Liver-ultrasound based motion modelling to estimate 4D dose distributions for lung tumours in scanned proton therapy

Giger, Alina; Krieger, Miriam; Jud, Christoph; Duetschler, Alisha; Salomir, Rares Vincent; Bieri, Oliver; Bauman, Grzegorz; Nguyen, Damien; Weber, Damien C; Lomax, Antony J; Zhang, Ye; Cattin, Philippe C

How to cite

GIGER, Alina et al. Liver-ultrasound based motion modelling to estimate 4D dose distributions for lung tumours in scanned proton therapy. In: Physics in medicine and biology, 2020, vol. 65, n° 23, p. 235050. doi: 10.1088/1361-6560/abaa26

This publication URL: https://archive-ouverte.unige.ch/unige:164604

Publication DOI: 10.1088/1361-6560/abaa26

This document is the accepted manuscript version of the following article:
Giger, A., Krieger, M., Jud, C., Duetschler, A., Salomir, R., Bieri, O., … Cattin, P. C. (2020). Liver-ultrasound based motion modelling to estimate 4D dose distributions for lung tumours in scanned proton therapy. Physics in Medicine and Biology. https://doi.org/10.1088/1361-6560/abaa26

Liver-ultrasound based motion modelling to estimate 4D dose distributions for lung tumours in scanned proton therapy

Alina Giger[a,b*], Miriam Krieger[c,d*], Christoph Jud[a,b], Alisha Duetschler[c,d], Rares Salomir[e,f], Oliver Bieri[a,g], Grzegorz Bauman[a,g], Damien Nguyen[a,g], Damien C Weber[c,h,i], Antony J Lomax[c,d], Ye Zhang[c], Philippe C Cattin[a,b]

[a] Department of Biomedical Engineering, University of Basel, Allschwil, Switzerland
[b] Center for medical Image Analysis & Navigation, University of Basel, Allschwil, Switzerland
[c] Paul Scherrer Institute (PSI), Center for Proton Therapy, Villigen PSI, Switzerland
[d] Department of Physics, ETH Zurich, Zurich, Switzerland
[e] Image Guided Interventions Laboratory (GR-949), Faculty of Medicine, University of Geneva, Geneva, Switzerland
[f] Radiology Division, University Hospitals of Geneva, Geneva, Switzerland
[g] Department of Radiology, Division of Radiological Physics, University Hospital Basel, Basel, Switzerland
[h] Department of Radiation Oncology, University Hospital Zurich, Zurich, Switzerland
[i] Department of Radiation Oncology, Inselspital Bern, Bern, Switzerland
[*] Both authors contributed equally

E-mail: alina.giger@unibas.ch, miriam.krieger@psi.ch

July 2020

Abstract. Motion mitigation strategies are crucial for scanned particle therapy of mobile tumours in order to prevent geometrical target miss and interplay effects. We developed a patient-specific respiratory motion model based on simultaneously acquired time-resolved volumetric MRI and 2D abdominal ultrasound images. We present its effects on 4D pencil beam scanned treatment planning and simulated dose distributions. Given an ultrasound image of the liver and the diaphragm, principal component analysis and Gaussian process regression were applied to infer dense motion information of the lungs. 4D dose calculations for scanned proton therapy were performed using the estimated and the corresponding ground truth respiratory motion; the differences were compared by dose difference volume metrics. We performed this simulation study on 10 combined CT and 4DMRI data sets where the motion characteristics were extracted from 5 healthy volunteers and fused with the anatomical CT data of two lung cancer patients. Median geometrical estimation errors below 2 mm for all data sets and maximum dose differences of $V_{diff>5\,\%} = 43.2\,\%$ and $V_{diff>10\,\%} = 16.3\,\%$ were found. Moreover, it was shown that abdominal ultrasound imaging allows to monitor organ drift. This study demonstrated the feasibility of the proposed ultrasound-based motion modelling approach for its application in scanned proton therapy of lung tumours.

Keywords: 4DMRI, motion modelling, Gaussian process regression, proton therapy, 4D dose calculation, pencil beam scanning, lung tumour

Submitted to: *Phys. Med. Biol.*

1. Introduction

In recent years, research has focused on understanding and mitigating the effects of respiratory motion on scanned particle therapy. When moving targets such as lung or liver tumours are treated with pencil beam scanned (PBS) proton therapy, geometrical target miss can lead to underdosage at the lateral edge and thus reduce the conformity of the field dose. The interference of the sequential delivery of the proton dose and the motion of the tumour additionally leads to local hot and cold spots, known as the interplay effect (Phillips et al. 1992, Bert & Durante 2011). Over- or underdosage at the distal field occurs due to the sensitivity of the proton beam with respect to the motion induced density variations. All those effects together may lead to a sub-optimal treatment plan which requires detailed investigation for each patient specifically (Chang et al. 2017, Zhang et al. 2018). Integral tools to study the aforementioned effects are 4D dose calculations (4DDCs). These are expansions of conventional 3D dose calculations that include both patient motion and delivery dynamics into the calculations. Various 4DDCs have been developed, see e.g. Bert & Rietzel (2007), Richter et al. (2013), Ammazzalorso & Jelen (2014), Paganetti et al. (2005), Kang et al. (2005), Li et al. (2014), Engwall et al. (2016), Boye, Lomax & Knopf (2013), Krieger et al. (2018), Grassberger et al. (2015), in which patient-specific motion information is the prerequisite for time-resolved dose estimation. Usually, this information is extracted from a pre-treatment respiratory-correlated computed tomography (4DCT) assuming that this average breathing cycle is repeated periodically during the plan delivery. Indeed, it is clear that any single 4DCT cannot capture respiratory variations nor does it provide any real-time information reflecting the real motion scenario during the treatment. However, both above aspects are crucial inputs for accurately estimating or reconstructing the actual dose that a patient receives.

Time-resolved volumetric magnetic resonance imaging (4DMRI) as e.g. developed by von Siebenthal et al. (2007), Giger, Stadelmann, Preiswerk, Jud, De Luca, Celicanin, Bieri, Salomir & Cattin (2018), or Jud et al. (2018) solves the issue of capturing motion variabilities. However, to date no imaging modality can capture and process full 3D lung information in real-time, meaning that some surrogate data is needed in order to estimate the patient motion during dose delivery. An optimal surrogate signal should capture internal motion information which is correlated to the organ motion of interest, expose the patient to no additional radiation dose, and can provide multidimensional surrogate signals. A candidate modality that fulfils all mentioned requirements is abdominal ultrasound (US) imaging. The motion of the liver and diaphragm can be captured in real-time and provides two-dimensional internal information (Giger, Stadelmann, Preiswerk, Jud, De Luca, Celicanin, Bieri, Salomir & Cattin 2018, Preiswerk et al. 2014). Due to physical constraints, direct US imaging of lung tissue is not possible. However, with the help of respiratory motion models, motion characteristics extracted from liver US can be used to estimate lung tumour motion (Mostafaei et al. 2018) and lung deformation (Giger, Sandkühler, Jud, Bauman,

Bieri, Salomir & Cattin 2018, Giger et al. 2019). While Mostafaei et al. (2018) have demonstrated the correlation between diaphragm motion and lung tumour motion in superior-inferior (SI) direction, Giger, Sandkühler, Jud, Bauman, Bieri, Salomir & Cattin (2018) and Giger et al. (2019) have inferred dense lung motion information from 2D abdominal US images. Since we aim to investigate respiratory motion and its variabilities in the context of PBS proton treatment, monitoring 1D tumour motion only is not sufficient as motion information of the surrounding tissue needs to be additionally taken into account (Trnková et al. 2018, Bertholet et al. 2019). Moreover, in contrast to our previous work (Giger et al. 2019), the motion model presented here is based on non-linear methods and therefore potentially has better estimation power.

The aim of this study was to investigate the dosimetric impact of lung motion estimation with the use of online abdominal US imaging and respiratory motion modelling in PBS proton therapy, mainly for two reasons: (1) The performed 4DDCs, when combined with temporal data provided by delivery log files (see e.g. Krieger et al. (2018)), may be used to retrospectively estimate the dose delivered to a patient and, if needed, to adapt the plans accordingly for future fractions (Meijers et al. 2019). (2) The presented framework forms the basis for US-guided proton beam tracking deliveries and, thus, should be understood as one of several steps towards this goal. We simulated realistic treatment scenarios with intra-fractional respiratory motion variabilities on synthetic 4DCT(MRI) data sets. Although similar studies have been performed on 4D data of the liver (Zhang et al. 2013, Zhang et al. 2014, Zhang et al. 2015), this work is the first focusing on the motion model application for lung tumour treatments and the use of abdominal US imaging as a non-invasive and informative surrogate signal.

2. Methods

2.1. Data acquisition

Hybrid US and 4DMRI data sets of five healthy volunteers were acquired under free respiration in a 1.5 T MRI scanner (MAGNETOM Aera, Siemens Healthineers, Erlangen, Germany). Among those, two data sets were acquired with a navigator-based slice-stacking approach (Celicanin et al. 2017) based on an ultra-fast balanced steady-state free precession (ufbSSFP) pulse sequence (Bieri 2013). Multi-slice coronal 2D MR images were acquired in sequential order, alternating with coronal navigator images at a fixed slice position. The remaining three data sets were acquired with a recently presented spoiled gradient echo 4DMRI sequence where core and pseudo-randomly sampled peripheral k-space patches are acquired alternately (Jud et al. 2018). The core patches are used to estimate the spatial relative motion in order to correct the peripheral patches for motion. These motion-corrected patches are subsequently accumulated into a consistent k-space for each motion state. Detailed information about the sequence parameters is listed in table 1. Hereinafter the two 4DMRI sequences are referred to as *slice stacking* and *patch registration* approach, respectively.

The frame rate of the two 4DMRI methods are $f_{\mathrm{MRI}} = 1.25\,\mathrm{Hz}$ and $2.25\,\mathrm{Hz}$, respectively. Simultaneous abdominal US imaging was performed at $f_{\mathrm{US}} = 15\,\mathrm{Hz}$ with an Acuson clinical scanner (Antares, Siemens Healthineers, Mountain View, CA). The system was equipped with a specially designed MR-compatible US linear array transducer consisting of 192 elements (Fraunhofer IBMT, Sulzbach, Germany), equivalent to the Siemens' VF7-3 probe, with central frequency 5 MHz, and bandwidth 4 MHz. The US imaging probe was strapped to the subjects' abdomen and held in place using a home-built casting, as shown in the report of (Santini et al. 2020). The probe was operated in B-mode at 3.3 MHz and covered a maximum field depth of 16 cm and an angular sector of up to 90°. The transmitted acoustic power was adjusted in situ case-specifically and was in the range of −11 to 8 dBm. No interference occurred between the US and the MR acquisitions. Data acquisition was performed for a duration of up to 11 min.

To ensure temporal correspondence between the two imaging modalities, two different strategies were employed. For the 4DMRI based on slice stacking, a continuous US video was acquired whose starting point coincided with the starting point of MR data acquisition within 0.5 to 1.1 s. The synchronisation was achieved via the software user interface of the two systems using a unique button signal sent at the beginning of the acquisition. The two imaging modalities were further temporally aligned in a post-processing step to correct for the residual offset caused by the dissimilar system response times of their software interface command. To do so, both the US images and the MR navigator slices were cropped to a region around the diaphragm based on which the mean image intensity was computed. These respiratory surrogate signals were filtered using a moving average filter before automatic peak detection was performed. Finally, the offset was computed as the time difference between the first peaks associated with end-inhalation for both signals. For the 4DMRI approach based on patch registration, optical output triggers of the MRI scanner were used to automatically start the US acquisitions using a fast hardware interface. One independent trigger signal was sent after every 6.675 s, or 15 MR frames, yielding the record of an US video of 5 s, as offered by the US system. Visual inspection of the extracted mean image intensity signals, computed similarly as described above, revealed that the offset between the two image modalities was negligible. Therefore, no post-processing was required in this case. The time difference of 1.675 s between the MR and US recording windows was chosen deliberately to take into account the US system latency while storing the US files. As a consequence, some MR images have no corresponding US images and were excluded from the motion modelling. Nevertheless, it happened that a trigger was issued before the previous US video was stored resulting in a further loss of US data of up to 7 % as shown in figure 1.

Due to the different frame rates of the 4DMRI and the US video, the sampling points do not perfectly match. However, since the frame rate of the US stream f_{US} is substantially higher than the one of the 4DMRI f_{MRI}, corresponding image pairs were assumed to represent the lungs in a sufficiently similar respiratory state.

Table 1: 4DMRI sequence parameters. The field of view and the image matrix are given in SI × LR × AP direction.

	Slice stacking (Celicanin et al. 2017)	Patch registration (Jud et al. 2018)
Contrast	T2/T1	T1
Flip angle	35°/28°	5°
Echo time	0.86 ms	1.0 ms
Repetition time	1.91 ms	2.5 ms
Bandwidth	2365 Hz px^{-1}	1560 Hz px^{-1}
Field of view	400 × 395.8 × 174.2 mm^3	400 × 400 × 275 mm^3
Image matrix	192 × 190 × 32 px	128 × 128 × 88 px
Core/peripheral patch radii	–	6 px/5 px

2.2. Deformable image registration

To extract motion information from the 4DMRI, volumetric deformable image registration (DIR) between 3D motion 'frames' was applied. For the slice stacking approach, a 3D B-spline based registration method implemented in Plastimatch‡ was used. The reconstructed image volumes were registered to a reference volume at end-exhalation. Conversely, for the patch registration approach, DIR of the core patches transformed to the spatial domain is required for the reconstruction of the 4DMRI in the first place. The resulting deformation vector fields (DVF) were the primary focus for this study and directly used as motion information for the subsequent step. In this case, the image registration was performed using the image registration framework AIRLab, a B-spline transformation model, an isotropic total variation regulariser, and the mutual information similarity measure (Sandkühler et al. 2018). The registration was performed for the lungs only by applying a semi-automatically extracted sliding-organ mask (Vezhnevets & Konouchine 2005). Again, all respiratory states were registered to a reference volume at end-exhalation.

2.3. 4DCT(MRI)

In order to conduct proton dose calculations, information on the relative proton stopping power is needed. This information can be extracted from CT Hounsfield Units (HU) with the use of a calibration curve. However, 4DCT, although capable of capturing motion, only represents one averaged breathing cycle, and does not include motion variability. Motion variabilities can be acquired using 4DMRI as described in the previous subsection. In order to combine these two types of information, synthetic 4DCT(MRI) data sets as described by Boye, Samei, Schmidt, Székely & Tanner (2013), Zhang et al. (2016) and Krieger et al. (2020) were generated.

DVFs extracted from the five 4DMRI data sets were matched to the full-exhale CT scans

‡ `www.plastimatch.org`, accessed: 13.01.2020

of two lung cancer patients§ following the meshing approach described in detail by Boye, Samei, Schmidt, Székely & Tanner (2013). Using this procedure, the two geometries (volunteer MRI and patient CT) are brought to anatomical correspondence by definition and thus the MRI DVFs can be readily transferred to the CT geometry. In order to preserve the sliding boundaries between the lungs and the chest wall, the motion of the chest wall is set to zero. By warping two CT geometries with five MRI motion patterns, ten different geometry/motion combinations were generated, each including 99 to 159 full and variable breathing cycles.

The two CT geometries were deliberately chosen because of their significant difference in size and position: geometry 1 shows a small tumour not attached to neighbouring tissues, geometry 2 contains a larger tumour which is partially attached to the spine. Figure 1 shows the target motion patterns in SI direction for all data sets. For motion 5, a drift of the clinical target volume (CTV) is observable in both CT geometries. A schematic of the target amplitudes and periods is shown in figure 2.

2.4. Motion modelling

A motion model based on Gaussian process (GP) regression was employed in order to estimate the respiratory motion during dose delivery (Williams & Rasmussen 2006). Given the input US image, the model infers the corresponding DVF of the lungs. A schematic overview of the motion model is given in figure 3.

Let $\boldsymbol{x}_t \in \mathbb{R}^{mn}$ denote the vectorised US image of dimension $m \times n$ at time t and analogously $\boldsymbol{y}_t \in \mathbb{R}^{3pqr}$ the vectorised DVF of dimension $p \times q \times r \times 3$. The factor of 3 in $\boldsymbol{y}_t$ is introduced since the voxel values of the DVF are 3-dimensional deformation vectors. Principal component analysis (PCA) has been applied prior to fitting the motion model in order to reduce the dimensionality of $\boldsymbol{x}_t$ and $\boldsymbol{y}_t$ and to remove co-linearities and noise (McClelland et al. 2013). Let $\boldsymbol{\alpha}_t \in \mathbb{R}^u$ and $\boldsymbol{\beta}_t \in \mathbb{R}^v$ represent the principal components (PC) of the US image $\boldsymbol{x}_t$ and the DVF $\boldsymbol{y}_t$, respectively. The number of PCs u and v considered for motion modelling was chosen such that the explained variance is greater than 50 %.

Given a set of M corresponding PCs, $\mathcal{S} = \{(\boldsymbol{\alpha}_t, \boldsymbol{\beta}_t) \,|\, t = 1, \ldots, M\}$, the regression can now be formulated in the feature space. The objective is to find a function $f : \mathbb{R}^u \to \mathbb{R}^v$ which maps the input PCs $\boldsymbol{\alpha}_t$ to the output PCs $\boldsymbol{\beta}_t$ assuming noisy observations:

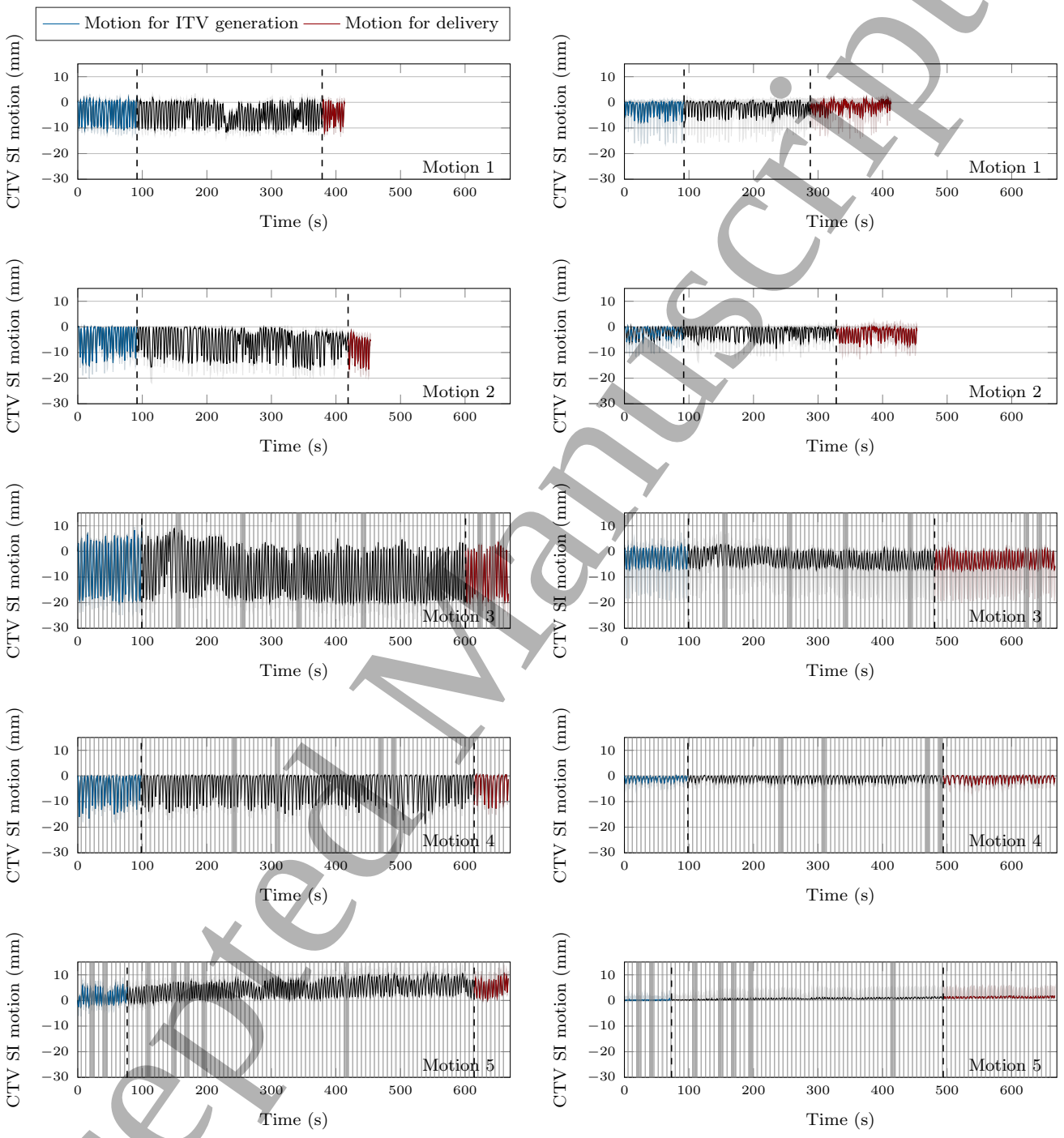
$$\boldsymbol{\beta}_t = f(\boldsymbol{\alpha}_t) + \boldsymbol{\epsilon}_t, \qquad \boldsymbol{\epsilon}_t \sim \mathcal{N}(\mathbf{0}, \sigma_n^2 I_v), \tag{1}$$

where σ_n^2 denotes the noise variance and I_v the v-dimensional identity matrix. Let $f(\boldsymbol{\alpha}_t)$ be a GP with zero mean and covariance function $k : \mathbb{R}^u \times \mathbb{R}^u \to \mathbb{R}^{v \times v}$, that is

$$f(\boldsymbol{\alpha}_t) \sim \mathcal{GP}(\mathbf{0}, k(\boldsymbol{\alpha}_t, \boldsymbol{\alpha}_{t'})). \tag{2}$$

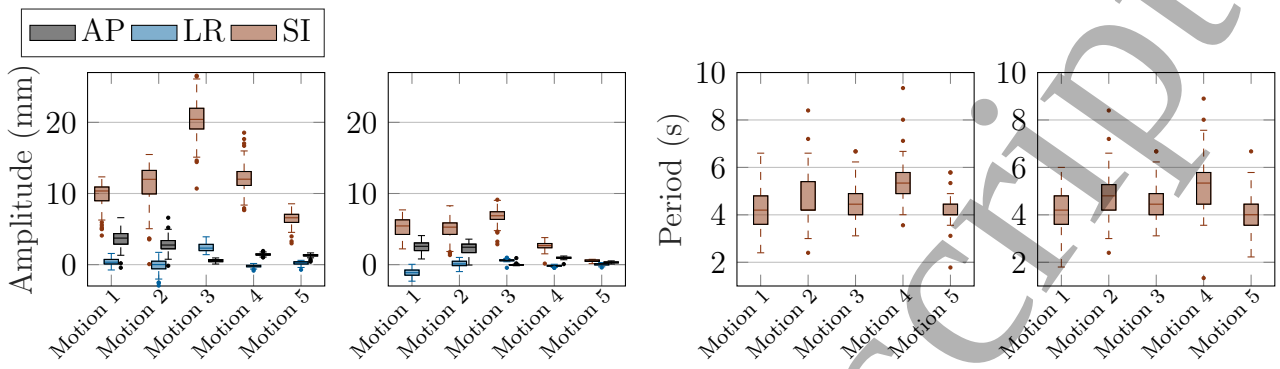
Then, given the input $\boldsymbol{\alpha}_*$ and the measurements $\mathcal{S}$, the estimate $\boldsymbol{\beta}_*$ can be calculated from the posterior distribution which is again a GP with mean $\boldsymbol{\mu}_S \in \mathbb{R}^v$ and covariance

§ Hugo et al. (2016), Hugo et al. (2017), Balik et al. (2013), Roman et al. (2012), Clark et al. (2013)



(a) CT geometry 1 (b) CT geometry 2

Figure 1: Motion patterns of the tumour in SI direction for all data sets. The grey shaded areas depict MR motion states without corresponding US frames. The red curve represents the motion for dose delivery and, therefore, the test set for motion modelling. The remaining data (blue and black curve) was split into a training and validation set for motion modelling while the blue curve was additionally used for ITV generation and treatment planning. Solid line: median, shades: 10th to 90th percentile.



(a) CT geometry 1 (b) CT geometry 2 (c) CT geometry 1 (d) CT geometry 2

Figure 2: Overview over the motion amplitudes in each motion direction (a), (b) and motion periods (c), (d). The boxplots include all breathing cycles in the respective data set. The whiskers extend to the most extreme values still within 1.5 time the inter-quartile range (IQR). SI: superior-inferior, LR: left-right, AP: anterior-posterior.

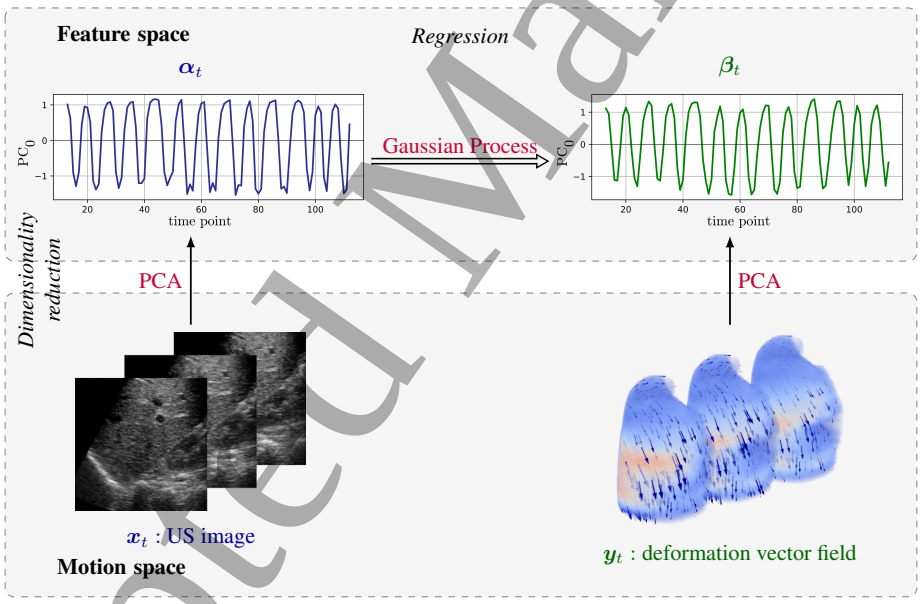


Figure 3: Schematic overview of the motion model.

$k_S \in \mathbb{R}^{v \times v}$ (Williams & Rasmussen 2006):

$$p(\boldsymbol{\beta}_*|\boldsymbol{\alpha}_*, \mathcal{S}) = \mathcal{GP}(\boldsymbol{\mu}_S, k_S). \tag{3}$$

Let $K \in \mathbb{R}^{vM \times vM}$ represent a matrix of the covariance $k(\boldsymbol{\alpha}_i, \boldsymbol{\alpha}_j)$ evaluated for all input training point pairs, $(\boldsymbol{\alpha}_i, \boldsymbol{\alpha}_j)$ with $i, j \in \{1, \ldots, M\}$. Similarly, let $K_* \in \mathbb{R}^{vM \times v}$ denote the matrix of the covariance between the test point $\boldsymbol{\alpha}_*$ and the M training points, $k(\boldsymbol{\alpha}_*, \boldsymbol{\alpha}_i)$ with $i \in \{1, \ldots, M\}$. Finally, the output training points are collected in a vector $\boldsymbol{b} = \left[\boldsymbol{\beta}_1^T \ldots \boldsymbol{\beta}_M^T\right]^T \in \mathbb{R}^{vM}$. The mean and covariance of the posterior can now be

written in closed form using the vM-dimensional identity matrix I_{vM}:

$$\boldsymbol{\mu}_S = K_*^T \left(K + \sigma_n^2 I_{vM}\right)^{-1} \boldsymbol{b}, \tag{4}$$

$$k_S = k\left(\boldsymbol{\alpha}_*, \boldsymbol{\alpha}_*\right) - K_*^T \left(K + \sigma_n^2 I_{vM}\right)^{-1} K_*. \tag{5}$$

Since the respiratory motion in the lungs is expected to be smooth, a Gaussian kernel was chosen as covariance function,

$$k\left(\boldsymbol{\alpha}_t, \boldsymbol{\alpha}_{t'}\right) = \theta_0^2 \exp\left(-\frac{\left\|\boldsymbol{\alpha}_t - \boldsymbol{\alpha}_{t'}\right\|^2}{2\theta_1^2}\right) I_v, \tag{6}$$

with scaling parameter θ_0 and characteristic length-scale θ_1. The mean of the posterior distribution $\boldsymbol{\mu}_S$ represents the best estimate of the GP regression. Moreover, the posterior variance k_S can be interpreted as estimation uncertainty and might serve as a confidence value or quality measure (Williams & Rasmussen 2006).

The hyperparameters were manually optimised and set to $\theta_0 = 30$, $\theta_1 = 35$ and $\sigma_n^2 = 1$ for all geometry/motion data sets. The motion data as shown in figure 1 was split into a training, validation and test set. The size of the test set (red curve) was defined as the minimum number of motion states needed for the dose delivery and is in the range of 59 to 88 states for geometry 1 and 210 to 286 states for geometry 2. The remaining data (blue and black curve) was further split into a constant validation set consisting of the last 55 states and a training set of 425 to 913 states.

2.5. 4D treatment planning and dose calculation

As a planning target, we use a previously presented probabilistic ITV concept (Krieger et al. 2020). The ITV50 of the first 20 breathing cycles of each geometry/motion case was used as the target for optimisation. For this, conventional ITVs were first created by unifying all CTV positions of one breathing cycle at a time. By summing up these 20 ITVs, a probability map of the tumour position was generated. The probabilistic ITV50 was then defined as all voxels with a probability of 50 % of being within any ITV calculated from any motion cycle. In order to achieve sufficient target coverage in the static case, a 2 mm margin was added to the ITV50 to define the technical planning target volume (tech-PTV). In this study, the set-up uncertainties were neither considered in the treatment planning nor simulated for the dynamic dose delivery. As planning CT, the definition of Botas et al. (2018) was applied. The density values in the whole CT minus the ITV50 were set to the average value within the first 20 breathing cycles, whereas the density values within the ITV50 were defined as the maximum intensity projection of the same 20 breathing cycles. Thus, every geometry/motion case had its own planning CT and tech-PTV. For both CT geometries, a single field, uniform dose (SFUD) plan with two fields was optimised on the tech-PTV and the planning CT of the respective geometry/motion case due to SFUD's superior stability in the presence of motion (Gorgisyan et al. 2019). The field directions and CTV volumes are shown in figure 4. For the optimisation and the consequent 4D dose calculations, the delivery model of the PSI gantry 2 was used (Pedroni et al. 2004).

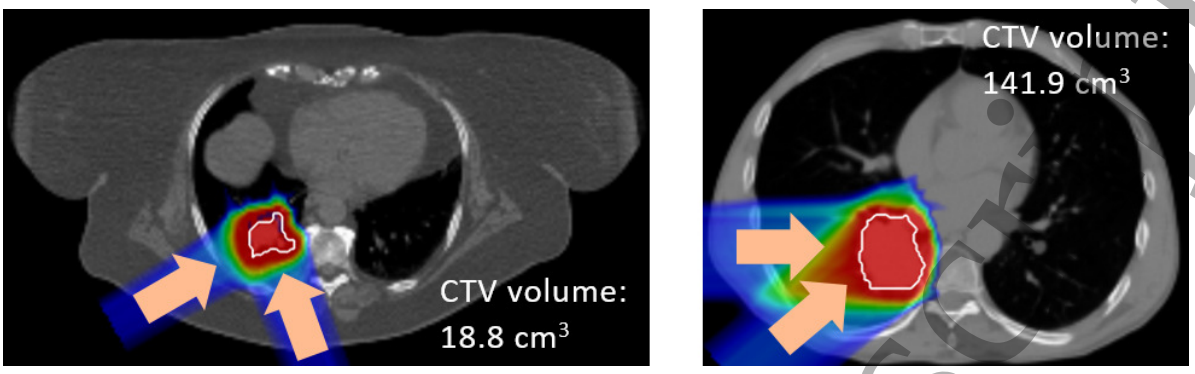


Figure 4: Illustration of the two CT geometries and the beam directions of the treatment plans. Left panel: CT geometry 1, right panel: CT geometry 2.

A ray casting based, deforming grid 4D dose calculation engine as described by Boye, Lomax & Knopf (2013) and validated experimentally by Krieger et al. (2018) was used in order to study the dosimetric effects of the motion estimation. 4DDCs were performed using either the ground truth or the corresponding estimated motion states as input. All motion states of the first *delivery* cycle were used as starting point for the simulation since it is usually unknown in which respiratory state the delivery would start. Depending on the motion pattern, the number of starting phases per field ranged from 5 to 13 (average: 9.2), resulting in 25 to 169 (average: 91.2) scenarios per two-field plan.

2.6. Evaluation

Geometrical error The geometrical estimation error $\boldsymbol{e}_t \in \mathbb{R}^{pqr}$ for motion state t is defined based on the voxel-wise difference between the ground truth and the estimated DVFs. Let $\boldsymbol{v}_t^i \in \mathbb{R}^3$ denote the ground truth deformation vector at voxel $i \in \{1, \ldots, pqr\}$. Similarly, let $\tilde{\boldsymbol{v}}_t^i \in \mathbb{R}^3$ be the estimated deformation vector at voxel i. Then, the i-th entry of the estimation error $\boldsymbol{e}_t$ is given as

$$e_t^i = \|\boldsymbol{v}_t^i - \tilde{\boldsymbol{v}}_t^i\|, \tag{7}$$

where $\|\cdot\|$ denotes the Euclidean norm. Note that the term ground truth refers to the target DVF used to train the model. Although the extracted DVFs do not necessarily represent the real patient motion due to errors introduced by the registration and the 4DCT(MRI) generation, they do still represent a valid ground truth for the motion model.

In order to get a qualitative overview of the error distributions in space, error maps were generated using a colour wash to indicate the spatial distribution of the estimation error, averaged over all motion states.

For a quantitative analysis, the 50th and 95th percentiles of the estimation error within a given volume of interest $\mathrm{VOI}_{\mathrm{geom}}$ were calculated for each respiratory state. Below, the results are reported as a distribution of these percentiles over all motion states. The

VOI_{geom} was defined as all voxels which receive any dose in the static dose calculation and which lie within the lung or the ITV.

Drift analysis The geometrical estimation error was equivalently evaluated as described above in order to investigate the influence of the organ drift on the performance of the motion model. To do so, the training sets for motion 5 and both geometries were divided into equally sized subsets consisting of 220 motion states or 97.9 s each. The quantitative analysis was conducted using each of the 4 subsets for geometry 1 and 3 subsets for geometry 2 as training set while the test set remained unchanged. Subset 1 represents the training motion data from respiratory states 1 to 220, subset 2 from states 221 to 440, and so on. Consequently, there is a decreasing time gap between the training set and the test set for an increasing number of the subset.

Dosimetric error The dose error was calculated as the voxel-wise absolute dose difference between ground truth and estimated 4DDCs. Absolute dose difference volume histograms (DDVHs) were calculated for the CTV and the volume of interest VOI_{dose} defined as all voxels within the CTV plus 20 mm margin, which lie within the lung. Additionally, the median volume percentages for which the absolute dose difference was more than 5 % or 10 % of the prescribed dose, respectively, were calculated ($V_{diff>5\,\%}$, $V_{diff>10\,\%}$).

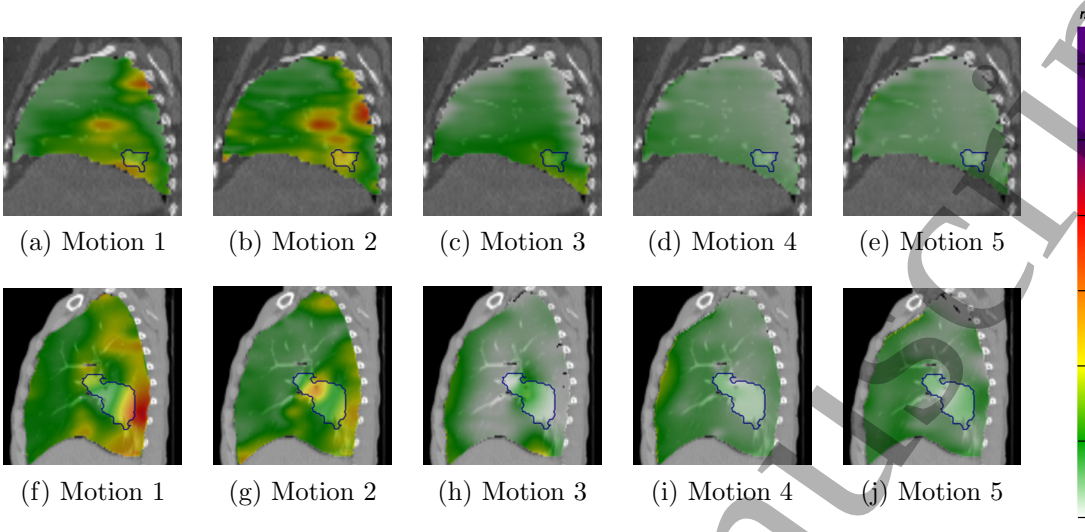
3. Results

3.1. Geometrical error

Figure 5 shows one sagittal slice through the CT geometries with an error map overlaid for each geometry/motion case. It can be seen that the error tends to become larger towards the edge of the lung. For all scenarios, the tumour lies partly within the high error region and partly within the lower error region. Motions 1 and 2 tend to have larger geometrical errors than motions 3–5. Boxplots of the geometrical estimation error including all motion states are shown in figure 6. It is seen that the 50th and 95th percentile error within the VOI_{geom} for one motion are similar for the two CT geometries. Again, the largest errors of up to 8 mm (95th percentile) are found for motions 1 and 2 (geometries 1 and 2). For motion 3, geometries 1 and 2, the error stays below 4 mm and 2 mm respectively except for a few outlying motion states. Motion 4 and 5 show errors below 2 mm for both CT geometries. When looking at the 50th percentile (median), the errors stay below 3 mm for all cases except for a few outliers. For motions 3–5 it is lower than 1 mm.

3.2. Drift analysis

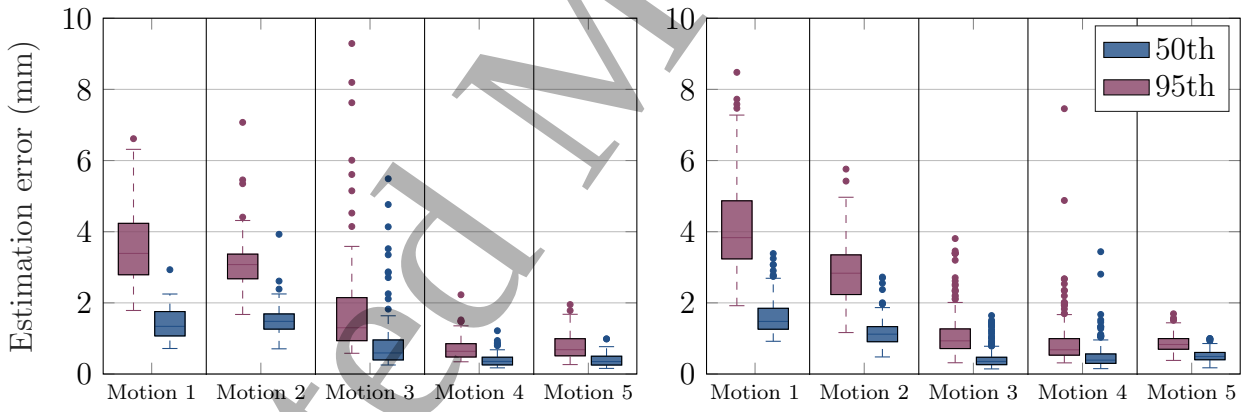
The drift analysis in figure 7 indicates that the estimation error increases the greater the time gap between the training data and the test data. For both CT geometries,



(a) Motion 1 (b) Motion 2 (c) Motion 3 (d) Motion 4 (e) Motion 5

(f) Motion 1 (g) Motion 2 (h) Motion 3 (i) Motion 4 (j) Motion 5

Figure 5: Example sagittal slices to illustrate the spatial error distribution, averaged over time. Top row: CT geometry 1, bottom row: CT geometry 2.



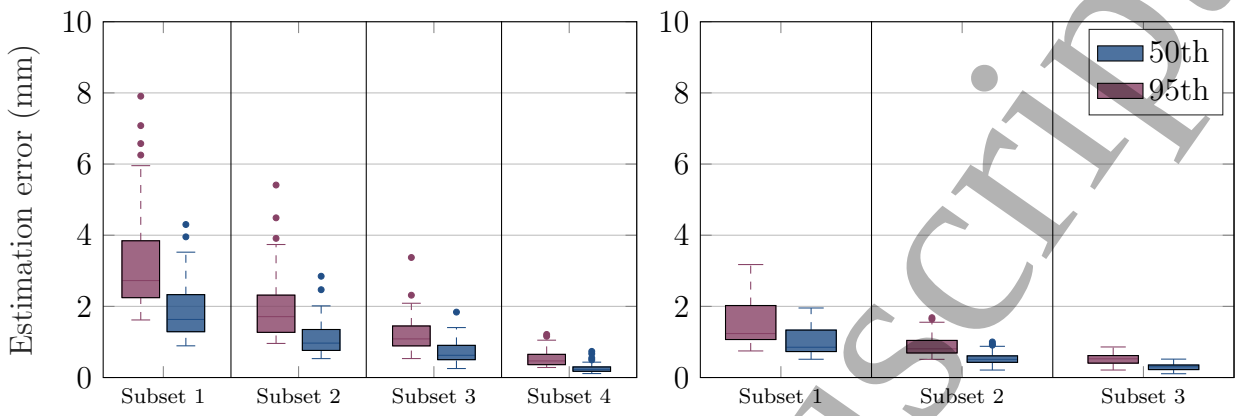
(a) CT geometry 1 (b) CT geometry 2

Figure 6: Motion estimation errors for all geometries and motions. The boxplots indicate the error percentiles of all voxels within the VOI_{geom}. The whiskers include all values within 1.5 IQR.

the estimation error is up to 4 times higher for subset 1 when compared to subset 4. The performance of the motion model is similar when it was trained on the complete training set or subset 4 only.

3.3. Dosimetric error

The influence of the estimation error on the dose distributions is shown in figure 8. The absolute dose difference volume histograms within the CTV and the VOI_{dose} are plotted for each 4DCT(MRI) data set. The solid lines display the median values, whereas the



(a) CT geometry 1, motion 5 (b) CT geometry 2, motion 5

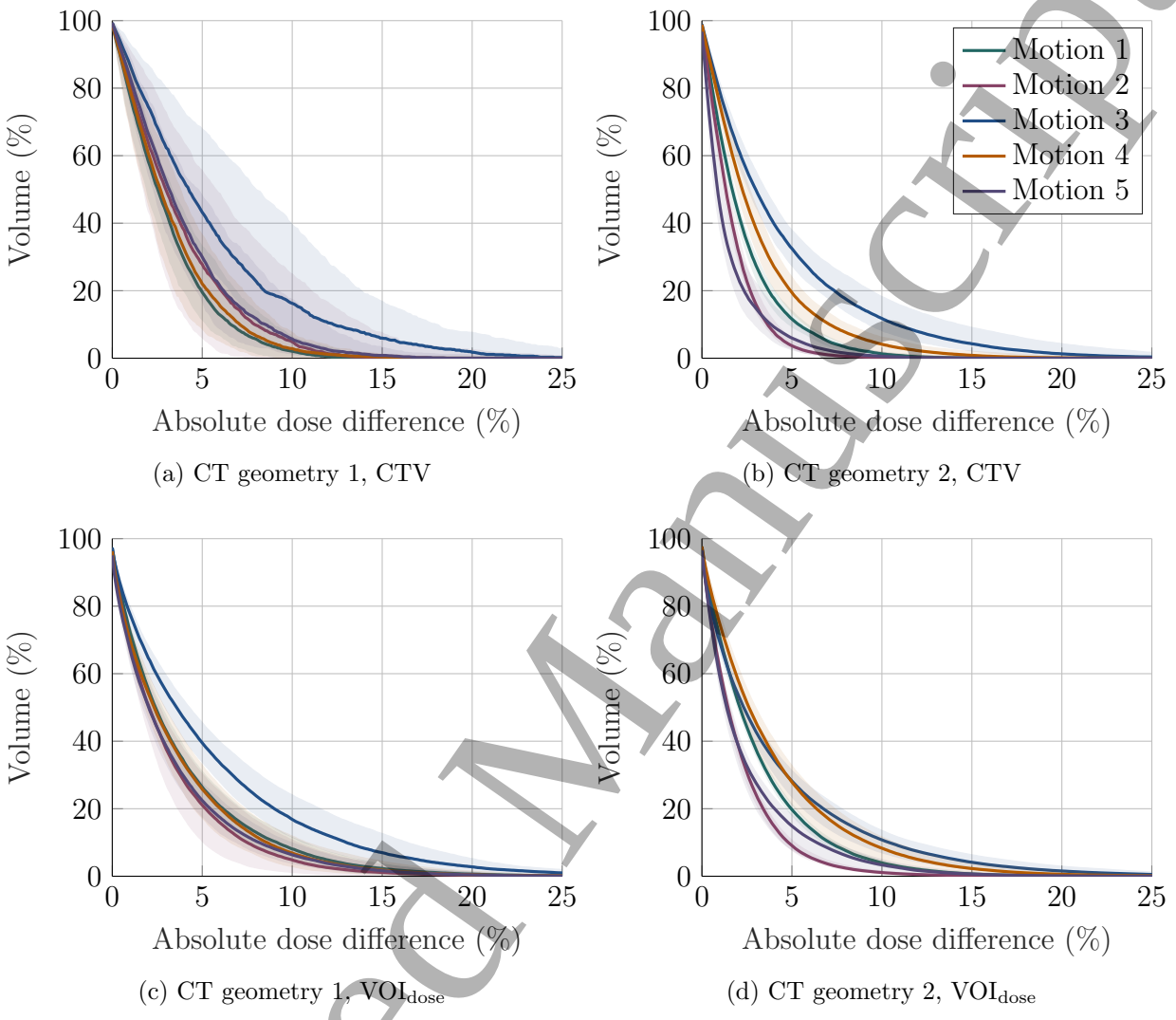
Figure 7: Drift analysis for motion 5 and different training subsets. The boxplots indicate the error percentiles of all voxels within the VOI_{geom}. The whiskers include all values within 1.5 IQR.

shaded bands include all possible starting phases ($5^2 - 13^2$ per two-field plan). It is seen that for CT geometry 2, the dose differences are somewhat lower and the spread due to different starting phases is less pronounced than for CT geometry 1. Motion 3 presents the highest dose differences, whereas the other motions present similar dose differences. When looking at the volumes with a difference of more than 5 % or 10 % (table 2), it is again seen that motion 3 leads to the highest percentages compared to other motion patterns. Again, geometry 2 shows lower values than geometry 1, except for the $V_{diff>10\,\%}$ of motion 4. All $V_{diff>5\,\%}$ values are below 44 %, and even below 30 % if motion 3 is not considered. Similarly, the $V_{diff>10\,\%}$ percentages are below 17 % for all cases and below 9 % when excluding motion motion 3.

4. Discussion

In this work, we have investigated the effects of respiratory motion estimation on the dose distribution in PBS proton therapy of lung tumours. Dense motion information in the lungs was estimated based on abdominal US imaging and patient-specific respiratory motion modelling. Time-resolved 4DCT(MRI) were employed to simulate motion variabilities over a comprehensive time duration of up to 11 min using two different 4DMRI approaches. To take these motion variabilities into account for treatment planning, a recently presented probabilistic ITV definition was applied (Krieger et al. 2020) and two-field SFUD plans were optimised on composite planning CTs.

Good geometrical and dosimetric agreement were achieved, however, with a tendency of higher geometrical errors for the 4DMRI based on slice stacking when compared to the patch registration approach. Due to the reconstruction properties of the latter method, which is based on DIR of low spatial resolution core patch



(a) CT geometry 1, CTV

(b) CT geometry 2, CTV

(c) CT geometry 1, VOI_{dose}

(d) CT geometry 2, VOI_{dose}

Figure 8: Absolute dose difference histograms for all geometries and motion cases, evaluated in the CTV and in VOI_{dose}.

reconstructions, the DVFs tend to be smoother and therefore potentially less prone to estimation errors. Additionally, the higher temporal resolution and larger training data sets available for this method may further influence the estimation accuracy positively. With 50th and 95th percentiles estimation errors in the range of 2 mm and 4 mm, respectively, for all geometry/motion cases, the presented respiratory motion model achieves clinically relevant results. It was also demonstrated that higher geometrical errors do not coincide with higher dosimetric errors. This suggests that the presented analysis is robust against the chosen 4DMRI method and, therefore, allows for viable conclusions regarding dosimetric errors. The dosimetric errors found in this study are in a clinically acceptable range, especially because for treatment, a motion mitigation technique such as rescanning can be combined. This has been shown to additionally reduce dosimetric uncertainties due to motion (Krieger et al. 2018, Zhang et al. 2014).

Table 2: Dose difference parameters, analysed in the VOI_{dose} and the CTV.

(a) VOI_{dose}

		Motion 1	Motion 2	Motion 3	Motion 4	Motion 5
$V_{diff>5\,\%}$	CT geom. 1	26.4 %	21.1 %	39.5 %	25.8 %	22.4 %
	CT geom. 2	19.7 %	9.1 %	28.2 %	27.9 %	14.9 %
$V_{diff>10\,\%}$	CT geom. 1	8.1 %	4.8 %	16.8 %	6.9 %	6.3 %
	CT geom. 2	4.0 %	1.1 %	10.8 %	8.3 %	3.4 %

(b) CTV

		Motion 1	Motion 2	Motion 3	Motion 4	Motion 5
$V_{diff>5\,\%}$	CT geom. 1	19.4 %	27.8 %	43.2 %	22.0 %	29.9 %
	CT geom. 2	11.6 %	3.7 %	32.5 %	19.5 %	6.0 %
$V_{diff>10\,\%}$	CT geom. 1	2.1 %	4.8 %	16.3 %	2.8 %	5.7 %
	CT geom. 2	1.3 %	0.2 %	11.8 %	4.1 %	0.6 %

Moreover, reasonable motion estimation performance was achieved for all geometry/motion cases even though the same hyperparameter set was chosen for all cases. This points to the conclusion that the presented motion model is robust with respect to different patient geometries and motion patterns. The fact that no extensive parameter tuning is required for every individual patient, makes this approach feasible for clinical applications. The drift analysis in figure 6 further revealed that the internal motion information provided by the abdominal US imaging was able to record organ drift. That does not necessarily mean that the motion model is able to cope with this long-term changes, however, it might still be used to monitor potential organ drift and to trigger an intervention of the clinicians if needed. Further investigations on prolonged data sets are required and the aim of future studies.

The 4DCT(MRI) data sets are synthetic in the sense that the respiratory motions of healthy volunteers were combined with the geometries of two cancer patients. This further implies that the motion of the CTV was simulated by healthy motion patterns and might not ideally represent pathological motion characteristics. Also, the US images used as motion surrogates do not match the anatomical geometry of the patients. However, we do not expect our findings to change substantially if pathological motion patterns and corresponding patient geometries were used for the analysis. Moreover, the current approach to generate 4DCT(MRI) data sets required the lung volume to be segmented. Consequently, the DVFs were applied to the lung tissue only while the surrounding structures such as the ribs and the skin remained static. However, we are working on a multi-organ numerical 4D phantom to fix this limitation.

For the motion model to be transferred to the clinics, it remains to be shown how the proposed approach performs in the case of inter-fractional motion variabilities. Besides anatomical changes on a day-to-day basis, this further implies dissimilar US imaging

planes due to the required repositioning of the US probe on the patient's abdominal wall. In a first study we have demonstrated the ability of a similar motion model to cope with US probe repositioning (Giger et al. 2019). However, the two data sets were acquired on the same day with only a few minutes in between, therefore not representing truly inter-fractional anatomical changes and motion variabilities.

Given the promising results of the presented study, further improvements are planned for future works. The synchronisation procedure between US and 4DMRI is currently being improved such that future data sets will not suffer from data loss as was the case for the 4DMRI based on patch registration. Instead, we aim for continuous US image acquisition as was the case for the 4DMRI based on slice stacking without, however, the need for subsequent temporal alignment. Further, we aim to extend the current analysis to the case where we simulate 3D tumour tracking by online proton beam adaptions. In this context, we will evaluate the use of the posterior variance of the GP regression as a quality measure for the motion estimation which could then potentially be used to combine tumour tracking and uncertainty-correlated gating.

5. Conclusion

Patient-specific motion modelling based on GP regression and abdominal US surrogates has shown to be a feasible and promising approach to estimate lung motion variabilities and their effects on dose distributions. It offers a possibility to take into account motion variabilities in 4D treatment planning, retrospective actual 4D dose reconstruction and online PBS beam tracking in future.

Acknowledgments

The authors thank Pauline Guillemin from the University of Geneva, Switzerland, for her support during the data acquisitions. This work was supported by the Swiss National Science Foundation, SNSF (project number 320030_163330/1).

References

Ammazzalorso F & Jelen U 2014 *Phys. Med. Biol.* **59**, N91–N99.

Balik S, Weiss E, Jan N, Roman N, Sleeman W C, Fatyga M, Christensen G E, Zhang C, Murphy M J, Lu J, Keall P, Williamson J F & Hugo G D 2013 *Int. J. Radiation Oncology Biol. Phys.* **86**, 372–379.

Bert C & Durante M 2011 *Physics in Medicine & Biology* **56**(16), R113.

Bert C & Rietzel E 2007 *Radiation Oncology* **2:24**.

Bertholet J, Knopf A, Eiben B, McClelland J, Grimwood A, Harris E, Menten M, Poulsen P, Nguyen D T, Keall P et al. 2019 *Physics in Medicine & Biology* **64**(15), 15TR01.

Bieri O 2013 *Magnetic resonance in medicine* **70**(3), 657–663.

Botas P, Grassberger C, Sharp G & Paganetti H 2018 *Phys. Med. Biol.* **63**, 035023.

Boye D, Lomax A J & Knopf A C 2013 *Med. Phys.* **40(6)**, 061702–1 – 061702–11.

Boye D, Samei G, Schmidt J, Székely G & Tanner C 2013 *in* 'Medical Imaging 2013: Image Processing' SPIE. doi: 10.1117/12.2007076.

Celicanin Z, Giger A, Bauman G, Cattin P & Bieri O 2017 *in* 'Proceedings of the ISMRM 25th Annual Meeting and Exhibition' Honolulu, Hawaii, USA.

Chang J Y, Zhang X, Knopf A, Li H, Mori S, Dong L, Lu H M, Liu W, Badiyan S N, Both S, Meijers A, Lin L, Flampouri S, Li Z, Umegaki K, Simone C B & Zhu X R 2017 *International Journal of Radiation Oncology*Biology*Physics* **99**(1), 41.

Clark K, Vendt B, Smith K, Freimann J, Kirby J, Koppe P, Moore S, Phillips S, Maffitt D, Pringle M, Tarbox L & Prior F 2013 *J. Digit. Imaging* **26(6)**, 1045–1057.

Engwall E, Hynning E, Janssens G & Glimelius L 2016 *Int. J. Particle Therapy: Summer 2016* **3(1)**, 71–276.

Giger A, Jud C, Nguyen D, Krieger M, Zhang Y, Lomax A J, Bieri O, Salomir R & Cattin P C 2019 *in* 'International Workshop on PRedictive Intelligence In MEdicine' Springer pp. 11–22.

Giger A, Sandkühler R, Jud C, Bauman G, Bieri O, Salomir R & Cattin P C 2018 *in* 'International Conference on Medical Image Computing and Computer-Assisted Intervention' Springer pp. 81–88.

Giger A, Stadelmann M, Preiswerk F, Jud C, De Luca V, Celicanin Z, Bieri O, Salomir R & Cattin P C 2018 *Physics in Medicine & Biology* **63**(14), 145015.

Gorgisyan J, Lomax A J, Munck af Rosenschöld P, Persson G F, Krieger M, Colvill E, Scherman J, Gagnon-Moisan F, Egloff M, Fattori G, Engelholm S A, Weber D C & Perrin R 2019 *Radiotherapy and Oncology* **134**, 135–142.

Grassberger C, Dowdell S, Sharp G & Paganetti H 2015 *Medical Physics* **42**(5), 2462–2469.

Hugo G D, Weiss E, Sleeman W C, Balik S, Keall P J, Lu J & Williamson J 2016 *Data from 4D Lung Imaging of NSCLC Patients. The Cancer Imaging Archive.*
URL: *http://doi.org/10.7937/K9/TCIA.2016.ELN8YGLE*

Hugo G D, Weiss E, Sleeman W C, Balik S, Keall P J, Lu J & Williamson J 2017 *Med. Phys.* **44**, 762–771.

Jud C, Nguyen D, Sandkühler R, Giger A, Bieri O & Cattin P C 2018 *in* 'International Conference on Medical Image Computing and Computer-Assisted Intervention' Springer pp. 198–205.

Kang Y, Zhang X, Chang J Y, Wang H, Wei X, Liao Z et al. 2005 *Int. J. Radiation Oncology Biol. Phys.* **67**, 906–914.

Krieger M, Giger A, Weber D C, Lomax A J & Zhang Y 2020 *Radiotherapy and Oncology* **145**, 154–161.

Krieger M, Klimpki G, Fattori G, Hrbacek J, Oxley D, Safai S, Weber D C, Lomax A J & Zhang Y 2018 *Physics in Medicine & Biology* **63**(5), 055005.

Li Y, Kardar L, Li H, Cao W, Chang J Y et al. 2014 *Med. Phys.* **41(2)**, 021721.

McClelland J R, Hawkes D J, Schaeffter T & King A P 2013 *Medical image analysis* **17**(1), 19–42.

Meijers A, Jakobi A, Sttzer K, Guterres Marmitt G, Both S, Langendijk J A, Richter C & Knopf A 2019 *Medical Physics* **46**(3), 1140–1149.

Mostafaei F, Tai A, Gore E, Johnstone C, Haase W, Ehlers C, Cooper D T, Lachaine M & Li X A 2018 *Medical physics* **45**(10), 4619–4626.

Paganetti H, Jiang H & Trofimov A 2005 *Phys. Med. Biol.* **50**, 983–990.

Pedroni E, Bearpark R, Böhringer T, Coray A, Duppich J, Forss S, George D, Grossmann M, Goitein G, Hilbes C, Jermann M, Lin S, Lomax A, Negrazus M, Schippers M & Kotrle G 2004 *Zeitschrift für Medizinische Physik* **14**, 25–34.

Phillips M H, Pedroni E, Blattmann H, Boehringer T, Coray A & Scheib S 1992 *Physics in Medicine & Biology* **37**(1), 223.

Preiswerk F, De Luca V, Arnold P, Celicanin Z, Petrusca L, Tanner C, Bieri O, Salomir R & Cattin P C 2014 *Medical image analysis* **18**(5), 740–751.

Richter D, Schwarzkopf A, Trautmann J, Kraemer M, Durante M, Jaekel O & Bert C 2013 *Med. Phys.* **40(5)**, 051722–1 – 051722–17.

Roman N O, Shepherd W, Mukhopadhyay N, Hugo G D & Weiss E 2012 *Int. J. Radiation Oncology Biol. Phys.* **83**, 1566–1572.

Sandkühler R, Jud C, Andermatt S & Cattin P C 2018 *arXiv preprint arXiv:1806.09907* .

Santini F, Gui L, Lorton O, Guillemin P C, Manasseh G, Roth M, Bieri O, Vallée J P, Salomir R & Crowe L A 2020 *Physica Medica* **70**, 161–168.

Trnková P, Knäusl B, Actis O, Bert C, Biegun A K, Boehlen T T, Furtado H, McClelland J, Mori S, Rinaldi I et al. 2018 *Physica Medica* **54**, 121–130.

Vezhnevets V & Konouchine V 2005 *in* 'proc. of Graphicon' Vol. 1 Citeseer pp. 150–156.

von Siebenthal M, Székely G, Gamper U, Boesiger P, Lomax A & Cattin P 2007 *Physics in Medicine & Biology* **52**(6), 1547.

Williams C K & Rasmussen C E 2006 *Gaussian processes for machine learning* Vol. 2 MIT press Cambridge, MA.

Zhang Y, Huth I, Weber D C & Lomax A J 2018 *Radiotherapy and Oncology* **128**(1), 182–188.

Zhang Y, Huth I, Wegner M, Weber D C & Lomax A J 2016 *Radiotherapy and oncology* **121**(2), 281–287.

Zhang Y, Knopf A C, Weber D C & Lomax A J 2015 *Physics in Medicine & Biology* **60**(20), 8141.

Zhang Y, Knopf A, Tanner C, Boye D & Lomax A J 2013 *Physics in Medicine & Biology* **58**(24), 8621.

Zhang Y, Knopf A, Tanner C & Lomax A 2014 *Physics in Medicine & Biology* **59**(24), 7793.