Physics in Medicine & Biology

IPEM Institute of Physics and Engineering in Medicine

PAPER • OPEN ACCESS

Limitations of phase-sorting based pencil beam scanned 4D proton dose calculations under irregular motion

To cite this article: A Duetschler *et al* 2023 *Phys. Med. Biol.* **68** 015015

IOP Publishing Phys. Med. Biol. 68 (2023) 015015 https://doi.org/10.1088/1361-6560/aca9b6

Physics in Medicine & Biology

IPEM Institute of Physics and Engineering in Medicine

RECEIVED
12 September 2022

REVISED
11 November 2022

ACCEPTED FOR PUBLICATION
7 December 2022

PUBLISHED
26 December 2022

PAPER

Limitations of phase-sorting based pencil beam scanned 4D proton dose calculations under irregular motion

A Duetschler[1,2,*], **J Prendi**[1,3], **S Safai**[1], **D C Weber**[1,4,5], **A J Lomax**[1,2] and **Ye Zhang**[1,*]

[1] Center for Proton Therapy, Paul Scherrer Institute, 5232 Villigen PSI, CH, Switzerland
[2] Department of Physics, ETH Zürich, 8092 Zürich, CH, Switzerland
[3] Department of Physics, University of Basel, 4056 Basel, CH, Switzerland
[4] Department of Radiation Oncology, University Hospital of Zürich, 8091 Zürich, CH, Switzerland
[5] Department of Radiation Oncology, Inselspital, Bern University Hospital, University of Bern, 3010 Bern, CH, Switzerland
[*] Authors to whom any correspondence should be addressed.

E-mail: alisha.duetschler@psi.ch **and** ye.zhang@psi.ch

Keywords: 4D dose calculation, motion irregularity, intra-fraction motion, proton therapy, lung tumor, liver tumors

Abstract

Objective. 4D dose calculation (4DDC) for pencil beam scanned (PBS) proton therapy is typically based on phase-sorting of individual pencil beams onto phases of a single breathing cycle 4DCT. Understanding the dosimetric limitations and uncertainties of this approach is essential, especially for the realistic treatment scenario with irregular free breathing motion. *Approach.* For three liver and three lung cancer patient CTs, the deformable multi-cycle motion from 4DMRIs was used to generate six synthetic 4DCT(MRI)s, providing irregular motion (11/15 cycles for liver/lung; tumor amplitudes ~4–18 mm). 4DDCs for two-field plans were performed, with the temporal resolution of the pencil beam delivery (4–200 ms) or with 8 phases per breathing cycle (500–1000 ms). For the phase-sorting approach, the tumor center motion was used to determine the phase assignment of each spot. The dose was calculated either using the full free breathing motion or individually repeating each single cycle. Additionally, the use of an irregular surrogate signal prior to 4DDC on a repeated cycle was simulated. The CTV volume with absolute dose differences >5% ($V_{\mathrm{dosediff}>5\%}$) and differences in CTV $V_{95\%}$ and $D_{5\%}$–$D_{95\%}$ compared to the free breathing scenario were evaluated. *Main results.* Compared to 4DDC considering the full free breathing motion with finer spot-wise temporal resolution, 4DDC based on a repeated single 4DCT resulted in $V_{\mathrm{dosediff}>5\%}$ of on average 34%, which resulted in an overestimation of $V_{95\%}$ up to 24%. However, surrogate based phase-sorting prior to 4DDC on a single cycle 4DCT, reduced the average $V_{\mathrm{dosediff}>5\%}$ to 16% (overestimation $V_{95\%}$ up to 19%). The 4DDC results were greatly influenced by the choice of reference cycle ($V_{\mathrm{dosediff}>5\%}$ up to 55%) and differences due to temporal resolution were much smaller ($V_{\mathrm{dosediff}>5\%}$ up to 10%). *Significance.* It is important to properly consider motion irregularity in 4D dosimetric evaluations of PBS proton treatments, as 4DDC based on a single 4DCT can lead to an underestimation of motion effects.

1. Introduction

Motion is a major challenge for radiotherapy delivery. Particularly respiratory-induced motion can lead to large displacements of tumors in the thorax and abdomen, which can result in highly inhomogeneous dose distributions for pencil beam scanned (PBS) proton therapy, due to the interplay effect with the dynamic beam delivery (Phillips *et al* 1992, Bert *et al* 2008). 4D dose calculations (4DDCs) are therefore an essential tool to understand and evaluate these dose deteriorating effects, such that an appropriate and sufficient motion management strategy can be determined at the treatment planning stage. Indeed, 4DDCs have repeatedly shown the effect of motion on planned dose distributions (Bert *et al* 2008, Seco *et al* 2009, Zhang *et al* 2012, Grassberger *et al* 2013, Ammazzalorso and Jelen 2014, Zou *et al* 2014, Dolde *et al* 2019, Meijers *et al* 2020), while the

effectiveness and efficiency of different motion mitigation approaches, such as rescanning, beam gating and tumor tracking have also been assessed (Li *et al* 2006, Bert and Rietzel 2007, Knopf *et al* 2011, Grassberger *et al* 2015, Zhang *et al* 2015, 2016, Ishihara *et al* 2017, Engwall *et al* 2018, Dolde *et al* 2019). In addition, 4DDCs are clearly indispensable for 4D optimization (Keall *et al* 2006, West *et al* 2007, Eley *et al* 2014, Graeff 2014, Bernatowicz *et al* 2017, Graeff 2017, Engwall *et al* 2018). Different 4DDC approaches have also been experimentally validated (Bert *et al* 2009, 2012, Testa *et al* 2013, Krieger *et al* 2018, Pfeiler *et al* 2018, Kostiukhina *et al* 2020, Spautz *et al* 2022).

As such, some form of 4DDC analysis is increasingly being recommended for quality assurance of 4D plans in a number of clinical trials (Lane *et al* 2014, Liao *et al* 2015) and has been defined as a standard QA tool in a recent report on proton treatments in the thorax (Chang *et al* 2017).

An essential component of 4DDC is a model of the (deformable) motion of the tumor and surrounding anatomy. For this, 4D computed tomography (4DCT) is the current standard, which typically represents an averaged breathing cycle, consisting of multiple (usually 8 or 10) bins/phases acquired over multiple actual breathing cycles. Imaging over longer time periods, for instance to capture motion variability, is hindered by concerns about imaging dose delivered to the patient. Although information about irregularities in the breathing period can be obtained using real-time monitoring of a surrogate signal (e.g. surface motion) acquired during treatment delivery (Richter *et al* 2013, 2014, Batista *et al* 2018, Meijers *et al* 2019, 2020), variabilities in motion amplitude, non-regular anatomical deformations and base-line shifts may be missed, as the 4DDC is still based on deformable motion from a single cycle 4DCT.

As such, and in order to also consider variable motion amplitudes and baseline shifts, Kraus *et al* (2011) and Wölfelschneider *et al* (2017) proposed scaling the deformation vector fields (DVFs) extracted from a single 4DCT. Indeed, as a part of their study Kraus *et al* (2011) investigated the effect of intra- and inter-fractional breathing changes for lung cancer patients and found pronounced dosimetric differences due to variations in period, amplitude, initial phase and baseline shifts in comparison to single 4DCTs. A simple linear scaling approach could, however, result in unrealistic deformations and does not account for deformable motion changes (e.g. deformable organ drift (Von Siebenthal *et al* 2007)). The impact on the treatment plan quality of varying amplitudes, periods and baseline shifts, assuming an ideal $\sin^4$ diaphragm motion, have also been investigated by Steinsberger *et al* (2021) using the numerical XCAT phantom (Segars *et al* 2008, 2010).

Alternatively, motion can be imaged with non-radiation based time-resolved imaging techniques such as 4DMRI (Von Siebenthal *et al* 2007, Jud *et al* 2018). 4DMRI captures the irregularity of respiratory motion, but lacks accurate density information necessary for dose calculations. Although significant work has been performed to enable MR only radiotherapy treatment planning for photon therapy (Edmund and Nyholm 2017), such methods are currently not accurate enough for proton therapy. As such, the use of synthetic and multiple cycle 4DCTs, whereby 4DMRI motion is mapped onto a CT image has been proposed (Boye *et al* 2013, Meschini *et al* 2019, Duetschler *et al* 2022), an approach referred to as 4DCT(MRI).

4DDC algorithms can then be performed using two main approaches: 'phase-sorting' or 'deforming dose grid' (DDG).

Phase-sorting approaches calculate multiple dose contributions, each calculated on a single phase of a 4DCT, which are then accumulated onto a single reference phase (Bert *et al* 2008, Grassberger *et al* 2013, Li *et al* 2014). For this, the delivery time of each proton pencil beam is first either pre-calculated (Bert *et al* 2008, Grassberger *et al* 2013) or extracted from a delivery log-file (Meijers *et al* 2019, 2020) and assigned to the closest phase of the input 4DCT. Intuitively, the accuracy of this approach is limited by the temporal resolution of the original 4DCT data (typically around 0.5 s), which is much coarser than time intervals between successive single pencil beams (4–200 ms).

For this reason, the DDG algorithm (Boye *et al* 2013, Krieger *et al* 2018, Zhang *et al* 2019) has been proposed. In contrast to the phase-sorting approach, this deforms the dose calculation grid according to DVFs extracted from the 4D image data set using deformable image registration (DIR). As described in Zhang *et al* (2019), with this approach, the motion induced DVFs can be interpolated to an arbitrarily fine temporal resolution, thus enabling 4DDC at a finer temporal resolution. The impact of this on 4D dose distributions has previously been studied in Zhang *et al* (2019) for liver cases, whilst the impact of temporal resolution generally on 4DDC has been studied from a theoretical aspect by Seo *et al* (2017).

Given the importance of 4DDC in the treatment of moving tumors with PBS proton therapy, understanding the associated uncertainties and limitations of commonly used 4DDC algorithms is essential, especially in the context of realistic motion scenarios where inter- and intra-fraction irregularities will generally be the norm. As such, in this study, we use previously defined multiple breathing cycle 4DCT(MRI) data sets (Zhang *et al* 2016, Duetschler *et al* 2022) to study the dosimetric impact of temporal resolution and intra-fractional breathing irregularity for PBS proton therapy. In short, we aim to:

(i) Expand on previous analysis on the importance of temporal resolution (Zhang *et al* 2019) by including more 4DCT data sets and the more challenging treatment site of lung.

(ii) Investigate how phase-sorting 4DDC performs under conditions of irregular free breathing.

(iii) Investigate the usefulness of partially incorporating irregularity in the breathing period for phase-sorting 4DDC, e.g. using a surrogate breathing signal.

2. Materials and methods

We first describe the 4DCT(MRI) numerical phantom data providing multiple breathing cycle 4D image data in section 2.1. We then introduce the DDG 4DDC algorithm in section 2.2 and how it was used to simulate different phase-sorting 4DDC scenarios in 2.3. The treatment planning and 4DDC parameters are described in section 2.4. Finally, in section 2.5 the methods for comparing the dose distributions resulting from the different 4DDC scenarios are described.

2.1. 4DCT(MRI) numerical phantom data

Clinical 3DCTs of three liver (unresectable hepatocellular cancer and cholangiocarcinoma) and three advanced-stage non-small-cell lung cancer patients (Josipovic *et al* 2016) provide the anatomy and density information for proton dose calculations. The 3DCTs for the liver patients were acquired in end exhalation (EE), whereas the lung 3DCTs were acquired in deep-inspiration breath-hold (DIBH). In order to study motion irregularity, the 3DCTs were animated with deformable motion extracted from multiple breathing cycle 4DMRIs from two volunteers to generate so-called synthetic 4DCT(MRI)s (liver: Boye *et al* (2013), Zhang *et al* (2016), lung: Duetschler *et al* (2022)). A navigator-based slice-stacking approach (Von Siebenthal *et al* 2007) was used for the liver 4DMRI, while for the lung 4DMRI a motion-aware acquisition technique based on a spoiled gradient echo sequence was used (Jud *et al* 2018). The volunteer 4DMRI used to generate the three liver 4DCT(MRI)s consists of 11 distinct breathing cycles with a temporal resolution of 2.77 Hz, whereas a 4DMRI with 15 breathing cycles with a temporal resolution of 2.25 Hz was utilized for the three lung 4DCT(MRI)s. Both 4DMRI motion patterns start in an EE phase. An illustration of the two free breathing motion patterns and the resulting superior-interior (SI) tumor center motion for one example liver (Liver1) and one lung (Lung1) case is depicted in figure 1(a). The motion variability in both period and amplitude among different cycles is summarized in figure 1(b), where differences in amplitude between the CT anatomies arise due to different tumor locations.

The clinical target volume (CTV) of the reference CT was extended by 5 mm in SI and anterior-posterior (AP) and 4 mm in lateral direction to obtain the planning target volume (PTV) (Josipovic *et al* 2016, Gorgisyan *et al* 2017). Using the 4DCT(MRI) DVFs, the contours of the PTV were then propagated to all other phases of the first breathing cycle and combined to form a geometric internal target volume (gITV).

2.2. PBS 4DDC

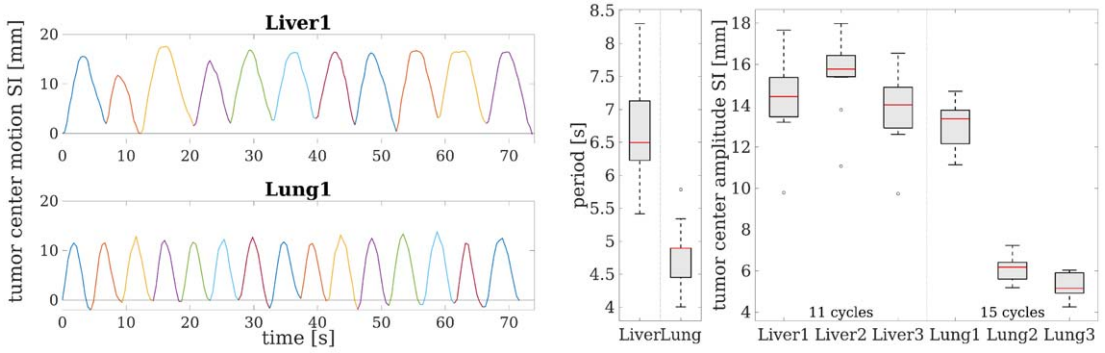
For 4DDC, the spots are ordered to determine the delivery sequence, which is used to estimate the delivery time t_{spot} of each pencil beam. The delivery times t_{spot} are calculated considering the time for lateral scanning of the beam, energy switching and dose delivery. Further, both deformations and density changes have to be taken into account. These are provided by 4D images, from which DVFs can be extracted using DIR.

Conventional 4DDC approaches assign each spot to one phase of the pre-treatment 4DCT based on the spot delivery time and the dose is calculated separately on each phase. The dose is then accumulated on one reference phase by warping the phase-wise dose contributions according to DVFs between the reference and all other phases. Such a phase-sorting (PS) approach is therefore restricted by the temporal resolution Δt_{image} of the 4DCT images (typically around 0.5 s dependent on breathing period and number of phases).

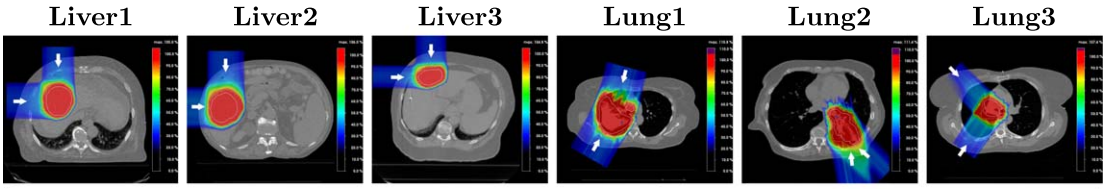
The DDG algorithm (Boye *et al* 2013, Krieger *et al* 2018, Zhang *et al* 2019) deforms the dose calculation grid as a function of time according to the DVFs. Linear interpolation of the DVFs between successive phases, previously extracted from the 4D images using DIR, allows a 4DDC at any given temporal resolution. More details can be found in appendix A and Zhang *et al* (2019). It is therefore possible to use the delivery time t_{spot} of single pencil beams for a spot-wise (SW) calculation (typically 4–200 ms between spots).

2.3. Simulation of phase-sorting 4DDC under different motion scenarios

By extracting the SI motion of the tumor center, the breathing patterns were separated into individual breathing cycles (starting with EE). We performed 4DDC either considering the full multiple cycle free breathing motion ('free') or repeating each single cycle individually ('rep'). In the latter scenarios, each single breathing cycle 4DCT



(a) Motion Liver1 and Lung1 (b) Motion period and amplitude



CTV volume [ml]					
164	260	61	398	152	89

(c) Static dose distributions and CTV volumes

Figure 1. Top: tumor center SI motion for Liver1 and Lung1 (a) and boxplot of periods (b) of the two volunteer MRI motion patterns used for liver and lung and resulting amplitudes extracted from the different liver and lung 4DCT(MRI)s. Bottom: axial slice showing the static dose distribution for all six patient cases with white arrows indicating the two field directions. The CTV (full) and gITV (dashed) are marked in white and their volumes, as contoured on the reference CT, are listed.

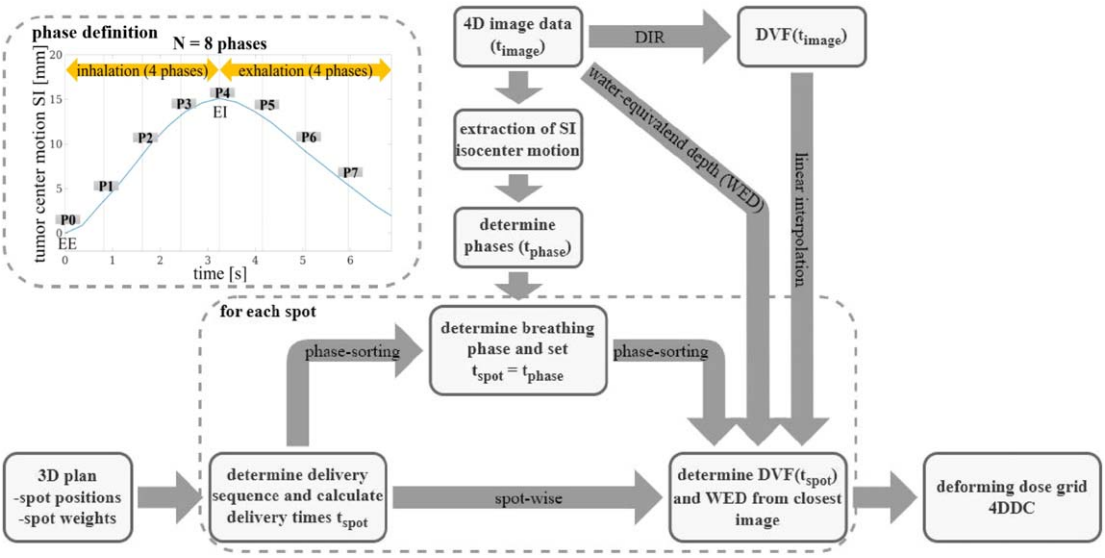


Figure 2. Schematic representation of the workflow for phase-sorting and spot-wise 4DDC. Top left: definition of 8 phases (P0–P7) for a single cycle. First end exhalation (EE) and end inhalation (EI) are determined to split cycle into inhalation and exhalation. Inhalation and exhalation are then both equally divided.

can be considered as a possible pre-treatment 4DCT in a clinical setup, providing no information on motion irregularity.

For spot-wise 4DDCs, the estimated spot delivery times were directly used either by repeating a single cycle (SW_{rep}) or for the multiple cycle free breathing motion (SW_{free}). For phase-sorting 4DDCs, each breathing cycle was first divided into 8 phases with an equal number of phases in inhalation and exhalation (see figure 2 with

Table 1. Summary of 4DDC scenarios. 4DDCs performed in this study either rely on phase-sorting (PS) of spots or they are performed in a spot-wise (SW) manner. The PS and the subsequent 4DDC either rely on a single repeated reference cycle ('rep') or use the full multiple cycle free breathing pattern ('free'). Alternatively, the PS can also rely on a surrogate breathing signal ('surr').

	Acronym	Temporal resolution 4DDC	Cycles PS	Cycles 4DDC
Repeated motion	**SW**$_{rep}$	pencil beam delivery times	—	single cycle
	PS$_{rep}$	8 phases per cycle	single cycle	single cycle
Free breathing	**SW**$_{free}$	pencil beam delivery times	—	multiple cycles
	PS$_{free}$	8 phases per cycle	multiple cycles	multiple cycles
Surrogate motion	**PS**$_{surr}$	8 phases per cycle	multiple cycles	single cycle

$N = 8$ phases). Each spot was assigned to the phase occurring at time t_{phase} closest to the spot delivery time t_{spot}. Phase-sorting 4DDCs were also performed for repeated single cycles (PS_{rep}) and free breathing (PS_{free}).

Besides the four above scenarios, we also simulated the retrospective 4DDC scenario based on surrogate motion traces (similar to the approach by Meijers *et al* (2019, 2020)), referred here to as PS_{surr}. In this scenario, the SI tumor center breathing trace of the full free breathing pattern extracted from each 4DCT(MRI) was used as a surrogate for the phase-sorting, whereas the 4DDC was only conducted on a repeated single reference cycle 4DCT. Although these surrogate motion traces provide information on motion variability during the treatment, for 4DDC only the variations in the breathing period are considered, while amplitude changes and baseline drifts are neglected as the patient geometry is modeled by a single cycle 4DCT.

The complete workflow for phase-sorting and spot-wise 4DDCs is illustrated in figure 2 and the different scenarios are summarized in table 1. It is worth mentioning here that, the PS_{rep} scenario is the classical approach for prospectively estimating 4D dose distribution at the pre-treatment stage after obtaining the patient-specific 4DCT.

2.4. Treatment planning and 4DDC parameters

Single-field uniform-dose (SFUD) plans with two field directions were optimized on the reference 3DCT (liver: EE, lung: DIBH) (see bottom of figure 1). A prescription dose of 2 Gy_{RBE} to the gITV (1 Gy_{RBE} per field) was used. 4DDCs were performed using a dose calculation grid of 4 mm × 4 mm (lateral) × 2.5 mm (distal). The DDG 4DDC algorithm relies on the ray-casting algorithm (Schaffner *et al* 1999) and a beam model of PSI-Gantry2 (Pedroni *et al* 2004, Zenklusen *et al* 2010, Safai *et al* 2012) was used with a 4 mm spot spacing (lateral and distal). Based on the delivery dynamics of PSI-Gantry2, lateral scanning of the beam was assumed to take 3.5 ms and 4 ms for the 'fast' (aligned with SI axis of patient) and 'slow' scanning direction, respectively. An energy switching time of 80 ms was adopted. The irradiation time of each spot assumed a constant beam current but varies slightly as a function of the beam energy mainly due to transmission losses in the beamline (Zhang *et al* 2019).

For the results presented in this paper, the beam delivery was assumed to start at the beginning of the first breathing cycle (EE starting phase). On the other hand, we also explicitly studied the impact of the starting phase, simulating a delivery starting during each phase of the first cycle. The results are presented in appendix B. In addition to the scenario without any motion mitigation, the results of simulations for 8-times volumetric rescanning are presented in appendix C. This scenario, representative of an extreme scenario of what can be achieved with rescanning, was introduced to study whether the dosimetric limitations of phase-sorting based 4DDC are still significant when motion effects are mitigated.

2.5. Dose evaluation and comparison

We first investigated the impact of the temporal resolution of the 4DDC. The results serve as a baseline for further results in this paper and the reader is referred to Zhang *et al* (2019) for a more detailed analysis. The impact of the temporal resolution was separately investigated for the repeated motions, through comparison of PS_{rep} to SW_{rep}, and for free breathing motion, by comparing PS_{free} to SW_{free}. In our comparisons, we then focused on the impact of motion variability by comparing the results of 4DDC with regular and irregular motion. Considering the free breathing scenario SW_{free} to be most representative of the actual 4D dose, all 4D dose distributions were further compared to this scenario. As such, by comparing simulations of PS_{rep} to SW_{free} the accuracy of phase-sorting based on a single cycle 4DCT under conditions of breathing irregularity were investigated. In this context, the accuracy of using a breathing surrogate for phase-sorting to partially incorporate motion irregularity was further studied through the comparison of PS_{surr} to SW_{free}. Lastly, we evaluated the dosimetric differences when different reference motion cycles were used as pre-treatment 4DCT for both PS_{rep} and PS_{surr}.

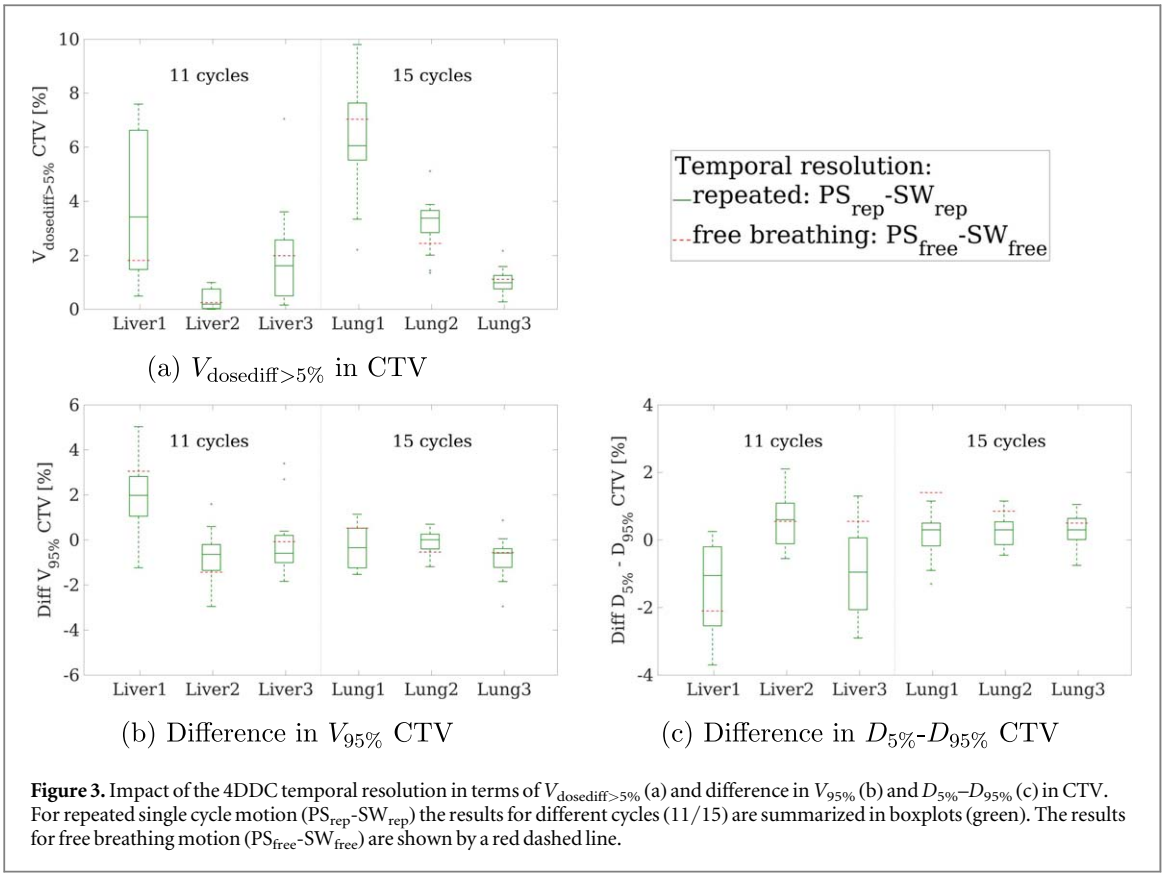


Figure 3. Impact of the 4DDC temporal resolution in terms of $V_{\text{dosediff}>5\%}$ (a) and difference in $V_{95\%}$ (b) and $D_{5\%}$–$D_{95\%}$ (c) in CTV. For repeated single cycle motion (PS_{rep}-SW_{rep}) the results for different cycles (11/15) are summarized in boxplots (green). The results for free breathing motion (PS_{free}-SW_{free}) are shown by a red dashed line.

The resulting 4D dose distributions were compared by looking at dose differences. The percentage of voxels in the CTV with an absolute point dose difference larger than 5% of the prescribed dose ($V_{\text{dosediff}>5\%}$) was calculated. To quantify the effects on the plan quality, dose volume histograms (DVHs) for the CTV and surrounding liver/lung were used. Furthermore, we focused on the analysis of $D_{5\%}$–$D_{95\%}$ in the CTV as a measure of the dose homogeneity in the tumor as well as CTV $V_{95\%}$ for tumor coverage and differences to the free breathing scenario SW_{free} were calculated (PS_{rep}-SW_{rep} for impact of temporal resolution for repeated motion). For organs at risk (OAR), the mean dose to the healthy surrounding liver or lung was assessed.

3. Results

3.1. Impact of temporal resolution

3.1.1. Repeated motion

Figure 3(a) shows the percentage of voxels in the CTV with an absolute dose difference larger than 5% ($V_{\text{dosediff}>5\%}$) due to the different 4DDC temporal resolution of spot-wise and phase-sorting 4DDC approaches (i.e. PS_{rep}-SW_{rep} and PS_{free}-SW_{free}). The results simulating regular breathing for different single breathing cycles as reference 4DCT are presented as green boxplots. The temporal resolution difference between PS_{rep} and SW_{rep} can result in $V_{\text{dosediff}>5\%}$ up to 10% (mean: 2.8%), highly depending on the reference cycle. The resulting impact of the temporal resolution on the dose coverage and homogeneity in the CTV was found to be less pronounced with differences in $V_{95\%}$ and $D_{5\%}$–$D_{95\%}$ in the CTV typically around ±1% (no more than 5%) (see figures 3(b) and (c)). Furthermore, almost no differences in the mean dose to the surrounding healthy liver/lung due to the different temporal resolution can be observed (differences below 0.15% of the prescribed dose).

3.1.2. Free breathing motion

The resulting differences due to the temporal resolution for irregular free breathing motion (PS_{free}-SW_{free}) are shown by red dashed lines in figure 3. The temporal resolution results in $V_{\text{dosediff}>5\%}$ of on average 2.7% and differences in CTV $V_{95\%}$ and $D_{5\%}$–$D_{95\%}$ also around ±1%. Overall, the effect of the temporal resolution is very similar for repeated and free breathing motion.

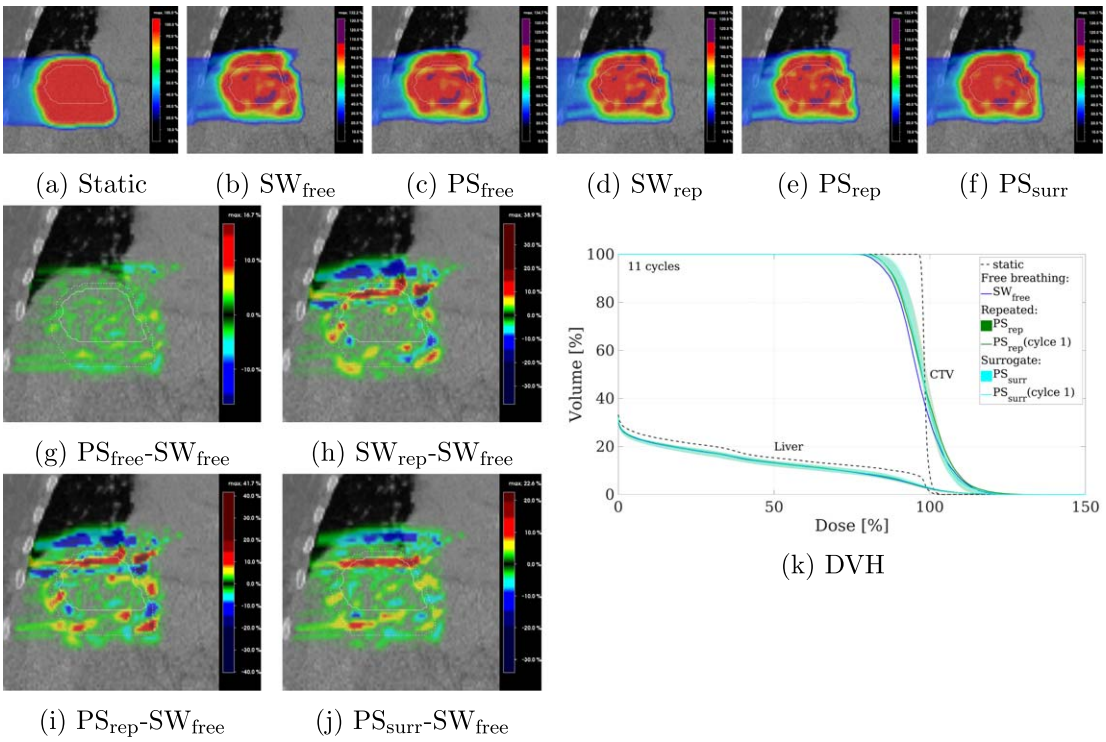


Figure 4. Dose (a)–(f) and dose differences (g)–(j) for Liver1. Static, SW_{free}, PS_{free}, SW_{rep}, PS_{rep} and PS_{surr} 4D dose distributions and dose differences to SW_{free} are shown. The first cycle was selected for SW_{rep}, PS_{rep} and PS_{surr}. All dose distributions/differences are visualized using the same colormap and the maximum dose (difference) value is indicated at the top of each colorbar. The CTV (full) and gITV (dashed) are marked in white. DVH (k) for static (black-dashed), free breathing SW_{free} (blue), repeated PS_{rep} (green) and surrogate motion PS_{surr} (cyan). For PS_{rep} and PS_{surr} the DVH for cycle 1 (line) and a band for the different cycles is shown.

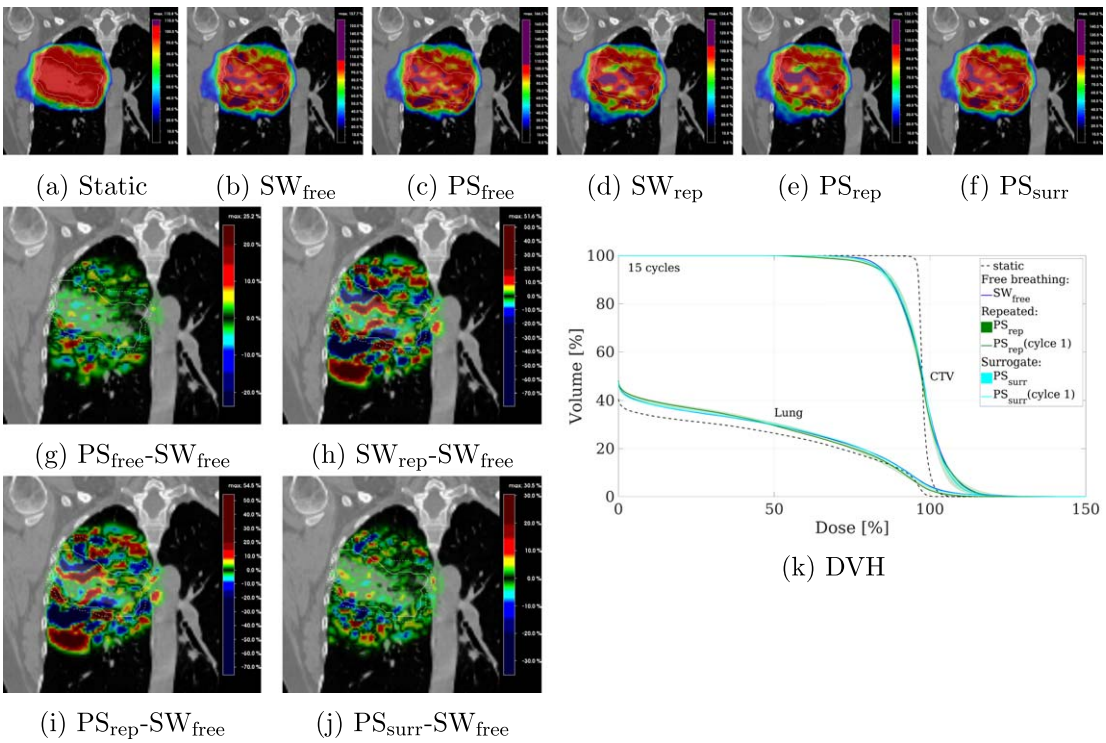
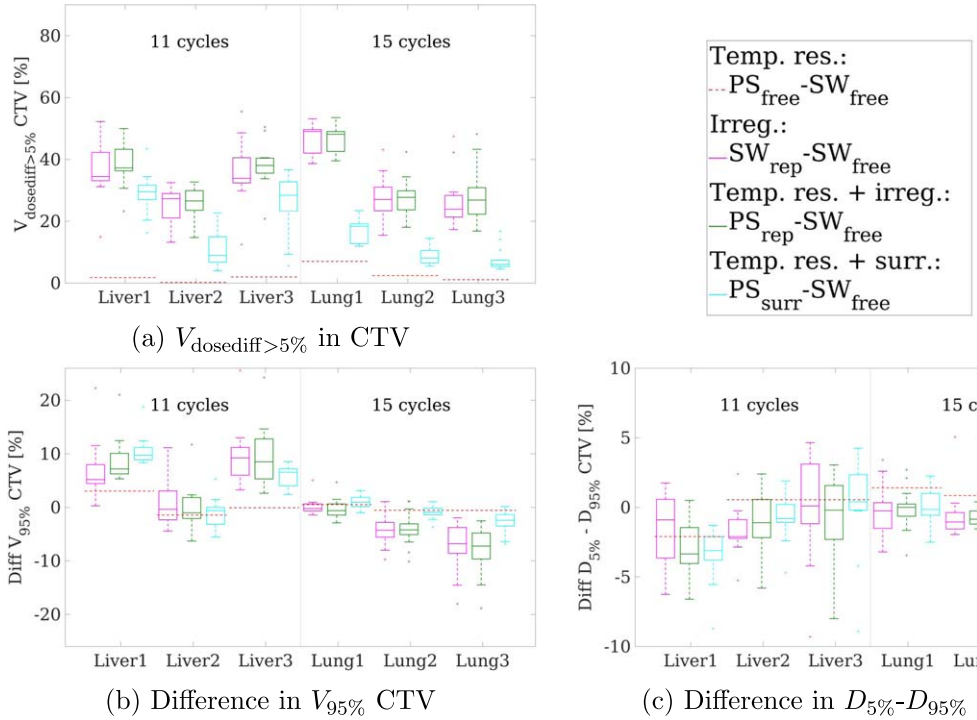


Figure 5. Dose (a)–(f) and dose differences (g)–(j) for Lung1. Static, SW_{free}, PS_{free}, SW_{rep}, PS_{rep} and PS_{surr} 4D dose distributions and dose differences to SW_{free} are shown. The first cycle was selected for SW_{rep}, PS_{rep} and PS_{surr}. The CTV (full) and gITV (dashed) are marked in white. DVH (k) for static (black-dashed), free breathing SW_{free} (blue), repeated PS_{rep} (green) and surrogate motion PS_{surr} (cyan). For PS_{rep} and PS_{surr} the DVH for cycle 1 (line) and a band for the different cycles is shown.



(a) $V_{\text{dosediff}>5\%}$ in CTV

(b) Difference in $V_{95\%}$ CTV

(c) Difference in $D_{5\%}$-$D_{95\%}$ CTV

Figure 6. Impact of motion irregularity in terms of $V_{\text{dosediff}>5\%}$ (a) and difference in $V_{95\%}$ (b) and $D_{5\%}$–$D_{95\%}$ (c) in CTV compared to SW_{free}. Differences due to motion irregularity (SW_{rep}-SW_{free} and PS_{rep}-SW_{free}) and with a surrogate (PS_{surr}-SW_{free}) for different cycles (11/15) are summarized in boxplots (magenta, green and cyan, respectively). The differences due to temporal resolution for free breathing motion (PS_{free}-SW_{free}) are shown by a red dashed line.

3.2. Impact of motion irregularity

3.2.1. Repeated versus free breathing motion

Static and 4D dose distributions for two example cases (Liver1 and Lung1) considering varying degrees of motion irregularity and different temporal resolutions are shown in figures 4 and 5. Dose differences to the SW_{free} scenario, considering motion irregularity at the higher spot-wise temporal resolution, are also displayed. Only small dose differences are visible due to the different temporal resolution for PS_{free}-SW_{free} (figures 4(g) and 5(g)). Larger dose differences can be observed for PS_{rep}-SW_{free} (and SW_{rep}-SW_{free}) due to motion irregularity.

These dose differences are quantified in terms of $V_{\text{dosediff}>5\%}$ in the CTV in figure 6(a). The average $V_{\text{dosediff}>5\%}$ value over the different cases and breathing cycles of 34.1% for PS_{rep}-SW_{free} (green boxplots) due to motion irregularity is much larger than the impact of the temporal resolution (average $V_{\text{dosediff}>5\%}$ of 2.7% for PS_{free}-SW_{free}, red dashed lines). Comparable $V_{\text{dosediff}>5\%}$ values can be observed for the liver and lung cases.

DVHs for Liver1 and Lung1 are displayed in figures 4(k) and 5(k), showing the results for free breathing (SW_{free}) as blue bands. The DVH for phase-sorting using a repeated cycle (PS_{rep}) is shown in green using lines for the first breathing cycle and a band for the different cycles. For Liver1, only small differences in the dose to the healthy surrounding liver occur, while a clear overestimation of the CTV coverage for PS_{rep} compared to SW_{free} can be observed. The DVH for Lung1 shows only small differences for the different scenarios.

Figures 6(b) and (c) show the differences in CTV coverage ($V_{95\%}$) and homogeneity ($D_{5\%}$–$D_{95\%}$) compared to the free breathing scenario SW_{free} for all six cases. The CTV coverage and homogeneity can both be over- or underestimated when neglecting the irregular motion. Depending on the choice of reference cycle, $V_{95\%}$ can be overestimated by up to 24% and $D_{5\%}$–$D_{95\%}$ underestimated by as much as 8% for PS_{rep} compared to SW_{free}. No substantial differences in the mean dose to the healthy surrounding liver/lung could be observed for the studied 4DDCs.

The results here demonstrate that the effect of irregular motion is much larger than the effect of the temporal resolution.

3.2.2. Usefulness of surrogate signal for phase-sorting

Having established the impact of not considering motion irregularity in the 4DDC in the previous section, the effectiveness of using a surrogate signal for phase-sorting applied to a single 4DCT, such as to incorporate variability in the breathing period, is investigated next.

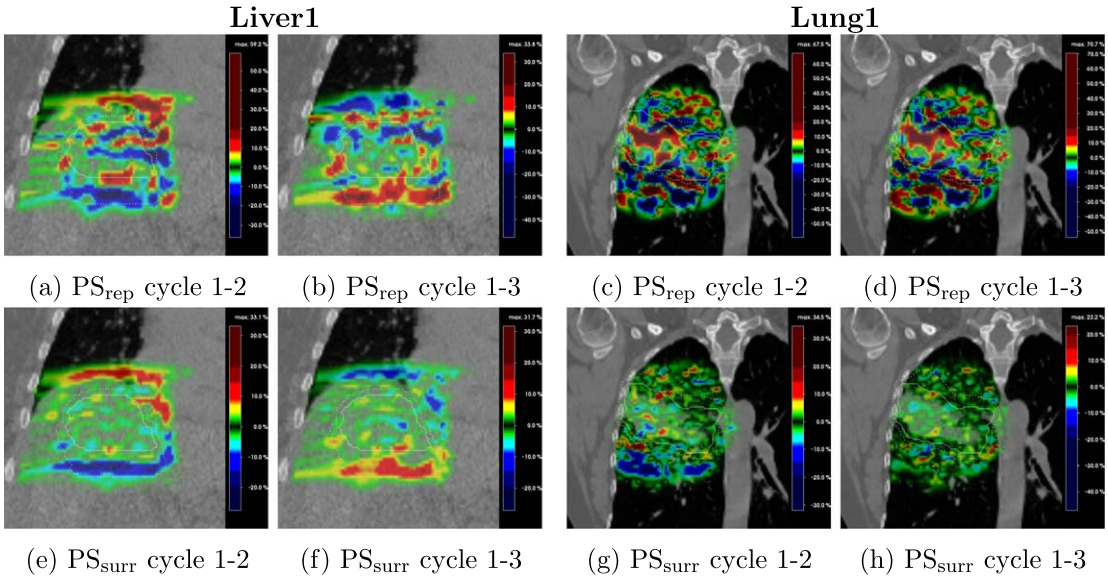


Figure 7. Dose differences for Liver1 (left) and Lung1 (right) due to different reference cycles for PS_{rep} and PS_{surr} 4DDCs. Dose differences to cycle 1 for cycle 2 and cycle 3 are shown for repeated PS_{rep} and surrogate motion PS_{surr}.

Dose differences between the scenario considering free breathing motion (SW_{free}) or their estimation using a surrogate signal (PS_{surr}) are presented in figures 4(j) and 5(j). These dose differences are reduced compared to the scenario assuming regular repeated breathing (figures 4(i) and 5(i)). For example, for Lung1 the maximum point dose difference to SW_{free} of 75.5% for PS_{rep} is reduced to 35.1% for PS_{surr}. Moreover, this reduction of the dose differences (w.r.t. SW_{free}) for PS_{surr} (cyan) and PS_{free} (red) is also shown by $V_{dosediff>5\%}$ in figure 6(a) for all six studied patient cases. The average $V_{dosediff>5\%}$ value of 34.1% for PS_{rep}-SW_{free} (green) is roughly halved to 16.4% for PS_{surr}-SW_{free} (cyan). The use of a surrogate signal to incorporate irregularity is more effective for the three lung cases, where $V_{dosediff>5\%}$ for PS_{surr}-SW_{free} is on average around 11% smaller than for the studied liver cases. As visible in figure 1, the motion used for the lung cases exhibits smaller variations in the breathing amplitude, which is more appropriate for surrogate based phase-sorting.

The DVHs for the Liver1 and Lung1 cases are also displayed in figures 4(k) and 5(k). Also with a surrogate signal, the CTV dose coverage for Liver1 is overestimated compared to SW_{free}. The differences in $V_{95\%}$ and $D_{5\%}$–$D_{95\%}$ for all six cases are shown in figures 6(b) and (c). While the dose differences (PS_{surr}-SW_{free} (cyan) and PS_{rep}-SW_{free} (green) in figure 6(a)) are clearly reduced by using a surrogate, similar $V_{95\%}$ and $D_{5\%}$–$D_{95\%}$ differences are reached. The maximum difference in $V_{95\%}$, however, is reduced for PS_{surr}-SW_{free} (18.8%) compared to PS_{rep}-SW_{free} (24.3%).

3.2.3. Impact of choice of reference cycle

In clinical practice, the reference 4DCT used for 4DDC, i.e. that used for PS_{rep}/SW_{rep} based calculations, is essentially arbitrary, being taken at some time during the planning CT session. As such any single breathing cycle 4DCT data set could be the reference. Figure 7 therefore shows dose differences due to the choice of reference cycle for repeated motion without (PS_{rep}) and with surrogate (PS_{surr}). As expected, for Liver1 case, different motion amplitudes of the reference 4DCT can lead to increased or reduced dose below and above the tumor, which can result in overdosage of nearby OARs. Smaller differences, especially also within the CTV, occur for PS_{surr} compared to PS_{rep}. For the Lung1 case, the inter-cycle motion variations are less pronounced, but large density heterogeneities in the high dose region can introduce pronounced but less localized dose differences. $V_{dosediff>5\%}$ in the CTV for different reference cycles compared to calculations using cycle 1 as reference is shown in figure 8 for PS_{rep} and PS_{surr}. For the repeated breathing scenario PS_{rep}, the use of a different reference than cycle 1 can result in more than half of the CTV with a dose difference greater than 5% (average/max $V_{dosediff>5\%}$ of 34%/56%). The impact of the reference cycle is reduced by the use of a surrogate for phase-sorting PS_{surr}, resulting in $V_{dosediff>5\%}$ values smaller than 35% (average $V_{dosediff>5\%}$ of 11%).

The dosimetric impact is visible in figures 6(b) and (c) with differences in $V_{95\%}$ and $D_{5\%}$–$D_{95\%}$ up to 22% and 14%, respectively, for different reference cycles. The variations in $V_{95\%}$ due to different reference cycles (i.e. spread of boxplots) are roughly halved for PS_{surr} (cyan) compared to SW_{rep} (magenta) and PS_{rep} (green). The variations both in $V_{95\%}$ and $D_{5\%}$–$D_{95\%}$ are smaller for the three lung cases compared to the liver cases, which show larger inter-cycle breathing variations (figure 1).

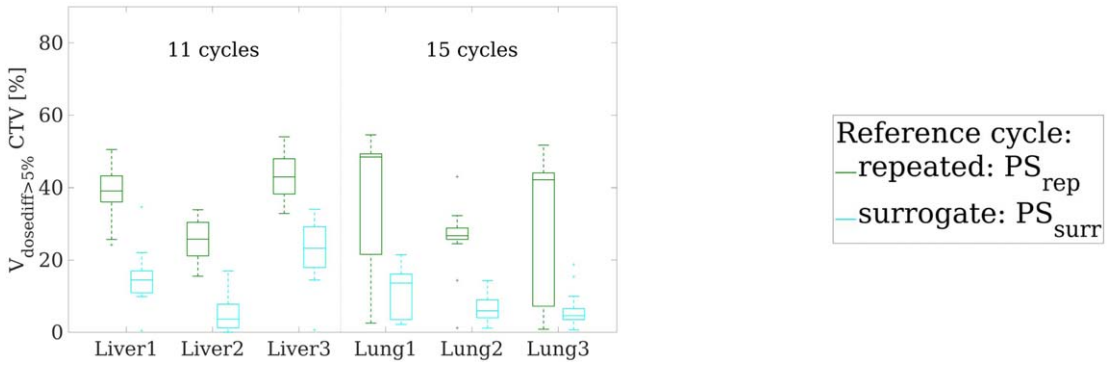


Figure 8. $V_{dosediff>5\%}$ due to impact of reference cycle for PS_{rep} (green) and PS_{surr} (cyan). Dose differences to cycle 1 were calculated for all other cycles.

4. Discussion

In clinical routine, 4DDC is typically based on a single 4DCT. A 4DCT, usually acquired before the start of the treatment, represents only one average breathing cycle, while in reality breathing motion can vary substantially both intra- and inter-fractionally. The delivery time of each proton pencil beam, either extracted from a log-file or estimated based on known machine parameters, is used to assign every pencil beam to a phase of the 4DCT. At the treatment planning stage, the phase-sorting is typically done assuming regular motion (i.e. repetition of the 4DCT motion) for a prospective 4DDC of one or multiple scenarios. Additionally, retrospective 4D dose reconstruction based on phase-sorting using a breathing motion surrogate acquired during the treatment can also be performed. While the prospective 4DDC based on a single pre-treatment 4DCT does not consider breathing motion irregularity at all, the retrospective surrogate based 4D dose reconstruction only accounts for changes in the breathing period. The goal of this comprehensive simulation study was therefore to investigate the limitations of these phase-sorting based 4DDC approaches and quantify the associated dosimetric uncertainties in the presence of irregular breathing motion for PBS proton therapy and evaluate the usefulness of a surrogate signal for retrospective phase-sorting. For our work, we have used the deforming dose grid algorithm and machine parameters of PSI-Gantry2. The investigation of variable patient CT geometries (3 liver, 3 lung), multiple breathing cycles (11 for liver, 15 for lung), different temporal resolutions (spot-wise, 8 phases), different starting phases and rescanning (VS1 and VS8) resulted in several thousand 4D dose distributions.

We took full advantage of previously developed synthetic 4DCT(MRI) data sets, which contain varied deformable motion with irregular respiratory patterns, which enable this study. These synthetic 4D images were obtained by combining the density information from CTs of cancer patients with the respiratory motion extracted from 4DMRIs of two healthy volunteers. While this provides realistic and plausible motion, the actual tumor motion would have been different for each patient and the motion of cancerous tissue could be different from healthy tissue. Further, one limitation of the used 4DCT(MRI)s is that their body surface and ribs remain static. A moving ribcage is being considered in an ongoing update of the numerical 4DCT(MRI) phantoms. As all our 4D dose comparisons, however, relied on the same data sets we believe our findings would also be transferable to real 4D patient data. Further, a dosimetric study in an anthropomorphic thorax phantom with a moving target did not show any significant difference with and without moving ribs (Lebbink *et al* 2022).

For the results presented in this paper, we focused on phase-sorting with 8 phases per breathing cycle. However, all phase-sorting 4DDCs were additionally also performed for 10 phases. Similar results were obtained for phase-sorting with 8 and 10 phases. In general, differences in dose and dose-volume indices due to different temporal resolutions of the 4DDCs were found to be much smaller than differences due to variable breathing motion or different delivery starting times. A more detailed investigation of the influence of the temporal resolution used in PBS proton 4DDC was performed by Zhang *et al* (2019). As also seen throughout our studies in this paper, they found that different 4DDC scenarios can result in pronounced local dose differences, which however have a less pronounced impact on dose-volume indices. Here, we focused only on the conventional fractionation scheme (2 Gy per fraction), more significant dosimetric differences due to the temporal resolution were observed for hypo-fractionated treatments (>4 Gy per fraction) and short breathing periods (<4 s), which were not investigated in the present work.

The interplay effect can lead to considerable local dose differences due to irregular motion, which are not captured with 4DDC based on a repeated single cycle 4DCT. For our studied cases, the percentage of voxels with local dose differences larger than 5% could on average be roughly halved using retrospective surrogate based

phase-sorting 4DDC (PS_{surr}-SW_{free}) compared to prospective phase-sorting 4DDC (PS_{rep}-SW_{free}) on a repeated pre-treatment 4DCT. Thus, the use of a motion surrogate for phase-sorting with a single cycle 4DCT for 4DDC was shown to be partially effective in reducing local dosimetric uncertainties due to motion irregularity (especially if rescanning is applied, as discussed in appendix C). The use of a motion surrogate for phase-sorting also made the 4D dose distributions slightly less dependent on the reference cycle, that is the 4DCT used for 4DDC. However, even with the use of a motion surrogate for phase-sorting (PS_{surr}) $V_{95\%}$ could be overestimated by as much as 18.8 % (8.1%) and $D_{5\%}$–$D_{95\%}$ underestimated by up to 8.9% (4.8%) without rescanning (with 8-times rescanning) compared to SW_{free}, which also considers changes in motion amplitude and base-line shifts. It is interesting to mention that all effects were more pronounced for local dose differences than for statistical dose-volume indices. In this study we focused on the mean dose to the liver/lung, which did not show any substantial differences for the different motion scenarios. On the other hand, the observed local dose differences could lead to more pronounced differences in e.g. the maximal dose to nearby OAR, which could be detrimental for serial OARs.

As presented in appendix B, large dose differences were also observed due to different starting phases of the delivery. For prospective 4DDCs (PS_{rep}), different starting phases resulted in $V_{dosediff>5\%}$ as high as 85%, with on average almost half of the CTV subject to dose differences larger than 5% (average $V_{dosediff>5\%}$ of 49%). Different starting phases also resulted in non-negligible differences in $V_{95\%}$ and $D_{5\%}$–$D_{95\%}$ up to 35% and 17%, respectively. As such, for the investigated cases, the dosimetric differences due to different starting phases were in fact larger than due to motion irregularity (PS_{rep}-SW_{free}). While fractionation might slightly average out these effects, we believe they should still be taken into consideration, especially also for hypo-fractionated treatments. Alternatively, the uncertainties related to the start of the treatment could easily be reduced by synchronizing the treatment delivery with a certain breathing phase. This demonstrated the large uncertainties associated with prospective 4DDC, in comparison to retrospective dose reconstruction based on log-files synchronized with surrogate motion traces.

As currently no online 3D motion information is available, the effectiveness of incorporating information from surrogate motion traces through phase-sorting 4DDC onto a single pre-treatment 4DCT, which only considers changes in the breathing period but not variations in amplitude or baseline shifts, has been discussed in this paper. Baseline shifts as well as inter-fractional anatomical variations could be incorporated using repeated 4DCTs, this, however, would still not consider intra-fractional variations in the breathing amplitude and leads to additional imaging dose to the patient. Another approach is to scale the deformation vector fields extracted from a 4DCT according to observed motion traces (Kraus *et al* 2011, Wölfelschneider *et al* 2017). However, this approach will not be able to accurately consider complicated deformable motion changes (e.g. deformable organ drift (Von Siebenthal *et al* 2007)). Lacking online 3D imaging, variations in the breathing pattern could potentially be represented most accurately by the use of a surrogate driven 3D deformable motion-model (Zhang *et al* 2007, Arnold *et al* 2011, Li *et al* 2011, Preiswerk *et al* 2012, McClelland *et al* 2013, Zhang *et al* 2013, Giger *et al* 2018, 2020, Krieger *et al* 2021). Combined with 4DDC based on synchronized treatment log-files, this could potentially reduce the dosimetric uncertainties related to irregular motion. We are currently studying the effectiveness of this approach.

The results presented in this paper once again showed the complexity of accurate 4DDC for PBS proton therapy and further investigations with different machine parameters, more patient data and diverse respiratory motion data are certainly necessary to get a better understanding of the dosimetric uncertainties associated with respiratory motion in PBS proton therapy. Further, the initial treatment planning decisions, such as the use of an average or mid-position planning CT (Kang *et al* 2007, Wang *et al* 2013), the target definition (Krieger *et al* 2020) or an intensity modulated planning approach with inhomogeneous field doses, will also influence the 4DDC results and should be further investigated and compared.

5. Conclusion

The dosimetric impact of the temporal resolution of the 4DDC was found to be much smaller than dosimetric differences due to variations in the breathing motion. Understanding the effects of motion irregularity is essential for evaluating PBS proton therapy treatments in mobile regions of the body. 4DDC based on a single cycle 4DCT can lead to an underestimation of motion effects and is highly dependent on the acquired pre-treatment 4DCT. Incorporating information from an online measured breathing signal during the dose delivery can consider motion period irregularity for a phase-sorting based 4DDC approach, therefore reducing the dosimetric uncertainties. However, pronounced dosimetric differences still remain due to neglecting variations in breathing amplitude. With this approach, the starting phase of the delivery could also easily be recorded or monitored, which would reduce the considerable dosimetric uncertainties due to the typically unknown start of the treatment delivery.

Acknowledgments

This project is funded by the Swiss Cancer Research foundation (KFS-4517-08-2018). The authors thank Giovanni Fattori for his explanations on 4DCT acquisition and reconstruction. The authors further thank Arturs Meijers for reviewing the manuscript and his valuable comments.

Conflict of interest

The authors report no conflict of interest.

Author contributions

AD developed the workflow, implemented the different phase-sorting scenarios for the 4D dose calculation, performed the computations, analysed the results and wrote the manuscript. JP provided a first phase-sorting implementation and performed and analyzed initial calculations. YZ designed the study, drafted the manuscript and supervised the project as principal investigator together with AJL, SS and DCW. All authors have read, significantly revised and approved the final manuscript.

Appendix A. PBS proton 4D dose calculation

A.1. 3D dose calculation

Based on the initial energy $E_{0,n}$, spot position $(v_{0,n}, u_{0,n})$ and optimized beam weight $\omega_{0,n}$ of N pencil beams $(n \in \{1, 2,\ldots,N\})$ the 3D dose to a dose calculation grid point with coordinate (s, v, u) can be calculated as:

$$D_{3D}(s,\, v,\, u) = \sum_{n=1}^{N} \omega_{n,0}\mathrm{ID}_{0,n}(\mathrm{WER}_0)\frac{1}{2\pi\sigma_{0,u}\sigma_{0,v}}\exp\left(-\frac{(v_{0,n}-v)^2}{2\sigma_{0,v}^2}\right)\exp\left(-\frac{(u_{0,n}-u)^2}{2\sigma_{0,u}^2}\right)$$
$$\text{where } \mathrm{WER}_0 = \mathrm{WER}_0(s,\, v,\, u)$$
$$\text{and } \sigma_{0,i} = \sigma_{0,i}(n,\, \mathrm{WER}_0(s,\, v,\, u)) \text{ with } i = v,\, u. \tag{A1}$$

$\mathrm{ID}_{0,n}$ is the integral depth dose dependent on the initial beam energy and energy spectrum and WER_0 the water-equivalent range of the grid point (s, v, u). The dose of each pencil beam is modeled by a single Gaussian in both lateral directions with standard deviations of $\sigma_{0,u}$ and $\sigma_{0,v}$.

4D dose calculation

For the 4D dose calculation (4DDC) the continuous delivery time $\mathcal{T}$ is discretized into $\widehat{\mathcal{T}}$ (see Zhang *et al* (2019) for more details):

$$D_{4D}(s,\, v,\, u) = \int_{\mathcal{T}} d_t(s,\, v,\, u)\,dt$$
$$\approx \sum_{t\in\widehat{\mathcal{T}}} \omega_t \mathrm{ID}_t(\mathrm{WER}_t)\frac{1}{2\pi\sigma_u\sigma_v}\exp\left(-\frac{(v_{0,t}-(v+\Delta v_t))^2}{2\sigma_v^2}\right)$$
$$\times \exp\left(-\frac{(u_{0,t}-(u+\Delta u_t))^2}{2\sigma_u^2}\right)$$
$$\text{where } \mathrm{WER}_t = \mathrm{WER}_t(s,\, v,\, u)$$
$$\text{and } \sigma_i = \sigma_i(t,\, \mathrm{WER}_t(s,\, v,\, u)) \text{with } i = v,\, u. \tag{A2}$$

The displacements Δv_t and Δu_t orthogonal to the beam direction at each dose grid point (s, v, u) can be obtained through linear interpolation of the deformation vector fields obtained from deformable image registration of the 4D images. Density information is provided by the 4DCT images and to calculate the water-equivalent range $\mathrm{WER}_t(s, v, u)$ the closest 4DCT image is taken. Using the deforming dose grid algorithm (equation (A2)) an arbitrary temporal resolution Δt can be used in the discretization of the delivery time $\mathcal{T}$. A typical choice is to take the spot the delivery time t_{spot} of each pencil beam for the 4DDC, resulting in a spot-wise 4DDC.

Appendix B. Impact of starting phase of delivery

The start of the treatment is often not synchronized with the patient's breathing and the beam delivery of each field can start during an arbitrary breathing phase. However, for 4DDC the beam delivery of each field is often assumed to start at the beginning of the first breathing cycle (end exhalation (EE) in this work). In this section,

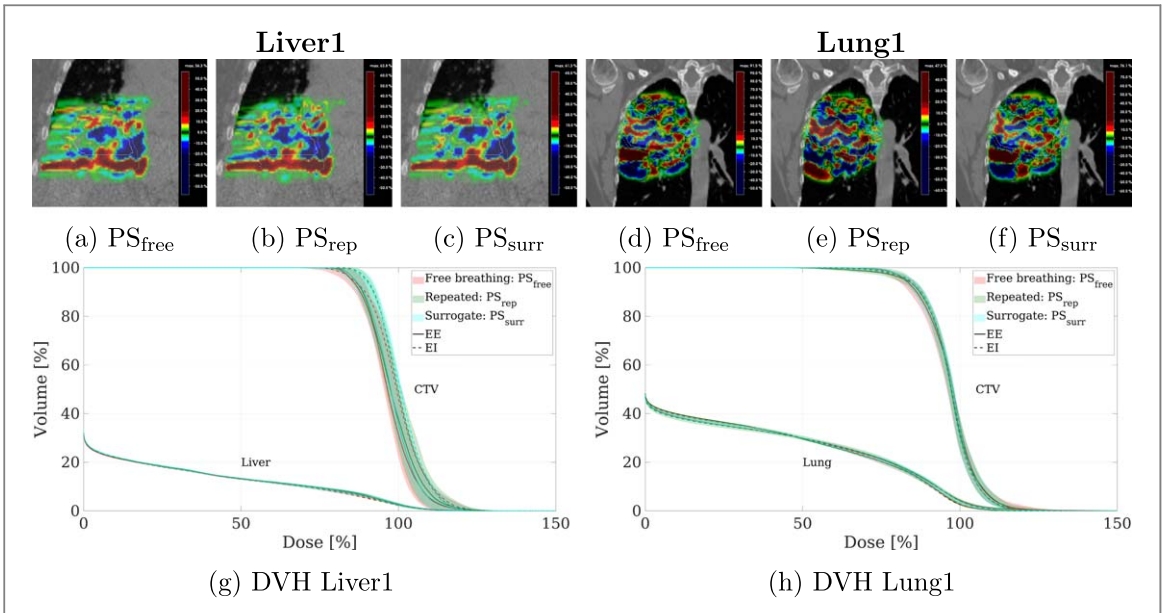


Figure B1. Top: dose differences for Liver1 (left) and Lung1 (right) due to end exhalation and end inhalation (EE-EI) starting phase are shown for PS_{free}, PS_{rep} and PS_{surr}. Bottom: DVHs for Liver1 (g) and Lung1 (h) for different starting phases for the two fields (band, 64 combinations). The solid and dashed lines show DVHs for EE and EI starting phases, respectively. Results for PS_{free} (red), PS_{rep} (green) and PS_{surr} (cyan) are shown.

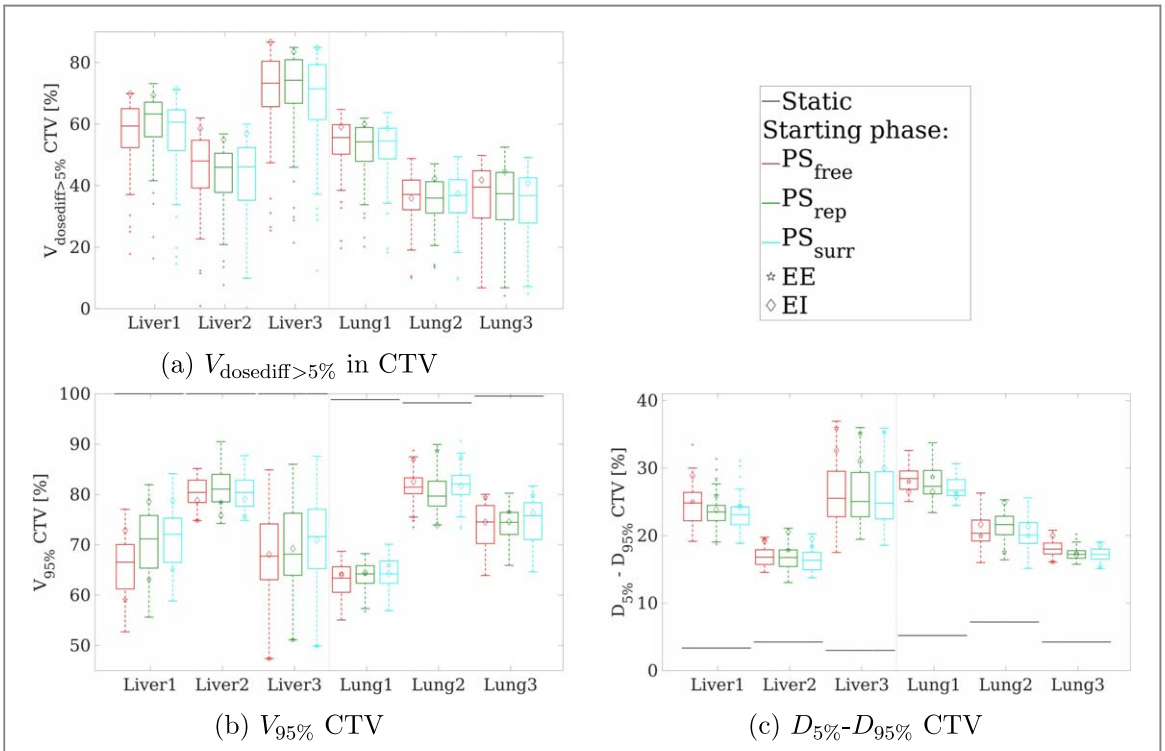


Figure B2. Impact of starting phases used for 4DDC (64 combinations) in terms of $V_{dosediff>5\%}$ for dose differences to EE starting (a), $V_{95\%}$ (b) and $D_{5\%}$–$D_{95\%}$ (c) in CTV. Results for free breathing PS_{free} (red), repeated motion PS_{rep} (green) and motion surrogate PS_{surr} (cyan) are shown. Results for EE and EI starting phases are marked by a star (*) and diamond (◊), respectively.

we, therefore, compare 4DDC results with an EE starting phase for each field to all other starting phase combinations. With 8 phases per breathing cycle, 64 combinations of starting phases for the two fields of each plan were simulated. For repeated motion PS_{rep} and breathing surrogate PS_{surr} 4DDCs were limited to calculations on the first breathing cycle. Dose differences between EE and end inhalation (EI) starting phase (for both fields) are displayed in figures B1(a)–(f), showing pronounced dose differences for all three motion scenarios. $V_{dosediff>5\%}$ in figure B2(a) shows absolute dose differences above 5% for on average almost half of the CTV due to a different starting phase than EE for all scenarios. The largest differences can be observed for the

opposite extreme EI starting phases. The dosimetric impact is visible in the DVHs in figure B1(g) for Liver1 and figure B1(h) for Lung1. Different starting phases can lead to differences up to 38% in $V_{95\%}$ (figure B2(b)) and 20% in $D_{5\%}$–$D_{95\%}$ (figure B2(c)) with similar $V_{\text{dosediff}>5\%}$ values for PS_{free}, PS_{rep} and PS_{surr}. The largest dose differences between EE and EI starting phase did not necessarily result in the largest differences in tumor coverage and homogeneity. The smaller dose differences for the liver patient did, however, result in slightly smaller differences in CTV $V_{95\%}$ and $D_{5\%}$–$D_{95\%}$. In general, the influence of the starting phase is very similar for the different 4DDC scenarios.

Appendix C. Simulations including rescanning

C.1. Impact of motion irregularity

We first study the impact of motion irregularity, when 8-times volumetric rescanning (VS8) is applied. As can be seen by comparing figure C1(a) to figure 6(a), $V_{\text{dosediff}>5\%}$ due to motion irregularity is reduced by the use of rescanning for PS_{rep}-SW_{free} (apart from Liver3). On average, $V_{\text{dosediff}>5\%}$ for PS_{rep}-SW_{free} is more than halved for VS8 compared to no rescanning (VS1) (see also figure 6(a) for VS1). However, a comparable impact of the reference cycle (spread of boxplots) on $V_{\text{dosediff}>5\%}$ for PS_{rep}-SW_{free} can be observed for VS1 and VS8.

However, by the use of a surrogate signal for phase-sorting when rescanning is applied, the dose differences between PS_{surr} and SW_{free} become almost negligible (see figure C1(a)) and the influence of the reference cycle becomes very small for PS_{surr}.

The impact of motion irregularity on the dose coverage $V_{95\%}$ and homogeneity $D_{5\%}$–$D_{95\%}$ is displayed in figures C1(b) and (c) for VS8 (see also figures 6(b) and (c) for VS1) and shows comparable differences as for VS1. Slightly larger differences between VS1 and VS8 occur for the liver cases. Differences in $V_{95\%}$ and $D_{5\%}$–$D_{95\%}$ due to different reference cycles are also similar for VS1 and VS8 and are reduced for PS_{surr} compared to PS_{rep}.

C.2. Impact of starting phase of delivery

A study of the impact of the starting phase on 4DDCs with 8-times rescanning shows reduced dose differences compared to the case without any motion mitigation. For example, the mean (max) $V_{\text{dosediff}>5\%}$ value for free breathing PS_{free} 4DDCs with a starting phase other than EE is reduced from 49.7% (86.7%) for VS1 (figure B2(a)) to 13.1% (58.9%) for VS8 (figure C2(a)). Rescanning also lead to a slight reduction in the dependency of CTV $V_{95\%}$ and $D_{5\%}$–$D_{95\%}$ on the starting phase (figures C2(b) and (c)). For example for PS_{free},

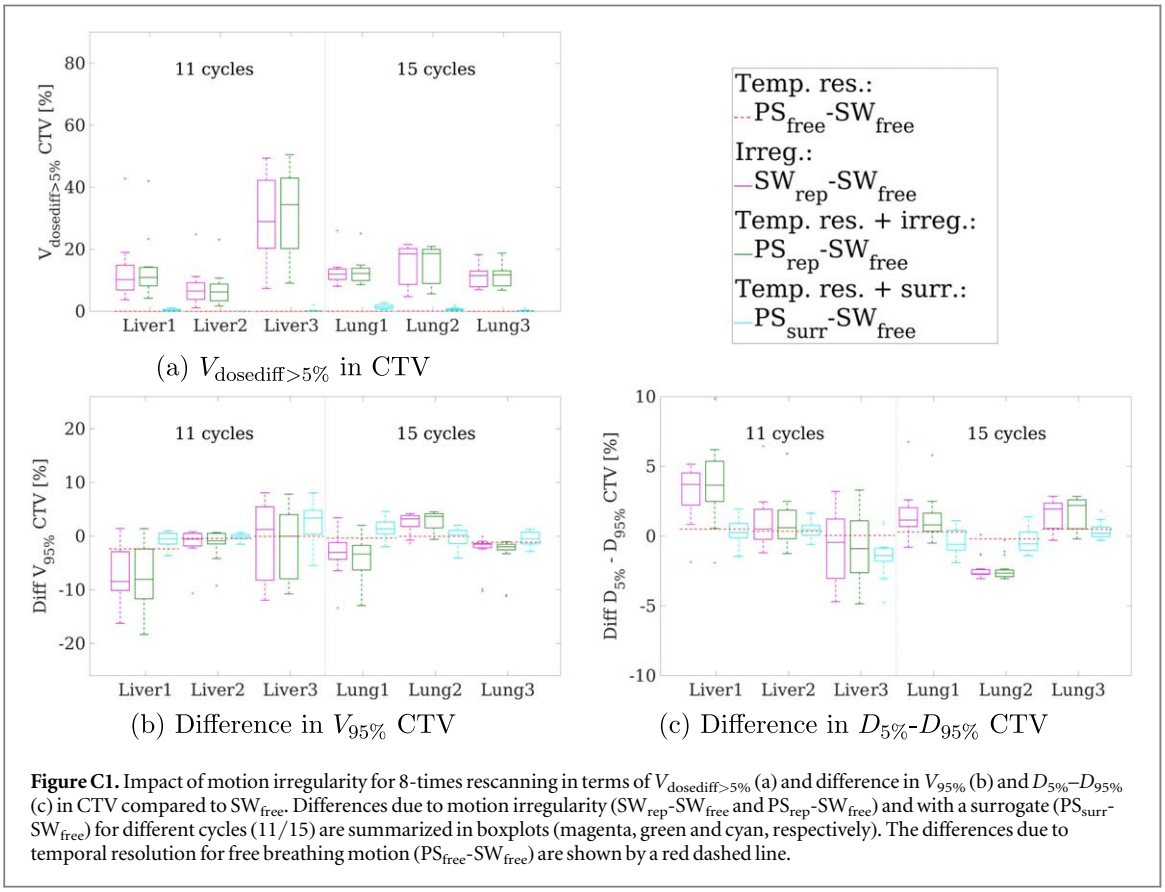


Figure C1. Impact of motion irregularity for 8-times rescanning in terms of $V_{\text{dosediff}>5\%}$ (a) and difference in $V_{95\%}$ (b) and $D_{5\%}$–$D_{95\%}$ (c) in CTV compared to SW_{free}. Differences due to motion irregularity (SW_{rep}-SW_{free} and PS_{rep}-SW_{free}) and with a surrogate (PS_{surr}-SW_{free}) for different cycles (11/15) are summarized in boxplots (magenta, green and cyan, respectively). The differences due to temporal resolution for free breathing motion (PS_{free}-SW_{free}) are shown by a red dashed line.

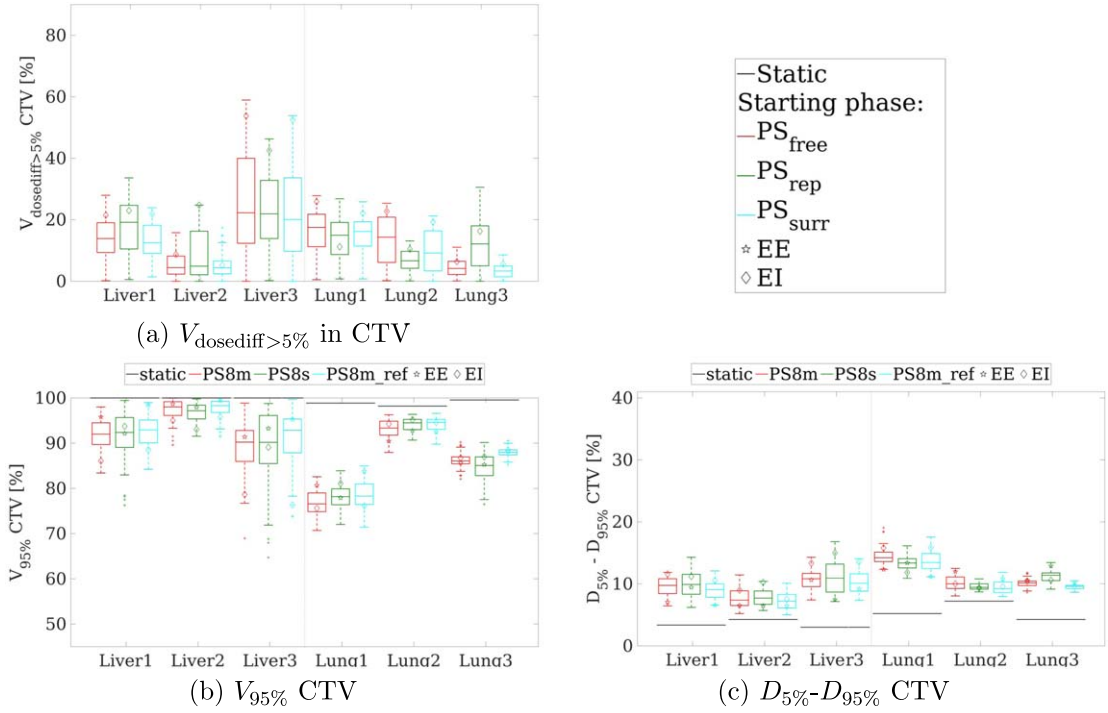


Figure C2. Impact of starting phases used for 4DDC (64 combinations) for 8-times rescanning in terms of $V_{\text{dosediff}>5\%}$ for dose differences to EE starting (a), $V_{95\%}$ (b) and $D_{5\%}$–$D_{95\%}$ (c) in CTV. Results for free breathing PS_{free} (red), repeated motion PS_{rep} (green) and motion surrogate PS_{surr} (cyan) are shown. Results for EE and EI starting phases are marked by a star (*) and diamond (◊), respectively.

differences in $D_{5\%}$–$D_{95\%}$ up to 20% occur for VS1, while for VS8 the differences due to different starting phases never exceed 7%. All phase-sorting 4DDC scenarios show a similar dependency on the starting phase and no reduction for PS_{surr} compared to the other scenarios could be observed.

ORCID iDs

A Duetschler https://orcid.org/0000-0001-5121-8032
Ye Zhang https://orcid.org/0000-0003-1608-4467

References

Ammazzalorso F and Jelen U 2014 A 4D dose computation method to investigate motion interplay effects in scanned ion beam prostate therapy *Phys. Med. Biol.* **59** N91

Arnold P, Preiswerk F, Fasel B, Salomir R, Scheffler K and Cattin P C 2011 3D organ motion prediction for MR-guided high intensity focused ultrasound *Lecture Notes in Computer Science,Medical Image Computing and Computer-Assisted Intervention – MICCAI 2011* (Berlin, Heidelberg: Springer) (*Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics*) vol 6892, 623–30

Batista V, Richter D, Chaudhri N, Naumann P, Herfarth K and Jäkel O 2018 Significance of intra-fractional motion for pancreatic patients treated with charged particles *Radiat. Oncol.* **13** 1–11

Bernatowicz K, Zhang Y, Perrin R, Weber D C and Lomax A J 2017 Advanced treatment planning using direct 4D optimisation for pencil-beam scanned particle therapy *Phys. Med. Biol.* **62** 6595–609

Bert C, Gemmel A, Saito N and Rietzel E 2009 Gated irradiation with scanned particle beams *Int. J. Radiat. Oncol. Biol. Phys.* **73** 1270–5

Bert C, Grözinger S O and Rietzel E 2008 Quantification of interplay effects of scanned particle beams and moving targets *Phys. Med. Biol.* **53** 2253–65

Bert C, Richter D, Durante M and Rietzel E 2012 Scanned carbon beam irradiation of moving films: Comparison of measured and calculated response *Radiat. Oncol.* **7** 1–13

Bert C and Rietzel E 2007 4D treatment planning for scanned ion beams *Radiat. Oncol.* **2** 1–10

Boye D, Lomax T and Knopf A 2013 Mapping motion from 4D-MRI to 3D-CT for use in 4D dose calculations: a technical feasibility study. *Med. Phys.* **40** 61702

Chang J Y *et al* 2017 Consensus guidelines for implementing pencil-beam scanning proton therapy for thoracic malignancies on behalf of the ptcog thoracic and lymphoma subcommittee *Int. J. Radiat. Oncol. Biol. Phys.* **99** 41–50

Dolde K, Zhang Y, Chaudhri N, Dávid C, Kachelrieß M, Lomax A J, Naumann P, Saito N, Weber D C and Pfaffenberger A 2019 4DMRI-based investigation on the interplay effect for pencil beam scanning proton therapy of pancreatic cancer patients *Radiat. Oncol.* **14** 30

Duetschler A *et al* 2022 Synthetic 4DCT(MRI) lung phantom generation for 4D radiotherapy and image guidance investigations *Med. Phys.* **49** 2890–903

Edmund J M and Nyholm T 2017 A review of substitute CT generation for MRI-only radiation therapy *Radiat. Oncol.* **12** 1–15

Eley J G, Newhauser W D, Lüchtenborg R, Graeff C and Bert C 2014 4D optimization of scanned ion beam tracking therapy for moving tumors *Phys. Med. Biol.* **59** 3431–52
Engwall E, Fredriksson A and Glimelius L 2018 4D robust optimization including uncertainties in time structures can reduce the interplay effect in proton pencil beam scanning radiation therapy *Med. Phys.* **45** 4020–9
Engwall E, Glimelius L and Hynning E 2018 Effectiveness of different rescanning techniques for scanned proton radiotherapy in lung cancer patients *Phys. Med. Biol.* **63** 9
Giger A, Sandkühler R, Jud C, Bauman G, Bieri O, Salomir R and Cattin P C 2018 Respiratory motion modelling using cGANs *Lect. Notes Comput. Sci.* **11073** 81–8 (including Subser. Lect. Notes Artif. Intell. Lect. Notes Bioinformatics)
Giger A *et al* 2020 Liver-ultrasound based motion modelling to estimate 4D dose distributions for lung tumours in scanned proton therapy *Phys. Med. Biol.* **65** 235050
Gorgisyan J, Munck af Rosenschold P, Perrin R, Persson G F, Josipovic M, Belosi M F, Engelholm S A, Weber D C and Lomax A J 2017 Feasibility of pencil beam scanned intensity modulated proton therapy in breath-hold for locally advanced non-small cell lung cancer *Int. J. Radiat. Oncol. Biol. Phys.* **99** 1121–8
Graeff C 2014 Motion mitigation in scanned ion beam therapy through 4D-optimization *Phys. Med.* **30** 570–7
Graeff C 2017 Robustness of 4D-optimized scanned carbon ion beam therapy against interfractional changes in lung cancer *Radiother. Oncol.* **122** 387–92
Grassberger C, Dowdell S, Lomax A, Sharp G, Shackleford J, Choi N, Willers H and Paganetti H 2013 Motion interplay as a function of patient parameters and spot size in spot scanning proton therapy for lung cancer *Int. J. Radiat. Oncol. Biol. Phys.* **86** 380–6
Grassberger C, Dowdell S, Sharp G and Paganetti H 2015 Motion mitigation for lung cancer patients treated with active scanning proton therapy *Med. Phys.* **42** 2462–9
Ishihara Y, Nakamura M, Miyabe Y, Mukumoto N, Matsuo Y, Sawada A, Kokubo M, Mizowaki T and Hiraoka M 2017 Development of a four-dimensional Monte Carlo dose calculation system for real-time tumor-tracking irradiation with a gimbaled X-ray head *Phys. Med.* **35** 59–65
Josipovic M, Persson G F, Dueck J, Bangsgaard J P, Westman G, Specht L and Aznar M C 2016 Geometric uncertainties in voluntary deep inspiration breath hold radiotherapy for locally advanced lung cancer *Radiotherapy and Oncology* **118** 510–4
Jud C, Nguyen D, Sandkühler R, Giger A, Bieri O and Cattin P C 2018 Motion aware MR imaging via spatial core correspondence *Lecture Notes in Computer Science* ed A F Frangi *et al* (Cham: LNCS, Springer Verlag) (*Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics*) 11 070, 198–205 11070
Kang Y *et al* 2007 4D Proton treatment planning strategy for mobile lung tumors *Int. J. Radiat. Oncol.* **67** 906–14
Keall P J *et al* 2006 The management of respiratory motion in radiation oncology report of AAPM Task Group 76 *Med. Phys.* **33** 3874–900
Knopf A C, Hong T S and Lomax A 2011 Scanned proton radiotherapy for mobile targets—the effectiveness of re-scanning in the context of different treatment planning approaches and for different motion characteristics *Phys. Med. Biol.* **56** 7257–71
Kostiukhina N, Palmans H, Stock M, Knopf A, Georg D and Kn usl B 2020 Time-resolved dosimetry for validation of 4D dose calculation in PBS proton therapy *Phys. Med. Biol.* **65** 125015
Kraus K M, Heath E and Oelfke U 2011 Dosimetric consequences of tumour motion due to respiration for a scanned proton beam *Phys. Med. Biol. Biol.* **56** 6563–81
Krieger M, Giger A, Salomir R, Bieri O, Celicanin Z, Cattin P C, Lomax A J, Weber D C and Zhang Y 2020 Impact of internal target volume definition for pencil beam scanned proton treatment planning in the presence of respiratory motion variability for lung cancer: a proof of concept *Radiother. Oncol.* **145** 154–61
Krieger M, Klimpki G, Fattori G, Hrbacek J, Oxley D, Safai S, Weber D C, Lomax A J and Zhang Y 2018 Experimental validation of a deforming grid 4D dose calculation for PBS proton therapy *Phys. Med. Biol.* **63** 055005
Krieger M *et al* 2021 Liver-ultrasound-guided lung tumour tracking for scanned proton therapy: a feasibility study *Phys. Med. Biol.* **66** 035011
Lane J A *et al* 2014 Active monitoring, radical prostatectomy, or radiotherapy for localised prostate cancer: study design and diagnostic and baseline results of the ProtecT randomised phase 3 trial *Lancet Oncol.* **15** 1109–18
Lebbink F, Stock M, Georg D and Knäusl B 2022 The influence of motion on the delivery accuracy when comparing actively scanned carbon ions *versus* protons at a synchrotron-based radiotherapy facility *Cancers* **14** 1788
Li R, Lewis J H, Jia X, Zhao T, Liu W, Wuenschel S, Lamb J, Yang D, Low D A and Jiang S B 2011 On a PCA-based lung motion model *Phys. Med. Biol.* **56** 6009–30
Li X A, Stepaniak C and Gore E 2006 Technical and dosimetric aspects of respiratory gating using a pressure-sensor motion monitoring system *Med. Phys.* **33** 145–54
Li Y *et al* 2014 On the interplay effects with proton scanning beams in stage III lung cancer *Med. Phys.* **41** 1–7
Liao Z *et al* 2015 RTOG 1308 protocol information: phase III randomized trial comparing overall survival after photon *versus* proton chemoradiotherapy for inoperable stage II-IIIB NSCLC (https://clinicaltrials.gov/ct2/show/NCT01993810)
McClelland J, Hawkes D, Schaeffter T and King A 2013 Respiratory motion models: a review *Med. Image Anal.* **17** 19–42
Meijers A, Jakobi A, Stützer K, Guterres Marmitt G, Both S, Langendijk J A, Richter C and Knopf A 2019 Log file-based dose reconstruction and accumulation for 4D adaptive pencil beam scanned proton therapy in a clinical treatment planning system: Implementation and proof-of-concept *Med. Phys.* **46** 1140–9
Meijers A, Knopf A-C, Crijns A P G, Ubbels J F, Niezink A G H, Langendijk J A, Wijsman R and Both S 2020 Evaluation of interplay and organ motion effects by means of 4D dose reconstruction and accumulation *Radiother. Oncol.* **150** 268–274
Meschini G *et al* 2019 Virtual 4DCT from 4DMRI for the management of respiratory motion in carbon ion therapy of abdominal tumors *Med. Phys.* **47** 909–16
Pedroni E *et al* 2004 The PSI Gantry 2: a second generation proton scanning gantry *Z. Med. Phys.* **14** 25–34
Pfeiler T, Bäumer C, Engwall E, Geismar D, Spaan B and Timmermann B 2018 Experimental validation of a 4D dose calculation routine for pencil beam scanning proton therapy *Z. Med. Phys.* **28** 121–33
Phillips M H, Pedroni E, Blattmann H, Boehringer T, Coray A and Scheib S 1992 Effects of respiratory motion on dose uniformity with a charged particle scanning method *Phys. Med. Biol.* **37** 223–34
Preiswerk F, Arnold P, Fasel B and Cattin P C 2012 Towards more precise, minimally-invasive tumour treatment under free breathing *Proc. Annu. Int. Conf. IEEE Eng. Med. Biol. Soc.* (EMBS) pp 3748–51
Richter D, Schwarzkopf A, Trautmann J, Krämer M, Durante M, Jäkel O and Bert C 2013 Upgrade and benchmarking of a 4D treatment planning system for scanned ion beam therapy *Med. Phys.* **40** 1–17
Richter D *et al* 2014 Four-dimensional patient dose reconstruction for scanned ion beam therapy of moving liver tumors *Int. J. Radiat. Oncol. Biol. Phys.* **89** 175–81

Safai S, Bula C, Meer D and Pedroni E 2012 Improving the precision and performance of proton pencil beam scanning *Transl. Cancer Res.* **1** 196–206
Schaffner B, Pedroni E and Lomax A 1999 Dose calculation models for proton treatment planning using a dynamic beam delivery system: an attempt to include density heterogeneity effects in the analytical dose calculation *Phys. Med. Biol.* **44** 27–41
Seco J, Robertson D, Trofimov A and Paganetti H 2009 Breathing interplay effects during proton beam scanning: simulation and statistical analysis *Phys. Med. Biol.* **54** N283
Segars W P, Mahesh M, Beck T J, Frey E C and Tsui B M 2008 Realistic CT simulation using the 4D XCAT phantom *Med. Phys.* **35** 3800–8
Segars W P, Sturgeon G, Mendonca S, Grimes J and Tsui B M 2010 4D XCAT phantom for multimodality imaging research *Med. Phys.* **37** 4902–15
Seo J, Han M C, Yeom S, Lee H S, Kim C H, Jeong J H and Kim S 2017 Temporal resolution required for accurate evaluation of the interplay effect in spot scanning proton therapy *J. Korean Phys. Soc.* **70** 720–5
Spautz S, Jakobi A, Meijers A, Peters N, Löck S, Knopf A-C, Troost E G, Richter C and Stützer K 2022 Experimental validation of 4d log file-based proton dose reconstruction for interplay assessment considering amplitude-sorted 4dcts *Med. Phys.* **49** 3538–49
Steinsberger T, Alliger C, Donetti M, Krämer M, Lis M, Paz A, Wolf M and Graeff C 2021 Extension of RBE-weighted 4D particle dose calculation for non-periodic motion *Phys. Med.* **91** 62–72
Testa M, Schümann J, Lu H M, Shin J, Faddegon B, Perl J and Paganetti H 2013 Experimental validation of the TOPAS Monte Carlo system for passive scattering proton therapy *Med. Phys.* **40** 1–16
Von Siebenthal M, Székely G, Gamper U, Boesiger P, Lomax A and Cattin P 2007 4D MR imaging of respiratory organ motion and its variability *Phys. Med. Biol.* **52** 1547–64
Wang N, Patyal B, Ghebremedhin A and Bush D 2013 Evaluation and comparison of New 4DCT based strategies for proton treatment planning for lung tumors *Radiat. Oncol.* **8** 1–10
West J B, Park J, Dooley J R and Maurer C R 2007 4D treatment optimization and planning for radiosurgery with respiratory motion tracking *Treating Tumors that Move with Respiration* ed H C Urschel *et al* (Berlin, Heidelberg: Springer Berlin Heidelberg) pp 249–64
Wölfelschneider J, Seregni M, Fassi A, Ziegler M, Baroni G, Fietkau R, Riboldi M and Bert C 2017 Examination of a deformable motion model for respiratory movements and 4D dose calculations using different driving surrogates *Med. Phys.* **44** 2066–76
Zenklusen S M, Pedroni E and Meer D 2010 A study on repainting strategies for treating moderately moving targets with proton pencil beam scanning at the new gantry 2 at PSI *Phys. Med. Biol.* **55** 5103–21
Zhang Q, Pevsner A, Hertanto A, Hu Y C, Rosenzweig K E, Ling C C and Mageras G S 2007 A patient-specific respiratory model of anatomical motion for radiation treatment planning *Med. Phys.* **34** 4772–81
Zhang Y, Boye D, Tanner C, Lomax A J and Knopf A 2012 Respiratory liver motion estimation and its effect on scanned proton beam therapy *Phys. Med. Biol.* **57** 1779–95
Zhang Y, Huth I, Weber D C and Lomax A J 2019 Dosimetric uncertainties as a result of temporal resolution in 4D dose calculations for PBS proton therapy *Phys. Med. Biol.* **64** 125005
Zhang Y, Huth I, Wegner M, Weber D C and Lomax A J 2016 An evaluation of rescanning technique for liver tumour treatments using a commercial PBS proton therapy system *Radiother. Oncol.* **121** 281–7
Zhang Y, Knopf A, Tanner C, Boye D and Lomax A J 2013 Deformable motion reconstruction for scanned proton beam therapy using on-line x-ray imaging *Phys. Med. Biol.* **58** 8621–45
Zhang Y, Knopf A C, Weber D C and Lomax A J 2015 Improving 4D plan quality for PBS-based liver tumour treatments by combining online image guided beam gating with rescanning *Phys. Med. Biol.* **60** 8141–59
Zou W *et al* 2014 Dynamic simulation of motion effects in IMAT lung SBRT *Radiat. Oncol.* **9** 1–9