Dosimetric robustness of lung tumor photon radiotherapy evaluated from multiple event CT imaging

Nils Olofsson a,∗, Kenneth Wikström a,b, Anna Flejmer a,b,c, Anders Ahnesjö a, Alexandru Dasu a,c

a Department of Immunology, Genetics and Pathology, Uppsala University, Uppsala, Sweden
b Uppsala University Hospital, Uppsala, Sweden
c The Skandion Clinic, Uppsala, Sweden

A B S T R A C T

Purpose: Intrafractional respiratory motion is a concern for lung tumor radiotherapy but full evaluation of its impact is hampered by the lack of images representing the true motion. This study presents a novel evaluation using free-breathing images acquired over realistic treatment times to study the dosimetric impact of respiratory motion in photon radiotherapy.

Methods: Cine-CT images of 14 patients with lung cancer acquired during eight minutes of free-breathing at three occasions were used to simulate dose tracking of four different planning methods. These methods aimed to deliver 54 Gy in three fractions to \( D_{95\%} \) of the target and were denoted as robust 4D (RB4), homogeneous fluence to the ITV (FLU) and an isodose prescription to the ITV with a high central dose (ISD), concurrently renormalized (IRN). Differences in dose coverage probability and homogeneity between the methods were quantified. Correlations between underdosage and attributes regarding the tumor and its motion were investigated.

Results: Despite tumor motion amplitudes being larger than in the 4DCT all but FLU achieved the intended CTV \( D_{95\%} \) for the cohort average. For all methods but IRN at least 93% of the patients would have received 95% of the intended dose. No differences in \( D_{95\%} \) were found between RB4 and ISD nor IRN. However, RB4 led to better homogeneity.

Conclusions: Tumor motion in free-breathing not covered by the 4DCT had a small impact on dose. The RB4 is recommended for planning of free-breathing treatments. No factor was found that consistently correlated dose degradation with patient or motion attributes.

1. Introduction

Intrafractional respiratory motion is a concern for radiation therapy of lung tumors. Full evaluation of its dosimetric impact depends however on the availability of images describing the full spectrum of motion states during treatment. Real-time images for direct evaluation of the dosimetric impact of respiratory motion at the time of treatment delivery are not generally available. Respiratory phase-binned computed tomography (4DCT) images, that typically contain 10 pseudo-respiratory phases captured during a few breathing cycles, are commonly used for treatment planning [1]. They allow for better geometric estimation of the tumor motion and also contribute to improving clinical results compared with earlier methods [2].

However, 4DCT images have been shown to underestimate the full extent of the amplitude of the tumor motion during free breathing [3–5]. Additionally, such images do not contain information regarding the relative prevalence of each motion phase. Finally, the interfractional variability that the patient may have between treatment sessions is not captured by a single, standard, 4DCT.

The geometrical aspects of these shortcomings have been studied using snapshot CT acquisitions (cine-CT) [5,6], cone-beam CT [7], magnetic resonance imaging (MRI) [8–10], fluoroscopy [4,11], implanted electromagnetic (EM) marker tracking [12] and also with dedicated respiratory phantoms [13] but with limited analysis of their dosimetric impact.

Treatment planning for lung tumors mitigates the impact of the tumor motion through different strategies [14]. Active strategies [15] imply that the breathing of the patient is controlled or additional technological devices for real-time monitoring and motion compensation are used. For example, surface scanning, imaging using ultrasound, MRI or fluoroscopy or tracking of implanted EM markers are deployed at the time of treatment in order to coach the patient, gate or perform active

Keywords: Advanced lung radiotherapy Robust optimization Breathing motion evaluation

https://doi.org/10.1016/j.ejmp.2022.09.007

Received 2 April 2022; Received in revised form 2 September 2022; Accepted 13 September 2022

Available online 28 September 2022

© 2022 Associazione Italiana di Fisica Medica e Sanitaria. Published by Elsevier Ltd. This is an open access article under the CC BY license (http://creativecommons.org/licenses/by/4.0/).
tracking of the delivery [16]. The use of such methods can improve
the dose to organs at risk (OAR), e.g. by lowering the dose to the
surrounding lung, but implies a more advanced clinical workflow.

The alternative consists of methods in which 4D treatment plans
take the motion into account through additional margins around the
target volume with the delivery given using conventional techniques
[1]. An example of this is by using an internal target volume (ITV) to
encapsulate the motion of the clinical target volume (CTV). The ITV can be
delineated from a time average CT, from a 4DCT using the maximum
intensity projection, average intensity projection or constructed as the
union of CTVs defined per reconstructed phase. The ITV concept is also
used in cardiac radioablation where the optimal choice of cardiac
motion imaging in order to create the ITV has been investigated [17].

Another planning strategy is robust 4D planning where each phase in
the 4DCT is treated as a possible treatment scenario with all scenarios
taken into account by the optimization method [18,19].

Particle therapy has become an option for the treatment of lung
tumors and the dosimetric consequences of respiratory motion are
expected to be even larger than for photon therapy. These have been
studied in this context using computer simulations [20] and phan-
toms [21,22] with the clinical implications being reviewed as well [23].

The incorporation of scenarios for uncertainties has been investigated
by sampling values based on isoprobability hypersurfaces [24].

Published dosimetric assessments of lung cancer photon treatments
based on simulations and/or planning studies typically only use the
4DCT image data to simulate respiratory motion [25–27]. This is a
limitation, since the 4DCT may lack motion data, both in terms of
amplitude and frequency, and hence the calculated dose distribution
can be different from the actually delivered one.

Compared to studies based only on 4DCT, the present investigation
uses a more elaborate approach where tumor motion data gathered
during imaging sessions with realistic lengths for treatment delivery
were used to create a large set of evaluation images. These were used
to simulate dose tracking and to perform an extensive evaluation of the
dosimetric robustness of target coverage for respiratory motion of lung
tumor photon radiotherapy.

The aims of this study were threefold.

1. Quantify the dosimetric effects stemming from tumor motion
during free breathing.
2. Compare different stereotactic body radiation therapy (SBRT) 4D
planning methods with regards to the potential loss of dose to
target.
3. Investigate patient and motion analysis attributes as a predictor
for respiratory-induced loss of dose to target.

2. Materials and methods

2.1. Image data and structure delineations

Image data from 14 patients with non-small cell lung cancer eligible
for radiation therapy were used in this study. Details of the patient
data are given in a previous publication by Wikström et al. [5] and
are summarized in Table 1. For each patient a breath-hold CT (BHCT),
a 4DCT with ten respiratory phases and 300 cine-CT images had been
acquired using a Philips Brilliance® CT Big Bore 16 slice. The BHCT
and 4DCT acquisition volumes, in plane resolution and slice thickness
varied between patients.

The 300 cine-CT images were acquired evenly distributed over three
different occasions, as to emulate a treatment of three fractions. As
such, 100 cine-CT images were acquired per imaging session, each such
session lasting approximately eight minutes. The imaging sessions were
on average seven days and at the most 21 days apart. Potential base-
line shift was accounted for during the image registration process as
described in Section 2.3. The patient positioning was the same as would
have been during treatment. The first three patients were coached
during the second halves of each imaging session. The acquisitions
were triggered at regular intervals with a random offset [5]. Cine-CT
images were all acquired as sixteen slices, with a thickness of 1.5 mm
for a total volume thickness of 24 mm, and an in plane resolution of
approximately 0.5 mm × 0.5 mm per voxel. The limited cine-CT volume
thickness is due to the coverage of the detector arrays in the CT scanner
which determines the geometry that can be captured during a single
rotation. Additional total dose from cine-CT was estimated as to not
exceed 27 mGy for any patient, which is approximately equal to the
dose received during clinical 4DCT examinations.

The use of patient images for simulation studies was approved by
the Swedish Ethical Review Authority (Etikprövningsmyndigheten) under
the number 2020-00816.

The CTV was delineated on the BHCT by an experienced physician,
then mapped onto the phase images of the 4DCT using deformable
image registration in the treatment planning system (TPS), RayStation
10b (RaySearch Laboratories AB, Stockholm, Sweden). The relevant
organs at risk were delineated by another experienced operator using
assistive tools in the TPS. Before acceptance all delineations mapped
between image sets were visually inspected and, if necessary, manually
adjusted.

An ITV was constructed as the union of the CTV in all phases of the
4DCT.

For each patient, BHCT and 4DCT were used for treatment planning
with additional cine-CT image data being used to create the images for
the evaluation, as depicted by the schematic workflow in Fig. 1. Four
treatment plans were created per patient and motion information from
cine-CT, captured during free breathing sessions, were used to simulate
and evaluate the plans.

2.2. Treatment planning

Four methods of 4D planning, using prescription methods based on
recent trends for hypo-fractionated SBRT treatments, were compared
using the proposed extensive dose evaluation [28,29]. The four plan-
ing methods were robust 4D (RB4), a conformal method based on
homogeneous fluence to the ITV (FLU), one based on a central higher
dose and an ITV isodose (ISD) and a renormalization of the latter to a
dose to the CTV (IRN).

In this dosimetric context we differentiate between intended and
prescribed dose where the prescribed dose is a means of achieving the
intended dose to the gross tumor or the CTV. The prescription dose in
SBRT is often defined as a dose that should cover a larger irradiated
volume, such as a planning target volume (PTV), which implicitly will
cause the intended dose to be delivered to the tumor itself.
This, among other factors such as dose computation algorithm accuracy, has contributed to a demonstrable large inter-institutional variation in dose levels for the treatment of lung tumors [28,29]. Prescribing a median or median dose to the CTV has emerged as a way of adapting prescriptions derived using poor accuracy dose computation algorithms with known clinical outcome [30]. Such adapted prescriptions have shown good clinical results [31–33].

All plans in the present study had the same intention of delivering 18 Gy per fraction for a total dose of 54 Gy to 50% of the CTV, i.e. $D_{50\%} = 3 \times 18$ Gy. The prescription dose and target volume used to attempt to achieve $D_{50\%} = 54$ Gy for the CTV depended on the planning method and are further described in Sections 2.2.1 to 2.2.4. In RB4 and FLU the prescription and the intended dose were the same while ISD and IRN the prescription and the intended dose were different.

In this work the respiratory motion has been selected as the primary subject of study and zero setup errors were assumed both during planning and evaluation. Therefore no additional margins or scenarios for patient shifts were used to accommodate the dosimetric impact of setup errors. The 30% phase in the 4DCT was considered an approximation of the mid-ventilation and was used for planning in FLU, ISD and IRN. Furthermore, it was used for computing the nominal plan dose used to assess and compare the dose to the ipsilateral lung for all four planning methods.

Dose limits [34,35] for organs at risk were taken into account for all plans and could to a large degree be accounted for by choice of gantry angles. The dose constraints per fraction to OARs were $D_{\text{max}} \leq 10$ Gy for aorta, $D_{\text{max}} \leq 8.4$ Gy and $5$ cm$^3$ at 5.9 Gy for esophagus, $D_{\text{max}} \leq 10$ Gy and $15$ cm$^3$ at 8 Gy for the heart, $D_{\text{max}} \leq 7.3$ Gy and 10% of total volume at 6 Gy for the spinal cord and $D_{\text{max}} \leq 11$ Gy and $10$ cm$^3$ at 10 Gy for the skin.

Depending on the patient geometry, five or six coplanar 6 MV beams with static gantry angles were used in each plan. Gantry angles were the same for all plans for the same patient. All dose calculations were performed using collapsed cone convolution (RayStation 10B, collapsed cone v5.4) with dose grid resolution of $1.5$ mm $\times 1.5$ mm $\times 1.5$ mm. The methods RB4 and ISD used inverse planning, IRN only implicitly as it was a renormalization of the ISD plans. No information from the cine-CT was taken into account during treatment planning, as they were used only for evaluation.

### 2.2.4. ITV isodose renormalized (IRN)

The IRN planning method was the same as ISD but renormalized to the intended CTV dose, $D_{\text{SRT}} = 18$ Gy per fraction. This for the purpose of reducing the variability in the results between patients [30,38].
2.3. Creation of evaluation images through image registration

The cine-CT images covered too little of the patients' anatomies along the craniocaudal direction to be directly useful for dose computations. Tumor position information from the cine-CT images were used in preparatory steps to compute the evaluation doses. For this purpose rigid image registration was used to find the position of the tumor in each cine-CT in terms of CTV centroid and evaluation images with these tumor positions were created based on the 4DCT.

First the tumor with its nearby tissues was cropped from the BHCT and registered to each cine-CT image, \( i \), with a rotation and translation \( T_i^c \). Second, a frame of reference registration was performed between each cine-CT image and the reference image, \( i.e. \) the 4DCT phase labeled 30%, using a volume of interest around the spine and the back of the patient in the 4DCT image, resulting in a second rotation and translation \( T_i^b \).

The registrations were done with six degrees of freedom using normalized cross correlation as similarity metric and a simplex optimizer. This conservative approach was chosen to remove shifts due to patient setup differences between imaging sessions still present after the imaging fraction, \( f \), was set to the mean tumor position in the 4DCT.

The mean of all tumor positions in the cine-CT images for a patient per each cine-CT in terms of CTV centroid and evaluation images with these tumor positions were created based on the 4DCT.

The complete process, shown in Fig. 2, was performed for each of the 300 cine-CT images for all 14 patients.

2.4. Dose evaluation

2.4.1. Evaluation dose to target

For each patient all plans were recomputed for each evaluation image, as seen in Fig. 2, resulting in \( 14 \times 4 \times 300 = 16800 \) additional dose computations. The recomputation of the plans was facilitated by the scripting interface in the TPS and the dose summation and analysis were performed using custom software.

The accumulated dose to the CTV was computed for each patient and planning method from the sum of all evaluation doses to the transformed CTV for that particular case, divided by the number of evaluation doses (\( i.e. \) 300).

Dose volume histograms (DVH) for each accumulated evaluation dose to the CTV were created and dose-to-volume metrics \( D_{95\%}, D_{98\%}, D_{99\%}, D_{p90\%}, D_{p95\%}, D_{p99\%} \) and \( D_{2\%} \) were calculated.

Inhomogeneous dose distribution in the target is often a necessity for radiation therapy of lung tumors and is largely allowed and occasionally even desirable during SBRT in general. However, for this study we considered a good plan to have a homogeneous dose in the CTV, although it was not an explicit goal for any of the planning methods. The homogeneity index (HI) was calculated according to the ICRU recommendations [39].

\[
HI = \frac{D_{2\%} - D_{98\%}}{D_{95\%}} \tag{5}
\]

The metrics \( D_{95\%} \) and HI were used when comparing the planning methods according to their evaluation doses. In order to identify patients with worse dose coverage, relative underdosage of a patient, \( p \), compared with the other patients in the cohort was defined as the binary outcome where \( D_{p95\%}^{\text{new}} \) was lower than the average \( D_{95\%} \) across all patients, \( D_{95\%}^{\text{avg}} \) for that planning method, \( m \).

\[
\text{Underdosage}(p, m) = \begin{cases} 1 & \text{if } D_{p95\%}^{\text{new}} < D_{95\%}^{\text{avg}} \\ 0 & \text{else} \end{cases} \tag{6}
\]

This differs from the definition of underdosage used in clinical settings where dose relative to the prescription to the target volume is used [40].

2.4.2. Nominal dose to ipsilateral lung

Dose to the ipsilateral lung was assessed using the planned nominal dose computed on the 4DCT 30% phase image. The absolute volumes receiving at least 20 Gy and 5 Gy (\( V_{20\text{ Gy}} \) and \( V_{5\text{ Gy}} \)) were used when comparing the planning methods. Both \( V_{20\text{ Gy}} \) and \( V_{5\text{ Gy}} \) have been suggested as indicators of radiation induced fatal pneumonitis [41].

2.4.3. Dose volume coverage and avoidance maps

The DVH results were summarized for the target using dose volume coverage maps (DVCM) [42,43] which gave the probability to achieve a certain dose level to a certain fraction of the volume. Similarly the DVHs were summarized for the ipsilateral lung using dose volume avoidance maps (DVAM) which gave the probability to avoid a certain dose level to a certain volume.

Percentile dose volume histograms (PDVH) were defined as isocurves where the dose coverage probability is constant such that for a DVAM it represents a curve that lies at least at some value for \( N\% \) of the population. Analogously it is defined for a DVAM, using Eq. (4), such that the curve represents an upper bound. PDVHs were computed at the 90th and 50th percentile (PDVH_{90\%} and PDVH_{50\%}).
Fig. 2. Schematic illustration of the registration process to create an evaluation image. Presented for a sagittal section with a tumor shown relative to lung and part of the liver. (a) The tumor in the BHCT is registered and transformed, \( T_i^a \), to the cine-CT. (b) The cine-CT is registered and transformed, \( T_i^b \), to the 30% phase of the 4DCT. (c) A correctional translation, \( T_f^c \), is applied. (d) A matching phase image in the 4DCT is selected based on closest tumor position. (e) The tumor along with some additional lung volume is copied into the best matching 4DCT phase image. (f) Dose is recomputed on this new image geometry.

Fig. 3. (a) Volumetric rendering of an evaluation image example with imposed ITV. The tumor has been shifted according to its position in one of the cine-CT images and is no longer entirely covered by the ITV derived from the 4DCT. Part of the body and the ITV has been cropped from right to left to avoid occlusion. Results of the motion analyses are shown on the right where the maximum tumor displacement from the 30% phase in the 4DCT is shown in (b) as a dashed gray line, the displacement for all evaluation images are shown in blue and the illustrated example number has been highlighted by orange. (c) Analogously the CTV coverage by the ITV is shown in green. Both (b) and (c) have a percentile indicated by a dotted line, which for the displacement implies that 90% and for the coverage implies that 10% of all values are below the line.

2.5. Statistics

The population average outcome for each planning method was compared to the intended dose level at \( D_{50\%} \) using the t-test.

The four planning methods were compared with each other by examining dose metrics, \( D_{50\%} \) and HI, for the target and by examining volume at dose, \( V_{20\ Gy} \) and \( V_5\ Gy \), for the ipsilateral lung. The comparisons of the planning methods were performed using the Wilcoxon signed ranked test. Significance for these tests are shown in Figs. 5 and 7 and are indicated at three levels of p-values, \( \star \star \star \) (p<0.001), \( \star \star \) (p<0.01) and \( \star \) (p<0.05). The distributions of the compared dose metrics were illustrated using box plots where the boxes show the first and the third quartiles with the median indicated inside and whiskers offset by one and a half of the inter-quartile range.

Correlation between relative underdosage and attributes regarding the tumor and its motion was investigated using Spearman rank correlation.

Tumor displacement in the 4DCT was compared with the 90th percentile of the cine-CT derived images using the Wilcoxon signed ranked test.

3. Results

3.1. Patient attributes and motion analysis

Information about patient sex, ipsilateral lung and size of the CTV and ITV are summarized in Table 1 along with the displacement and coverage results.

The free-breathing momentary displacements, \( d_{\text{cine-CT}} \), and their 90th percentile, \( P_{\leq 90\%}(d_{\text{cine-CT}}) \), and CTV coverage by ITV and their 10th percentile, \( P_{\leq 10\%}(d_{\text{cine-CT}}) \), can be seen for patient 2 in Fig. 3 and for all patients in the supplementary material.

Amplitudes derived from the cine-CT images, \( A_{\text{cine-CT}} \), were larger than those observed in the 4DCT, \( A_{\text{4DCT}} \), for all patients. In the 4DCTs the amplitudes varied between 1.4 mm to 16.8 mm across patients while in the cine-CT derived images it varied between 7.4 mm and 41.6 mm.

The 90th percentile of momentary tumor displacements observed in the cine-CT derived evaluation images, \( P_{\leq 90\%}(d_{\text{cine-CT}}) \), was greater than the maximum tumor displacement observed in the 4DCT, \( \max(d_{\text{4DCT}}) \), in some patients while for others the opposite was true. In the 4DCTs the maximum displacement varied between 1.1 mm to 11.6 mm across patients.
patients while in the cine-CT derived images the 90th percentile of the displacement varied between 2.5 mm and 13.4 mm. However, no overall difference was observed ($p > 0.05$).

The CTV coverage at the 10th percentile, $P_{\leq 10\%}(CTV \cap ITV)$, varied between 99.1% and 76.9% across patients.

### 3.2. Evaluation dose to target

The accumulated evaluation doses to the CTV for the simulated treatments are summarized in Fig. 4 which shows how likely each planning method was to achieve a certain dose to a certain percentage of the volume.

In the simulated treatments the intended dose at $D_{50\%}$ was likely to be reached in about half of the cases, by inspection of PDVH$_{50\%}$, for RB4, ISD and IRN while FLU had a lower probability. Indeed, when comparing the population average outcomes at $D_{50\%}$ only FLU was lower than the intended dose level ($p < 0.05$). Average $D_{50\%}$ for FLU was 52.7 Gy compared with 53.7 Gy for RB4, 54.1 Gy for ISD and 53.3 Gy for IRN.

Furthermore, FLU had probabilities of much larger fluctuations in dose coverage at both high and low relative volumes.

The planning methods were compared at $D_{50\%}$ and HI using the Wilcoxon signed rank test, see Fig. 5.

No differences in $D_{50\%}$ were found between RB4 to ISD nor IRN, nor between FLU and IRN. The RB4 method was closer to the intended dose than FLU ($p < 0.05$). ISD was closer to the intended dose than FLU ($p < 0.05$) and had higher dose than IRN ($p < 0.05$).

The RB4 method had better homogeneity than the other three methods ($p < 0.001$).

Tabulated results including additional dose metrics can be found in Table 2 and in the supplementary material.

Correlation between relative underdosage, as defined in Eq. (6), and CTV size and CTV coverage by the ITV was investigated. CTV coverage had a correlation to relative underdosage for FLU ($r = -0.54$, $p < 0.05$) and IRN ($r = -0.67$, $p < 0.01$) where $r$ is the correlation coefficient.
3.3. Nominal dose to ipsilateral lung

Nominal dose to the ipsilateral lung was computed for each planning method to the reference phase, 30%, in the 4DCT and DVAMs were created. These illustrate how likely each planning method was to avoid a certain dose to an absolute volume of the lung, as illustrated in Fig. 6.

When comparing \( V_{5\text{ Gy}} \) FLU had lower volume than the other three \( (p < 0.001) \) and IRN had lower volume than ISD \( (p < 0.05) \). Average \( V_{5\text{ Gy}} \) for FLU was 297 cm\(^3\) compared with 391 cm\(^3\) for RB4, 387 cm\(^3\) for ISD and 384 cm\(^3\) for IRN. For \( V_{20\text{ Gy}} \) FLU had lower volume than the other three \( (p < 0.001) \) and IRN had lower volume than RB4 and ISD \( (p < 0.05) \). The distributions for \( V_{5\text{ Gy}} \) and \( V_{20\text{ Gy}} \) can be seen in Fig. 7. Average \( V_{20\text{ Gy}} \) for FLU was 51 cm\(^3\) compared with 86 cm\(^3\) for RB4, 74 cm\(^3\) for ISD and 71 cm\(^3\) for IRN.

4. Discussion

The impact of irregular intrafractional motion has previously been investigated geometrically and the found differences have been postulated to lead to dosimetric consequences \[3–9,11\]. Accurately quantifying the difference between the planned dose and what is more realistically delivered to the patient is warranted, both in order to identify potential candidates for additional motion management besides 4D planning and also to find relevant correlation between estimated delivered dose to the target and the clinical outcome.

The present study used snapshot cine-CT images to evaluate the effects of the underestimation of motion in 4DCT, compared to real positions during free breathing, on the actual physical dose delivered to the tumor. Four treatment planning strategies were investigated.

The source of the underestimation of the tumor motion amplitude in 4DCT compared to the cine-CT is two-fold. First, the 4DCT represents an arbitrary respiration at the 4DCT imaging session and temporary, more extreme, tumor positions may not be captured in the 4DCT reconstruction. Second, the respiratory pattern of patients may vary between different imaging and treatment sessions. However, the motion estimation differences were only derived from the respective images, and not confirmed using e.g. a phantom.

When comparing the population averages at \( D_{50\%} \) to the CTV only FLU differed from the intended dose level. The RB4, ISD and IRN had higher probability of a \( D_{50\%} \) dose closer to the intended dose than FLU. Further, RB4 had a higher probability of more homogeneous dose than the other three. The size of this difference was especially distinct when compared to FLU. These advantages in favor of RB4 regarding dose coverage to the target came at the cost of slightly higher dose to the ipsilateral lung compared to FLU and IRN. Overall dose to ipsilateral lung was lowest for FLU compared to the other planning methods but at the cost of worse dose to target.

Even though the tumor motion amplitudes in the evaluation images, seen in Table 1, were larger than in the 4DCTs used for treatment planning, the dose was not lower than 95% of the intended dose for more than one patient for RB4, FLU and ISD and for two patients for IRN. This corresponds to 95% of the intended dose to at least 93% and 86%, respectively, of the patients which can be compared to historical population-based dosimetric treatment goals, e.g. 95% of the intended dose to 90% of the patients \[44\].
Overall the dosimetric effect of the motion was small which implies that all four planning methods compensate, more or less, well for the respiratory motion during free breathing.

We only found a correlation between relative underdosage and one of the relevant investigated parameters, CTV coverage, for two of the planning methods, FLU and IRN. It seems reasonable that how well the CTV stays inside the ITV should impact the dose such that there should be a correlation, however we found that this presumed effect was, at most, small.

The consistently poorer results across all four methods for a certain patient, number 7, could possibly be explained by a lower tissue density in the CTV cropped from the BHCT compared with the CTVs in the 4DCT, see the supplementary material for CTV CT value histogram comparisons.

The overall worse $D_{50\%}$ to the CTV for FLU compared to the dose in the other three methods can be explained by the smaller field openings created by the MLCs. The other three planning methods used the optimizer to shape the openings which resulted in leaf edges further from the ITV in BEV. When comparing ISD with IRN the renormalization resulted in lower doses for all but three patients where the dose was slightly increased. As such the evaluated dosimetric results were different when compared using the non-parametric statistical test, see ISD and IRN in Fig. 7, even though the differences were very small.

Furthermore, when comparing ISD to IRN the renormalization indeed harmonized the results at $D_{50\%}$ as proposed by de Jong et al. [28]. The spread decreased which can be seen as the reduced $D_{50\%}$ range and inter quartile ranges when comparing ISD with IRN in Fig. 5. However we note that a scaling of the dose in this manner will not change the HI-values.

The dose degradation effect can be compared to other sources of uncertainty such as interplay effects during treatment delivery and discrepancies between different dose computation algorithms.

The dynamic properties of the beam delivery were not taken into account in our study. The independent motion of the tumor and the rotation of the gantry, as it moves between the selected treatment angles, and the motion of wedges and MLC leaves give rise to a potential interplay effect. This interplay effect has been studied using lung phantoms with tumor insertions moving according to 4DCT motion amplitude [45] directly or by perturbing a computed dose [46]. It is also studied using dose computed on a 4DCT [47,48]. Potential interplay effects between breathing motion and dynamic dose delivery techniques for volumetric arc therapy (VMAT) and intensity modulated radiotherapy (IMRT) are typically regarded as small, <0.5%, and without clinical relevance [46].

The present study used dose calculations with collapsed cone convolution [49] for all dose computations and therefore the calculated doses are considered representative of those employed in current clinical practice in lung radiotherapy.

The dosimetric effect of motion during free breathing was smaller than the effect seen when comparing type A to type B dose computation algorithms in lung, as reviewed by Lacornerie et al. [30]. However it was larger than reported differences between type B and C [50] and also larger than the reported dosimetric interplay effect in IMRT and VMAT studied using only 4DCT [47,48], phantom [45] or perturbed dose [46].

Schwarz et al. [51] have reviewed the status of dosimetric uncertainties in lung SBRT, both with regards to geometrical uncertainties and from other sources.

### 4.1. Limitations and strengths

We have purposefully studied the effect of the respiratory motion in isolation without e.g. setup uncertainties. Additional data, method development and/or analysis are needed to determine clinical dose prescription levels and total treatment margins when considering other uncertainties.

The dosimetric effects of tumor shifts due to respiratory motion and setup uncertainties in SBRT have been analyzed by Karlsson et al. [52] using a dose shift approximation with only a small observed impact on dose. This approximation is supposedly adequate for motion that is covered by the PTV margins but not for the potentially larger motions and out-of-field shifts present in our study.

The typical incorporation of setup uncertainties, as a positive and a negative shift along each cardinal direction, would for our presented methodology require a multiplication by seven regarding the amount of dose computations needed for a total of $7 \times 4 \times 4 \times 300 = 117,600$ computations. The prospect is computationally feasible for the size of our current patient cohort, but it was deliberately left out as combining margins fulfilling different functions may have hampered the analysis of the effects of intrafractional motion. It could however be addressed in a separate study.

Creating evaluation images using deformable image registration between the cine-CT and 4DCT images was found to be challenging due to the limited anatomical information in the cine-CT. The main limitation stems from the limited volume captured by the cine-CT in the craniocaudal direction, while the in-plane field of view was of same or of comparable size as for the BHCT and 4DCT. Since the motion in the lungs generally is the largest in the craniocaudal direction using deformable image registration, that uses many degrees of freedom, to move parts of the source image containing the tumor outside of the cine-CT acquisition volume in the craniocaudal direction without any further regularization in the registration method is to be considered an ill-posed problem and indeed gave very inconsistent and unsatisfactory registration results. Without compromising the results regarding the reported doses, dose to target and nominal dose to ipsilateral lung, we settled for a schema using rigid image registrations. In order to analyze accumulated doses to structures other than the CTV this approximation will not be sufficient as full deformations of the anatomy are required. Such deformed images could potentially be created, based on the same...
image data, using either bio-mechanical, statistical deformation or motion models. For this reason the dose to ipsilateral lung was analyzed using the nominal plan dose.

Visual inspection, re-initialization and confirmation of the registration results proved to be necessary since many of the registrations were challenging due to the restricted total volume thickness of the cine-CT images. In quite a few cine-CT images it was not visible at all, forcing the registrations to rely on contrast differences due to other anatomical features such as the bronchial tree or blood vessels.

The correctional translation, \( T_{\text{corr}} \) in Eq. (1), was added as a way of removing a systematic difference in tumor position between imaging sessions. It is our assumption that it is mostly an artifact of the difficult process of accurately registering the relatively thin cine-CT volumes to the reference 4DCT phase image in a consistent manner. For future studies, we recommend a full length BHCT both in the beginning and the end of every imaging session, for frame of reference alignment to account for potential patient shifts or use surface guidance to assess and compensate for the shifts. Alternatively a similar study design could be achieved using MRI which is a non-ionizing image modality [10]. Dose reconstruction using MRI in this manner has been demonstrated for prostate radiotherapy [53].

All plans were created for dosimetric assessment purposes and not as viable clinical plans. As such no additional margins apart from the ITV were used and the dose levels to organs at risk such as the ipsilateral lung could be lower than if the planning methods had been adapted for clinical use.

We have set the intended dose to the same value for all tumors regardless of size but as suggested by Bibault et al. [31] this could be a factor to take into consideration during clinical plan design.

Nevertheless, the planning methods can still be compared with each other in this regard but caution should be taken when extrapolating the conclusions for plans that target a larger PTV. Furthermore, some of the patients that exhibited very large amplitudes, \( e.g. > 10 \text{ mm} \), would probably have been candidates for additional motion mitigation strategies such as abdominal compression when implementing clinical plans.

5. Conclusion

While motion analysis showed that some instantaneous tumor displacements could be much larger than in the 4DCT, the dosimetric effects of motion during free breathing was small when using 4D planning techniques. If treated in free breathing, 93% of the patients would have received at least 95% of the intended dose level for the three planning methods RB4, FLU and ISD and for 86% of the patients for IRN.

However, homogeneity of the dose varied between the methods in favor of RB4. Using robust 4D planning and prescribing to the CTV thus gave the best result for target coverage.

A correlation between the risk of relative underdosage and CTV coverage by ITV was found for two of the planning methods, FLU and IRN. We conclude that stronger statistical evidence is needed to be used for predicting patients with higher risk of poor target coverage.

If planning is to be done following recent trends in prescription of \( D_{\text{mean}} \) or \( D_{\text{95s}} \) to CTV we recommend using robust 4D planning, making sure the prescription level is achieved at each phase, in order to minimize the risk of underdosage while achieving good homogeneity. The increased availability of computational tools and resources can facilitate the easy implementation of the technique in clinical practice.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgments

The authors would like to acknowledge financial support from Cancerfonden, Sweden (project CAN 2017-616) as well as from the research department of clinical oncology at Uppsala University hospital, Sweden. The authors would also like to thank Erik Traneus at RaySearch for technical assistance and Kristina Nilssom for the original delineations of the CTVs in the patient cohort.

Patient image data is not made publicly available due to ethical reasons.

Uppsala University has a research agreement with RaySearch Laboratories AB.

Appendix A. Supplementary data

Supplementary material related to this article can be found online at https://doi.org/10.1016/j.epjemp.2022.09.007.

References

N. Olofsson et al.