfractional gastric motion. *Strahlenther Onkol*. 2016;192:322–332.