A novel concept to include uncertainties in the evaluation of stereotactic body radiation therapy after 4D dose accumulation using deformable image registration

Juan Diego Azcona[a)]
Service of Radiation Physics and Radiation Protection, Clínica Universidad de Navarra, Avda. Pío XII, 31008 Pamplona, Navarra, Spain

Carlos Huesa-Berral
Service of Radiation Physics and Radiation Protection, Clínica Universidad de Navarra, Avda. Pío XII, 31008 Pamplona, Navarra, Spain
Department of Physics and Applied Mathematics, School of Sciences, Universidad de Navarra. C/ Irunlarrea, 31008 Pamplona, Navarra, Spain

Marta Moreno-Jiménez
Service of Radiation Oncology, Clínica Universidad de Navarra, Avda. Pío XII, 31008 Pamplona, Navarra, Spain

Benigno Barbés
Service of Radiation Physics and Radiation Protection, Clínica Universidad de Navarra, Avda. Pío XII, 31008 Pamplona, Navarra, Spain

José Javier Aristu
Service of Radiation Oncology, Clínica Universidad de Navarra, Avda. Pío XII, 31008 Pamplona, Navarra, Spain

Javier Burguete
Department of Physics and Applied Mathematics, School of Sciences, Universidad de Navarra. C/ Irunlarrea, 31008 Pamplona, Navarra, Spain

Purpose: To use four-dimensional (4D) dose accumulation based on deformable image registration (DIR) to assess dosimetric uncertainty in lung stereotactic body radiation therapy (SBRT) treatment planning. A novel concept, the *Evaluation Target Volume* (ETV), was introduced to achieve this goal.

Methods: The internal target volume (ITV) approach was used for treatment planning for 11 patients receiving lung SBRT. Retrospectively, 4D dose calculation was done in Pinnacle v9.10. Total dose was accumulated in the reference phase using DIR with MIM. DIR was validated using landmarks introduced by an expert radiation oncologist. The 4D and three-dimensional (3D) dose distributions were compared within the gross tumor volume (GTV) and the planning target volume (PTV) using the D_{95} and D_{min} (calculated as $D_{min,0.035cc}$) metrics. For lung involvement, the mean dose and V_{20}, V_{10}, and V_5 were used in the 3D to 4D dose comparison, and D_{max} ($D_{0.1cc}$) was used for all other organs at risk (OAR). The new evaluation target volume (ETV) was calculated by expanding the GTV in the reference phase in order to include geometrical uncertainties of the DIR, interobserver variability in the definition of the tumor, and uncertainties of imaging and delivery systems. D_{95} and $D_{min,0.035cc}$ metrics were then calculated on the basis of the ETV for 4D accumulated dose distributions, and these metrics were compared with those calculated from the PTV for 3D planned dose distributions.

Results: The target registration error (TRE) per spatial component was below 0.5 ± 2.1mm for all our patients. For five patients, dose degradation above 2% (>4% in 2 patients) was found in the PTV after 4D accumulation and attributed to anatomical variations due to breathing. Comparison of D_{95} and $D_{min,0.035cc}$ metrics showed that the ETV (4D accumulated dose) estimated substantially lower coverage than the PTV (3D planning dose): in six out of the 11 cases, and for at least for one of the two metrics, coverage estimated by ETV was at least 5% lower than that estimated by PTV. Furthermore, the ETV approach revealed hot and cold spots within its boundaries.

Conclusions: A workflow for 4D dose accumulation based on DIR has been devised. Dose degradation was attributed to respiratory motion. To overcome limitations in the PTV for the purposes of evaluating DIR-based 4D accumulated dose distributions, a new concept, the ETV, was proposed. This concept appears to facilitate more reliable dose evaluation and a better understanding of dosimetric uncertainties due to motion and deformation. © *2019 American Association of Physicists in Medicine* [https://doi.org/10.1002/mp.13759]

Key words: deformable image registration, dose accumulation, uncertainty

1. INTRODUCTION

Intrafraction lung tumor motion[1] and deformation[2] are a cause of uncertainties in the calculation and delivery of the absorbed dose[3,4] in lung stereotactic body radiation therapy (SBRT). These uncertainties can be partially accounted for by using four-dimensional (4D) computed tomography (CT) treatment planning.[1,5–8] Gross tumor volume (GTV) is depicted on each respiration phase on a 4D CT and projected to the mid-ventilation (reference) phase. The internal target volume (ITV) is established as the volume that encompasses all GTV projections. The planning target volume (PTV) includes the ITV plus a margin to account for patient setup and treatment uncertainties due to the equipment used. Current treatment planning systems perform these dose calculations only in one phase.[6,9] Although this is a practical approach, uncertainties covered by a full 4D dose calculation are not taken into account. Specifically, single-phase calculations do not cover how dose distributions can be effected by breathing, which can change the position and shape of the tumor and the tissue around the tumor.

Another important source of uncertainty in the absorbed dose that affects patient treatment planning lies in the definition of the tumor volume to be treated, which depends on the physician.[2,10] Such uncertainty can be large in lung treatments[11] and arguably represents the most important factor contributing to geometric inaccuracy.[12] However, in lung tumors treated with SBRT, where the GTV is well-defined, this interobserver delineation variability has been estimated to be mainly in the range of 1.2–3 mm.[10,13]

With respect to patient treatment, although uncertainty in breathing amplitude during treatment are small for most patients,[8,14] their treatment may still benefit from real-time correction to prevent possible underdosages.[1,3,15–17] In addition to these intrafractional variations, patient respiratory patterns change during the course of the treatment.[10,18–21] The use of 4D CBCT in clinical practice[10,22] attempts to limit partially the geometrical uncertainties in tumor targeting by minimizing the systematic error in patient positioning. This IGRT technique matches the tumor's trajectory-of-the-day — as determined by 4D CBCT — with the planned trajectory identified through the ITV as defined by means of planning 4D CT. In this way, any interfractional positioning error is minimized.[22,23]

Uncertainties in the relationship between planned and delivered dose also include the interplay effect when using IMRT or VMAT as the delivery technique. These uncertainties seem to be small[24] but, in some cases, may not be negligible.[25,26] Several methods have been recently developed in an attempt to assess this effect.[26,27]

Deformable image registration (DIR) is an image processing technique that is gaining acceptance in treatment planning. The output of DIR is the so-called displacement vector field (DVF), a set of vectors that makes it possible to establish a spatial correlation between voxels in different respiration phases that correspond to the same anatomical area.[2] DIR performed for the various respiration phases on the planning 4D CT scan makes it possible to perform the accumulation of the absorbed dose on a single phase, once a full dose calculation has been done. Mexner et al.[28] calculated accumulated doses after 4D dose calculation for patients with large lung tumor motion, concluding that the effect of motion on the accumulated dose was very small. Valdes et al.[29] compared full 4D dose distributions accumulated on a single phase with projections to other phases of dose distributions calculated in one phase. They concluded that the accuracy of such projected dose distributions "might" be good enough for clinical purposes.

The Task Group 132 of the AAPM has recently published a report[2] which reviews the use of DIR in the clinical environment and provides recommendations for quality assurance (QA). The performance of registration algorithms should be tested before being used in clinical practice.[2,30] Current achievable accuracy of these algorithms for lung tumors can be quantified by the target registration error (TRE). The overall accuracy of good registration in lung nodules per spatial direction should be comparable to the voxel size in that direction.[2,31] The effect of the uncertainties associated with DIR in its application to treatment planning in SBRT is still largely unknown.[32,33] Another relevant study is that of Samavati et al.[32] who looked at the role of DIR in lung SBRT dose accumulation and the impact of geometrical uncertainties in several dose metrics for tumor and lung used in clinical practice. They showed how the combination of the quantitative measurement of DIR uncertainty with patient-specific properties such as tumor volume and planning heterogeneity have an important effect on the dosimetric uncertainty, concluding that a 1.6 mm average reduction in DIR uncertainty may have clinical impact.

The objective of the current paper was to use DIR to assess uncertainties in determination of absorbed dose during the planning process for patients with lung tumors to be treated with SBRT. Geometrical uncertainties in the DIR were evaluated in first place by validating the software with real-patient images. Uncertainty assessment in planning was performed by 4D dose calculation followed by DIR-based dose accumulation in the reference phase. Several dose metrics for tumor and organs at risk (OAR) were evaluated in the reference phase for the planned and the accumulated dose distributions. In this way we evaluated dose coverage in the GTV and PTV, as well as dose in OAR. In view of the observed limitations of the PTV concept to evaluate 4D accumulated dose distributions, we introduced the new concept of the Evaluation Target Volume (ETV), a volume that is defined for dose evaluation after DIR-based 4D dose accumulation. The ETV is calculated on the basis of the GTV for the reference phase, and the outcomes of DIR transformation of each phase with the reference phase, plus the uncertainties included in PTV expansions: delineation, patient setup including image systems, and delivery systems uncertainties. In this way the ETV takes into account the effect of respiratory motion.

2. MATERIALS AND METHODS

2.A. Patient treatment planning and delivery

Eleven patients with lung tumors were treated by SBRT using 6 and 10 MV flattening filter free (FFF) beams. An abdominal compression belt was used in six patients to limit breathing motion. Treatment planning was done in Pinnacle v9.10 with a 4D CT scan. The dose calculation algorithm was the collapsed-cone convolution. The extension of tumor motion was calculated by the position of the GTV centroid on each phase. The ITV was built-up by superposition of all the GTVs defined in each respiratory phase. The expansion of ITV-to-PTV margins was 3mm in anterior–posterior and lateral directions, and 5mm in the superior–inferior direction. In some cases (patients 4, 7, and 8), a trade-off was sought between target coverage and OAR sparing (i.e., rib sparing). In those cases, the PTV was further modified manually by the medical doctor. ITV and PTV were depicted in the reference phase for treatment planning and dose calculation. The mid-ventilation phase was selected as the reference phase. It represents the respiration phase closest to the time-averaged tumor position and it has the potential to reduce margins for treatment of SBRT lung tumors using 4D CT.[6] No density override was done in the ITV-to-PTV expansion. OAR were only contoured for the reference phase.

Patients were treated in 3, 5, or 8 fractions, according to three different SBRT protocols depending on the size and location of the tumor. Several dosimetric indicators (D_{95} and D_{min}) were recorded to assess the tumor coverage (GTV and PTV) and involvement of OAR. $D_{min,0.035cc}$, which is the minimum dose to the tumor without taking into consideration the 0.035 cm^3 volume with minimum dose values, was chosen as a surrogate for D_{min}. For OAR, we evaluated V_{20}, V_{10}, and V_5 metrics for the lung. Mean dose to the lung was also evaluated, as were maximum doses (D_{max}) to other OAR. Similar to D_{min}, D_{max} was expressed in terms of $D_{0.1cc}$, which represents the maximum dose to the organ without taking into consideration 0.1 cm^3 of the organ's volume with maximum dose values. Table I includes clinical and dosimetric data for each patient.

Patients were scanned in a Siemens Somatom Plus CT, which uses "phase binning" to assign each CT slice to its corresponding respiration phase. The slices are assigned to one of the eight different phases. With this method, patient phases are sorted at equal acquisition times. Patients were treated on a Versa (Elekta Medical, Crawley, UK) linear accelerator (linac), after 4D CBCT-based setup. 4D CBCT calculates the time-average tumor position. It minimizes patient setup error by registering all 10 CBCT phases with the mid-ventilation phase of the 4D CT.

2.B. Validation of DIR

We used MIM commercial software (MIM Software Inc., Cleveland, OH) to perform all the deformable registrations between 4D CT respiration phases in this work. MIM uses a proprietary constrained, intensity-based, free-form algorithm that performs a DIR based on the intensity values of two images. MIM enables users to check and refine the initial DIR through the introduction of locked alignments (using the *RegReveal* and *RegRefine* tools). In these cases, MIM runs a second intensity-based deformation using the locks to steer the final deformation results.[34] The selection of the locked alignments is to some extent operator dependent. To ensure homogeneity among cases, all the registrations in this work were done by the same person, a radiation oncologist with expertise in lung SBRT. The effectiveness of the RegReveal and RegRefine tools in the lung has been shown by Johnson et al.[35]

We validated MIM software in lung DIR in two steps. First, we used the "point-validated pixel-based breathing thorax model" (POPI model,[36,37] https://www.creatis.insa-lyon.fr/rio/popi-model). The POPI model is a set of real-patient images in 4D CT with 10 phases of respiration, in which a set of 100 landmarks have been identified in respiration phases by an expert observer. In three patients, all respiration phases are covered; in two patients, just two phases are covered by landmarks. The POPI data are accepted — with justification — as ground truth. The landmarks cover the full extension of both lungs and identify the same anatomical point in the different phases of respiration.

Our second MIM validation step was to validate registration for the images of the patients in our study. We first performed deformable registration between the phases of full expiration and full inspiration for each one of our 11 patients. Two sets of 30 and 65 pairs of landmarks — each pair comprising the same anatomical point in both phases — were identified by a radiation oncologist with expertise in lung SBRT. Landmarks were distributed throughout the whole volume of the two lungs. In the first set, 30 pairs of landmarks (15 on each lung) were placed allowing an approximately equal spacing on each lung on both sets. In the second set, another 35 pairs of landmarks were placed in the lung that contains the tumor and more concentrated in the area where the tumor was, which was the area with higher absorbed dose. In this way, the outcome when evaluating the DIR is weighted toward the area of the tumor, which is most relevant. The goodness of fit of registration was assessed through the target registration error (TRE).[2]

2.C. 4D dose calculation

To retrospectively assess the geometrical and dosimetrical uncertainties in tumor-dose delivery when using the ITV approach, we did a full 4D dose calculation of the treatment plan using Pinnacle v9.10 with collapsed-cone convolution.

In-house computer code was programmed in MATLAB to calculate the dose-volume histograms (DVH) and dosimetric indicators needed to assess the tumor (GTV) coverage (D_{95}) from the dose matrices for each respiratory phase. Metrics for the PTV were calculated on the reference phase. We also used Pinnacle to calculate dosimetric indicators (D_{95}) for the GTV on each respiratory phase (in accordance with the medical doctor's determination of the GTV on each phase) and for

TABLE I. Patients with their tumor localization, motion, volume change, and treatment planning characteristics.

| Pt. # | Loc. | GTV motion amplitude (mm) | | | GTV volume (cm^3) | | | Tech. | Energy (MV) | # fields | # segs | Total dose to PTV (Gy) | # fx | Belt |
		LAT	AP	SI	Ref.	Min.	Max.							
1	RLL	1.0	1.5	4.8	4.64	3.45	5.31	IMRT	6	9	27	52.0	3	N
2	LLL	0.8	2.0	12.5	1.68	1.01	1.68	IMRT	10	7	27	51.3	3	Y
3	RUL	0.6	4.1	4.1	3.12	2.01	3.23	IMRT	10	8	29	52.0	3	Y
4	LUL	0.3	3.7	2.0	1.86	1.63	1.86	3D	10	10	10	47.2	5	N
5	RML	2.3	0.5	1.1	1.75	1.70	1.96	3D	10	8	8	47.6	3	Y
6	RUL	1.6	2.8	2.7	0.53	0.46	0.62	3D	6	8	8	50.8	3	Y
7	LUL	0.0	0.0	0.0	1.89	1.89	1.89	3D	10	9	9	48.1	5	N
8	RLL	2.2	2.1	1.9	1.34	1.10	1.44	3D	10	9	9	48.7	5	N
9	LUL	1.0	1.1	1.3	6.37	5.52	6.50	IMRT	10	8	28	63.3	8	N
10	LUL	1.7	3.7	4.8	1.00	0.76	1.22	IMRT	10	9	19	56.5	5	Y
11	RUL	0.3	0.6	0.5	32.67	29.27	32.67	IMRT	6	8	28	28.9	3	Y

RLL: right lower lobe, LLL: left lower lobe, RUL: right upper lobe, LUL: left upper lobe, RML: right middle lobe.

the PTV on the reference phase. MIM was also used to calculate these metrics (D_{95}) after DICOM exportation from Pinnacle.

2.D. Dose accumulation and DIR transformation of each of the 4D CT phases with the reference phase

A workflow schema was designed in MIM in order to register the phases of 4D CT and accumulate the dose. The schema is represented in Fig. 1. As a starting point, it uses the 4D CT scan, the dose matrices calculated for each of the eight phases, and the set of contours of the reference phase. The workflow steps are as follows: (a) Select the mid-ventilation phase in the 4D CT scan as the reference phase, (b) Calculate the 3D dose distribution on each CT phase, (c) Perform DIR between each phase and the reference CT phase, (d) Check, and, if necessary, tweak DIR using MIM tools (*RegReveal* and *RegRefine*), (e) Transfer all 3D dose matrices from each CT phase to the reference CT phase using the DIR results, (f) Assign weights to the dose matrix of each of the respiration phases, and, (g) Accumulate the total dose received in the reference phase.

The weight for each dose matrix is set to 1/8 because the 4D CT acquisition is "phase-binned." This is an exact approach for 3D dose plans. In the case of IMRT plans, it is an approximation because interplay effects are not taken into account in this approach.

GTV and OAR were determined in the reference CT phase by an expert radiation oncologist. The GTV in each of the remaining CT phases was also similarly demarcated. Furthermore, all GTV and OAR structures were propagated by MIM from the reference phase to all of the other CT phases using the DIR results. DVH and dosimetric indicators for tumor volumes on the reference phase (D_{95}, $D_{min,0.035cc}$) after dose accumulation were calculated with our MATLAB code. MIM was also used to calculate DVH and metrics from the accumulated dose matrices and structures on the reference phase.
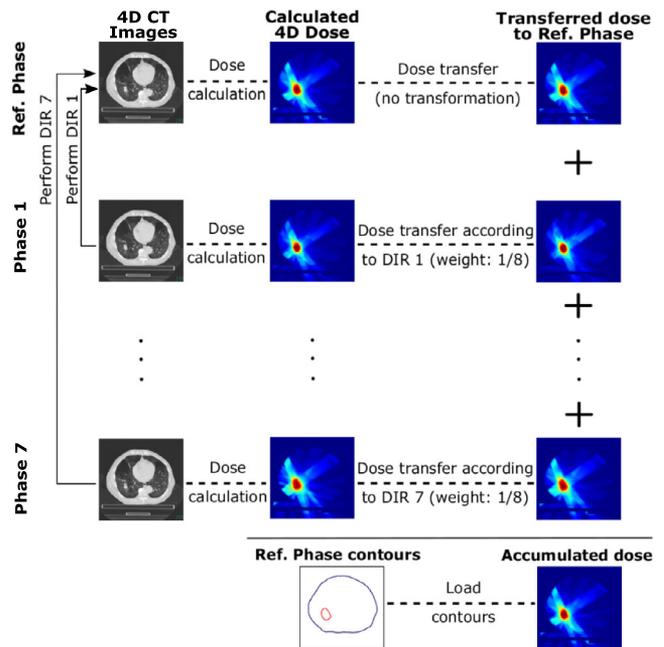
FIG. 1. Workflow for four-dimensional (4D) dose accumulation. Deformable image registrations are done for all respiration phases with respect to the reference phase. By means of registration, dose calculations previously performed on each respiration phase are transferred to the reference phase, and summed up to obtain the 4D accumulated dose. [Color figure can be viewed at wileyonlinelibrary.com]

2.E. Introduction of the Evaluation Target Volume (ETV) concept

The novel concept of ETV is introduced in this section. Its purpose is to surmount the limitations of the PTV when evaluating 4D dose accumulated distributions. The PTV is used for treatment planning in the reference phase to ensure that the GTV receives the prescribed dose despite image and delivery system uncertainties and despite motion effect. Motion effect is taken into account by the ITV.

When DIR is used afterward to evaluate the plan, the dose matrices for the patient's anatomy as defined in all breathing phases are projected to the reference phase. In this way, they can be summed up to calculate the 4D accumulated dose.

The PTV, when derived from the ITV as described above, is appropriate for the treatment planning but is inadequate for our evaluation purposes here because it incorporates the effect of tumor motion and deformation, which are also considered by the DIR used in evaluation. Due to this shortcoming, a new volume (the ETV) is needed. The ETV is used for dose assessment once a DIR-based 4D dose accumulation has been performed. It is defined by expanding the GTV in the reference phase by the appropriate margin to include: (a) geometrical uncertainties of DIR, (b) interobserver variability in the definition of the tumor, and, (c) uncertainties associated with the patient setup and imaging and delivery systems. Uncertainties summarized in points (b) and (c) above are those included in PTV expansions.

Dose metrics calculated on the basis of the PTV in this cohort of patients were also calculated on the basis of the ETV.

3. RESULTS

3.A. DIR validation

For the purposes of assessing DIR, the accuracy of landmark placement done by the expert physician was assumed to be the voxel size in each spatial direction. The TRE was calculated for each of the three spatial coordinates and the 3D vector. The data are displayed in Table II with their standard deviation (SD), for the 5 POPI phantom cases and our 11 patients. Error components are provided after refining the registration with the MIM tools *RegReveal* and *RegRefine*. We used five points to steer the final registration outcome, namely sternum, lung hilum (one on each side), and diaphragm (one on each side). The use of these points improved registration for those cases in which the motion and deformation was relatively large, whereas for cases with less motion and deformation the effect of the refinement was negligible. With respect to the number of landmarks employed to validate the registration, the second set, which has more landmarks, results in a lower TRE. The reduction in 3D TRE is <0.5 mm for both mean and SD in the majority of cases.

3.B. Dose accumulation metrics for the tumor and OARs

We compared the following two sets of the GTV and PTV dose metrics in the reference phase: the retrospectively accumulated metrics and the patient treatment planning metrics. Table III shows the D_{95} and $D_{min,0.035cc}$ for the GTV and PTV, calculated with our MATLAB code. Dose differences in GTV coverage were small, always below or equal to $\pm 1\%$. However, differences in PTV were greater: in one case, the retrospectively accumulated metrics were diminished relative

to the planning metrics by -6.6% for the D_{95} and -18.9% for the $D_{min,0.035cc}$; in two out of 11 cases (cases 2 and 3), the difference in the dose coverage for both metrics was above 4%; in another five cases, at least one of the indicators was decreased by more than 3%.

Table IV shows the comparison of reference phase OAR metrics calculated, by means of MIM, for the distributions of treatment planning dose and retrospectively accumulated dose. Accumulated and reference phase planned doses were, in general, far below OAR tolerances.

3.C. Tumor and lung metrics for individual phases

Results comparing metrics for GTV coverage through D_{95} are shown in Fig. 2. The GTV metrics displayed show (upper box), for patients 1, 5, and 8, the largest changes in GTV coverage in different breathing phases (SD is of the order of 1 Gy for D_{95} in both sets of GTV contours). GTV contours were validated by an expert radiation oncologist. In addition, GTV contours were propagated by MIM from the reference phase contour validated by the expert radiation oncologist. The metric most sensitive to lung movement was V_5 (lower box).

3.D. Assessment of treatment plans on the basis of the ETV

To accommodate the uncertainty of DIR, we expanded the ETV by the voxel size in each spatial direction. This expansion set a lower limit for the amount of uncertainty to be included. We assessed that in the majority of cases, mean TRE values plus one SD were lower than voxel size.

Figure 3 shows the dosimetric indicators D_{95} and $D_{min,0.035cc}$ calculated on the basis of ETV from accumulated dose distributions and compared with those indicators calculated on the basis of PTV from treatment planning dose distributions. In the lower box of Fig. 3, the differences in the metrics (Diff. ETV – PTV) values for D_{95} and $D_{min,0.035cc}$ are displayed. All volumes are in the reference phase. In case 2, for which dose degradation in the PTV was large after 4D accumulation, differences in coverage between the ETV and planned PTV were small. In contrast, differences in cases 1, 3, 8, and 11 were greater than 10% for the $D_{min,0.035cc}$. In these four cases, and in most cases, ETV metrics based on accumulated dose were smaller than PTV metrics for use in treatment planning doses.

Finally, in cases 4, 7, and 8 the medical doctor chose to modify the PTV for treatment planning purposes. In these cases, we estimated the ETV on the basis of the modified volumes.

4. DISCUSSION

The registration accuracy of MIM for the lung was evaluated and the TREs were found to be small. Registration error is greatest in the z direction, where voxel size is largest, but

TABLE II. Target registration error (TRE) expressed as the three-dimensional (3D) mean error and the mean error per spatial coordinate for the landmarks considered, along with TRE SD. Spatial directions are x (lateral), y (anterior–posterior), and z (cranio–caudal). The maximum 3D registration error found is also reported (Max. RE). For the POPI phantom cases 1 to 3, the values displayed correspond to the DIR with the highest TRE among the 9 deformable image registration performed for those cases. POPI validation employed sets of 100 landmarks distributed throughout both lungs. Patient validation consists of two sets: one with 30 pairs of landmarks distributed throughout both lungs (15 on each) and another with 65 pairs of landmarks, 50 of them distributed in the lung that contains the tumor and 15 in the other lung.

	TRE 3D (mm)		TRE x (mm)		TRE y (mm)		TRE z (mm)		Max RE (mm)	Resolution (mm)
POPI										
POPI 1	1.1 ± 0.6		−0.1 ± 0.5		−0.2 ± 0.5		0.4 ± 0.9		4.3	0.98 × 0.98 × 2
POPI 2	1.2 ± 0.9		0.1 ± 0.7		0.0 ± 0.8		0.1 ± 1.1		5.9	0.98 × 0.98 × 2
POPI 3	1.1 ± 2.1		0.0 ± 0.9		0.1 ± 1.0		−0.7 ± 1.8		10.5	0.88 × 0.88 × 2
POPI 4	1.5 ± 3.6		0.2 ± 1.6		−0.1 ± 1.7		−0.9 ± 3.0		20.2	0.78 × 0.78 × 2
POPI 5	0.6 ± 1.1		−0.1 ± 0.7		0.0 ± 0.5		−0.3 ± 0.9		8.9	1.17 × 1.17 × 2
# landmarks	15 + 15	15 + 50	15 + 15	15 + 50	15 + 15	15 + 50	15 + 15	15 + 50	15 + 50	
PATIENTS										
Patient 1	0.8 ± 1.0	0.2 ± 0.2	0.1 ± 0.6	−0.0 ± 0.1	−0.1 ± 0.7	0.0 ± 0.2	0.0 ± 0.9	−0.0 ± 0.1	1.1	0.97 × 0.97 × 2.1
Patient 2	1.1 ± 0.7	0.2 ± 0.3	0.0 ± 0.8	−0.0 ± 0.1	0.1 ± 0.7	0.0 ± 0.1	−0.1 ± 0.8	−0.0 ± 0.3	2.2	0.96 × 0.96 × 2.1
Patient 3	1.3 ± 1.5	1.2 ± 1.6	0.1 ± 0.6	0.1 ± 0.4	−0.3 ± 1.2	−0.2 ± 1.2	0.1 ± 1.5	0.3 ± 1.5	6.1	0.92 × 0.92 × 2.1
Patient 4	1.4 ± 2.2	1.0 ± 1.6	−0.4 ± 1.1	−0.2 ± 0.8	0.0 ± 1.1	−0.1 ± 0.8	0.2 ± 2.1	0.1 ± 1.5	10.0	0.96 × 0.96 × 2.1
Patient 5	0.5 ± 0.7	0.4 ± 0.7	0.0 ± 0.2	0.0 ± 0.2	0.1 ± 0.7	0.1 ± 0.5	0.0 ± 0.5	0.1 ± 0.5	2.9	0.96 × 0.96 × 2.1
Patient 6	0.3 ± 0.6	0.3 ± 0.7	0.0 ± 0.1	0.0 ± 0.2	0.1 ± 0.3	−0.0 ± 0.3	0.1 ± 0.6	−0.0 ± 0.6	4.2	0.98 × 0.98 × 2.1
Patient 7	0.6 ± 1.0	0.5 ± 0.9	0.1 ± 0.7	0.1 ± 0.6	−0.1 ± 0.5	−0.1 ± 0.5	0.1 ± 0.8	0.1 ± 0.6	4.9	0.97 × 0.97 × 2.1
Patient 8	0.3 ± 0.3	0.2 ± 0.3	0.0 ± 0.1	−0.0 ± 0.1	0.0 ± 0.4	−0.0 ± 0.3	−0.1 ± 0.3	−0.0 ± 0.2	1.4	0.98 × 0.98 × 1.5
Patient 9	0.2 ± 0.3	0.1 ± 0.2	−0.1 ± 0.2	−0.0 ± 0.1	0.0 ± 0.2	−0.0 ± 0.1	0.0 ± 0.2	−0.0 ± 0.1	1.2	0.96 × 0.96 × 2.1
Patient 10	0.3 ± 0.6	0.2 ± 0.3	−0.1 ± 0.2	0.0 ± 0.1	0.0 ± 0.2	0.0 ± 0.1	−0.1 ± 0.7	0.1 ± 0.3	1.9	0.96 × 0.96 × 2.1
Patient 11	0.3 ± 0.3	0.2 ± 0.2	0.0 ± 0.2	−0.0 ± 0.1	0.0 ± 0.1	0.0 ± 0.1	0.0 ± 0.3	−0.0 ± 0.2	1.3	0.96 × 0.96 × 2.1

in the field of SBRT, errors are smaller than the mean differences in delimitation of lung tumors by expert observers[10,13] (mainly in the range of 1.2–3 mm).

Dose evaluation after 4D accumulation revealed adequate GTV coverage. These results are in line with the conclusions of Mexner et al.[28]. However, in several patients, we observed differences in PTV coverage when used to assess 4D accumulated when compared to PTV coverage in 3D treatment planning dose matrices in the reference phase (Table III). This indicates the possibility of GTV underdosage due to treatment uncertainties. For those patients in which the PTV underdosage was greatest (Table III, cases 1, 2, 3, and 5), we attributed the degradation to changes in anatomical shape and/or position due to respiration; our results demonstrate the benefit of performing a 4D dose calculation. In addition, for patients 1, 5, and 8, the SD in the D_{95} metrics was higher (in the order of 1 Gy, data are summarized in Fig. 2), a finding that further supports the role of respiration as a main cause of dose degradation.

It is worth noting that $D_{min,0.035cc}$ is very sensitive to contour definition. It was calculated by removing a few voxels with the lowest dose values within the target volume, which are usually located in the tumor's surface. For this reason, we did not include $D_{min,0.035cc}$ values when comparing metrics for tumor volumes within different respiratory phases.

What is needed is a way to identify when discrepancies from PTV coverage are clinically relevant. By taking into account lung movement and deformation, the DIR-based ETV approach evaluated here can help to identify true over- and underdosages due to breathing. The margins of the ETV take into account the uncertainties in the DIR process. To define this margin extension, we used the voxel size. Note that the volume expansion need not be the mean TRE plus just one SD, it could include two or three SDs. In this way, three ETVs can be demarcated, and the dose metrics for each of them can be determined with a probability of 68.3%, 95.5%, and 99.7% of including all DIR geometrical uncertainties (under the assumption that such geometrical uncertainties follow a Gaussian distribution).

To better illustrate the concept of ETV vs PTV for assessing the dose distribution, the 11 cases are shown in Fig. 4 (cases 1 to 10) and Fig. 5 (case 11), where the planned dose and the accumulated dose in the reference phase are compared and the differences for each voxel are expressed in Gy. The figures show several overlapping contours: the ETV (blue) as expanded from the GTV (red) and, by contrast, the ITV (pink) with its related PTV (black). It can be seen that the ETV is, in some cases, smaller than the PTV. It is worth stressing that this fact roots in how the ETV and PTV are built. In 4D planning, the PTV is obtained by expanding the ITV. The proposed ETV concept is defined to evaluate 4D DIR-accumulated dose distributions, which already include the effect of tumor motion. Thus, the ETV is obtained by expansion of the GTV. For case 2, the underdosage in the

TABLE III. Comparison between D_{95} and $D_{min,0.035cc}$, as calculated by accumulation and for the reference phase, for the gross tumor volume (GTV) and for the planning target volume (PTV).

	GTV reference phase dose and accumulated dose						PTV reference phase dose and accumulated dose					
	3D Ref. phase dose (Gy)		4D accumulated dose (Gy)		*Differences (%)*		3D Ref. phase dose (Gy)		4D accumulated dose (Gy)		*Differences (%)*	
	D_{95}	$D_{min,0.035cc}$	D_{95}	$D_{min,0.035cc}$	D_{95}	$D_{min,0.035cc}$	D_{95}	$D_{min,0.035cc}$	D_{95}	$D_{min,0.035cc}$	D_{95}	$D_{min,0.035cc}$
1	55.8	55.1	56.2	55.5	0.7	0.8	52.0	46.8	52.0	43.9	0.0	−6.1
2	55.8	55.1	56.1	55.4	0.5	0.5	51.2	44.4	47.8	36.0	−6.6	−18.9
3	55.2	53.3	55.3	53.4	0.2	0.0	51.9	49.5	49.6	43.8	−4.4	−11.5
4	48.5	47.1	48.3	46.6	−0.4	−1.0	46.4	45.0	45.5	43.3	−1.9	−3.8
5	51.6	51.0	51.3	50.9	−0.6	−0.3	46.3	44.0	44.9	41.9	−3.0	−4.8
6	56.7	57.3	56.9	57.4	0.4	0.1	50.4	48.2	50.8	48.7	0.8	1.0
7	51.5	51.1	51.5	51.0	0.0	−0.2	48.4	46.9	48.4	46.8	0.0	−0.3
8	48.9	48.3	49.1	48.8	0.4	1.0	48.8	47.2	47.9	46.6	−1.8	−1.2
9	65.3	64.6	65.5	64.8	0.3	0.4	63.2	61.0	63.4	60.7	0.3	−0.6
10	62.3	62.2	62.5	62.5	0.3	0.5	56.0	53.7	54.9	51.7	−2.0	−3.8
11	30.5	29.6	30.5	29.6	0.0	0.0	28.8	26.2	28.7	25.4	−0.3	−3.2

TABLE IV. Comparison of organs at risk metrics for accumulated doses and doses calculated in the reference phase. A blank field indicates that the estimated dose to the organ(s) was so low that it was assumed to be zero.

	Lung								Spinal cord		Esophagus		Heart		Ribs	
	Mean dose (Gy)		V_{20} (%)		V_{10} (%)		V_5 (%)		$D_{0.1cc}$ (Gy)		$D_{0.1cc}$ (Gy)		$D_{0.1cc}$ (Gy)		$D_{0.1cc}$ (Gy)	
	Ref.	Accum.	Ref.	Accum.	Ref.	Accum.	Ref.	Accum.	Ref.	Accum.	Ref.	Accum.	Ref.	Accum.	Ref.	Accum.
1	3.1	3.1	4.4	4.3	9.0	8.9	13.5	13.4	13.5	13.6	7.5	7.7	15.8	15.5	27.8	27.3
2	3.8	3.5	4.3	3.8	7.8	7.1	20.7	19.0	14.1	14.1	–	–	8.9	8.6	26.9	26.7
3	3.7	3.6	5.2	5.0	9.8	9.5	17.0	16.7	1.1	1.1	10.8	10.6	–	–	27.1	27.0
4	0.7	0.7	1.0	1.0	1.6	1.5	2.4	2.4	10.6	10.7	11.8	12.1	0.0	0.0	43.4	43.1
5	2.6	2.6	2.6	2.5	6.1	6.1	14.5	14.6	0.3	0.3	–	–	–	–	28.5	28.3
6	1.9	1.9	2.4	2.5	4.7	4.8	9.0	9.4	7.2	7.1	10.9	10.9	0.3	0.3	28.1	28.0
7	1.5	1.5	2.3	2.4	4.3	4.4	6.2	6.3	5.2	5.2	–	–	–	–	46.1	47.3
8	2.4	2.5	2.9	3.0	6.4	6.7	11.7	12.1	2.1	2.1	–	–	3.6	3.6	46.7	47.1
9	1.9	1.9	3.2	3.2	4.5	4.5	6.3	6.3	5.1	5.2	15.3	15.4	0.0	0.2	64.6	64.7
10	4.8	4.9	5.9	6.2	13.8	14.4	21.9	22.5	7.3	7.2	8.6	8.4	11.8	11.7	60.0	59.0
11	3.1	3.1	3.4	3.4	10.7	10.4	18.5	18.1	2.9	2.9	2.3	2.3	0.8	0.8	30.0	30.0

PTV after dose accumulation does not affect the ETV, and so the ETV provides a better way to assess 4D dose accumulation in this case in which there was considerable anatomical deformation. Differences between dose metrics for accumulated doses in the ETV and planned doses in the PTV (reference phase) were small.

After 4D dose accumulation, hot (case 3) and cold (case 8) spots were found inside the GTV. These dose differences support the use of 4D calculations to get a better assessment of the dose to patients. Cases 1 and 6 had similar dose metrics. For case 6, ETV coverage was bigger than PTV coverage. Case 3 had underdosage in the ETV despite having a hot spot in the GTV; the cold spots after dose accumulation affected the lower part of the ETV. In summary, the ETV concept appears to facilitate accurate evaluation of dose distribution.

For clinicians, the ETV might provide a way to adjust margins while reducing radiation of OARs that are close to the GTV.[9] In practice and in most situations, however, differences in estimates of OAR exposure to radiation as a result of 4D dose calculation and accumulation have been found to be small.

It is interesting to note that we used pairwise registration methods in our DIR, which need of the registration of an image to a reference image among the same set. Groupwise registration methods, on the contrary, simultaneously register all images of a 4D set to a common reference frame, thus
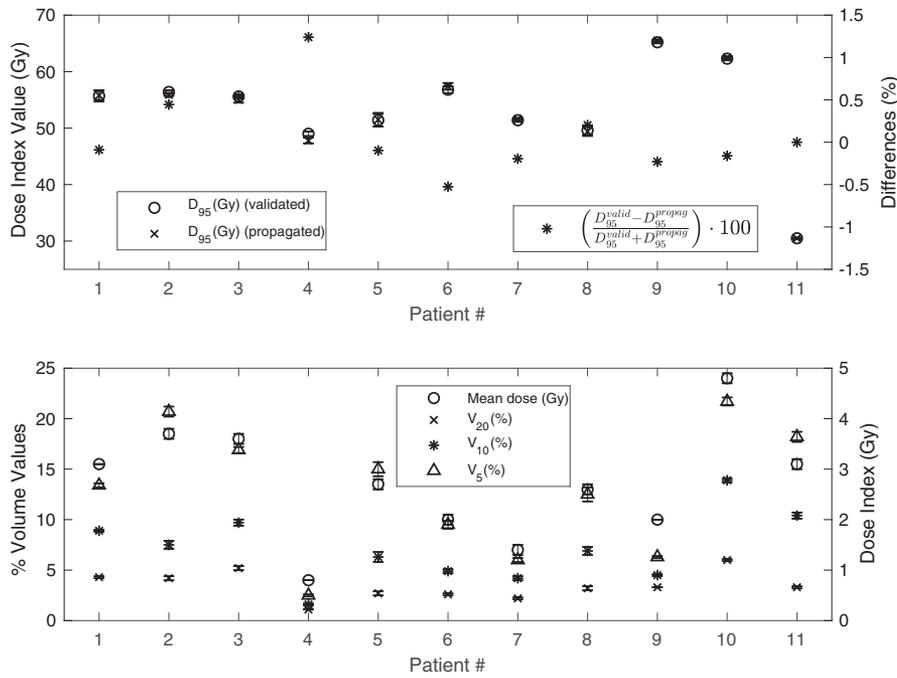
FIG. 2. Gross tumor volume (GTV) coverage dose metrics (D_{95}) calculated from GTVs determined by a radiation oncologist and propagated using deformable image registration (upper box) and lung involvement dose and volume metrics (lower box) for dose calculations on each four-dimensional computed tomography phase. Each point represents the mean of the metrics for all phases; the bar represents the standard deviation.
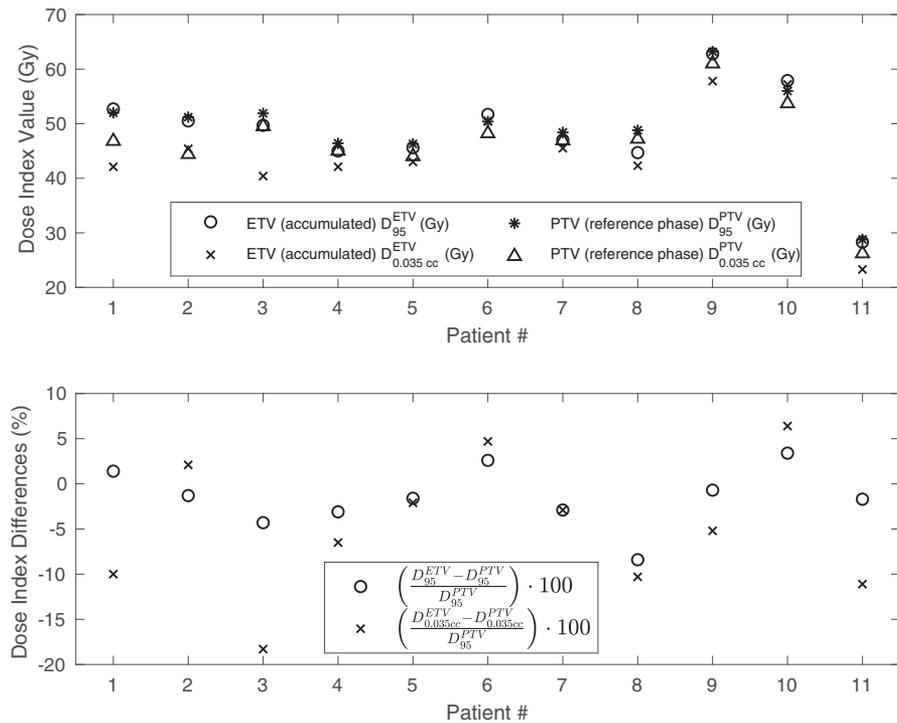
FIG. 3. Comparison between D_{95} and $D_{min,0.035cc}$ estimated from the evaluation target volume over the accumulated four-dimensional doses and the same metrics estimated from the planning target volume over the planned three-dimensional dose, with both volumes as defined in the reference phase. Metric values are plotted in the upper box and differences in the lower box.

minimizing the influence of artifacts on any particular image representing the patient at any specific time on the final outcome.[37,38] Registration is thus expected to be more robust and accurate. In the context of medical physics in radiation oncology, groupwise registration has been used to improve quality in 4D CT and 4D CBCT reconstruction.[37,38] This new method for DIR in the context may potentially benefit the accuracy of 4D dose accumulation.
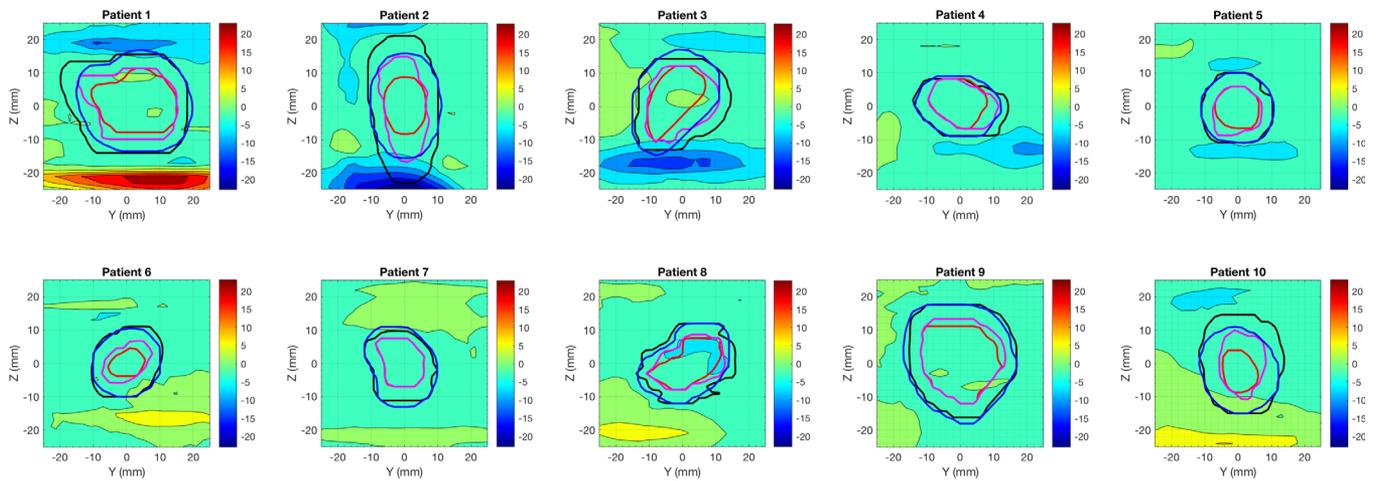
FIG. 4. Comparison of three-dimensional (3D) dose and four-dimensional (4D) accumulated dose distributions in the reference phase, for patients 1 to 10. Different volumes are marked: gross tumor volume (GTV) (thick red), internal target volume (thick pink), planning target volume (thick black) and evaluation target volume (thick blue line). The isocontour of the difference between the planning and accumulated doses is shown in Gy units. Negative values mean that the 4D accumulated dose is lower than the 3D dose calculated in the reference phase. All plots are presented for the sagittal plane that splits the GTV into two equal volumes (the middle plane). [Color figure can be viewed at wileyonlinelibrary.com]
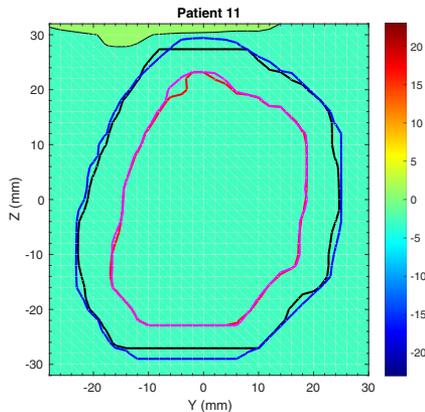
FIG. 5. Comparison of three-dimensional dose and four-dimensional accumulated dose distributions in the reference phase, for patient 11. Colors, units, and contours are similar to those in Fig. 4. The length scale for this case is different. [Color figure can be viewed at wileyonlinelibrary.com]

A final discussion can be established in terms of the energy used for the treatment. Most patients were treated with 10 MV FFF photons. The dose rate achievable approximately doubles that with 6 MV FFF photons. Lung SBRT patients can be treated faster, improving their comfort and potentially reducing their positioning uncertainty throughout their whole treatment irradiation. However, dose calculation based on 3D convolution/superposition algorithms, and in particular using the collapsed-cone convolution implementation, has some uncertainties. These increase with the range of the secondary electrons, that is, with the photon beam energy, and could be manifested in the calculation on several respiration phases. Only three (out of 11) patients were treated with 6 MV photons. However, when looking at Table III, the patients with larger differences in dose metrics between 4D accumulated and 3D dose calculations were all of them treated with 10 MV photons. However, there are other possible causes that could have an influence on these results, such as tumor size, location, and motion. Although not conclusive due to the size

of the cohort, this fact gives a note of caution on the use of larger energies in lung SBRT and its potential impact on 4D dose accumulation, and suggests the use of 6 MV photons to limit dose calculation uncertainties in these cases.

The objective of this study was to assess dose delivery uncertainties on the planned dose distribution by comparing it with the distribution after dose accumulation. Future work might assess, using 4D CBCT, the actual tumor motion and deformation during treatment.[39] Breathing can change from day-to-day[9] and, with respect to planned dose, such changes will induce larger uncertainties in the patient delivered dose. This highlights the ideal for the future of real-time tumor tracking, accumulating the dose based on the actual way a patient breathes during treatment, and, ultimately, adaptive replanning[40] and real-time treatment adaptation.[41]

5. CONCLUSION

A workflow for 4D dose accumulation based on DIR has been devised and applied to a cohort of 11 lung SBRT patients. Dose degradation was found in the PTV after accumulation in several patients and was attributed to respiratory motion. A new concept, ETV, was proposed for evaluating DIR-based 4D accumulated dose distributions. This concept appears to facilitate more reliable dose evaluation, a better understanding of dosimetric uncertainties due to motion and deformation, and is thus of potential value in helping a clinician to adjust margins and treatment.

ACKNOWLEDGMENTS

CONFLICT OF INTEREST

All the authors above report no conflict of interest.

[a] Author to whom correspondence should be addressed. Electronic mail: jazcona@unav.es; Telephone: +34 948 25 54 00 ext. 4913.

REFERENCES

1. Sonke JJ, Belderbos J. Adaptive radiotherapy for lung cancer. *Semin Radiat Oncol*. 2010;20:94–106.
2. Brock KK, Mutic S, McNutt TR, Li H, Kessler M. Use of image registration and fusion algorithms and techniques in radiotherapy: Report of the AAPM Radiation Therapy Committee Task Group No. 132. *Med Phys*. 2017;44:e43–e376.
3. Takao S, Miyamoto N, Matsuura T, et al. *Int J Radiat Oncol Biol Phys*. 2016;94:172–180.
4. Bortfeld T, Jiang SB, Rietzel E. Effects of motion on the total dose distribution. *Semin Radiat Oncol*. 2004;14:41–51.
5. Korreman S. Motion in radiotherapy: photon therapy. *Phys Med Biol*. 2012;57:R161–R191.
6. Wolthaus JWH, 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.
7. Hof H, Rhein B, Haering P, Kopp-Schneider A, Debus J, Herfarth K. 4D-CT-based target volumen definition in stereotactic radiotherapy of lung tumors: comparison with a conventional technique using individual margins. *Radiother Oncol*. 2009;93:419–423.
8. Bissonnette JP, Franks KN, Purdie TG, et al. Quantifying interfraction and intrafraction tumor motion in lung stereotactic body radiotherapy using respiration-correlated cone beam computed tomography. *Int J Radiat Oncol Biol Phys*. 2009;75:688–695.
9. Li H, Chang JY. Accounting for, mitigating, and choice of margins for moving tumors. *Semin Radiat Oncol*. 2018;28:194–200.
10. Yang M, Timmerman R. Stereotactic ablative radiotherapy uncertainties: delineation, setup and motion. *Semin Radiat Oncol*. 2018;28:207–217.
11. Van de Steene J, Linthout L, de Mey J, et al. Definition of gross tumor volumen in lung cancer: inter-observer variability. *Radiother Oncol*. 2002;62:37–49.
12. Weiss E, Hess CF. The impact of gross tumor volumen (GTV) and clinical target volumen (CTV)definition on the total accuracy in radiotherapy. *Strahlenter Onkol*. 2003;179:21–30.
13. Persson GF, Nygaard DE, Hollensen H, et al. Interobserver delineation variation in lung tumour stereotactic body radiotherapy. *Br J Radiol*. 2012;85:e654–e660.
14. Sonke JJ, Rossi M, Wolthaus J, Van Herk M, Damen E, Belderbos J. Frameless stereotactic body radiotherapy for lung cancer using four-dimensional cone beam CT guidance. *Int J Radiat Oncol Biol Phys*. 2009;74:567–574.
15. Li R, Mok E, Chang DT, et al. Intrafraction verification of gated RapidArc by using beam-level kilovoltage x-ray images. *Int J Radiat Oncol Biol Phys*. 2012;83:e709–e715.
16. Shirato H, Suzuki K, Sharp GC, et al. Speed and amplitude of lung tumor motion precisely detected in four-dimensional setup and in real-time tumor-tracking radiotherapy. *Int J Radiat Oncol Biol Phys*. 2006;64:1229–1336.
17. Shirato H, Onimaru R, Ishikawa M, et al. Real-time 4-D radiotherapy for lung cancer. *Cancer Sci*. 2012;103:1–6.
18. Hugo G, Vargas C, Liang J, Kestin L, Wong JW, Yan D. Changes in the respiratory pattern during radiotherapy for cancer in the lung. *Radiother Oncol*. 2006;78:326–331.
19. Redmond KJ, Song DY, Fox JL, Zhou J, Rosenzweig CN, Ford E. Respiratory motion changes of lung tumors over the course of radiation therapy on respiration-correlated four-dimensional computed tomography scans. *Int J Radiat Oncol Biol Phys*. 2009;75:1605–1612.
20. Shah A, Kupelian PA, Waghorn BJ, et al. *Int J Radiat Oncol Biol Phys*. 2013;86:477–483.
21. Sonke JJ, Lebesque J, Van Herk M. Variability of four-dimensional computed tomography patient models. *Int J Radiat Oncol Biol Phys*. 2008;70:590–598.
22. Sonke JJ, Zijp L, Remeijer P, Van Herk M. Respiratory correlated cone beam CT. *Med Phys*. 2005;32:1176–1186.
23. Korreman S, Persson G, Nygaard D, Brink C, Juhler-Notrupp T. Respiration-correlated image guidance is the most important radiotherapy motion management strategy for most lung cancer patients. *Int J Radiat Oncol Biol Phys*. 2012;83:1338–1343.
24. Stambaugh C, Nelms BE, Dilling T, et al. Experimentally studied dynamic dose interplay does not meaningfully affect target dose in VMAT SBRT lung treatments. *Med. Phys*. 2013;40:091710.
25. Li H, Park P, Liu W, et al. Patient-specific quantification of respiratory motion-induced dose uncertainty for step-and-shoot IMRT of lung cancer. *Med Phys*. 2013;40:121712.
26. Edvardsson A, Nordstrom F, Ceberg C, Ceberg S. Motion induced interplay effects for VMAT radiotherapy. *Phys Med Biol*. 2018;63:1–15.
27. Netherton T, Li Y, Nitsch P, et al. Interplay effect on a 6-MV flattening-filter-free linear accelerator with high dose rate and fast multi-leaf collimator motion treating breast and lung phantoms. *Med Phys*. 2018;45:2369–2376.
28. Mexner V, Wolthaus JWH, van Herk M, Damen EMF, Sonke JJ. Effects of respiration-induced density variations on dose distributions in radiotherapy of lung cancer. *Int J Radiat Oncol Biol Phys*. 2009;74:1266–1275.
29. Valdes G, Lee C, Tenn S, et al. The relative accuracy of 4D dose accumulation for lung radiotherapy using rigid dose projection versus dose recalculation on every breathing phase. *Med Phys*. 2017;44:1120–1127.
30. Fatyga M, Dogan N, Weiss E, et al. A voxel-by-voxel comparison of deformable vector fields obtained by three deformable image registration algorithms applied to 4DCT lung studies. *Front Oncol*. 2015;5:1–9.
31. Brock KK. Results of a multi-institution deformable registration accuracy study (MIDRAS). *Int J Radiat Oncol Biol Phys*. 2010;76:583–596.
32. Samavati N, Velec M, Brock KK. Effect of deformable registration on lung SBRT dose accumulation. *Med Phys*. 2016;43:233–240.
33. Chan MKH, Kwong DLW, Ng SCY, et al. Experimental evaluations of the accuracy of 3D and 4D planning in robotic tracking stereotactic body radiotherapy for lung cancers. *Med Phys*. 2013;40:041712.
34. Piper J, Nelson A, Harper J. Deformable image registration in MIM MaestroTM evaluation and description. MIM Software: White paper; 2013.
35. Johnson PB, Padgett KR, Chen KL, et al. Evaluation of the tool "RegRefine" for user-guided deformable image registration. *J Appl Clin Med Phys*. 2016;17:158–170.
36. Vandemeulebroucke J, Sarrut D, Clarysse P. The POPI-model, a point-validated pixel-based breathing thorax model. Proceedings of the XVth ICCR Conference, Toronto, Canada; 2007.
37. Vandemeulebroucke J, Rit S, Kybic J, et al. Spatiotemporal motion estimation for respiratory-correlated imaging of the lungs. *Med Phys*. 2011;38:166–178.
38. Riblett MJ, Christensen GE, Weiss E, Hugo GD. Data-driven respiratory motion compensation for four-dimensional cone-beam computed tomography (4D-CBCT) using groupwise deformable registration. *Med Phys*. 2018;45:4471–4482.
39. Shepherd A, St. James S, Rengan R. The practicality of ICRU and considerations for future ICRU definitions. *Semin Radiat Oncol*. 2018;28:210–206.
40. Dial C, Weiss E, Siebers J, et al. Benefits of adaptive radiation therapy in lung cancer as a function of replanning frequency. *Med Phys*. 2016;43:1787–1794.
41. Colvill E, Booth J, Nill S, et al. A dosimetric comparison or real-time adaptive and non-adaptive radiotherapy: a multi-institutional study encompassing robotic, gimbaled, multileaf collimator and couch tracking. *Radiother Oncol*. 2016;119:159–165.