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

Krieger, Miriam; Giger, Alina; Salomir, Rares Vincent; Bieri, Oliver; Celicanin, Zarko; Cattin, Philippe C; Lomax, Antony J; Weber, Damien C; Zhang, Ye

How to cite

KRIEGER, Miriam et al. 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. In: Radiotherapy and oncology, 2020, vol. 145, p. 154–161. doi: 10.1016/j.radonc.2019.12.001

This publication URL: https://archive-ouverte.unige.ch/unige:164603
Publication DOI: 10.1016/j.radonc.2019.12.001

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

Miriam Krieger[a,b], Alina Giger[c,d], Rares Salomir[e,f], Oliver Bieri[c,g], Zarko Celicanin[c,g,1], Philippe C Cattin[c,d], Antony J Lomax[a,b], Damien C Weber[a,h,j], Ye Zhang[a]

[a]Center for Proton Therapy, Paul Scherrer Institute (PSI), 5232 Villigen PSI, Switzerland

[b]Department of Physics, ETH Zurich, 8092 Zurich, Switzerland

[c]Department of Biomedical Engineering, University of Basel, 4123 Allschwil, Switzerland

[d]Center for medical Image Analysis & Navigation, University of Basel, 4123 Allschwil, Switzerland

[e]Image Guided Interventions Laboratory (949), Faculty of Medicine, University of Geneva, 1211 Geneva, Switzerland

[f]Radiology Division, University Hospitals of Geneva, 1205 Geneva, Switzerland

[g]Department of Radiology, Division of Radiological Physics, University Hospital Basel, 4031 Basel, Switzerland

[h]Department of Radiation Oncology, University Hospital Zurich, 8091 Zurich, Switzerland

[j]Department of Radiation Oncology, Inselspital Bern, 3010 Bern, Switzerland

[1]Current affiliation: Intuitive Therapeutics SA, 1025 Saint-Sulpice, Switzerland

Corresponding author:
Miriam Krieger
Paul Scherrer Institute
Forschungsstrasse 111
WBBB/105
5232 Villigen PSI
Switzerland
Phone: +41 56 310 58 14
E-mail: miriam.krieger@psi.ch

Keywords: proton therapy; motion variability; pencil beam scanning; probabilistic planning; internal target volume; lung cancer

Highlights:

- Conventional 4DCT based ITV volumes vary substantially in presence of respiration variability
- 4D plan quality (target coverage and dose homogeneity) strongly depends on selected 4D images which plan is optimised on
- Variable respiration from 4DMRI can be considered in the treatment planning process as Probabilistic ITVs.
- Probabilistic ITVs significantly improve 4D plan quality without increasing the healthy tissue dose

Acknowledgements

This work was funded by the Swiss National Science Foundation (SNF) grant No. 320030_163330/1.

Abstract

Purpose: Motion management is crucial in scanned proton therapy for mobile tumours. Current motion mitigation approaches rely on single 4DCTs before treatment, ignoring respiratory variability. We investigate the consequences of respiratory variations on internal target volumes (ITV) definition and motion mitigation efficacy, and propose a probabilistic ITV based on 4DMRI.

Materials and Methods: Four 4DCT(MRI) datasets, each containing 40 variable cycles of synthetic 4DCTs, were generated by warping single-phase CTs of two lung patients with motion fields extracted from two 4DMRI datasets. Two-field proton treatment plans were optimised on ITVs based on different parts of the 4DCT(MRI)s. 4D dose distributions were calculated by considering variable respiratory patterns. Different probabilistic ITVs were created by incorporating the voxels covered by the CTV in at least 25%, 50%, or 75% (ITV25, ITV50, ITV75) of the cycles, and compared with the conservative ITV encompassing all possible CTV positions.

Results: Depending on the selected planning 4DCT, ITV volumes vary up to 20%, resulting in significant variation in CTV coverage for 4D treatments. Target coverage and homogeneity improved with the conservative ITV, but was associated with significantly increased lung dose (~1%). ITV25 and ITV50 led to acceptable plan quality in most cases without lung dose increments. ITV75 best minimised lung dose, but was insufficient to ensure coverage under all motion scenarios.

Conclusion: Irregular respiration significantly affects CTV coverage when ITVs are only defined by single 4DCTs. A probabilistic ITV50 provides an adequate compromise between target coverage and lung dose for most motion and patient scenarios investigated.

Introduction

Respiratory motion is a challenge for pencil beam scanned (PBS) proton therapy for lung tumours. The physical advantages of protons can be compromised because of geometrical target misses at the field edge (due to out of field tumour movements) and interplay effects (due to the interferences between tumour motion and sequential dynamic PBS delivery). Both can lead to unacceptable dose distributions if motion mitigation is not adopted during treatment delivery [1]. Many mitigation approaches have been proposed and applied clinically [2,3], including breath-hold [4-7], gating [8],tumour tracking [9,10] or 4D optimisation [11]. All these, to a greater or lesser extent, rely on some form of online motion monitoring, plan adaptation and patient cooperation. An alternative approach is rescanning, which can be applied during free breathing treatments. For this, the treatment delivery is divided into a number of rescans, delivered consecutively, such that dose inhomogeneities are statistically smeared out [12]. Its efficacy strongly depends on both machine parameters and the characteristics of the motion pattern [13-15]. However, expanding the target volume is a pre-requisite in order to avoid insufficient dose coverage at the target periphery due to edge-of-field smoothing. Consequently, appropriate definition of the internal target volume (ITV) is important for PBS treatments to mobile tumours delivered using rescanning.

Based on the recommendations in ICRU 78 [16], an ITV needs to encompass all possible positions of the clinical target volume (CTV) during the *beam-on* time. However, in current clinical practice, the ITV is typically defined considering motions extracted from 4DCT, which provides information on the averaged motion over the duration of the 4D-image acquisition. Clearly, this does not necessarily reflect the respiratory situation at treatment or any potential respiratory variabilities. The efficacy of an ITV to best mitigate motion-induced dose corruption when defined from a single 4DCT is therefore questionable.

In this study, we aim to explore the efficacy of the current ITV approach for PBS based lung tumour treatments under conditions of respiratory variability, and to investigate alternative ITV definitions based on breathing variability estimated via pre-treatment 4D imaging. For this, motion data sets for multiple breathing cycles have been acquired using 4DMRI [17], and combined with single phase CTs to generate multiple, simulated 4DCT data sets (4DCT(MRI), see e.g. [18]). Based on these, we propose and validate a probabilistic ITV definition, which takes into consideration motion variability during the treatment planning process.

Materials and methods

Data acquisition

Motion data from volunteer 4DMRI studies

Ten minutes of free-breathing 4DMRIs were acquired on two volunteers using a Siemens Aera 1.5T scanner. Both volunteers signed an informed consent according to the local IRB regulations. A navigator-based, retrospectively sorted, 4DMRI approach [19] was adapted and optimised for lung imaging [17], and allows to capture variable respiratory patterns in contrast to conventional

4DCT [20]. This provided ~80 complete and variable respiratory cycles for each volunteer, consisting of over 700 motion states per volunteer at a temporal resolution of 0.6 s.

Density data from patient CT

For patient specific densities and tumour geometries, full-exhale CTs of two lung cancer patients were used (shown in Figure 1a). Case 1 is a small tumour (CTV volume: 18.8cm^3) located in the lower part of the right lung lobe, unattached to anatomical structures. Case 2 was a larger right lower lobe tumour (CTV size: 141.9cm^3) located next to the spine, extracted from the Cancer Imaging Archive [21-25]. These cases were deliberately chosen because of their significantly different volumes and locations.

Generation of 4DCT(MRI) data sets

4DCT(MRI) data sets were generated by warping the end-exhale phase of each patient using deformation vector fields extracted from the 4DMRI datasets for both volunteers, by first geometrically correlating between the volunteers and patients using meshing (Figure 2, see [18] for details). For this, the lungs were first segmented on the end-exhale phase in both CT and MRI geometries, and landmarks (e.g. spine) defined to determine the subjects' orientation. Lung surfaces were then matched by converting them to triangular surface meshes with an equidistant 3D grid of vertices representing the lung volume. Deformation vector fields (DVFs) of the 4DMRI datasets were extracted using B-spline based deformable image registration (DIR) in Plastimatch (www.plastimatch.org). Each CT mesh vertex was then warped using MRI motion vectors extracted at the corresponding vertex of the MRI mesh, such that a deformed CT mesh resulted for each corresponding MRI motion state. Finally, these CT meshes were used to create the voxel-wise DVF for each motion state, by which the end-exhale CT was warped for each motion state. To avoid sliding boundary issues, the motion outside of the thorax, in the contralateral lung and at the inferior end of the CT scan were set to zero, causing the DVF interpolation to have a smooth transition to the non-moving geometry. In this work, 160 such 4DCT(MRI) data sets were thus generated, each consisting of ~9 phases, resulting in overall ~1400 motion states for all CT geometries and motion patterns together. Examples are shown in Figure 1b). The amplitudes of the target motion were (mean(SD) in mm): CT1 MRI1: 9.44(2.29), CT1 MRI 2: 11.41(2.86), CT2 MRI1: 5.06(1.42), CT2 MRI2: 5.13(1.45).

ITV definitions

Three methods of defining ITVs have been here investigated, based on different combinations of single-phase CTVs. An overview is given in Table 1.

Conventional ITV

From 20 cycles of each 4DCT(MRI) data set, every 5th cycle was used to define a conventional, single breathing-cycle ITV, leading to four different ITVs per patient. This scenario simulates the

conventional approach of defining an ITV from one 4DCT acquisition.

Conservative ITV

For this, the union of CTVs from all 20 planning respiratory cycles of each patient was used for the ITV, thus providing the most conservative (largest) volume which encompasses all 20 motion states.

Probabilistic ITVs

This final strategy defines ITVs in a probabilistic way. First, conventional ITVs were generated for all $n_{cyc} = 20$ cycles of the planning motion. These ITV masks (ITV_i) were then summed up, resulting in a probability map, stating for how many cycles the respective voxel v is included in the ITV:

$$p(v) = \frac{1}{n_{cyc}} \sum_{i=1}^{n_{cyc}} ITV_i(v).$$

Using a percentage threshold (25%, 50%, or 75%), the ITVx can then be defined as:

$$ITVx = \{v | p(v) \geq x\%\}.$$

Thus, ITV25 tends towards the conservative ITV ('ITV0') and ITV75 towards the conventional ITVs.

4D planning

Treatment planning was performed using the PSI in-house planning system for PSI Gantry 2 [26]. All ITVs were calculated as described above and a 2mm margin added for the 'PTV' in order to ensure target coverage in the static case. However, no consideration of patient set-up errors have been included for either treatment planning or simulation of delivery. The planning CT was then defined according to [27], with voxels outside the ITV being assigned average values extracted from all 4DCT(MRI) phases included in the respective scenario (mean intensity projection). Within the ITV, the maximum voxel intensity, as assessed from the same phases, was assigned (maximum intensity projection), ensuring that Bragg peaks were placed in the whole ITV. Two-field SFUD (single field uniform dose) plans were optimised on these synthetic planning CTs due to its superior robustness [28]. The chosen beam angles are shown in Figure 1a.

To assess the efficacy of each scenario, 4D dose calculations were performed using the ray-casting based [29] deforming grid algorithm [30,31] and accumulated on the reference full-exhale CT using motions extracted from 4 *different* batches of 5 consecutive respiratory cycles from the same 4DMRI acquisition (Figure 1b). As such, each 5-breathing-cycle batch was assumed to represent a different treatment day. Simulations were performed for 9x volumetric and layered rescanning [8] to minimise the influence of the interplay effect.

Plan evaluation and statistical analysis

For all ITVs, both patients and all MR motion data sets, dose volume histograms (DVHs) of the 4D dose distributions were calculated for the CTV and ipsilateral lung excluding the CTV on the reference phase (full-exhale). Analysed parameters included CTV V95% (target coverage), CTV D5-D95% (target homogeneity), absolute mean lung dose and lung V20Gy, assuming a prescribed dose of 66 Gy(RBE). Fractionation was simulated by summing up the 4D dose distributions over 30 fractions per treatment, with the start of delivery being simulated at random starting phases each fraction. For each ITV scenario, 100 such simulated treatments were calculated.

To detect significant differences between ITV definitions, Wilcoxon sum rank tests for non-Gaussian distributions were performed with a p-value <0.05 being considered significant. Results are displayed as violin plots, where the width of the plots indicate the density of data points within a certain band[1].

Results

ITV volumes are reported in the supplementary material. Of note, the conventional ITVs (defined from single breathing cycles) for the same case varied in volume by up to +/- 20% depending on the breathing pattern used for their calculation. For dose coverage, results for V95% and D5-D95% for the CTV, all ITVs and 9x volumetric rescanning are shown in Figures 3 and 4 for cases CT1/MRI1 and CT2/MRI1 respectively. Corresponding results for the other two scenarios can be found in the supplement, as well as for 9x layered rescanning.

For single fractions, 4D dose calculations for the conventional ITV predict considerable spreads of dose indices for the target, with V95% as low as 60% for the small tumour (median <95%) and 85% for the large tumour, depending on the 4DCT(MRI)s used for planning or delivery. Similarly, D5-D95% varied from 5% to 30% (20% for the large tumour), indicating the potential limitations of the conventional ITV in individual fractions when respiratory variability is present. For CT1, the median values for the conservative ITV, ITV25 and ITV50 are however significantly improved in comparison to the conventional ITV, whereas the ITV75 was too small to achieve a reasonable target coverage. Nevertheless, even for the conservative ITV, ITV25 and ITV50, coverage was only 80%, with a similar trend for dose homogeneity (D5-D95%). For CT2/MRI1, the conservative ITV, ITV25 and ITV50 achieved coverages of V95>90% for all starting phases and had significantly higher median values than the conventional ITVs. CTV D5-D95% shows a similar pattern, with the conservative ITV, ITV25 or ITV50 all having significantly lower D5-D95% than conventional ITVs, even if none can guarantee D5-D95% below 10%. On the other hand, CT2/MRI2 shows little dependence of target dose on ITV definition.

---

[1] https://ch.mathworks.com/matlabcentral/fileexchange/45134-violin-plot, last accessed on 27.08.2019

Fractionation however helps smear out uncertainties and reduces variations in coverage for the conventional ITV, resulting in V95% minima of 80% for CT1 and 95% for CT2, but D5-D95% is reduced to at most 15% for all cases. The conservative and probabilistic ITVs (except ITV75) show a significantly smaller variation in both coverage and homogeneity, leading to clinically acceptable coverages of V95 > 95% for all cases and D5-D95% <15%. Without rescanning, the motion effects were in the order of 60-90% for V95% and 15-40% for D5-D95, which are less dependent on the ITV definition (see supplementary Figures S9 and S10).

For single fractions, the mean dose and V20Gy to the ipsilateral lung (excluding the CTV), clearly reflects the differing ITV volumes. For the conservative ITV and ITV25, lung doses are highest, whereas ITV50 and ITV75 lead to lung doses comparable to the conventional ITVs, but with a significantly reduced spread of values. Fractionation reduces these variations for all scenarios, but with the same overall trend between scenarios, with the use of ITV50 resulting in mean lung doses comparable to the conventional ITVs.

Discussion

In this study, we investigated the consequences of respiratory variability for the treatment of lung tumours when conventional ITV definitions, based on single 4DCT studies, are used. In addition, we proposed different probabilistic ITV definitions that may help to mitigate dosimetric deterioration due to such variabilities, and have compared their efficacy to conventionally defined ITVs. We have demonstrated that, depending on the selected respiratory cycle, the volume of the ITV can vary significantly, with direct consequences on target coverage and dose to healthy tissue. Indeed, the coverage (V95%) can be as low as 80% even for fractionated conventional treatments, whereas the conservative (covering all imaged CTV positions) and probabilistic ITVs stay above 90% for all cases. However, although the conservative approach provides the best target coverage and dose homogeneity, it substantially increases dose to healthy tissue, whilst the use of a probabilistic ITV incorporating all voxels covered in at least 25% or 50% of respiratory cycles restores clinically acceptable dose distributions with minimal increase in lung dose. Similarly, whereas target dose homogeneity varies by 10% for conventional ITVs, it stays within 5% for all conservative and probabilistic ITV definitions. Thus, probabilistic ITVs (particularly ITV50) provide an effective approach for reducing the effect of breathing variability on target coverage and lung dose. Similar results have been found for hypo-fractionated treatments (i.e. 10 fractions). The results are however only valid if the planning and delivery motion amplitudes are in the same range. In the case of significantly larger amplitudes during delivery than in planning, target coverage will clearly be negatively affected independently of the ITV approach. Therefore, motion monitoring during plan delivery is of vital importance.

In clinical practice, motion information for probabilistic ITVs could be obtained using the same technique as used to model respiratory variability in this work, e.g. by acquiring 4DMRI data as part of the pre-treatment imaging work-up. For instance, 10 minutes of 4DMRI acquisition provides data on about 80 respiratory cycles per volunteer - not a major additional time penalty in the treatment planning process. Additionally, 4DMRIs could be acquired periodically during the

treatment course to asses if such ITV volumes are still valid, allowing ITV adaptation as necessary.

There are however, some limitations of this study. First, we did not consider positioning errors in the definition of the PTV/ITV and second, have used motion patterns obtained from healthy volunteers as we currently have no datasets available that include single-phase CT and 4DMR acquisitions of the same subject. Although one could argue that the motion patterns of a healthy lung differ from those of a lung including a lesion, we nevertheless believe that the variability included in this study is in the clinically relevant range. The same argument holds for the fact that the 4DMRI datasets were acquired within 10 minutes which clearly cannot represent day-to-day variations in respiration. Nevertheless, they present a relatively large range of possible variations and help to highlight the potential limitations of the conventional ITV approach under conditions of breathing variability. In addition, the process of generating 4DCT(MRI) is not perfect. Due to different contrasts and spatial resolutions of CT and MR images, the generated surface meshes and their correspondence cannot be exact, leading to uncertainties in the transformation of the deformation fields, and even the deformable image registration itself is subject to uncertainties. Consequently, the pseudo-4DCT datasets cannot perfectly represent a moving CT. Nevertheless, 4D datasets that are clinically plausible were generated using such procedures, and which provide a valuable source of potential motion variations with which to study the efficacy of the ITV concept, even with the limitations outlined above. The ray-casting algorithm on which our 4D dose calculation is based may not be as accurate as a full Monte Carlo simulation. However, a recent study [32] has shown only minor differences between the two algorithms, confirming the validity of our results.

Finally, the two lung cancer cases have been chosen to represent different anatomies: a small tumour with no restriction in motion and a larger tumour, which is located in a part of the lung with limited motion. Although this could be viewed as a rather restricted set of clinical tumour types, together with the many different motion patterns, and the resulting >1400 motion states, we believe our conclusions are representative for a broader variety of patients and are thus clinically relevant. Nevertheless, our results should not be applied blindly, but should be independently validated, where possible using 4DMRI and CT data from the same subject. Motion effects are very patient specific, and as such, it is important to have patient specific 4D dose calculations before any treatment decision is taken.

sWe have evaluated and analysed the consequences of respiratory motion variabilities on PBS 4D treatments, and have shown that ITV volume, and resulting target coverage, is extremely variable as a function of respiratory irregularities if defined on a single, example respiratory cycle. As such, we have proposed a probabilistic method to define the ITV, which is more robust to motion variations. The results show that our probabilistic approach (specifically the ITV50) can help ensure target coverage and homogeneity while keeping the healthy lung dose low.

Figure and table captions

Figure 1. a) Overview over the CT geometries in axial, coronal, and sagittal view. Top row: CT 1, bottom row: CT 2. The arrows indicate the chosen beam directions. b) CTV motion in superior-inferior direction for two example combinations of CT/MRI. Top row: CT 1, MRI 1, bottom row: CT 2, MRI 1. Solid line: median motion, shade: 10%-90% range of all CTV voxels.

Figure 2. a) Illustration of the generation process of a 4DCT(MRI). Blue (left): MRI related steps, yellow (middle): CT related steps, green (right): MRI and CT are combined. b) Example of an MR mesh (left) and a CT mesh (right) which are in correspondence.

Figure 3. Results for CT 1, MRI 1 and 9x volumetric rescanning. Left: single fractions, right: fractionated treatments. The numbers at the bottom indicate the number of data points included in the respective plot. a) CTV coverage in terms of V95%, b) dose homogeneity in D5-D95%, c) mean lung dose, d) lung V20Gy, assuming a prescribed dose of 66 Gy (RBE).

Figure 4. Results for CT 2, MRI 1 and 9x volumetric rescanning. Left: single fractions, right: fractionated treatments. The numbers indicate the number of data points included in the respective plot. a) CTV coverage in terms of V95%, b) dose homogeneity in D5-D95%, c) mean lung dose, d) lung V20Gy, assuming a prescribed dose of 66 Gy(RBE).

Table 1. Overview of the scenarios for ITV definitions and delivery validation.

References

[1] 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 Y R 2017 Consensus guidelines for implementing pencil-beam scanning proton therapy for thoracic malignancies on behalf of the PTCOG thoracic and lymphoma subcommittee *Int. J. Radiation Oncology Biol. Phys.* **99(1)**, 41-50.

[2] Fracchiolla F, Dionisi F, Giacomelli I, Hild S, Esposito P G, Lorentini S, Engwall E, Amichetti M, Schwarz M 2019 Implementation of proton therapy treatments with pencil beam scanning of targets with limited intrafraction motion *Physica Medica* **57**, 215-220.

[3] Zeng C, Plastaras J P, Tochner Z A, White B M, Hill-Kayser C E, Hahn S M, Both S 2015 Proton pencil beam scanning for mediastinal lymphoma: the impact of interplay between target motion and beam scanning *Phys. Med. Biol.* **60**, 3013-3029.

[4] Mageras G S & Yorke E 2004 Deep inspiration breath hold and respiratory gating strategies for reducing organ motion in radiation treatment *Seminars in Radiation Oncology* **14(1)**, 65-75.

[5] Dueck J, Knopf A C, Lomax A J, Albertini F, Persson G F, Josipovic M, Aznar M, Weber D C

& Munck af Rosenschöld P 2016 Robustness of the voluntary breath-hold approach for the treatment of peripheral lung tumors using hypofractionated pencil beam scanning proton therapy *Int. J. Radiation Oncology Biol. Phys.* **95(1)**, 534-541.

[6] Eccles C, Brock K K, Bissonnette J P, Hawkins M & Dawson L A 2006 Reproducibility of liver position using active breathing coordinator for liver cancer radiotherapy *Int. J. Radiation Oncology Biol. Phys.* **64(3)**, 751-759.

[7] Den Otter L A, Kauza E, Kierkels R G J, Meijers A, Ubbels F J F, Leach M O, Collins D J, Langendijk J A, Knopf A C 2018 Reproducibility of the lung anatomy under active breathing coordinator control: dosimetric consequences for scanned proton treatments *Med. Phys.* **45(12)**, 5525-5534.

[8] Zhang Y, Knopf A C, Weber D C & 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.

[9] Bert C, Saito N, Schmidt A, Chaudhri N, Schardt D & Rietzel E 2007 Target motion tracking with a scanned particle beam *Phys. Med. Biol.* **34**, 4768-4771.

[10] Grözinger S O, Rietzel E, Li O, Bert C, Haberer T & Kraft G 2006 Simulations to design an online motion compensation system for scanned particle beams *Phys. Med. Biol.* **51**, 3517.

[11] Bernatowicz K, Zhang Y, Perrin R, Weber D C, Lomax A J 2017 Advanced treatment planning using direct 4D optimisation for pencil-beam scanned particle therapy *Phys. Med. Biol.* **62**, 6595-6609.

[12] Bert C & Durante M 2011 Motion in radiotherapy: particle therapy *Phys. Med. Biol.* **56**, R113-R144.

[13] Zhang Y, Huth I, Weber D C, Lomax A J 2018 A statistical comparison of motion mitigation performances and robustness of various pencil beam scanned proton systems for liver tumour treatments *Radiotherapy and Oncology* **128(1)**, 182-188.

[14] Bernatowicz K, Lomax A J & Knopf A 2013 Comparative study of layered and volumetric rescanning for different scanning speeds of proton beam in liver patients *Phys. Med. Biol.* **58**, 7905-7920.

[15] Knopf A C, Hong T S & Lomax A J 2011 Scanned proton radiotherapy for mobile targets – the effectiveness of re-scanning in the context of different planning approaches and for different motion characteristics *Phys. Med. Biol.* **56(22)**, 7257-7271.

[16] ICRU Report 78 2007 Prescribing, recording, and reporting proton-beam therapy. Bethesda, MD: International Commission on Radiation Units & Measurements.

[17] Celicanin Z, Giger A, Bauman G, Cattin P C, Bieri O 2017 Temporally-resolved volumetric imaging (4DMRI) of the lungs. Proc. 25th Scientific Meeting, ISMRM.

[18] Boye D, Samei G, Schmidt J, Székely G, Tanner C 2013b Population based modeling of respiratory lung motion and prediction from partial information *Medical Imaging 2013: Image Processing SPIE* **8669**, 86690U-1.

[19] von Siebenthal M, Szekely G, Gamper U, Bösiger P, Lomax A J & Cattin P 2007 4D MR imaging of respiratory organ motion and its variability *Phys. Med. Biol.* **52(6)**, 1547-1564.

[20] Block T K, Handarana C, Milla S 2014 Towards routine clinical use of radial stack-of-stars 3D gradient-echo sequences for reducing motion sensitivity. *J. Korean Soc. Magn. Reson. Med.* **18**, 87-106.

[21] 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' http://doi.org/10.7937/K9/TCIA.2016.ELN8YGLE. The Cancer Imaging Archive.

[22] Hugo G D, Weiss E, Sleeman W C, Balik S, Keall P J, Lu J & Williamson J F 2017 A longitudinal four-dimensional computed tomography and cone beam computed tomography dataset for image-guided radiation therapy research in lung cancer *Med. Phys.* **44**, 762-771.

[23] Balik S et al. 2013 Evaluation of 4-dimensional computed tomography to 4-dimensional cone-beam computed tomography deformable image registration for lung cancer adaptive radiation therapy Int. J. Radiation Oncology Biol. Phys. 86, 372-379.

[24] Roman N O, Shepherd W, Mukhopadhyay N, Hugo G D & Weiss E 2012 Interfractional positional variability of fiducial markers and primary tumors in locally advanced non-small-cell lung cancer during audiovisual biofeedback radiotherapy. *Int. J. Radiation Oncology Biol. Phys.* **83**, 1566-1572.

[25] 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 The cancer imaging archive (tcia): Maintaining and operating a public information repository *J. Digit. Imaging* **26(6)**, 1045-1057.

[26] Pedroni E, Bearpark R, Böhringer T, Coray A, Duppich J, Forss S, George D, Grossmann M, Goitein G, Hilbes C, Jermann M, Lind S, Lomax A J, Negrazus M, Schippers M & Kotrle G 2004 The PSI Gantry 2: a second generation proton scanning gantry *Z. Med. Phys.* **14**, 25-34.

[27] Botas P, Grassberger C, Sharp G & Paganetti H 2018 Density overwrites of internal tumor volumes in intensity modulated proton therapy plans for mobile lung tumors *Phys. Med. Biol.* **63**, 035023.

[28] 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 The dosimetric effect of residual breath-hold motion in pencil beam scanned proton therapy – an experimental study *Radiotherapy and Oncology* **134**, 135-142.

[29] Schaffner B, Pedroni E, 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.

[30] Boye D, Lomax A J & Knopf A C 2013a Mapping motion from 4D-MRI to 3D-CT for use in 4D dose calculations: A technical feasibility study *Med. Phys.* **40(6)**, 061702-1-061702-11.[31] Krieger M, Klimpki G, Fattori G, Hrbacek J, Oxley D, Safai S, Weber D C, Lomax A J & Zhang Y 2018 Experimental validation of a deforming grid 4D dose calculation for PBS proton therapy *Phys. Med. Biol.* **63**, 055005.

[32] Winterhalter C, Zepter S, Shim S, Meier G, Bolsi A, Fredh A, Hrbacek J, Oxley D, Zhang Y, Weber D C, Lomax A, Safai S 2019 Evaluation of the ray-casting analytical algorithm for pencil beam scanning proton therapy *Phys. Med. Biol.* **64**, 065021.

Figure 1

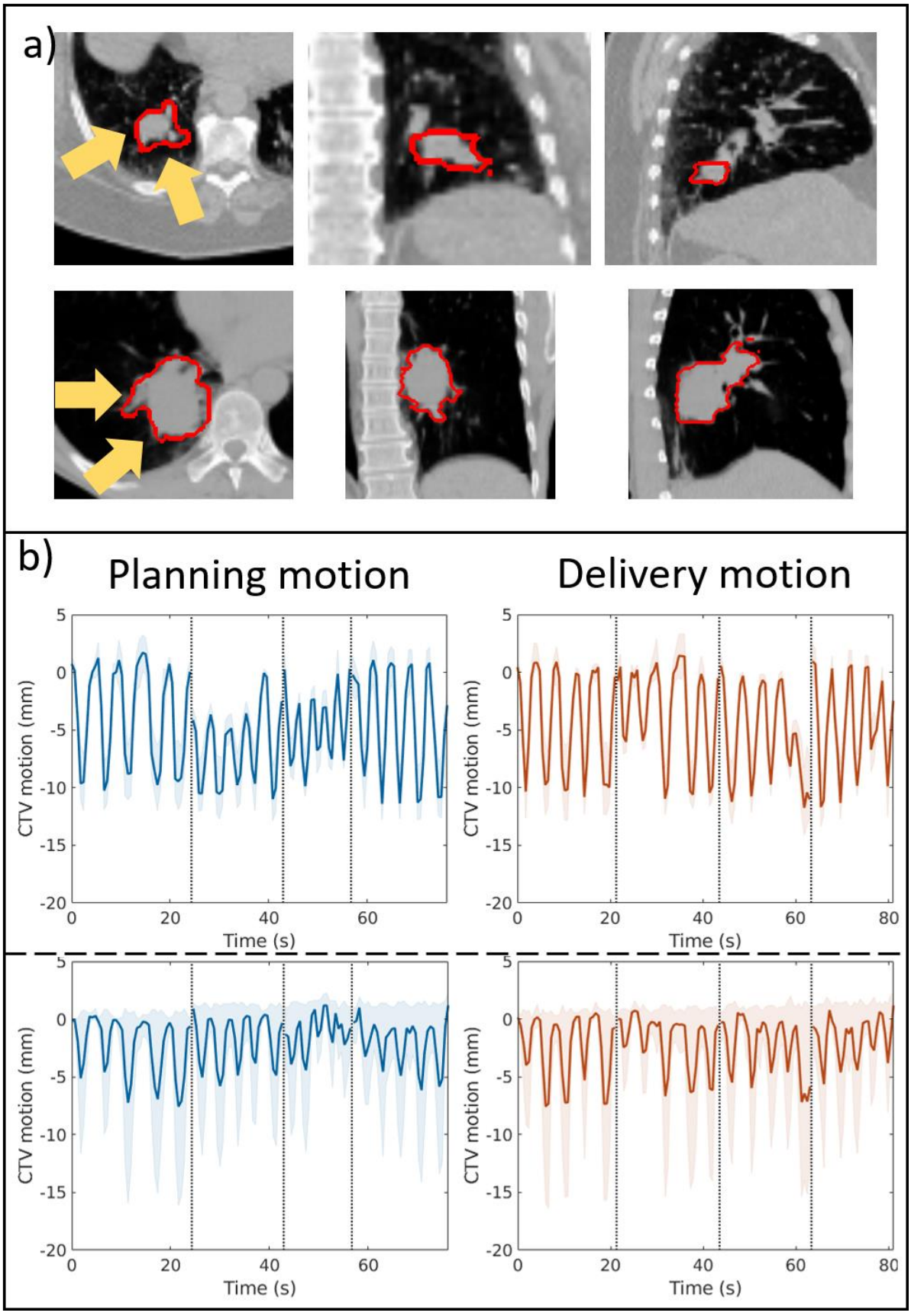
a)
b)
Planning motion
Delivery motion
CTV motion (mm)
Time (s)
5
0
-5
-10
-15
-20
0
20
40
60
80

Figure 2

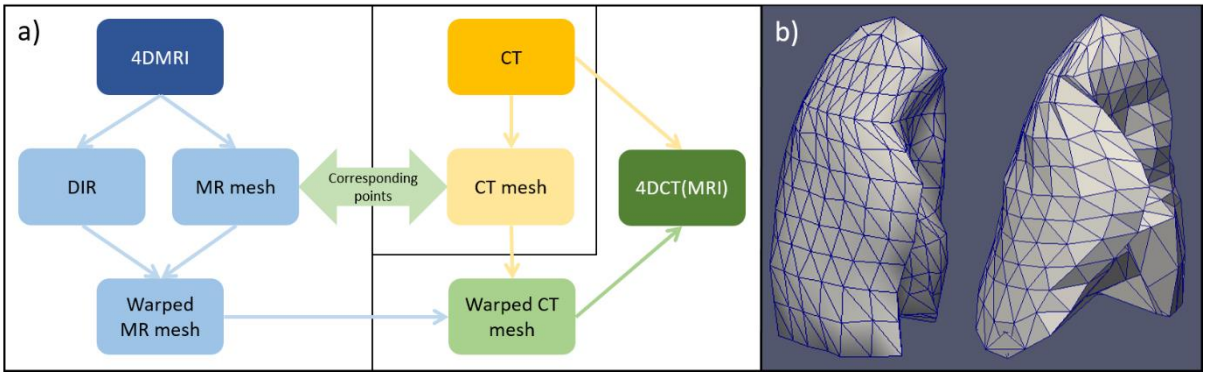
a)
4DMRI
DIR
MR mesh
Corresponding points
Warped MR mesh
CT
CT mesh
4DCT(MRI)
Warped CT mesh
b)

Figure 3

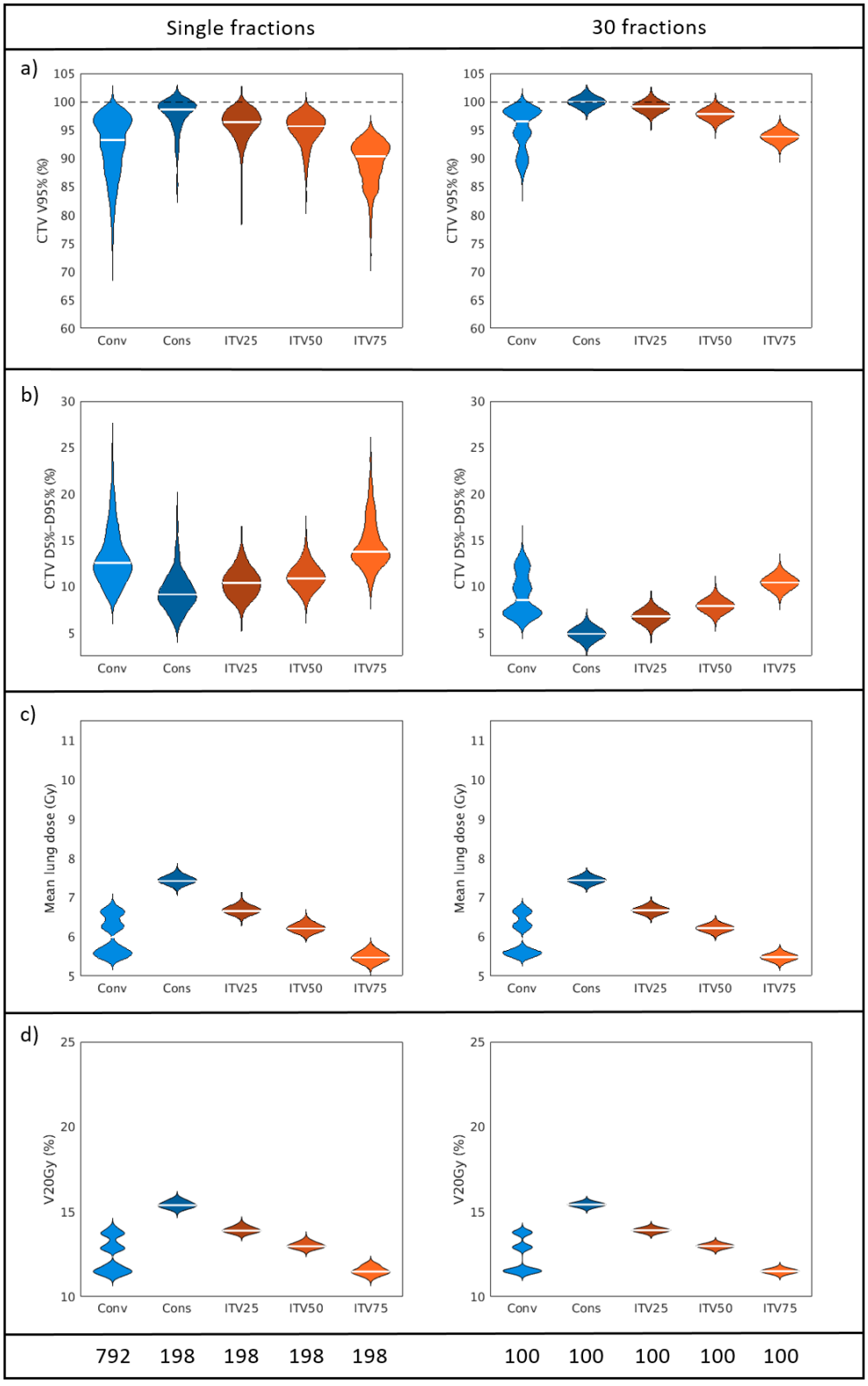
Single fractions
30 fractions
a)
CTV V95% (%)
b)
CTV D5%–D95% (%)
c)
Mean lung dose (Gy)
d)
V20Gy (%)
Conv
Cons
ITV25
ITV50
ITV75
792 198 198 198 198
100 100 100 100 100

Figure 4

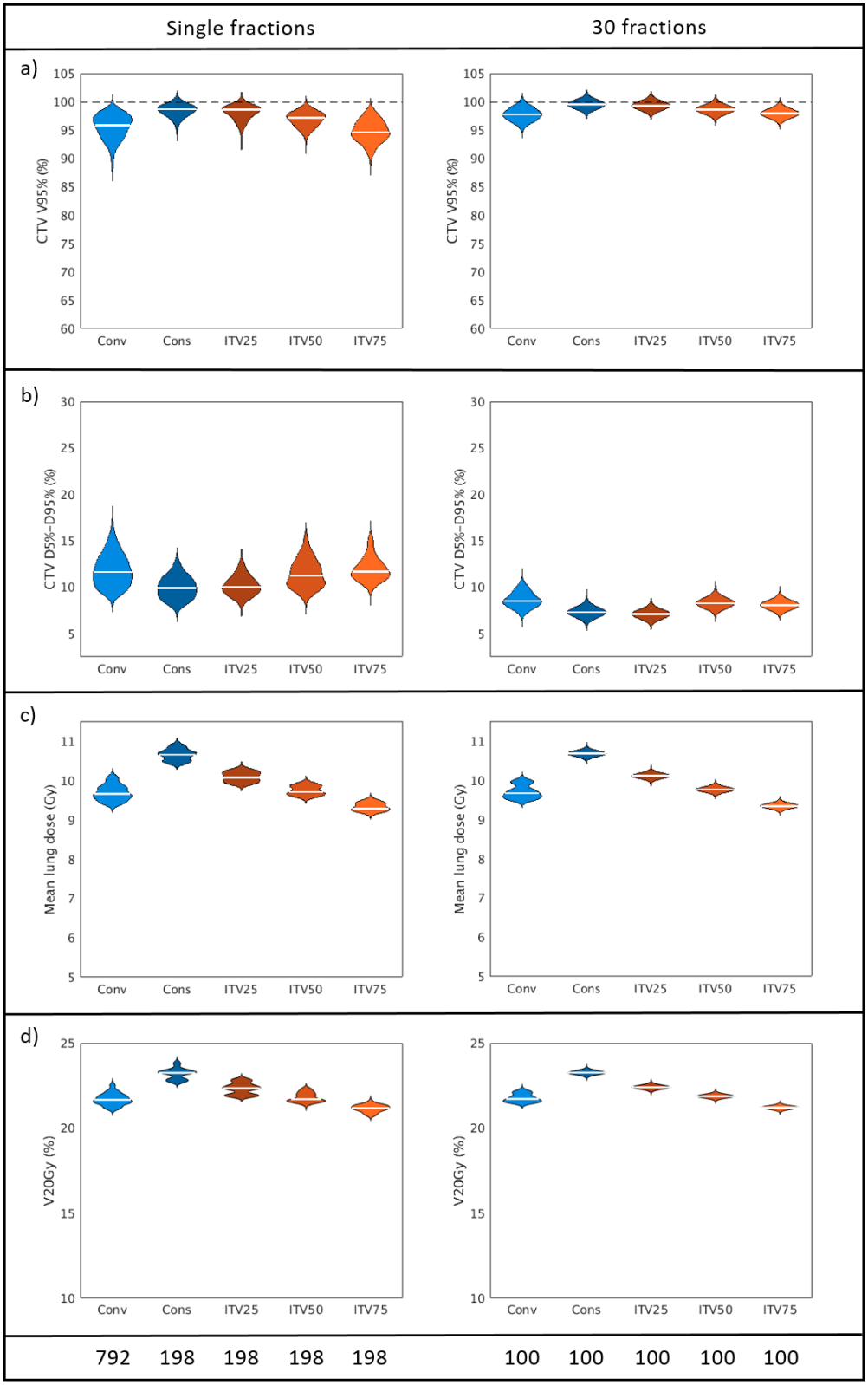
Single fractions
30 fractions
a)
CTV V95% (%)
b)
CTV D5%–D95% (%)
c)
Mean lung dose (Gy)
d)
V20Gy (%)
Conv
Cons
ITV25
ITV50
ITV75
792 198 198 198 198
100 100 100 100 100

Table 1

	Motion patterns used for ITV definitions	Delivery sessions	Total ITV and delivery combinations tested
Conventional (single-cycle) ITV	Single-cycle ITVs based on 4 different respiratory cycles	20 *different* breathing cycles, broken up into 4 treatment sessions, each of 5 breathing cycles	4x4
Conservative ITV	Single ITV based on all CTV positions from 20 breathing cycles		1x4
Probabilistic ITVs	3 different probabilistic ITVs (25%, 50%, 75% threshold) based on all CTV positions from 20 breathing cycles		3x4