Radiotherapy and Oncology 147 (2020) 178–185

Contents lists available at ScienceDirect

Radiotherapy and Oncology

journal homepage: www.thegreenjournal.com

Original Article

Deformable image registration uncertainty for inter-fractional dose accumulation of lung cancer proton therapy

Lena Nenoff [a,b,*], Cássia O. Ribeiro [c], Michael Matter [a,b], Luana Hafner [a,b], Mirjana Josipovic [d], Johannes A. Langendijk [c], Gitte F. Persson [d,e,f], Marc Walser [a], Damien Charles Weber [a,g,h], Antony John Lomax [a,b], Antje-Christin Knopf [c,i], Francesca Albertini [a], Ye Zhang [a]

[a] Paul Scherrer Institute, Center for Proton Therapy; [b] Department of Physics, ETH Zurich, Switzerland; [c] Department of Radiation Oncology, University Medical Center Groningen, University of Groningen, The Netherlands; [d] Department of Oncology, Rigshospitalet Copenhagen University Hospital; [e] Department of Oncology, Herlev-Gentofte Hospital Copenhagen University Hospital; [f] Department of Clinical Medicine, Faculty of Medical Sciences, University of Copenhagen, Denmark; [g] Department of Radiation Oncology, University Hospital Zurich; [h] Department of Radiation Oncology, University Hospital Bern, Switzerland; [i] Division for Medical Radiation Physics, Carl von Ossietzky University Oldenburg, Germany

ARTICLE INFO

Article history:
Received 24 January 2020
Received in revised form 22 April 2020
Accepted 25 April 2020
Available online 5 May 2020

Keywords:
Deformable image registration
Proton therapy
Dose accumulation
NSCLC

ABSTRACT

Background and purpose: Non-small cell lung cancer (NSCLC) patients show typically large anatomical changes during treatment, making recalculation or adaption necessary. For report and review, the applied treatment dose can be accumulated on the reference planning CT using deformable image registration (DIR). We investigated the dosimetric impact of using six different clinically available DIR algorithms for dose accumulation in presence of inter-fractional anatomy variations.

Materials and methods: For seven NSCLC patients, proton treatment plans with 66 Gy-RBE to the planning target volume (PTV) were optimised. Nine repeated CTs were registered to the planning CT using six DIR algorithms each. All CTs were acquired in visually guided deep-inspiration breath-hold. The plans were recalculated on the repeated CTs and warped back to the planning CT using the corresponding DIRs. Fraction doses warped with the same DIR were summed up to six different accumulated dose distributions per patient, and compared to the initial dose.

Results: The PTV-V95 of accumulated doses decreased by 16% on average over all patients, with variations due to DIR selection of 8.7%. A separation of the dose effects caused by anatomical changes and DIR uncertainty showed a good agreement between the dose degradation caused by anatomical changes and the dose predicted from the average of all DIRs (differences of only 1.6%).

Conclusion: The dose degradation caused by anatomical changes was more pronounced than the uncertainty of employing different DIRs for dose accumulation, with averaged results from several DIRs providing a good representation of dose degradation caused by anatomy. However, accumulated dose variations between DIRs can be substantial, leading to an additional dose uncertainty.

With proton therapy, high target coverage can be achieved, while sparing dose to organs-at-risk (OARs) [1,2]. This makes it especially attractive for tumours with many surrounding OARs, such as cancers in the brain [3], skull base [4–6], head and neck [7,8] or lung [9]. Recently, the potential of proton therapy has been assessed for non-small cell lung cancer (NSCLC) treatments [10–13], with the main concern being *intra*-fractional variability. To mitigate these effects, rescanning [14], gating [10], tracking [15] or 4D-optimisation [16,17] have all been investigated. Alternatively, deep-inspiration breath-hold (DIBH) to minimize intra-fraction motion has also been proposed [18].

Due to the finite range of protons however, *inter*-fractional anatomical changes in the entrance path of the beam can also play a major role, substantially distorting the planned dose even when intra-fraction motion is minimised [19–21]. As such, and even more so than for conventional therapy, regular re-imaging of the patient is required, on which the delivered dose can either be re-calculated, or adapted by reoptimising the plan [22–24]. With or without adaption however, substantially different dose distributions for each anatomical instance will result [22], making the reporting of the total dose distribution delivered to the patient over the whole treatment course challenging. For this, the calculation of the accumulated dose distribution on a reference (e.g. planning) CT is invaluable and is particularly important if dosimetric parameters such as maximum dose or D2, V95 etc. need to be reported for the whole treatment. Such parameters can only be

* Corresponding author at: WBBB 105, Forschungsstrasse 111, 5232 Villigen PSI, Switzerland.
E-mail address: lena.nenoff@psi.ch (L. Nenoff).

https://doi.org/10.1016/j.radonc.2020.04.046
0167-8140/

correctly calculated by accumulating the different dose distributions onto a common anatomical representation of the patient, which in the thorax requires deformable image registration (DIR) to warp each individual dose distribution back to the reference patient geometry [25].

It is recognised however that different DIRs tend to give different results, which can lead to pronounced differences in the warped and accumulated doses [26,27]. Especially in the case of large changes in tumour mass [28], as typically present in the lung, these uncertainties can be large. The handling of mass changes in DIR is challenging. From a clinical point of view, disappearing tissue (e.g. weight loss) requires an adequate shrinkage of structures and volume for dose accumulation. Other changes however (e.g. tumour shrinkage in the lung) do not necessarily imply a reduction of the volume with microscopic tumour spread, so a reduction of the clinical target volume (CTV) might be inadvisable. From a mathematical point of view, this separation, as well as the handling of sliding organ interfaces, are difficult. Modern algorithms however, have different ways of implementing these, which are reviewed elsewhere [29]. Previous studies have compared the dosimetric differences caused by the use of different DIRs in 4D dose accumulation for liver tumours planned with pencil beam scanned proton therapy [30,31]. For lung-stereotactic body radiotherapy, uncertainties have been reviewed previously [32], and the effect of different DIR uncertainties has been evaluated for intra-fractional motion [33] and complete treatments [34]. For proton therapy, DIR has also been used for 4D dose accumulation during treatment [35], but up to now no quantification about the influence of DIR uncertainty on dose accumulation after inter-fractional anatomical changes has been performed.

In this study, we evaluate the impact of using different DIR algorithms in the presence of inter-fractional anatomical changes on accumulated dose distributions for NSCLC patients treated in DIBH with intensity modulated proton therapy (IMPT). We first investigated the spatial distribution of the dosimetric variations of accumulated doses. Secondly, we compared the treatment doses accumulated with different DIRs to the initial planning dose. Finally, we evaluated how well the dose degradation caused by anatomical changes was represented by doses warped back to the planning CT.

Materials and methods

Patient data and treatment plans

In this retrospective study, seven NSCLC patients, previously treated with photon radiotherapy, each with a planning CT and nine repeated CTs acquired during treatment (three repeated off-line CT acquisitions each on day 2, 16 and 31 of treatment) were included in this study. To mitigate intra-fractional motion, all CTs were acquired with visually guided voluntary DIBH. In this study, each one of these nine CTs was assumed to represent the anatomy of one fraction. We also assumed the whole fraction can be delivered within one breath-hold. IMPT treatment plans with a prescribed dose of 66 Gy-relative biological effectiveness (RBE) in 2 Gy per fraction to the planning target volume (PTV) with three individually selected fields were designed using a fast in-house developed optimiser [36] and analytical dose calculation [37]. A PTV margin of 5 mm in the cranio-caudal and antero-posterior, and 4 mm in the lateral directions was used, derived from clinical breath-hold data [13,38].

Image registration

Repeated CTs were registered to the planning CT (reference CT) following a two-step process. They were first aligned rigidly in Velocity (Varian Medical Systems, Palo Alto, USA) with focus on the vertebra. Then DIR was applied using six different algorithms – two open access algorithms from Plastimatch (Demons and B-splines) and four commercial approaches from Velocity, Mirada (Mirada Medical, Oxford, UK) and RayStation (RaySearch Laboratories, Stockholm, Sweden) (Anaconda and Morfeus).

The B-splines algorithm implemented in Plastimatch models the deformation with a grid of B-splines control points and optimises mean square difference as the cost function [39]. Demons algorithms use the image intensity-based gradient force between the fixed and moving image for deformation [40], and then the deformation is smoothed by a Gaussian filter. Velocity has implemented an elastic B-splines algorithm which uses mutual information [41]. The 'CT deformable' algorithm provided by Mirada uses (similar to Demons) a gradient of the image intensity, but instead of a Gaussian smoothing, diffusion partial differential equations [42]. The RayStation Anaconda is an intensity-based algorithm that accounts for image similarity and a grid regularization for smoothing [43]. RayStation Morfeus is a feature-based biomechanical modelling DIR method [41,44]. All DIR algorithms, except Morfeus, were applied without a focus or controlling region of interest (ROI). For Morfeus, the external contour was used as the controlling ROI [45]. The output from all algorithms is a voxel specific displacement vector field (DVF), corresponding to the vector pointing from the planning CT to the repeated CT. The detailed settings of each DIR algorithm are summarised in Supplement 1.

Structure propagation

Although GTV volumes of the investigated patients changed on average by −16% (ranging from +1% to −18%) between the planning CT and the average of the three repeated CTs, in this study, the CTV and PTV have been propagated rigidly to each repeat CT, as recommended by Sonke et al. [26]. This is a conservative approach, assuming that a change in visible gross tumour volume (GTV) does not necessarily reduce the microscopic spread in the CTV. A visual check of the rigid PTV assured that the visible GTV was still encompassed by the PTV in each repeated CT. Note however that despite this approach, any substantial loss of mass of the tumour can still have a profound effect on the delivered proton dose distribution due to the residual range changes resulting from such losses.

Calculating 'fraction' and 'treatment' doses

Each plan was recalculated with the in-hose developed software on the previously rigidly registered repeated CTs. The resulting 'fraction doses' (differences caused by anatomy and patient misalignments, not by deformation) were then warped with each DVF (extracted from the different clinical DIR systems) using the dose warping function from Plastimatch. This results in six 'warped fraction doses' (with combined uncertainties from anatomy, misalignment and DIR) per repeated CT. Doses warped with the same algorithm were accumulated on the planning CT in Matlab (MathWorks, Natic, USA), resulting in an estimation of six different 'accumulated treatment doses' per patient (also containing uncertainties by anatomy and DIR). Fig. 1 shows a schematic representation of the workflow of this study.

Dosimetric evaluation

Evaluation of fraction specific doses

To separate the effects of anatomical changes and DIR uncertainties, we compared the PTV-V95 of the recalculated doses directly on the repeated CT ('fraction doses') with the fraction doses warped back to the planning CT, without accumulation

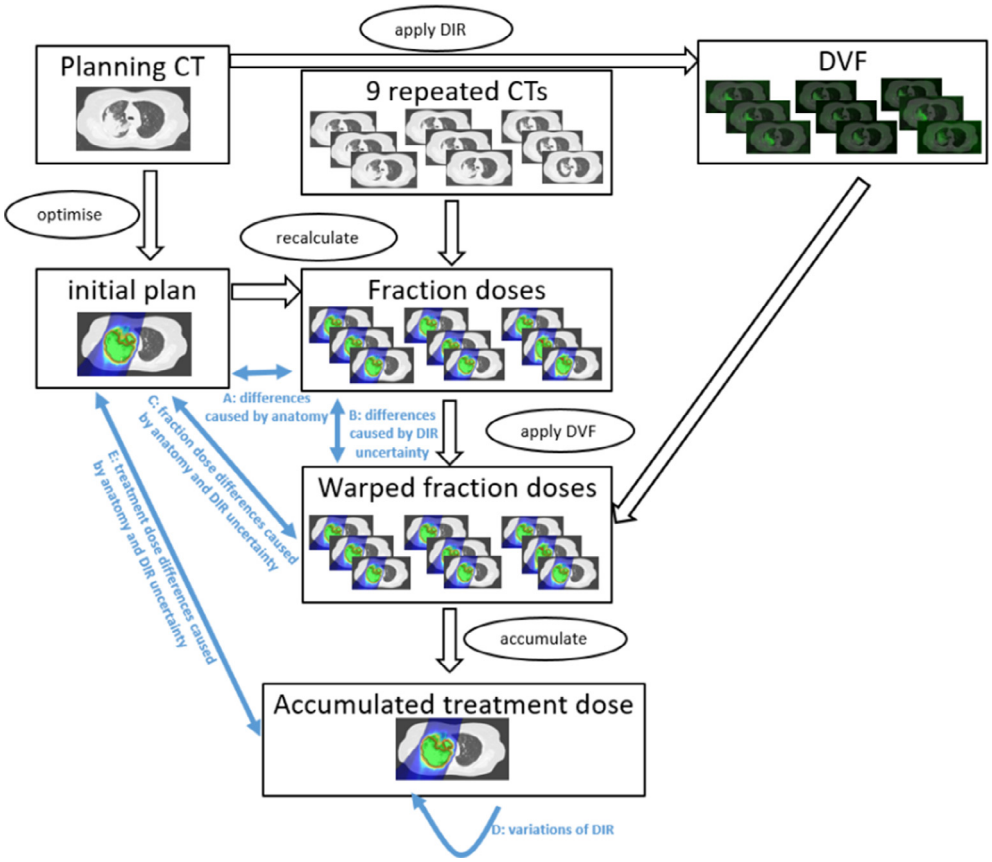

Fig. 1. Scheme of the workflow of this study for one example DIR. Examples of deformation vector fields (DVFs), the initial planned dose, fraction doses, warped fraction doses and the accumulated treatment doses warped with one DIR are given. The obtained dose distributions are compared with each other (blue arrows A–E). (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

('warped fraction doses'). For assessment of the fraction doses we used the rigidly propagated PTV, whereas for the warped fraction doses, the original PTV on the planning CT was used. In this way, comparisons of the fraction doses show differences caused by anatomical changes only (Fig. 1, comparison A), whereas differences between the fraction doses and the warped fraction doses add the uncertainty introduced by DIRs (Fig. 1, comparison B). Finally, differences between planned dose and warped fraction doses contain uncertainties from both anatomical changes and DIRs (Fig. 1, comparison C). As voxel positions change between the repeat CTs, voxel-wise dose differences could not be evaluated, only DVH parameters.

Evaluation of accumulated doses

To estimate the dosimetric effects of different DIRs during treatment, we compared differences in accumulated doses with all DIRs (Fig. 1, comparison D). For this, we calculated the voxel specific maximum and minimum in treatment dose accumulated with all six algorithms. This provides an estimate of the (non-physical) voxel-wise max-to-min dose-deviations caused by the use of different DIR algorithms. From this, dose-deviation-volume histograms (DDVHs) were calculated for selected structures (PTV, CTV, ipsilateral lung, heart and spinal cord).

In addition, dose-volume-histograms (DVHs) of the six accumulated treatment doses were compared to the initial plan, optimised on the planning CT (Fig. 1, comparison E). Also, selected DVH parameters, such as PTV-V95 and mean dose to ipsilateral lung and heart, were evaluated. These differences also contain the effects of both anatomical changes and DIR uncertainty, but now accumulated over all repeated CTs.

Results

The PTV-V95 of each fraction dose (changes caused by anatomical changes only, comparison A) decreased compared to the planned dose over all patients and fractions by 14% on average, ranging from 1.5% to 40.5% for single fractions (Fig. 2). Additionally, variations between the warped fraction doses with the six DIRs were on average 7.9% (between 1.7% for patient 1 and 23.3% for patient 6). The mean agreement was high, PTV-V95 differences between fraction doses and warped fraction doses were on average 1.6% (range 0.8% to 4.1%, comparison B). This good agreement is also seen in the OAR doses, with differences to the mean heart dose between the fraction doses and the average of the warped fraction doses being 3.4% (range 1.0% to 9.5%, Fig. 3/Supplement 2, comparison B). This indicates that the dose degradation caused by anatomical changes is well represented by the mean of all DIR algorithms, even if variations between different DIRs can be high (comparison C).

Inspecting the fraction doses obtained with different DIR algorithms, we found that RayStation Morfeus differed substantially for two patients (6 and 7), compared to the other DIR algorithms. For other patients only minor differences were observed. Excluding Morfeus from the analysis of all patients reduced the variation of the warped fraction doses to 3.2% (range 1.0–7.9%) compared to a variation of 7.9% (range 1.7–23.3%) when all six DIRs were included. Furthermore, the agreement between the 'fraction doses' recalculated on the repeated CT and the corresponding doses warped back to the planning CT also improved (average difference in the PTV-V95 of 0.9% (range 0.6–4.2%) vs. 1.6% (0.8–4.1%), comparison B).

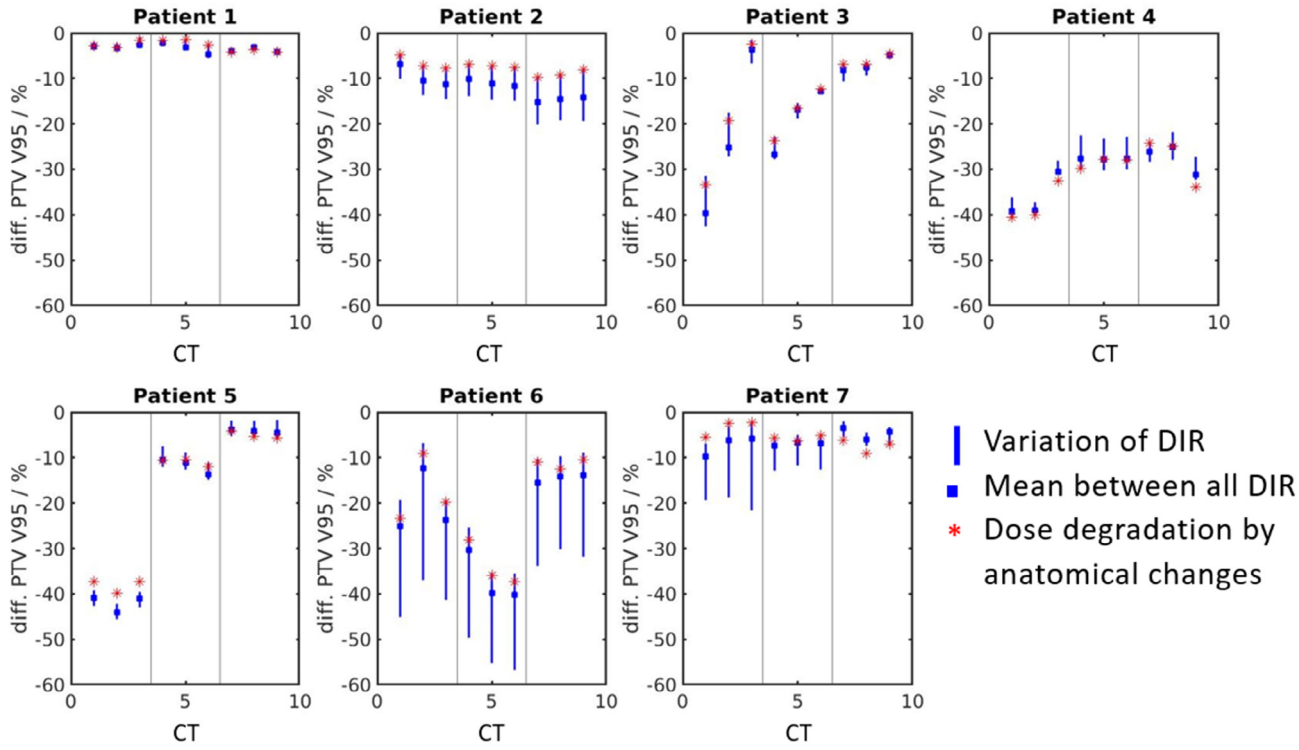

Fig. 2. The PTV-V95 differences between the initial plan and the fraction doses, evaluated before dose warping (red stars), as well as warped fraction doses (range: blue bars, mean: blue box). The mean of all red stars represents the dose degradation caused by anatomical changes only. The range of the blue bars is the variation caused by the DIRs. The difference between the mean of all warped fraction doses (blue box) and the fraction doses (red stars) shows how well the anatomical dose degradation is represented by the warped fraction doses. Different CT acquisition days are separated by vertical lines. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

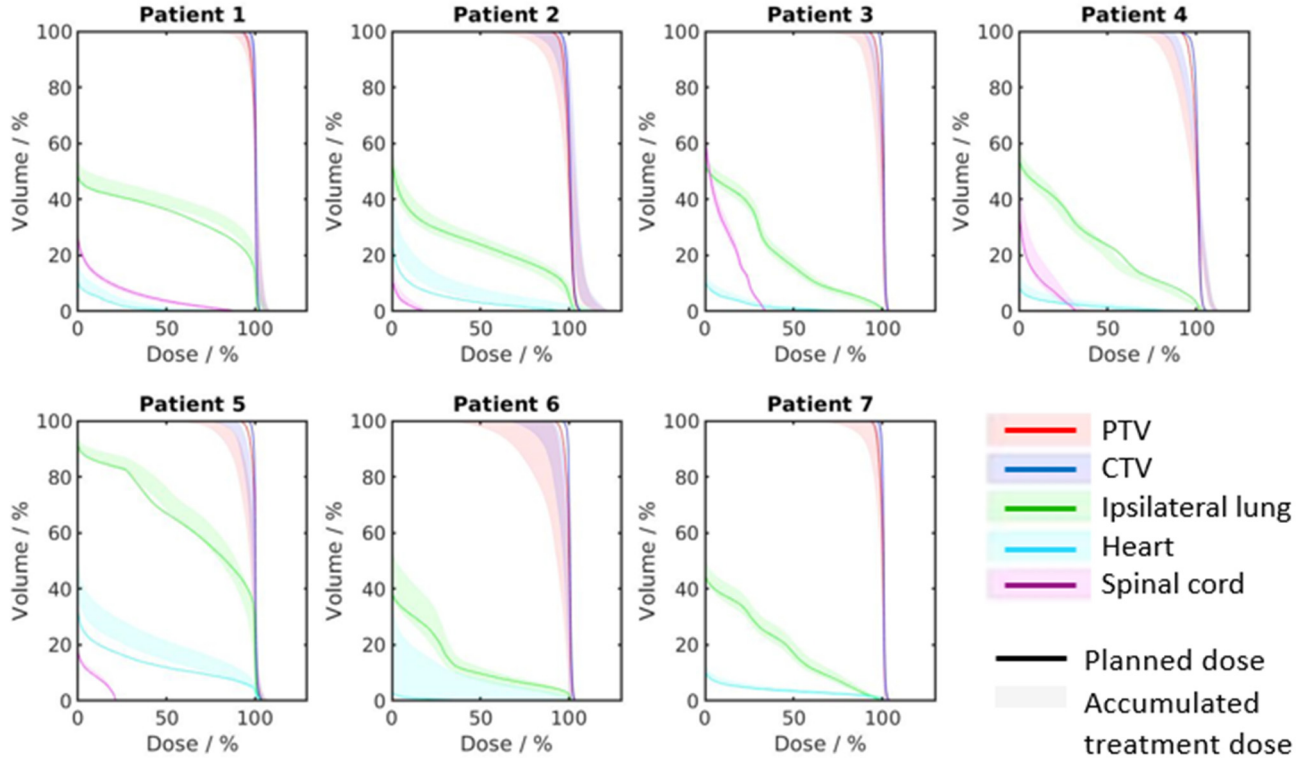

Fig. 3. DVHs of CTV, PTV, ipsilateral lung, heart and spinal cord of the initial treatment plan (solid line) and the accumulated treatment dose (light coloured band), warped with different DIRs. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

Fig. 4 reports the variation between accumulated treatment doses resulting from all six DIR algorithms (comparison D). Voxel-wise max-to-min dose distributions show that the largest treatment dose differences accumulated with different DIRs were found in the high dose gradient region. Consequently, the DDVHs show often large variations in neighbouring OARs, such as ipsilateral lung or heart. In particular, the mean dose to the ipsilateral lung can vary up to 3% (patient 5 and 7), and the mean heart dose up to 9.5% (patient 6). For some patients, large variations between different DIR algorithms in the PTV-V95 were observed (up to 26.3%, patient 6).

The comparison between the DVHs of the initial plan and the DVH uncertainty-band of the accumulated doses is shown in Fig. 3 (comparison E). The decrease of treatment dose quality compared to the initial plan is caused by both anatomical changes and DIR uncertainties. More specifically, the PTV-V95 of the treatment doses decreased by 16% on average over all patients (range 2.3–28.8%, Supplement 2). The variations in PTV-V95 caused by DIR in the accumulated treatment doses were on average 8.7%, ranging from 1.0% (patient 1) to 26.3% (patient 6). Moreover, the OAR doses have pronounced differences compared to the initial plan. The mean doses to the ipsilateral lung and heart showed variations of 1.8% and 8.5% due to DIR, and an increased value compared to the planned mean doses of on average 2.3% and 3.4%.

Discussion

We have evaluated the treatment doses of seven locally advanced NSCLC patients accumulated with six different DIR algorithms. An average PTV-V95 variation of 8.7% was measured between the accumulated treatment doses resulting from different DIRs. In total, the average reduction in PTV-V95 was 16% (Fig. 3,

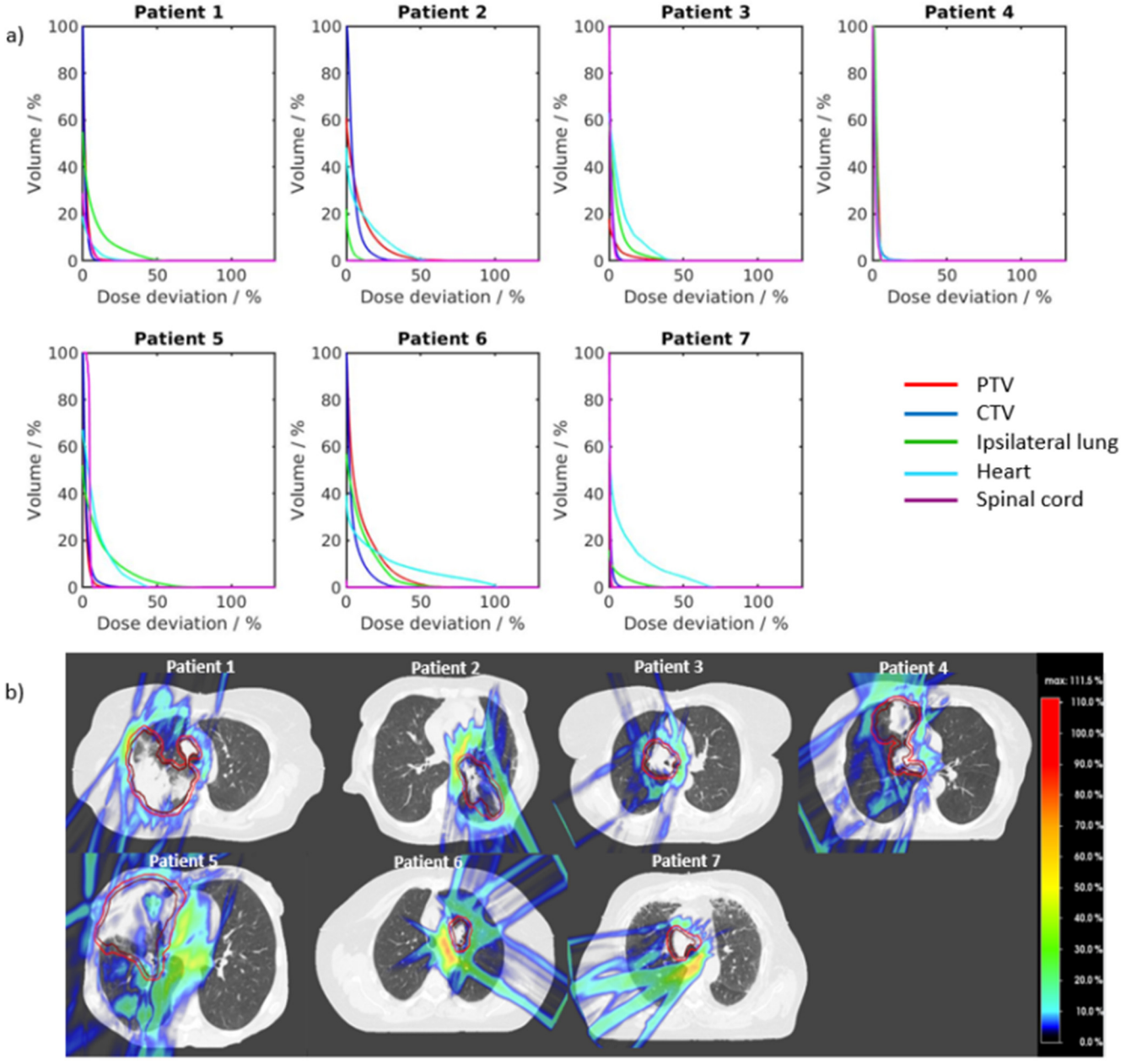

Fig. 4. (a) Dose-deviation-volume histograms (DDVHs) of the accumulated treatment dose difference warped with the six DIR algorithms. (b) An example slice of the max-to-min dose distribution difference, calculated as the voxel-wise difference between the maximum and minimum treatment dose, accumulated with the six DIRs.

Supplement 2), caused by a combination of anatomical changes and DIR uncertainty.

For each repeated CT, we compared DVH parameters for recalculated fraction doses with the planned dose. An average underdosage of 14% was measured in the PTV-V95, representing the dose degradation caused by anatomical changes only. However, for single fractions, a PTV-V95 reduction of up to 40.5% was found. This shows the extreme sensitivity of IMPT proton plans to density changes in the beam path, which are mainly caused by anatomical changes and by the patient set-up. The patient set-up was simulated here by rigidly registering the repeated CTs onto the planning CT, by focusing on the alignment of the vertebral body in the proximity of the target volume. Additionally, we observed PTV-V95 variations of 7.9% caused by DIR uncertainty alone (Fig. 4). This indicates that for these patients, the dosimetric impact of anatomical changes was larger than the variations caused by DIR uncertainty. This analysis is influenced by the fact that, despite the tumour shrinkage, the PTV was transferred rigidly, which is a conservative approach. The rationale is that the PTV is initially drawn on the planning CT to include uncertainties during treatment (setup, range, delineation uncertainties, typical anatomical changes [46]). However, it is debatable if this approach is the best. The mix of tissue displacement (for which the target structure should be changed) and shrinkage (where the microscopic disease should be treated, even if not visible anymore) makes a careful review necessary before reducing any target structure. This is challenging and still an open question in the community [47]. In our study, we did not adapt the treatment, but used this rigid target concept for the evaluation of the fraction doses before warping (comparison A and B). With this rigid target concept we assume that the CTV microscopic spread (and consequently the corresponding PTV margin) is not reduced even if the GTV has shrunk. We do not expect major changes of the overall results if the target contours were deformed instead. Especially the evaluation of the variations of the different DIRs (comparison D) and warped and accumulated doses (comparison C and E) do not use the fraction doses with the rigid target concept, and are therefore not affected at all.

Interestingly, the difference between the dose recalculated on each repeated CT and the average of the six doses warped back to the reference CT matched well (differences of only 1.6% in the PTV and 3.4% for the heart, see Fig. 2 and Supplement 2). This suggests that using multiple DIRs is a valid approach to estimate dose uncertainties caused by anatomical changes during treatment and to have a more realistic representation of the delivered dose. Indeed, if only one DIR algorithm would be used, DVH differences of more than 10% can be propagated into the accumulated treatment dose (Fig. 4, Supplement 2), clearly having an impact on clinical decisions. In addition, as there is no way of knowing the ground-truth deformations of the patient, the use of multiple DIRs provides an estimation of the error-bars on the accumulated dose at any particular anatomical point (c.f. Fig. 4b) in a way akin to robustness analysis of treatment plans [48,49]. Thus, this provides a 'map' indicating where dose accumulation can be trusted, or where uncertainty is expected and thus care should be taken in interpreting sensitive dosimetric parameters such as single point dose minima or maxima. The use of several DIR algorithms in clinical practice is however only possible if multiple DIRs are efficiently implemented in a treatment planning system, with fast calculation times and a high degree of automation.

For the patients evaluated here we used DIBH to suppress intrafractional motion. We calculated the fraction dose on each repeated CT, assuming that the complete fraction could be applied in one breath-hold. This is clearly a simplification, as in clinical practice it typically takes two to three breath-holds to deliver a field. However, this is a valid approach to evaluate the dosimetric variation of using different DIR algorithms in the same patient images. Additionally, previous studies with these patient images showed a high geometrical reproducibility of DIBH from the same day [38,50].

In this study, the total accumulated dose is based on the results from nine repeated CTs only. We assume that these are representative for the anatomy during treatment because they have been acquired in the beginning, middle and end of treatment, and were not triggered by considerable visible external changes. Also, some clinical trials recently used hypofractionated particle therapy [51] with even less fractions for treating NSCLC patients.

Another limitation of this study is the lack of a ground truth for the DVFs. This is an intrinsic problem when working with real patient data. One method to achieve a ground truth is to compare a variety of anatomical landmarks, as for example in DIR-lab or MIDRAS [52] for 4D lung registrations. The drawback is the substantial work required by a medical doctor to define relevant and meaningful reference points. It is anyway a method with its own uncertainties, especially when analysing images from different days. Another possibility is to generate a ground truth by warping the CT with a DIR algorithm [30], which is a good representation of the anatomical status of the patient (patient specific numerical phantom). However, this has the disadvantage that the dose calculation and warping is not done on the original patient image. As our main goal was to evaluate the variations of clinically used DIRs directly on real patient images, no ground truth was available.

The result of each DIR strongly depends on the specific settings [53]. It has been shown that the result of a DIR differs as much between the same algorithm with different settings as between different algorithms in head and neck cancer patients [54], and it is likely that this also applies for other anatomical areas. For intra-fractional lung motion, Kadoya et al. [45] found differing DIR results between clinics in 4D-CTs of the lung even if the same software was used, which underlines the dependency of settings and procedures in the DIR process.

In our study, we did not use a focus or controlling ROI for all intensity-based DIR algorithms. To be consistent and comparable with all algorithms, we used the external contour as controlling ROI for RayStation Morfeus. The external is the easiest contour to get automatically in RayStation, without any manual contouring. This makes it a likely approach in clinical practice, as has been described by other institutions [55]. Nevertheless, the large variations between Morfeus and the other algorithms we found for patients 6 and 7 might be improved if different DIR settings are used. Indeed, the developers presented this DIR algorithm with multiple controlling ROIs, such as external, lungs, trachea and tumour [44,56], the delineation of which would be time consuming in clinical practice. However, the goal of this work was to quantify the dosimetric variation introduced by using different DIR algorithms and not rank the different DIR algorithms. For this, a fine tuning of the input parameters would be needed, which is not realistic in a standard clinical application and would be highly user dependent. Also, for most patients we did not see a prominent deviation between Morfeus and other algorithms. This shows the challenges in the tuning of individual DIR algorithms. If the algorithm was validated on a subset of these patients where it had a good agreement with other algorithms, or even a ground truth, it does not ensure that it will work out for all patients with the same diagnoses and in this same anatomical area. A fast and automated QA of DIR is therefore needed. Such QA methods have been proposed by analysing some properties of the DVF [57]. Additionally, also a QA on the image or dose level is desirable. These should not only check the principal applicability of an algorithm to an anatomical site, but also estimate the correctness of this DIR for each individual patient.

The variation of PTV-V95 degradations was 8.7% for the accumulated treatment doses and 7.9% for the individual warped frac-

tion doses. This shows that the dosimetric uncertainties introduced by DIR were not reduced over several fractions. The systematic character of these uncertainties might be specific for the type of anatomical changes we observed in our patient cohort. The dominating anatomical changes we observed were differences in the breath-hold position and tumour shrinkage, as an effect of treatment response. In particular, the latter is handled quite differently by the DIR algorithms. As the change is usually uni-directional (only shrinkage), DIR uncertainty here has a systematic character.

Finally, we like to emphasise that the uncertainty of DIR is only one of many uncertainties in proton therapy. Range and setup uncertainties are well quantifiable and can be included in the optimisation process [58]. RBE uncertainties, for instance, are much harder to quantify and it is still an ongoing discussion if a homogeneous RBE approximation is a good approach for protons [59]. Dose inaccuracies due to analytical dose calculation (as performed here) should also be considered, but have been shown to have a smaller impact on the dose distribution than anatomical changes [22]. The high impact of anatomical changes on the dose during treatment underlines the importance of fast plan adaptions and a correct dose accumulation during therapy.

In conclusion, we have analysed dosimetric uncertainties of different DIR algorithms for dose accumulation in lung cancer proton therapy. For the patients investigated here, the IMPT dose degradations caused by anatomical changes are larger than the variations introduced by different DIR algorithms. Nevertheless, we found substantial differences between different DIR algorithms of the fraction and accumulated doses. Using multiple DIR algorithms is a valuable approach to reduce DIR uncertainty for estimating the dosimetric differences caused by anatomical changes during proton treatment.

Conflict of interest

We have no conflicts of interest to declare.

Acknowledgements

We acknowledge the SNF (project:165961) and the ESTRO mobility grant. Also, we kindly acknowledge Djamal Boukerroui for data conversion and Robert Poel for support with checking structures. Finally, we thank the Plastimatch community for their fast support.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.radonc.2020.04.046.

References

[1] Baumann M, Krause M, Overgaard J, Debus J, Bentzen SM, Daartz J, et al. Radiation oncology in the era of precision medicine. Nat Rev Cancer 2016. https://doi.org/10.1038/nrc.2016.18.
[2] Hill-Kayser CE, Both S, Tochner Z, Greenberger JS. Proton therapy: ever shifting sands and the opportunities and obligations within. Front Oncol 2011. https://doi.org/10.3389/fonc.2011.00024.
[3] Weber DC, Schneider R, Goitein G, Koch T, Ares C, Geismar JH, et al. Spot scanning-based proton therapy for intracranial meningioma: long-term results from the Paul Scherrer Institute. Int J Radiat Oncol 2012;83:865–71. https://doi.org/10.1016/j.ijrobp.2011.08.027.
[4] Weber DC, Badiyan S, Malyapa R, Albertini F, Bolsi A, Lomax AJ, et al. Long-term outcomes and prognostic factors of skull-base chondrosarcoma patients treated with pencil-beam scanning proton therapy at the Paul Scherrer Institute. Neuro Oncol 2016;18:236–43. https://doi.org/10.1093/neuonc/nov154.
[5] Weber DC, Rutz HP, Pedroni ES, Bolsi A, Timmermann B, Verwey J, et al. Results of spot-scanning proton radiation therapy for chordoma and chondrosarcoma of the skull base: The Paul Scherrer Institut experience. Int J Radiat Oncol Biol Phys 2005;63:401–9. https://doi.org/10.1016/J.IJROBP.2005.02.023.
[6] Pehlivan B, Ares C, Lomax AJ, Stadelmann O, Goitein G, Timmermann B, et al. Temporal lobe toxicity analysis after proton radiation therapy for skull base tumors. Int J Radiat Oncol Biol Phys 2012;83:1432–40. https://doi.org/10.1016/j.ijrobp.2011.10.042.
[7] Mendenhall NP, Malyapa RS, Su Z, Yeung D, Mendenhall WM, Li Z. Proton therapy for head and neck cancer: rationale, potential indications, practical considerations, and current clinical evidence. Acta Oncol 2011;50:763–71. https://doi.org/10.3109/0284186X.2011.590147.
[8] Holliday EB, Frank SJ. Proton radiation therapy for head and neck cancer: a review of the clinical experience to date. Int J Radiat Oncol 2014;89:292–302. https://doi.org/10.1016/j.ijrobp.2014.02.029.
[9] Nguyen Q-N, Ly NB, Komaki R, Levy LB, Gomez DR, Chang JY, et al. Long-term outcomes after proton therapy, with concurrent chemotherapy, for stage II–III inoperable non-small cell lung cancer. Radiother Oncol 2015;115:367–72. https://doi.org/10.1016/j.radonc.2015.05.014.
[10] Wink KCJ, Roelofs E, Solberg T, Lin L, Simone CB, Jakobi A, et al. Particle therapy for non-small cell lung tumors: where do we stand? A systematic review of the literature. Front Oncol 2014;4. https://doi.org/10.3389/fonc.2014.00292.
[11] Harada H, Murayama S. Proton beam therapy in non-small cell lung cancer: state of the art. Dovepress 2017;8:141–5. https://doi.org/10.2147/LCTT.S117647.
[12] Gomez DR, Li H, Chang JY. Proton therapy for early-stage non-small cell lung cancer (NSCLC). Transl Lung Cancer Res 2018;7:199–204. https://doi.org/10.21037/tlcr.2018.04.12.
[13] Gorgisyan J, Munck af Rosenschold P, Perrin R, Persson GF, Josipovic M, Belosi MF, et al. Feasibility of pencil beam scanned intensity modulated proton therapy in breath-hold for locally advanced non-small cell lung cancer. Int J Radiat Oncol 2017;99:1121–8. https://doi.org/10.1016/j.ijrobp.2017.08.023.
[14] Grassberger C, Dowdell S, Sharp G, Paganetti H. Motion mitigation for lung cancer patients treated with active scanning proton therapy. Med Phys 2015;42:2462–9. https://doi.org/10.1118/1.4916662.
[15] Krieger M, Giger A, Jud C, Cattin PC, Salomir RV, Bieri O, et al. Liver-ultrasound based motion model for lung tumour tracking in PBS proton therapy. ICCR Proc 2019.
[16] Boye D, Lomax T, Knopf A. Mapping motion from 4D-MRI to 3D-CT for use in 4D dose calculations: a technical feasibility study. Med Phys 2013. https://doi.org/10.1118/1.4801914.
[17] Graeff C. Robustness of 4D-optimized scanned carbon ion beam therapy against interfractional changes in lung cancer. Radiother Oncol 2017;122:387–92. https://doi.org/10.1016/J.RADONC.2016.12.017.
[18] Boda-Heggemann J, Knopf A-C, Simeonova-Chergou A, Wertz H, Stieler F, Jahnke A, et al. Deep inspiration breath hold—based radiation therapy: a clinical review. Int J Radiat Oncol 2016;94:478–92. https://doi.org/10.1016/j.ijrobp.2015.11.049.
[19] Knopf A-C, Lomax A. In vivo proton range verification: a review. Phys Med Biol 2013;58:R131–60. https://doi.org/10.1088/0031-9155/58/15/R131.
[20] Kraan AC. Range verification methods in particle therapy: underlying physics and Monte Carlo modeling. Front Oncol 2015;5:1–27. https://doi.org/10.3389/fonc.2015.00150.
[21] Parodi K, Polf JC. In vivo range verification in particle therapy. Med Phys 2018;45:e1036–50. https://doi.org/10.1002/mp.12960.
[22] Nenoff L, Matter M, Geetanjli AJ, Winterhalter C, Gorgysian J, Josipovic M, et al. Daily adaptive proton therapy: is it appropriate to use analytical dose calculations for plan adaption? Submitt to Int J Radiat Oncol Biol Phys 2020. https://doi.org/10.1016/j.ijrobp.2020.03.036.
[23] Stützer K, Jakobi A, Bandurska-Luque A, Barczyk S, Arnsmeyer C, Löck S, et al. Potential proton and photon dose degradation in advanced head and neck cancer patients by intratherapy changes. J Appl Clin Med Phys 2017;18:104–13. https://doi.org/10.1002/acm2.12189.
[24] Van De Water S, Albertini F, Weber DC, Heijmen BJM, Hoogeman MS, Lomax AJ. Anatomical robust optimization to account for nasal cavity filling variation during intensity- modulated proton therapy : a comparison with conventional and adaptive planning strategies. Phys Med Biol 2018;63.
[25] Jaffray DA, Lindsay PE, Brock KK, Deasy JO, Tomé WA. Accurate accumulation of dose for improved understanding of radiation effects in normal tissue. Radiat Oncol Biol 2010;76:S135–9. https://doi.org/10.1016/j.ijrobp.2009.06.093.
[26] Sonke J-J, Aznar M, Rasch C. Adaptive radiotherapy for anatomical changes. Semin Radiat Oncol 2019;29:245–57. https://doi.org/10.1016/j.semradonc.2019.02.007.
[27] Chetty IJ, Rosu-Bubulac M. Deformable registration for dose accumulation. Semin Radiat Oncol 2019;29:198–208. https://doi.org/10.1016/j.semradonc.2019.02.002.
[28] Zhong H, Chetty IJ. Caution must be exercised when performing deformable dose accumulation for tumors undergoing mass changes during fractionated radiation therapy. Int J Radiat Oncol Biol Phys 2017;97:182–3. https://doi.org/10.1016/j.ijrobp.2016.09.012.
[29] Brock KK, Mutic S, McNutt TR, Li H, Kessler ML. Use of image registration and fusion algorithms and techniques in radiotherapy: Report of the AAPM Radiation Therapy Committee Task Group No. 132. Med Phys 2017;44:e43–76. https://doi.org/10.1002/mp.12256.
[30] Ribeiro CO, Knopf A, Langendijk JA, Weber DC, Lomax AJ, Zhang Y. Assessment of dosimetric errors induced by deformable image registration methods in 4D

pencil beam scanned proton treatment planning for liver tumours. Radiother Oncol 2018;128:174–81. https://doi.org/10.1016/j.radonc.2018.03.001.
[31] Zhang Y, Boye D, Tanner C, Lomax AJ, Knopf A. Respiratory liver motion estimation and its effect on scanned proton beam therapy. Phys Med Biol 2012;57:1779–95. https://doi.org/10.1088/0031-9155/57/7/1779.
[32] Sarrut D, Baudier T, Ayadi M, Tanguy R, Rit S. Deformable image registration applied to lung SBRT: Usefulness and limitations. Phys Medica 2017;44:108–12. https://doi.org/10.1016/j.ejmp.2017.09.121.
[33] Mogadas N, Sothmann T, Knopp T, Gauer T, Petersen C, Werner R. Influence of deformable image registration on 4D dose simulation for extracranial SBRT: a multi-registration framework study. Radiother Oncol 2018;127:225–32. https://doi.org/10.1016/j.radonc.2018.03.015.
[34] Samavati N, Velec M, Brock KK. Effect of deformable registration uncertainty on lung SBRT dose accumulation. Med Phys 2015;43:233–40. https://doi.org/10.1118/1.4938412.
[35] Jakobi A, Perrin R, Knopf A, Richter C. Feasibility of proton pencil beam scanning treatment of free-breathing lung cancer patients. Acta Oncol 2018;57:203–10. https://doi.org/10.1080/0284186X.2017.1355107.
[36] Matter M, Nenoff L, Meier G, Weber DC, Lomax AJ, Albertini F. IMPT plan generation in under ten seconds on a GPU. Acta Oncol 2019.
[37] Schaffner B, Pedroni E, Lomax A. 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 1999;44:27–41.
[38] Josipovic M, Persson GF, Dueck J, Bangsgaard JP, Westman G, Specht L, et al. Geometric uncertainties in voluntary deep inspiration breath hold radiotherapy for locally advanced lung cancer. Radiother Oncol 2016;118:510–4. https://doi.org/10.1016/J.RADONC.2015.11.004.
[39] Unser MA, Aldroubi A, Gerfen CR. A multiresolution image registration procedure using spline pyramids. Math Imaging Wavelet Appl Signal Image Process 1993;2034:160–70. https://doi.org/10.1117/12.162061.
[40] Wang H, Dong L, O'Daniel J, Mohan R, Garden AS, Kian Ang K, et al. Validation of an accelerated "demons" algorithm for deformable image registration in radiation therapy. Phys Med Biol 2005;50:2887–905. https://doi.org/10.1088/0031-9155/50/12/011.
[41] Kadoya N, Fujita Y, Katsuta Y, Dobashi S, Takeda K, Kishi K, et al. Evaluation of various deformable image registration algorithms for thoracic images. J Radiat Res 2014;55:175–82. https://doi.org/10.1093/jrr/rrt093.
[42] Kessler M, Pouliot J. White paper: deformable registration: What to ask when assessing the options 2013.
[43] RayStation. Deformable registration in raystation (White Paper) 2017:4.
[44] Velec M, Moseley JL, Svensson S, Härdemark B, Jaffray DA, Brock KK. Validation of biomechanical deformable image registration in the abdomen, thorax, and pelvis in a commercial radiotherapy treatment planning system. Med Phys 2017;44:3407–17. https://doi.org/10.1002/mp.12307.
[45] Kadoya N, Nakajima Y, Saito M, Miyabe Y, Kurooka M, Kito S, et al. Multi-institutional validation study of commercially available deformable image registration software for thoracic images. Int J Radiat Oncol 2016;96:422–31. https://doi.org/10.1016/j.ijrobp.2016.05.012.
[46] Van Herk M. Errors and margins in radiotherapy. Semin Radiat Oncol 2004;14:52–64. https://doi.org/10.1053/j.semradonc.2003.10.003.
[47] Sonke J-J, Belderbos J. Adaptive radiotherapy for lung cancer. Semin Radiat Oncol 2010;20:94–106. https://doi.org/10.1016/J.SEMRADONC.2009.11.003.
[48] Lowe M, Albertini F, Aitkenhead A, Lomax AJ, MacKay RI. Incorporating the effect of fractionation in the evaluation of proton plan robustness to setup errors. Phys Med Biol 2016;61:413–29. https://doi.org/10.1088/0031-9155/61/1/413.
[49] Malyapa R, Lowe M, Bolsi A, Lomax AJ, Weber DC, Albertini F. Evaluation of robustness to setup and range uncertainties for head and neck patients treated with pencil beam scanning proton therapy. Int J Radiat Oncol Biol Phys 2016;95:154–62. https://doi.org/10.1016/j.ijrobp.2016.02.016.
[50] Josipovic M, Aznar MC, Thomsen JB, Scherman J, Damkjaer SMS, Nygård L, et al. Deep inspiration breath hold in locally advanced lung cancer radiotherapy: validation of intrafractional geometric uncertainties in the INHALE trial. Br J Radiol 2019;92. https://doi.org/10.1259/bjr.20190569.
[51] Shirai K, Kawashima M, Saitoh JI, Abe T, Fukata K, Shigeta Y, et al. Clinical outcomes using carbon-ion radiotherapy and dose-volume histogram comparison between carbon-ion radiotherapy and photon therapy for T2b–4N0M0 non-small cell lung cancer-a pilot study. PLoS ONE 2017;12:1–14. https://doi.org/10.1371/journal.pone.0175589.
[52] Brock KK. Results of a Multi-Institution Deformable Registration Accuracy Study (MIDRAS). Int J Radiat Oncol Biol Phys 2010;76:583–96. https://doi.org/10.1016/j.ijrobp.2009.06.031.
[53] Miura H, Ozawa S, Nakao M, Furukawa K, Doi Y, Kawabata H, et al. Impact of deformable image registration accuracy on thoracic images with different regularization weight parameter settings. Phys Medica 2017;42:108–11. https://doi.org/10.1016/j.ejmp.2017.09.122.
[54] Zhang L, Wang Z, Shi C, Pi Y, Long T, Luo W, et al. Validation of deformable image registration algorithms for head and neck adaptive radiotherapy in routine clinical setting. Med. Phys., vol. 43, John Wiley & Sons, Ltd; 2016, p. 3342–3342. doi:10.1118/1.4955644.
[55] Zhang L, Wang Z, Shi C, Long T, Xu XG. The impact of robustness of deformable image registration on contour propagation and dose accumulation for head and neck adaptive radiotherapy. J Appl Clin Med Phys 2018;19:185–94. https://doi.org/10.1002/acm2.12361.
[56] Brock KK, Sharpe MB, Dawson LA, Kim SM, Jaffray DA. Accuracy of finite element model-based multi-organ deformable image registration. Med Phys 2005;32:1647–59. https://doi.org/10.1118/1.1915012.
[57] Kierkels RGJ, Den Otter LA, Korevaar EW, Langendijk JA, Van Der Schaaf A, Knopf AC, et al. An automated, quantitative, and case-specific evaluation of deformable image registration in computed tomography images. Phys Med Biol 2018;63. https://doi.org/10.1088/1361-6560/aa9dc2.
[58] Engelsman M, Schwarz M, Dong L. Physics controversies in proton therapy. Semin Radiat Oncol 2013. https://doi.org/10.1016/j.semradonc.2012.11.003.
[59] Paganetti H. Relative biological effectiveness (RBE) values for proton beam therapy. Variations as a function of biological endpoint, dose, and linear energy transfer. Phys Med Biol 2014;59. https://doi.org/10.1088/0031-9155/59/22/R419.