NEW

UNIDOS® Tango & Romeo

The next generation of reference class electrometers

In 1922, PTW introduced the first therapy X-ray dosemeter that was a technological milestone. In 2020, PTW continues to lead the industry by releasing its third generation of UNIDOS electrometers — taking performance and usability to the next level.

UNIDOS® Tango

- Intelligent Detector Recognition
- Full remote control via LAN or WLAN connectivity using your smartphone, tablet or PC
- Industry-leading resolution (0.1 fA/1 fC)

UNIDOS® Romeo & Tango

- High-precision reference class electrometers
- Large, intuitive, multilingual touchscreen display
- Versatile detector database
- Auto start/stop with auto reset
- Standalone and/or remote operation

Intelligent Detector Recognition: Watch the video.

PTW THE DOSIMETRY COMPANY

Contact PTW today for all your reference dosimetry needs.
(516) 827-3181 • www.PTWdosimetry.com

Real-time 4DMRI-based internal target volume definition for moving lung tumors

Moritz Rabe[a)] and Christian Thieke
Department of Radiation Oncology, University Hospital, LMU Munich, Munich 81377, Germany

Mathias Düsberg
Department of Radiation Oncology, Klinikum rechts der Isar, Technical University, Munich 81675, Germany

Sebastian Neppl, Sabine Gerum, and Michael Reiner
Department of Radiation Oncology, University Hospital, LMU Munich, Munich 81377, Germany

Nils Henrik Nicolay
Medical Center, University of Freiburg, Freiburg 79106, Germany

Heinz-Peter Schlemmer
Department of Radiology, German Cancer Research Center, Heidelberg 69120, Germany

Jürgen Debus
Department of Radiation Oncology, University Hospital of Heidelberg, Heidelberg 69120, Germany
Heidelberg Institute of Radiation Oncology (HIRO), Heidelberg 69120, Germany

Julien Dinkel
Department of Radiology, University Hospital, LMU Munich, Munich 81377, Germany

Guillaume Landry
Department of Radiation Oncology, University Hospital, LMU Munich, Munich 81377, Germany
Department of Medical Physics, Ludwig-Maximilians-Universität München (LMU Munich), Garching 85748, Germany

Katia Parodi
Department of Medical Physics, Ludwig-Maximilians-Universität München (LMU Munich), Garching 85748, Germany

Claus Belka
Department of Radiation Oncology, University Hospital, LMU Munich, Munich 81377, Germany
German Cancer Consortium (DKTK), Munich 81377, Germany

Christopher Kurz*
Department of Radiation Oncology, University Hospital, LMU Munich, Munich 81377, Germany
Department of Medical Physics, Ludwig-Maximilians-Universität München (LMU Munich), Garching 85748, Germany

Florian Kamp*
Department of Radiation Oncology, University Hospital, LMU Munich, Munich 81377, Germany

(Received 2 August 2019; revised 20 December 2019; accepted for publication 7 January 2020; published 10 February 2020)

Purpose: In photon radiotherapy, respiratory-induced target motion can be accounted for by internal target volumes (ITV) or mid-ventilation target volumes (midV) defined on the basis of four-dimensional computed tomography (4D-CT). Intrinsic limitations of these approaches can result in target volumes that are not representative for the gross tumor volume (GTV) motion over the course of treatment. To address these limitations, we propose a novel patient-specific ITV definition method based on real-time 4D magnetic resonance imaging (rt-4DMRI).

Methods: Three lung cancer patients underwent weekly rt-4DMRI scans. A total of 24 datasets were included in this retrospective study. The GTV was contoured on breath-hold MR images and propagated to all rt-4DMRI images by deformable image registration. Different targets were created for the first (reference) imaging sessions: ITVs encompassing all GTV positions over the complete (ITV^{80s}) or partial acquisition time (ITV^{10s}), ITVs including only voxels with a GTV probability-of-presence (POP) of at least 5% ($ITV^{5\%}$) or 10% ($ITV^{10\%}$), and the mid-ventilation GTV position. Reference planning target volumes (PTV_r) were created by adding margins around the ITVs and midV target volumes. The geometrical overlap of the PTV_r with $ITV_n^{5\%}$ from the six to eight subsequent imaging sessions on days n was quantified in terms of the Dice similarity coefficient (DSC), sensitivity [SE: $(PTV_r \cap ITV_n^{5\%})/ITV_n^{5\%}$] and precision [PRE: $(PTV_r \cap ITV_n^{5\%})/PTV_r$] as surrogates for target coverage and normal tissue sparing.

Results: Patient-specific analysis yielded a high variance of the overlap values of PTV_r^{10s}, when different periods within the reference imaging session were sampled. The mid-ventilation-based

PTVs were smaller than the ITV-based PTVs. While the SE was high for patients with small breathing pattern variations, changes of the median breathing amplitudes in different imaging sessions led to inferior SE values for the mid-ventilation PTV for one patient. In contrast, $PTV_r^{5\%}$ and $PTV_r^{10\%}$ showed higher SE values with a higher robustness against interfractional changes, at the cost of larger target volumes.

Conclusions: The results indicate that rt-4DMRI could be valuable for the definition of target volumes based on the GTV POP to achieve a higher robustness against interfractional changes than feasible with today's 4D-CT-based target definition concepts. [https://doi.org/10.1002/mp.14023]

Key words: 4DMRI, interfractional changes, ITV, lung tumor, mid-ventilation, motion management

1. INTRODUCTION

In high-precision radiotherapy (RT) of lung tumors, target motion due to respiration remains a predominant challenge.[1] This motion can be substantial, is patient-specific, difficult to predict, can be irregular and change from one day to another.[2–8] Intrafractional motion, defined as any motion induced by physiological processes occurring within a treatment session such as breathing, and interfractional changes occurring between treatment sessions need to be accounted for in the planning and delivery process.[9]

To address intrafractional changes, passive and active motion management techniques have been developed over the last decades.[10] Passive methods include motion-encompassing margins and abdominal compression. Active methods include active breathing control (ABC), breath-hold techniques, gating, and tracking. While active approaches have a higher potential in reducing the integral dose to the patient, they can be invasive when fiducial markers are implanted, complex, costly, time-consuming, require specialized equipment, or are still in the research phase. The clinical benefit of many of these methods remains to be proven.[3,4] For these reasons, passive motion-management (PMM) techniques are still primarily used clinically, in particular for conventionally fractionated RT.[1,10] The use of internal target volumes (ITV) as a motion-encompassing method is described in Report 83 of the International Commission on Radiation Units and Measurements (ICRU).[11] Assessment of the range of motion by four-dimensional computed tomography (4D-CT) imaging has become the clinical standard-of-care.[4,12] The ITV is ideally obtained from the union of all gross tumor volumes (GTVs) delineated on the datasets at the different breathing phases.[7] It ideally includes all possible positions of the GTV throughout the course of treatment.[13] The ITV is then expanded by margins to account for interfractional changes and patient setup uncertainties to create the planning target volume (PTV). An alternative PMM technique is based on the definition of the mid-position[14] or mid-ventilation target volume (midV),[15], where the average position of the GTV is reconstructed from the 4D-CT and motion-dependent anisotropic margins are added to account for the respiration-induced target motion. The resulting PTVs are typically smaller than corresponding ITVs.[4,16] Although the midV concept has the potential to reduce the integral dose to the lungs, its clinical implementation is typically limited to academic RT centers.[17]

It is questionable whether the characterization of respiratory motion and the derivation of a corresponding PTV based on a single pretreatment 4D-CT is representative and adequate for treatment planning.[13,18–21] 4D-CT images are averaged over only a few breathing cycles and the target volume defined based on these images is therefore subject to random uncertainties,[22] since the breathing of the patient during this short acquisition time might not be representative for the patient's breathing pattern during treatment. Clinically relevant interfractional anatomical changes such as tumor shrinkage, weight loss or normal tissue alterations like pleural effusion, and onset or resolution of atelectasis are frequently observed during RT of lung tumors.[5,9,10] The patient's breathing pattern can change due to psychological factors, such as an increase of relaxation of the patient over the course of treatment.[23] These changes are in general not predictable but introduce systematic errors in the treatment that can compromise the quality of the RT treatment due to reduced target coverage or additional dose to organs at risk.[1,24]

Online-adaptive magnetic resonance imaging (MRI)-guided RT may deliver highly conformal doses to the tumor by adjusting the treatment plan just before the treatment session if interfractional changes occurred.[25] The equipment needed for this method is, however, not widespread.[26] Today, the majority of clinics use daily volumetric imaging with three-dimensional cone-beam CT (CBCT) for positioning in image-guided RT (IGRT) of lung cancer[27] through which relevant interfractional anatomical changes can be detected. Four-dimensional CBCT imaging can improve the accuracy of patient positioning compared to 3D-CBCT[28] and enables a motion assessment directly before treatment. This motion assessment is subject to random uncertainties related to the averaging of motion due to the reconstruction of a single breathing cycle from projections from several breathing cycles with potential inter-cyclic variations.[3] Therefore, the question whether the 4D-CT-based ITV used for treatment planning is still adequately representing the target motion cannot be answered.

The introduction of MRI in the RT workflow[29] has spurred research that could potentially contribute to the solution of this problem. The movement of lung tumors has been investigated with two-dimensional (2D) cine-MRI, which

enables the repeated acquisition of temporally resolved images of arbitrary duration with high soft tissue contrast without delivering dose to the patient.[5,6] Several studies have investigated intrafractional and interfractional changes with 2D cine-MRI.[6,30–35] Cai et al.[36] showed that the characterization of lung tumor motion with temporally resolved MRI is more representative than with 4D-CT, mainly due to improved statistics through longer acquisition times.[22] Hence, temporally resolved MRI can help to refine the ITV definition for treatment planning.

When 2D cine-MR images are used, the information about the target motion is limited to 2D planes and only indirect inferences of the out-of-plane motion of tumor and surrounding tissues can be drawn.[37] The extension to four-dimensional MRI (4DMRI) is therefore an active field of research,[38,39] through which the 3D motion of the target and surrounding tissues including translations, rotations and deformations can be directly captured. As pointed out in a recent review on 4DMRI,[39] research is mainly focused on respiratory-correlated 4DMRI (rc-4DMRI) opposed to real-time 4DMRI (rt-4DMRI).[39–42] In contrast to 4D-CT imaging and rc-4DMRI, no retrospective sorting of projections or 2D images from different breathing cycles based on a surrogate is needed for rt-4DMRI. Fast 3D gradient echo (GRE) or steady-state free precession (SSFP) sequences using parallel imaging techniques and echo sharing are usually used for rt-4DMRI.[42] The in-plane resolution is 3-4 mm and temporal resolution is typically limited to 2 volumes per second, depending on the spatial resolution.[39] rt-4DMRI is not routinely used clinically today, but a more widespread use due to the technological advances in this area is expected in the near future.[39]

The purpose of this proof-of-concept study is to demonstrate how rt-4DMRI could be used to reduce the random and systematic uncertainties associated with today's 4D-CT-based ITV definition approach. We describe how a 4DMRI-based ITV can be defined based on the probability-of-presence (POP) of the GTV to reduce random uncertainties. Additional PTV margins are added to reduce systematic uncertainties and prospectively account for potential interfractional changes. The new ITV definition concept is evaluated by analyzing the geometrical overlap of the obtained PTV with ITVs based on the GTV motion on different days. The method is compared to today's PMM concepts including PTVs that mimic a 4D-CT-based ITV definition and the mid-ventilation approach. We show how systematic errors due to interfractional changes could be quantified with regular 4D-MR imaging by metrics that are correlated to dosimetric quantities such as target coverage and normal tissue sparing. The potential integration of the proposed novel ITV definition into clinical practice is outlined and its limitations are discussed.

2. MATERIALS AND METHODS

2.A. Patient data and imaging protocols

Three patients with tumors in the right lung were included in this retrospective proof-of-concept study. The patients underwent regular MR imaging in treatment position with a 1.5 T scanner (Siemens Avanto) with 7-9 imaging sessions distributed over 11-12 weeks. A total of 24 datasets were acquired for all patients accumulated. A 3D image in breathhold (3DMRI) was acquired with a balanced steady-state free precession (bSSFP) sequence (TrueFISP; axial slices; slice thickness: 4-5 mm; in-plane resolution: 0.88×0.88 mm^2; TR/TE: 380/1.16 ms; flip angle: 63°; field-of-view (FOV): $45 \times 45 \times 24$ cm^3; receiver bandwidth: 1030 Hz/px) using parallel imaging (GRAPPA) at the beginning of each session. The acquisition time for each axial slice was 400 ms, resulting in a total acquisition time of 20-24 s for the whole 3D volume. An rt-4DMRI dataset (4DMRI) was subsequently acquired. A total of 157 3D volumes were collected over a period of 80 s with a temporal resolution of 500 ms and the patient breathing freely (TWIST; coronal slices; slice thickness: 10 mm; in-plane resolution: 3.91×3.91 mm^2; TR/TE: 1.47/0.61 ms; flip angle: 5°; FOV: $50 \times 50 \times 36$ cm^3; receiver bandwidth: 1565 Hz/px). TWIST (Time-resolved angiography With Interleaved Stochastic Trajectories)[44] is a dynamic 3D GRE sequence using view-sharing, where the center of *k*-space is sampled more frequently than the periphery in a semi-randomized fashion. Parallel imaging (GRAPPA) and partial Fourier imaging (sampling of 78% in frequency-encoding direction) were used as additional acceleration techniques to further shorten the image acquisition time. To account for geometrical distortions, the manufacturer's correction methods were applied to 3DMRI and 4DMRI. The GTV was contoured on 3DMRI and approved by a trained radiation oncologist.

The first imaging session for each patient was defined as the reference imaging session and taken as a surrogate for the imaging that would be acquired for RT planning. The consecutive imaging sessions $n \in [2,\ldots,N]$, where N is the total number of MRI sessions for the patient, were taken as surrogates for the treatment sessions over the course of therapy.

2.B. Study workflow

To generate 4DMRI-based target volumes, the GTV position at every point in time of 4DMRI is needed. An overview of the overall workflow performed for this purpose and the subsequent evaluation is given in Fig. 1. The main steps (indicated by the corresponding numbers in Fig. 1) are:

1. Determination of the 3D dataset within 4DMRI that closest resembles the breathing phase of 3DMRI, labeled 4DMRI(t'),
2. warping of the GTV from 3DMRI to 4DMRI(t') by deformable image registration (DIR),
3. propagation of the GTV to all 3D volumes within 4DMRI using DIR,
4. definition of 4DMRI-based ITVs and the midV,
5. creation of PTVs by expansion of the ITVs and midV,
6. creation of a time-averaged 4DMRI,
7. rigid registration (RR) focused on the tumor of time-averaged MR images acquired on day n to the dataset from the reference MRI session,

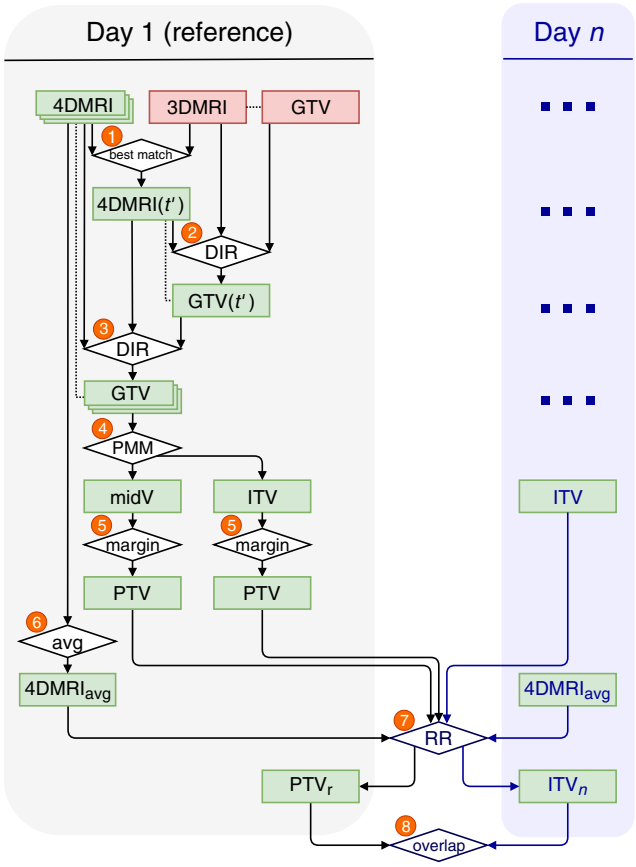

FIG. 1. Overview of data processing and analysis workflow. The individual steps, indicated by the numbers in this figure, are further described in Sections 2.C and 2.D. The dots in the field of day *n* indicate that the same workflow steps as shown for day 1 (reference) are performed. PTV_r represents the reference PTV and ITV_n the ITV on day *n*. GTV: gross tumor volume; DIR: deformable image registration; PMM: passive motion management; ITV: internal target volume; PTV: planning target volume; RR: rigid registration. [Color figure can be viewed at wileyonlinelibrary.com]

8. geometrical overlap evaluation of PTVs from the reference day (PTV_r) and ITVs from day *n* (ITV_n).

The individual steps are described in more detail in the following sections.

2.C. Definition of 4DMRI-based target volumes

2.C.1. Determination of breathing states

The superior-inferior (SI) diaphragm positions in each 3D volume of 4DMRI and 3DMRI were determined as surrogates for the breathing state. A region of interest around the transition between the lung and abdominal tissue in the nontumor-bearing hemithorax was manually selected and converted to a binary image by thresholding using Otsu's method.[45] The thresholded image was summed over the anterior-posterior (AP) and right-left (RL) image axes to create a one-dimensional (1D) signal in SI direction. The diaphragm position was defined at the SI position of the steepest gradient of the 1D signal. The point in time t' was defined as the time at which the difference of the absolute diaphragm positions of 3DMRI and any of the 3D images of 4DMRI was minimal (step 1 in Fig. 1). The corresponding 3D image at t' of 4DMRI, 4DMRI(t'), was used for all subsequent registration steps. This step was performed to find the image 4DMRI (t') that is as similar as possible to 3DMRI to reduce uncertainties of the DIR in the following step.

2.C.2. Deformable image registration

3DMRI and 4DMRI were resampled to a grid with an isotropic voxel size of $2 \times 2 \times 2$ mm^3 using linear 3D interpolation for all subsequent registration steps. To obtain the GTV on 4DMRI at t' (GTV(t')), 3DMRI was registered to 4DMRI(t') (step 2 in Fig. 1) in a multi-level b-spline DIR with mutual information as similarity metric using the software Plastimatch.[46] The DIR was focused on the GTV, expanded by isotropic margins of a few centimeters. Lower uncertainties were expected for the registration of images acquired with the same sequence and therefore a similar contrast (4DMRI(t') and 4DMRI(t_i)) compared to a registration of images acquired with two different sequences (3DMRI and 4DMRI(t_i)). Therefore, to determine the GTV at each point in time t_i in 4DMRI, GTV(t_i), with $i \in [1,\ldots,157]$, 4DMRI(t') was registered to 4DMRI(t_i), again using a multi-level b-spline DIR focused on the GTV with mutual information as similarity metric (step 3 in Fig. 1). The centroid positions of GTV(t_i) at all time steps t_i were measured to evaluate the motion amplitudes of the tumor in each breathing cycle. The amplitudes were calculated relative to the median exhale position of the GTV, since this position is expected to have a lower variance than the inhale position.[4,23]

2.C.3. Probability-of-presence ITV

The binary images GTV(t_i) were summed over all t_i and divided by the total number of time steps for normalization. The resulting 3D image has the same spatial resolution as the resampled 4DMRI ($2 \times 2 \times 2$ mm^3 voxel size). The voxel values correspond to the percentage of time in which the GTV was present at the voxels' positions over the whole acquisition time. If the acquisition time is long enough so that the respiratory-induced motion during 4DMRI is representative for the motion on the given day, this percentage becomes a POP. This 3D POP distribution was used to determine 4DMRI-based POP ITVs (step 4 in Fig. 1) for cutoffs of 5% and 10% ($ITV^{5\%}$ and $ITV^{10\%}$). All voxel values equal to or greater than the cutoff were set to 1, while all other voxel values were set to 0.

2.C.4. Conventional ITV and midV

To compare $ITV^{5\%}$ and $ITV^{10\%}$ with conventional PMM concepts, further target volumes were defined based on the rt-4DMRI dataset (step 4 in Fig. 1). This step was only performed for the reference imaging sessions:

1. An ITV encompassing all GTV positions over the whole acquisition time of 80 s (ITV_r^{80s}), which corresponds to a POP cutoff of 0%.
2. To mimic today's standard 4D-CT-based ITV definition workflow, the whole acquisition time was subdivided in eight 10 s periods. For each period, resembling a single 4D-CT scan, an ITV was created by including all GTV position within this period (ie, at 20 consecutive time points) to create eight ITV_r^{10s}. A similar approach was previously used by Thomas et al.[33]
3. The midV was defined by determining the centroid of the centroids of all 157 GTV positions and then selecting the GTV position at the point in time with the smallest distance of its centroid to this point. This method is equivalent to the approaches previously applied by Ehrbar et al.[47] and Thomas et al.[48]

2.C.5. PTV formation

For each patient, the ITVs and the midV of the reference imaging session were expanded by margins to create PTV_r (step 5 in Fig. 1). An isotropic margin of 5 mm was used for the expansion of the ITVs based on current clinical practice.[27] For the midV, the van Herk formula[43] was used to calculate the margins M_d in directions d (with $d \in$ [RL,SI,AP]):

$$M_d = \alpha\sqrt{\Sigma_{\text{setup}}^2 + \Sigma_{\text{BL}}^2 + \Sigma_{\text{del}}^2} + \beta\left(\sqrt{\sigma_{\text{setup}}^2 + \sigma_{\text{BL}}^2 + \sigma_{\text{br},d}^2 + \sigma_{\text{p}}^2} - \sigma_{\text{p}}\right), \quad (1)$$

including the systematic and random setup errors for patient positioning with CBCT ($\Sigma_{\text{setup}} = \sigma_{\text{setup}} = 0.8$ mm[49]), uncertainties due to baseline shifts over the course of treatment ($\Sigma_{\text{BL}} = 0.99$ mm and $\sigma_{\text{BL}} = 1.08$ mm[50]), delineation uncertainties ($\Sigma_{\text{del}} = 1.7$ mm[47]), the standard deviation of the breathing motion of the GTV in direction d ($\sigma_{\text{br},d}$) and the Gaussian beam penumbra width in lung ($\sigma_{\text{p}} = 6.4$ mm[16]). The approximation $\sigma_{\text{br},d} = A_{\text{cen}}^d/3$ was used, where A_{cen}^d is the median motion amplitude of the GTV centroid on the reference day in SI, AP and LR direction.[47] The values α=2.5 and β=1.64 were chosen to ensure a minimum of 95% of the prescribed dose to the target for 90% of the patients.[43] The margins M_d were rounded up to integer millimeter values. The ITVs and the midV were expanded by the margins on a $1 \times 1 \times 1$ mm^3 isovoxel grid and then resampled to the original $2 \times 2 \times 2$ mm^3 isovoxel grid. As a result of steps 1-5, 12 PTV_r were created for the reference imaging session of each patient: $PTV_r^{5\%}$, $PTV_r^{10\%}$, PTV_r^{80s}, eight PTV_r^{10s} and PTV_r^{midV}.

2.C.6. Rigid registration (RR)

Time-averaged 4DMRI images ($4DMRI_{avg}$) were calculated by averaging over the 3D volumes of all time steps t_i (step 6 in Fig. 1). To mimic patient positioning focused on the moving target during IGRT treatment, which is inevitably blurred on 3D-CBCT images,[3] $4DMRI_{n,avg}$ from day n, was rigidly registered to the reference image $4DMRI_{r,avg}$ (step 7 in Fig. 1) also focused on the target region. The RR ITK implementation in Plastimatch with mutual information as similarity metric was used.

2.D. Quantification of target volume overlaps

To evaluate the ability of the different PTVs to correctly predict the GTV positions during the course of treatment, a geometrical volume overlap analysis was performed. The reference PTVs of the first MRI sessions (PTV_r) were used as surrogates for the target volumes that would be used for RT treatment planning and the ITV of day n with a POP of 5% ($ITV_n^{5\%}$) as a surrogate for the real target position during treatment on day n. A cutoff POP of 5% was chosen with the goal to ensure a minimum dose of 95% to the GTV.[43]

2.D.1. Geometrical volume overlap analysis

The binary structures PTV_r and $ITV_n^{5\%}$ were compared on a voxel-by-voxel basis. A PTV_r voxel with a value of 1 is a prediction that parts of the GTV will be present during RT delivery at its position with a nonnegligible probability. An $ITV_n^{5\%}$ voxel with a value of 1 indicates that on day n, parts of the GTV are present at its position during at least 5% of the time. The total number of voxels that:

1. have a value of 1 in PTV_r and $ITV_n^{5\%}$ was labeled *True Positive* (TP),
2. have a value of 1 in PTV_r and a value of 0 in $ITV_n^{5\%}$ was labeled *False Positive* (FP) and
3. have a value of 0 in PTV_r and a value of 1 in $ITV_n^{5\%}$ was labeled *False Negative* (FN).

This is illustrated in Fig. 2. The geometrical volume overlap of PTV_r with $ITV_n^{5\%}$ was quantified in terms of sensitivity and precision, which are defined as following:

Sensitivity (SE; *true positive rate*):

$$\text{SE} = \frac{\text{TP}}{\text{TP} + \text{FN}} = \frac{\text{PTV}_r \cap \text{ITV}_n^{5\%}}{\text{ITV}_n^{5\%}} \quad (2)$$

Precision (PRE; *positive predictive value*):

$$\text{PRE} = \frac{\text{TP}}{\text{TP} + \text{FP}} = \frac{\text{PTV}_r \cap \text{ITV}_n^{5\%}}{\text{PTV}_r} \quad (3)$$

The Dice similarity coefficient (DSC) can be expressed as a function of SE and PRE:

$$\text{DSC} = \frac{2 \cdot \text{TP}}{2 \cdot \text{TP} + \text{FP} + \text{FN}} = \frac{2 \cdot (\text{PTV}_r \cap \text{ITV}_n^{5\%})}{\text{PTV}_r + \text{ITV}_n^{5\%}} = 2 \cdot \frac{\text{SE} \cdot \text{PRE}}{\text{SE} + \text{PRE}} \quad (4)$$

In the context of RT, FN can be interpreted as the extent of target miss, FP as the normal tissue damage and SE and

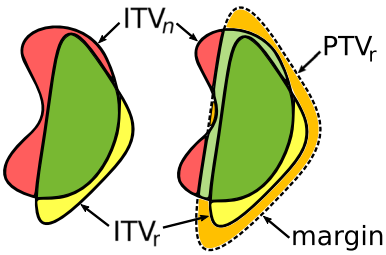

FIG. 2. Illustration of the overlap of PTV$_r$ and ITV$_n$. The assignment of attributes to the different regions is based on the inclusion or exclusion of these regions in PTV$_r$ and ITV$_n$. True positive (TP) areas are marked in dark green, false positive (FP) areas in yellow and false negative (FN) areas in red. In the right image, ITV$_r$ is expanded by an isotropic margin to obtain PTV$_r$, indicated by the dashed line, through which TP is increased and FN decreased (light green area) at the cost of an increased FP (orange area). [Color figure can be viewed at wileyonlinelibrary.com]

PRE as the scalar measures of target coverage and normal tissue sparing with respect to the target volume. In an ideal RT treatment, SE and PRE would both be 100%. The general goal of RT is to maximize SE (ie, the tumor coverage) while keeping PRE (ie, the normal tissue sparing) at an acceptable level. The margins around the ITVs and midV to create the PTV increase the SE while PRE decreases. The $(N-1)$ volume overlaps of each PTV$_r$ and ITV$_n^{5\%}$ were evaluated in terms of SE, PRE and DSC (step 8 in Fig. 1). For the PTV$_r^{10s}$, the "best" PTV$_r^{10s}$ (highest SE) and "worst" PTV$_r^{10s}$ (lowest SE) were determined for each patient.

3. RESULTS

3.A. Motion amplitudes

The median number of breathing cycles recorded in the 4DMRI sessions was 15.5 (range: [11.5, 24]) over an acquisition time of 80 s. This corresponds to a median breathing rate of approximately 12 cycles per minute (range: [9, 18]). MRI sessions with stable breathing amplitudes and frequency, as well as sessions with irregular breathing, were observed for all patients. Figure 3(a) shows the GTV centroid motion split up into the SI, RL and AP components for one exemplary MRI session. Figures 3(b)-3(d) depict the motion amplitudes with respect to the median exhale position for all MRI sessions for each patient. The median motion amplitudes are reported in Table I. For all patients and MRI sessions combined, the largest GTV centroid motion was observed in SI direction with a median motion amplitude of 8.8 mm, followed by RL direction (2.7 mm) and AP direction (2.2 mm).

3.B. Probability-of-presence ITV and midV

An exemplary POP distribution is shown in Fig. 4. The gradient of the POP distribution is steep in directions with small motion amplitudes (RL and AP) and shallow in the direction of large motion amplitudes (SI). The median volume ratio of ITV^{80s} to GTV is reported in Table I. Since this value correlates with the motion amplitude, the largest median value of ITV^{80s}/GTV was obtained for Patient 2 (3.25), the lowest for Patient 1 (1.29). Averaged over all patients, the mean 3D distance between the centroid of the centroids of all 157 GTV positions and the midV centroid was 0.5 mm.

3.C. PTV formation

All reference ITVs were expanded by isotropic 5 mm margins. For the reference midV, the direction-dependent term $\sigma_{\mathrm{br},d}$ in the van Herk formula yielded anisotropic margins. The median motion amplitudes during the reference imaging session were 3.5, 0.7, and 2.5 mm in RL direction and 1.8, 0.8, and 3.2 mm in AP direction for Patients 1, 2, and 3, respectively (cf. Fig. 3), which yielded 6 mm margins in both directions. The median SI motion amplitudes of 5.5, 13.2, and 15.7 mm for Patients 1, 2, and 3, respectively, yielded PTV margins of 6, 8, and 9 mm in SI direction.

3.D. Geometrical volume overlap analysis

The results of the geometrical overlap analysis of the different PTV$_r$ with the ITV$_n^{5\%}$ are depicted in Fig. 5. Table II summarizes the results and includes the volumes of the PTV$_r$ relative to the volume of PTV$_r^{80s}$.

By definition, the volume of PTV$_r^{80s}$ was the largest, representing the most conservative motion management approach of the analyzed PTV$_r$, where also single extreme GTV positions are included in the ITV. Consequently, the SE values were the largest (median SE between 98-100% for all patients) and the PRE values the lowest of all analyzed PTV$_r$.

The variance of the overlap values for the eight PTV$_r^{10s}$ of the different 10 s periods depends on the regularity of the breathing pattern during the reference imaging session. Since regular breathing patterns were measured for Patients 1 and 3 [cf. Fig. 3(a)], the differences between the overlap values of the best (highest SE) and worst (lowest SE) PTV$_r^{10s}$ were small ($\leq$2%). As the breathing pattern of Patient 2 was irregular during the reference imaging session, the variance of the different PTV$_r^{10s}$ was larger. A difference of 19% between the SE of the best and worst PTV$_r^{10s}$ was measured (99% vs 80%). Compared to PTV$_r^{80s}$, the same median SE values (differences <1%) at higher PRE (+4-6%) due to the PTV$_r$ volume reduction of up to 11% could be achieved for the best PTV$_r^{10s}$ for all patients. The differences in SE between the PTV$_r^{80s}$ and the worst PTV$_r^{10s}$ ranged from 0% (Patient 1) to 19% (Patient 2).

The volumes of PTV$_r^{\mathrm{midV}}$ were the smallest of all PTV$_r$ with a volume reduction between 16% and 56% compared to PTV$_r^{80s}$ which led to the highest observed median PRE values ($\approx$67% for all patients) that were 10-25% higher than for PTV$_r^{80s}$. The median SE values for PTV$_r^{\mathrm{midV}}$ were the smallest of all PTV$_r$ for the three patients, with the largest difference for Patient 2, where a median SE of 68% was measured (compared to 99% for PTV$_r^{80s}$).

For PTV$_r^{5\%}$, a median PTV volume reduction with respect to PTV$_r^{80s}$ of 9-31% was achieved at a similar median SE for Patients 1 and 3 (differences below 2%) but a reduced median

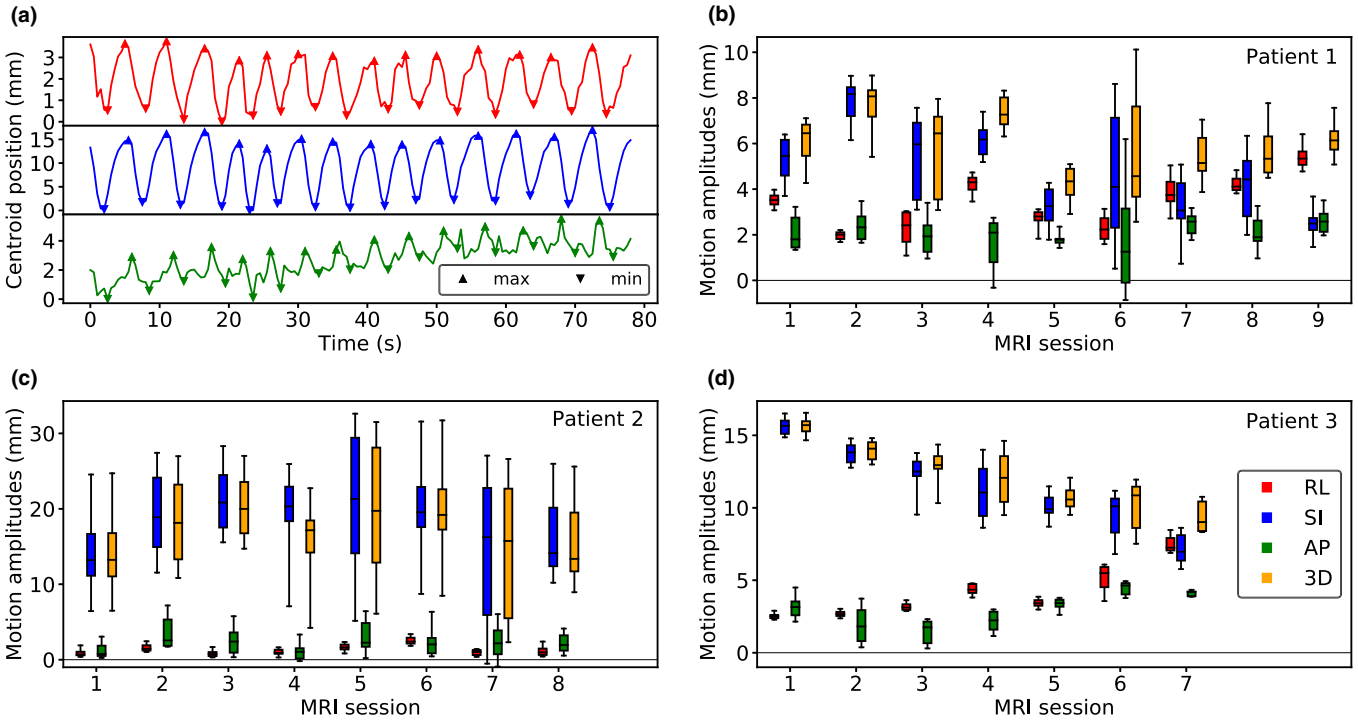

FIG. 3. Gross tumor volume centroid motion amplitudes. (a) Centroid position as a function of time for an exemplary magnetic resonance imaging (MRI) session. The triangular markers indicate the detected maxima and minima for each breathing cycle that were used to calculate the motion amplitudes in right–left (RL) (red), superior–inferior (SI) (blue), and anterior–posterior (AP) (green) directions. (b–d) Motion amplitudes in RL, SI and AP direction and in 3D for all MRI sessions and patients. The whiskers of the boxplots indicate the 5th and 95th percentiles. [Color figure can be viewed at wileyonlinelibrary.com]

TABLE I. Motion parameters for all patients. A^{d}_{cen} is the median motion amplitude of the gross tumor volume (GTV) centroid of all four-dimensional magnetic resonance imaging (4DMRI) sessions in three-dimensional (3D), right–left, superior–inferior, and anterior–posterior direction.

Patient	GTV position	$A^{\mathrm{3D}}_{\mathrm{cen}}$ (mm)	$A^{\mathrm{RL}}_{\mathrm{cen}}$ (mm)	$A^{\mathrm{SI}}_{\mathrm{cen}}$ (mm)	$A^{\mathrm{AP}}_{\mathrm{cen}}$ (mm)	GTV (ml)	$\frac{\mathrm{ITV}^{80s}}{\mathrm{GTV}}$
1	Middle right lobe	6.0	3.3	4.4	2.1	273	1.29
2	Lower right lobe	17.2	1.2	18.4	2.0	16	3.25
3	Lower right lobe	12.1	3.4	11.5	3.0	190	1.58

SE for Patient 2 (90% compared to 99%). The SE values for $\mathrm{PTV}_{\mathrm{r}}^{5\%}$ were higher than for $\mathrm{PTV}_{\mathrm{r}}^{\mathrm{midV}}$ by up to 22% at the cost of reduced PRE (ie, larger PTV volumes) by up to 11%. The use of $\mathrm{PTV}_{\mathrm{r}}^{10\%}$ instead of $\mathrm{PTV}_{\mathrm{r}}^{5\%}$ showed only small improvements of the PRE (≤4%) while the SE dropped by up to 7%.

The behavior of SE and PRE over time is mainly influenced by interfractional changes such as tumor shrinkage and different breathing patterns. The absolute variance of the median motion amplitudes for the different days n was low for Patient 1 (standard deviation of median 3D centroid motion amplitude was 1.3 mm) and the median motion amplitudes were decreasing over time for Patient 3. In combination with the regression or stagnation of the GTV size over time that was observed for these two patients, this led to a stable SE for all days n and all $\mathrm{PTV}_{\mathrm{r}}$ (standard deviation of SE values <4% for both patients and all $\mathrm{PTV}_{\mathrm{r}}$).

In contrast, the interfractional variance of the SE of the different $\mathrm{PTV}_{\mathrm{r}}$ for Patient 2 was considerably larger (up to 10% difference for $\mathrm{PTV}_{\mathrm{r}}^{\mathrm{midV}}$; cf. Fig. 6). This is a consequence of the high absolute variance of the median motion amplitudes for different days n [cf. Fig. 3(c)]. While the median GTV centroid motion amplitudes in the first and last MR imaging session were equivalent (13 mm), the amplitudes in the remaining imaging sessions were substantially larger (16-20 mm). This led to markedly higher SE values for the last compared to the other imaging sessions for most of the $\mathrm{PTV}_{\mathrm{r}}$ (cf. Fig. 6). The PRE for Patient 2 was gradually decreasing over time for all $\mathrm{PTV}_{\mathrm{r}}$, which can be explained by the tumor shrinkage that was observed for this patient.

4. DISCUSSION

The largest motion amplitudes were observed in SI direction. Considerably larger motion amplitudes were measured for GTVs in the lower lobe (Patients 2 and 3) than in the middle lobe (Patient 1). These findings are consistent with observations described in literature.[4,9,51,52]

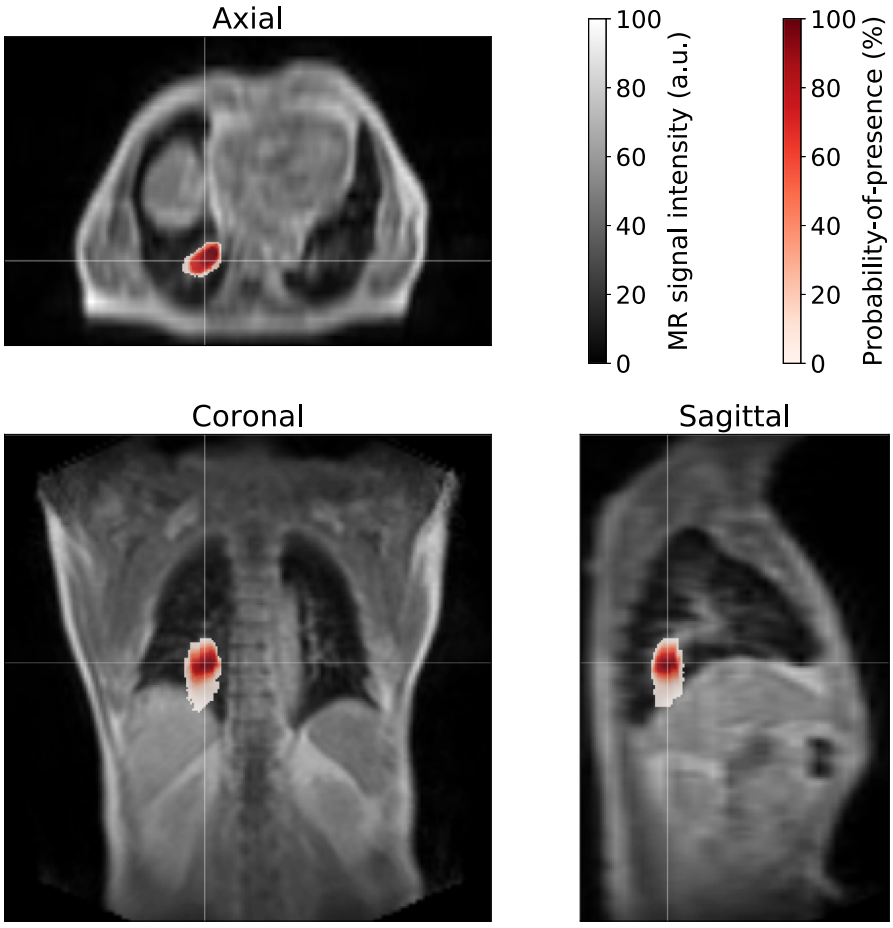

FIG. 4. Three-dimensional probability-of-presence (POP) distribution. The POP is color-encoded and superimposed on four-dimensional-magnetic resonance imaging (MRI) (t') for the reference MRI session of Patient 2. The crosshairs indicate the slice positions in the respective axial, coronal, and sagittal views. [Color figure can be viewed at wileyonlinelibrary.com]

The PTV_r^{80s}, as the most conservative PMM approach investigated in this study, yielded the highest SE values at the largest PTV volumes that might be unacceptably large for clinical application.

The large difference of the SE between the best and worst PTV_r^{10s} that was observed for Patient 2 demonstrates the risk of sampling the breathing motion over only a short time period. When an ITV is defined based on this motion sample—as is done routinely today in the 4D-CT-based ITV workflow—intrafractional uncertainties would directly translate to systematic errors impacting the whole RT treatment.

The SE values for Patient 1 were high for all PTV_r and all imaging sessions. This patient would likely benefit from being treated with the mid-ventilation approach, as the PTV_r^{midV} was smaller and hence the PRE was larger than for the ITV-based PTV_r. The low SE values of most PTV_r for Patient 2 indicate that the respiratory-induced GTV motion captured on the reference day was not representative for the motion of the GTV in the remaining imaging sessions. In particular, the SE of PTV_r^{midV} and $ITV_n^{5\%}$ was as low as 68%. A plausible explanation for this result can be inferred from Fig. 3(c). The variance of the SI GTV centroid motion amplitudes during the reference imaging session was high. The breathing motion was considered as a random uncertainty in the calculation of the PTV margin for the midV. However, since only the median motion amplitude was used for the margin determination, the rich information about the complex GTV motion trajectory was lost in this simplification step. Although large SI amplitudes of about twice the median value were observed during the reference imaging session, this was not appropriately accounted for by the PTV margins which in turn led to the poor SE. The SE values for Patient 3 were higher than 95% for all PTV_r with the exception of PTV_r^{midV} for which a SE of 90% was measured at a volume reduction of 23% relative to PTV_r^{80s}. Without a dosimetric analysis it cannot be concluded whether a SE value of 90% would be clinically acceptable.

The overall variance of the SE for all patients accumulated was second lowest for $PTV_r^{5\%}$ (after PTV_r^{80s}) at a high median SE value, indicating a higher robustness of this approach compared to PTV_r^{10s} and PTV_r^{midV} (cf. Fig. 5). While the SE for $PTV_r^{5\%}$ was reduced for Patient 2, it was not as low as for PTV_r^{midV}. By using the full information of the GTV motion over the whole duration of the reference imaging session, random uncertainties in the PTV definition could be reduced through the use of $ITV_r^{5\%}$.

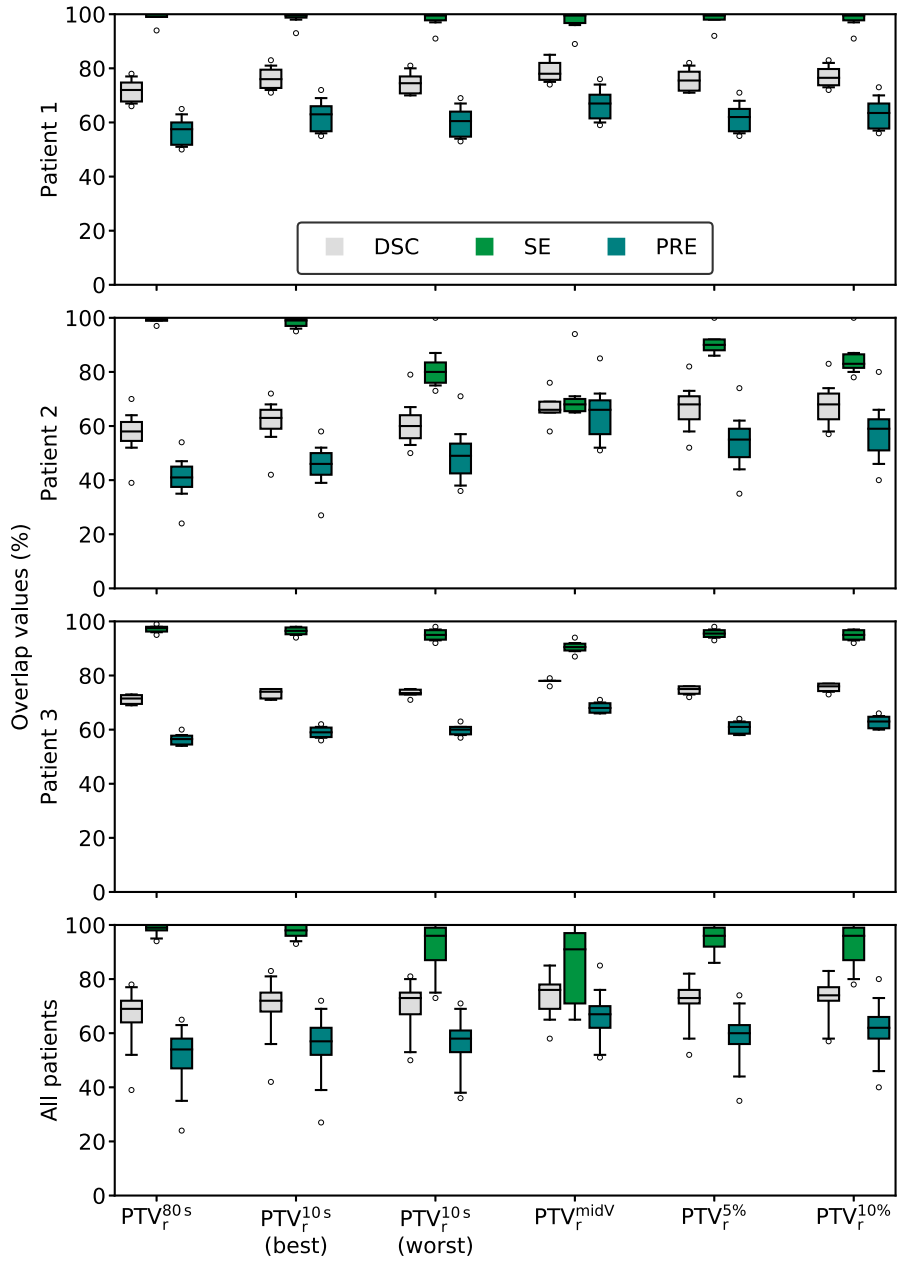

FIG. 5. Geometrical overlap analysis for different passive motion-management concepts. The overlap values are plotted for each of the three patients separately and for all patients accumulated. Each boxplot contains the values for all follow-up imaging sessions. PTV$_r^{10s}$ (best/worst) are defined as the PTV$_r^{10s}$ with the highest/lowest SE value. The whiskers of the boxplots indicate the 5th and 95th percentiles. [Color figure can be viewed at wileyonlinelibrary.com]

TABLE II. Summary of geometrical overlap analysis. The median overlap values over all follow-up imaging sessions are given in percent for the different PTV$_r$ and patients. V_{rel} describes the volume of the different PTV$_r$ relative to the volume of PTV$_r^{80s}$ in percent.

	Patient 1				Patient 2				Patient 3			
	DSC	SE	PRE	V_{rel}	DSC	SE	PRE	V_{rel}	DSC	SE	PRE	V_{rel}
PTV$_r^{80s}$	72	100	57	100	58	99	41	100	72	98	56	100
PTV$_r^{10s}$ (best)	76	100	63	90	63	99	46	89	75	97	60	93
PTV$_r^{10s}$ (worst)	75	100	61	94	60	80	49	67	74	95	60	92
PTV$_r^{midV}$	78	99	67	84	66	68	66	44	78	90	68	77
PTV$_r^{5\%}$	75	100	62	91	68	90	55	69	75	96	61	91
PTV$_r^{10\%}$	77	100	64	88	68	83	59	60	76	95	63	87

DSC, Dice similarity coefficient; SE, sensitivity; PRE, precision.

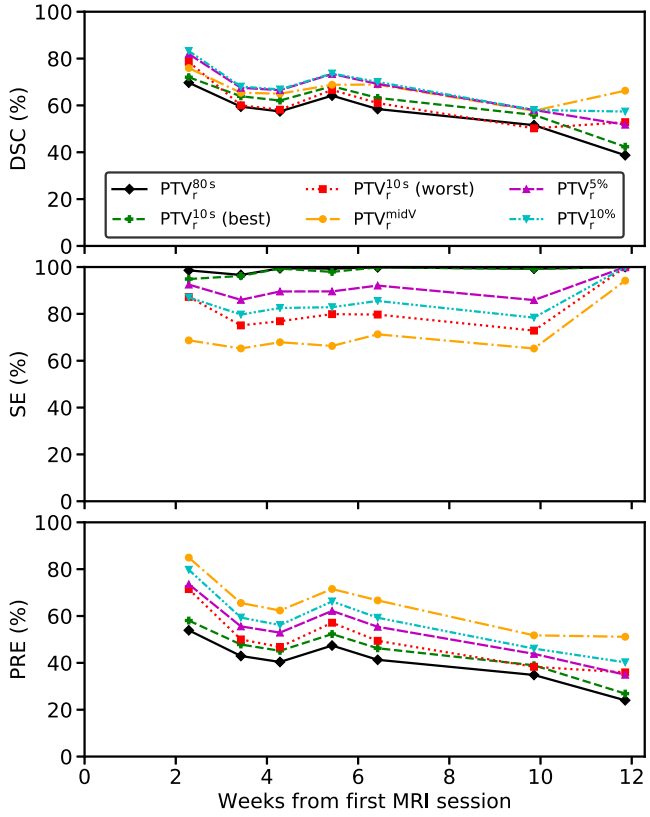

FIG. 6. Geometrical overlap values over time. For Patient 2, the overlap values for the six different PTV$_r$ with ITV$_n^{5\%}$ are plotted for all imaging sessions on days *n*. [Color figure can be viewed at wileyonlinelibrary.com]

Studies by Ehrbar et al.[47] and Thomas et al.[48] concluded in their dosimetric analyses that the mid-ventilation approach can reduce the dose to the lung with no or only slightly reduced tumor coverage with respect to the ITV approach. Both studies however, recalculated the dose on the same 4D-CT images that were used for definition of the target volumes. The results of the present study, however, suggest that the SE could be strongly affected by interfractional changes and the mid-ventilation concept could be less robust than the more conservative ITV$_r^{5\%}$ concept in these cases.

Keeping the volumetric effect in mind, setting a reasonable POP cutoff can lead to an over-proportionate improvement of PRE while the SE decreases less. However, setting this cutoff higher bears the risk of creating underdosed ("cold") spots at the edges of the GTV. Finding an ideal POP cutoff without an extensive dosimetric evaluation is not a trivial task and the optimum is expected to be patient-specific. For the patients included in this study, the use of a POP cutoff of 10% (PTV$_r^{10\%}$) instead of 5% resulted in only a slight improvement of PRE at a reduced SE. As a high SE (surrogate for target coverage) is of higher importance than a high PRE (surrogate for normal tissue sparing), the use of PTV$_r^{10\%}$ instead of PTV$_r^{5\%}$ to improve the PRE would not be justified in these cases.

For all patients, the differences of SE and PRE between the analyzed PTV$_r$ were larger than for the DSC. Due to the different importances of SE and PRE, the interpretation of the overlap results based on the DSC alone was found not to be sufficient to evaluate which PTV$_r$ represented the optimal target volume.

Different gradients of the POP distribution in different directions were observed depending on the motion amplitudes in the respective directions. This information could be used to create anisotropic safety margins around ITV$_r^{5\%}$ in a similar way as done for the margin expansion of the midV. Under the assumption that adequate image guidance techniques are used, this patient-specific margin could potentially increase the PRE while keeping the SE and hence the risk of underdosage of the GTV at a constant level.

A potential extension of the concepts presented in this study could be to directly incorporate the POP information in the treatment planning process. Following the work published by Shusharina et al.,[53] an "internal target distribution" based on the voxel-individual POP values could be defined instead of a binary ITV as input for a probabilistic treatment planning approach.

The results for Patient 2 indicate that most PTV$_r$ were not representative for the GTV motion in the following imaging sessions, resulting in a reduced SE for sessions 2-7. With regard to the use of SE and PRE in adaptive RT,[5,6] if the overlap parameters showed a clear deviation or decrease over time, this would indicate that the patient would benefit from replanning. These findings are supported by the conclusions drawn by St. James et al.[19,24] who demonstrated that a low SE (called "ITV coverage" in their study) could potentially lead to substantial underdosage of the target.

The presented method could be adapted for the use of retrospectively sorted rc-4DMRI instead of rt-4DMRI. rc-4DMRI provides higher spatial resolution than rt-4DMRI, but since only a single breathing cycle is reconstructed in rc-4DMRI, weighting factors for the probability of occurrence of the different breathing phases would have to be determined to reconstruct a 3D POP distribution. Uncertainties associated with the retrospective sorting would have to be accounted for. When 2D images are acquired in rc-4DMRI, instead of 3D images as in rt-4DMRI, less motion information per time is sampled. To achieve the same statistical uncertainty level compared to rt-4DMRI, longer acquisition times would be necessary.

A future workflow for the integration of the proposed 4DMRI-based POP ITV concept into clinical practice could be based on the following steps:

1. Acquisition of a 4DMRI scan instead or in addition to the 4D-CT scan at the planning stage,
2. definition of a 4DMRI-based ITV based on the POP of the GTV by choosing a suitable cutoff probability to overcome current limitations of a 4D-CT based ITV,
3. definition of additional margins around the ITV to prospectively account for potential interfractional changes,
4. use of image-guidance techniques for patient positioning (focused on the target region) and to detect relevant interfractional changes during RT treatment,

5. regular repetition of the 4DMRI scan and subsequent target volume overlap analysis with the parameters SE and PRE over the course of the RT treatment to assess interfractional changes in breathing motion and to verify the applied margin concept,
6. decision whether the original plan is still appropriate or replanning is necessary based on the overlap analysis, ideally in conjunction with a dosimetric evaluation.

This workflow would not necessarily replace the current 4D-CT-based workflow, but could be an enhancement for increased treatment accuracy and quality assurance, especially for patients who experience substantial interfractional changes.

In this proof-of-concept study, the patient cohort was limited to three patients, but compared to other studies that have evaluated lung tumor motion with MRI (cf. Table I in the publication by Thomas et al.[33]), the number of MRI sessions per patient is higher by a factor of 2-4. The acquisition time of the 4DMRI was limited to 80 s due to technical reasons, which is shorter than 2D cine-MR imaging studies reported in literature.[33–35] However, the proposed method could be easily extended to longer acquisition times in the future to improve statistics to assure that the representative target motion of the day is captured.[22] The DIR steps introduce uncertainties that are difficult to quantify but that are expected to be counterbalanced by the higher representativity of the target motion description compared to a 4D-CT-based ITV definition. Potential geometrical image distortions of the MR images have to be accounted for. The proposed workflow is more labor-intensive than today's clinical routine workflow but has the potential to reduce the integral dose to the patient and enhance target coverage. A dosimetric evaluation of the effect of the changing overlap values of the ITVs of different days was beyond the scope of this study. The spatial and temporal resolution of today's rt-4DMRI sequences is limited, including the sequence used in this study. A limited spatial resolution affects the accuracy of the image registration and the target delineation. To account for this, related errors could be estimated and absorbed in the PTV margin calculation.[47] Cai et al.[22] showed that a frame rate of fewer than two images per second could affect the reproducibility of the POP. A limited temporal resolution can furthermore lead to an apparent enlargement of moving structures and an underestimation of the inhale and exhale positions.[42] The former effect would lead to an overly conservative estimation of the GTV POP and therefore decrease the PRE. The latter effect could potentially lead to an underdosage of the GTV edges (lower SE). Since on average ten 3D images were acquired per breathing cycle, the underestimation of the range of motion is estimated to be similar or less pronounced compared to a 4D-CT scan that consists of ten phases. However, this effect has to be considered for patients with high breathing frequencies (>12 cycles/min). Improvements of the spatial and temporal resolution of rt-4DMRI are expected in the near future.[39]

5. CONCLUSIONS

We proposed and investigated a novel concept for ITV definition based on rt-4DMRI and pointed out its potentials and limitations. This ITV definition is based on the POP of the target in 3D, which is expanded by a PTV margin to account for interfractional anatomical and motion changes. In combination with image guidance techniques, the proposed method has the potential to reduce the statistical and systematic uncertainties associated with today's clinical standard-of-care workflow based on ITVs or midV defined on 4D-CT scans. Hereby, an improved target coverage, balanced against the dose to adjacent normal tissues, could be achieved. While this study was focused on the motion of lung tumors, the methods could be translated to other tumor sites which are strongly affected by intrafractional and interfractional motion, such as the liver or pancreas.

ACKNOWLEDGMENTS

This work was supported by the German Research Foundation (DFG) within the Research Training Group GRK 2274.

CONFLICT OF INTEREST

The authors have no conflict to disclose.

*Both authors share last authorship.
[a)]Author to whom correspondence should be addressed. Electronic mail: moritz.rabe@med.uni-muenchen.de.

REFERENCES

1. Li H, Chang JY. Accounting for, mitigating, and choice of margins for moving tumors. *Semin Radiat Oncol*. 2018;28:194–200.
2. Knybel L, Cvek J, Molenda L, Stieberova N, Feltl D. Analysis of lung tumor motion in a large sample: patterns and factors influencing precise delineation of internal target volume. *Int J Radiat Oncol Biol Phys*. 2016;96:751–758.
3. Caillet V, Booth JT, Keall P. IGRT and motion management during lung SBRT delivery. *Phys Med*. 2017;44:113–122.
4. Brandner ED, Chetty IJ, Giaddui TG, Xiao Y, Huq MS. Motion management strategies and technical issues associated with stereotactic body radiotherapy of thoracic and upper abdominal tumors: a review from NRG oncology. *Med Phys*. 2017;44:2595–2612.
5. Menten MJ, Wetscherek A, Fast MF. MRI-guided lung SBRT: present and future developments. *Phys Med*. 2017;44:139–149.
6. Paganelli C, Whelan B, Peroni M, et al. MRI-guidance for motion management in external beam radiotherapy: current status and future challenges. *Phys Med Biol*. 2018;63:22TR03.
7. Yang M, Timmerman R. Stereotactic ablative radiotherapy uncertainties: delineation, setup and motion. *Semin Radiat Oncol*. 2018;28:207–217.
8. Dhont J, Vandemeulebroucke J, Burghelea M, et al. The long- and short-term variability of breathing induced tumor motion in lung and liver over the course of a radiotherapy treatment. *Radiother Oncol*. 2018;126:339–346.
9. Keall PJ, Mageras GS, Balter JM, et al. The management of respiratory motion in radiation oncology report of AAPM Task Group 76. *Med Phys*. 2006;33:3874–3900.
10. De Ruysscher D, Faivre-Finn C, Moeller D, et al. European Organization for Research and Treatment of Cancer (EORTC) recommendations for planning and delivery of high-dose, high precision radiotherapy for lung cancer. *Radiother Oncol*. 2017;124:1–10.

11. Grégoire V, Mackie T, Neve W, et al. ICRU Report 83: prescribing, recording, and reporting photon-beam intensity-modulated radiation therapy (IMRT). *J ICRU*. 2010;10:1–112.
12. Pan T, Lee TY, Rietzel E, Chen GTY. 4D-CT imaging of a volume influenced by respiratory motion on multi-slice CT. *Med Phys*. 2004;31:333–340.
13. Tseng YD, Wootton L, Nyflot M, et al. 4D computed tomography scans for conformal thoracic treatment planning: is a single scan sufficient to capture thoracic tumor motion? *Phys Med Biol*. 2018;63:02NT03.
14. Wolthaus JWH, Sonke JJ, van Herk M, Damen EMF. Reconstruction of a time-averaged midposition CT scan for radiotherapy planning of lung cancer patients using deformable registration. *Med Phys*. 2008;35:3998–4011.
15. Wolthaus JW, Schneider C, Sonke JJ, et al. Mid-ventilation CT scan construction from four-dimensional respiration-correlated CT scans for radiotherapy planning of lung cancer patients. *Int J Radiat Oncol Biol Phys*. 2006;65:1560–1571.
16. Wolthaus JW, Sonke JJ, van Herk M, et al. Comparison of different strategies to use four-dimensional computed tomography in treatment planning for lung cancer patients. *Int J Radiat Oncol Biol Phys*. 2008;70:1229–1238.
17. Mercieca S, Belderbos JS, Jaeger KD, et al. Interobserver variability in the delineation of the primary lung cancer and lymph nodes on different four-dimensional computed tomography reconstructions. *Radiother Oncol*. 2018;126:325–332.
18. Guckenberger M, Wilbert J, Meyer J, Baier K, Richter A, Flentje M. Is a single respiratory correlated 4D-CT study sufficient for evaluation of breathing motion? *Int J Radiat Oncol Biol Phys*. 2007;67:1352–1359.
19. St. James S, Mishra P, Hacker F, Berbeco RI, Lewis JH. Quantifying ITV instabilities arising from 4DCT: a simulation study using patient data. *Phys Med Biol*. 2012;57:L1–L7.
20. Harada K, Katoh N, Suzuki R, et al. Evaluation of the motion of lung tumors during stereotactic body radiation therapy (SBRT) with four-dimensional computed tomography (4DCT) using real-time tumor-tracking radiotherapy system (RTRT). *Phys Med*. 2016;32:305–311.
21. Steiner E, Shieh CC, Caillet V, et al. Both four-dimensional computed tomography and four-dimensional cone beam computed tomography under-predict lung target motion during radiotherapy. *Radiother Oncol*. 2019;135:65–73.
22. Cai J, Read PW, Larner JM, Jones DR, Benedict SH, Sheng K. Reproducibility of interfraction lung motion probability distribution function using dynamic MRI: statistical analysis. *Int J Radiat Oncol Biol Phys*. 2008;72:1228–1235.
23. Seppenwoolde Y, Shirato H, Kitamura K, et al. Precise and real-time measurement of 3D tumor motion in lung due to breathing and heartbeat. *Int J Radiat Oncol Biol Phys*. 2002;53:822–834.
24. St. James S, Seco J, Mishra P, Lewis JH. Simulations using patient data to evaluate systematic errors that may occur in 4D treatment planning: a proof of concept study. *Med Phys*. 2013;40:091706.
25. Green OL, Rankine LJ, Cai B, et al. First clinical implementation of real-time, real anatomy tracking and radiation beam control. *Med Phys*. 2018;45:3728–3740.
26. van Herk M, McWilliam A, Dubec M, Faivre-Finn C, Choudhury A. Magnetic resonance imaging-guided radiation therapy: a short strengths, weaknesses, opportunities, and threats analysis. *Int J Radiat Oncol Biol Phys*. 2018;101:1057–1060.
27. Nabavizadeh N, Elliott DA, Chen Y, et al. Image guided radiation therapy (IGRT) practice patterns and IGRT's impact on workflow and treatment planning: results from a National Survey of American Society for Radiation Oncology members. *Int J Radiat Oncol Biol Phys*. 2016;94:850–857.
28. Sweeney RA, Seubert B, Stark S, et al. Accuracy and inter-observer variability of 3D versus 4D cone-beam CT based image-guidance in SBRT for lung tumors. *Radiat Oncol*. 2012;7:81.
29. Brock KK, Dawson LA. Point: principles of magnetic resonance imaging integration in a computed tomography-based radiotherapy workflow. *Semin Radiat Oncol*. 2014;24:169–174.
30. Koch N, Liu H, Starkschall G, et al. Evaluation of internal lung motion for respiratory-gated radiotherapy using MRI: part I – correlating internal lung motion with skin fiducial motion. *Int J Radiat Oncol Biol Phys*. 2004;60:1459–1472.
31. Blackall JM, Ahmad S, Miquel ME, McClelland JR, Landau DB, Hawkes DJ. MRI-based measurements of respiratory motion variability and assessment of imaging strategies for radiotherapy planning. *Phys Med Biol*. 2006;51:4147–4169.
32. Tryggestad E, Flammang A, Hales R, et al. 4D tumor centroid tracking using orthogonal 2D dynamic MRI: implications for radiotherapy planning. *Med Phys*. 2013;40:091712.
33. Thomas DH, Santhanam A, Kishan AU, et al. Initial clinical observations of intra- and interfractional motion variation in MR-guided lung SBRT. *Br J Radiol*. 2018;91:20170522.
34. Cusumano D, Dhont J, Boldrini L, et al. Predicting tumour motion during the whole radiotherapy treatment: a systematic approach for thoracic and abdominal lesions based on real time MR. *Radiother Oncol*. 2018;129:456–462.
35. van Sörnsen de Koste JR, Palacios MA, Bruynzeel AM, Slotman BJ, Senan S, Lagerwaard FJ. MR-guided gated stereotactic radiation therapy delivery for lung, adrenal, and pancreatic tumors: a geometric analysis. *Int J Radiat Oncol Biol Phys*. 2018;102:858–866.
36. Cai J, Read PW, Baisden JM, Larner JM, Benedict SH, Sheng K. Estimation of error in maximal intensity projection-based internal target volume of lung tumors: a simulation and comparison study using dynamic magnetic resonance imaging. *Int J Radiat Oncol Biol Phys*. 2007;69:895–902.
37. Paganelli C, Lee D, Kipritidis J, et al. Feasibility study on 3D image reconstruction from 2D orthogonal cine-MRI for MRI-guided radiotherapy. *J Med Imaging Radiat Oncol*. 2018;62:389–400.
38. von Siebenthal M, Székely G, Gamper U, Boesiger P, Lomax A, Cattin P. 4D MR imaging of respiratory organ motion and its variability. *Phys Med Biol*. 2007;52:1547–1564.
39. Stemkens B, Paulson ES, Tijssen RHN. Nuts and bolts of 4D-MRI for radiotherapy. *Phys Med Biol*. 2018;63:21TR01.
40. Kauczor H-U, Plathow C. Imaging tumour motion for radiotherapy planning using MRI. *Cancer Imaging*. 2006;6:S140–S144.
41. Dinkel J, Hintze C, Tetzlaff R, et al. 4D-MRI analysis of lung tumor motion in patients with hemidiaphragmatic paralysis. *Radiother Oncol*. 2009;91:449–454.
42. Biederer J, Hintze C, Fabel M, Dinkel J. Magnetic resonance imaging and computed tomography of respiratory mechanics. *J Magn Reson Imaging*. 2010;32:1388–1397.
43. van Herk M, Remeijer P, Rasch C, Lebesque JV. The probability of correct target dosage: dose-population histograms for deriving treatment margins in radiotherapy. *Int J Radiat Oncol Biol Phys*. 2000;47:1121–1135.
44. Laub G, Kroeker R. syngo TWIST for dynamic time-resolved MR angiography. *Magnetom Flash*. 2006;3:92–95.
45. Otsu N. A threshold selection method from gray-level histograms. *IEEE Trans Syst Man Cybern*. 1979;9:62–66.
46. Shackleford JA, Kandasamy N, Sharp GC. On developing B-spline registration algorithms for multi-core processors. *Phys Med Biol*. 2010;55:6329–6351.
47. Ehrbar S, Jöhl A, Tartas A, et al. ITV, mid-ventilation, gating or couch tracking – a comparison of respiratory motion-management techniques based on 4D dose calculations. *Radiother Oncol*. 2017;124:80–88.
48. Thomas SJ, Evans BJ, Harihar L, Chantler HJ, Martin AG, Harden SV. An evaluation of the mid-ventilation method for the planning of stereotactic lung plans. *Radiother Oncol*. 2019;137:110–116.
49. Garibaldi C, Piperno G, Ferrari A, et al. Translational and rotational localization errors in cone-beam CT based image-guided lung stereotactic radiotherapy. *Phys Med*. 2016;32:859–865.
50. Lang S, Shrestha B, Graydon S, et al. Clinical application of flattening filter free beams for extracranial stereotactic radiotherapy. *Radiother Oncol*. 2013;106:255–259.
51. Yu ZH, Lin SH, Balter P, Zhang L, Dong L. A comparison of tumor motion characteristics between early stage and locally advanced stage lung cancers. *Radiother Oncol*. 2012;104:33–38.
52. Schwarz M, Cattaneo GM, Marrazzo L. Geometrical and dosimetrical uncertainties in hypofractionated radiotherapy of the lung: a review. *Phys Med*. 2017;36:126–139.
53. Shusharina N, Craft D, Chen Y-L, Shih H, Bortfeld T. The clinical target distribution: a probabilistic alternative to the clinical target volume. *Phys Med Biol*. 2018;63:155001.