Effects of anatomical changes on pencil beam scanning proton plans in locally advanced NSCLC patients

DOI:
10.1016/j.radonc.2016.04.002

Document Version
Accepted author manuscript

Published in:
Radiotherapy and oncology : journal of the European Society for Therapeutic Radiology and Oncology

Effects of anatomical changes on pencil beam scanning proton plans in locally advanced NSCLC patients

Yenny Z. Szeto, Marnix G. Witte, Simon R. van Kranen, Jan-Jakob Sonke, José Belderbos, Marcel van Herk

Department of Radiation Oncology, The Netherlands Cancer Institute, Amsterdam, The Netherlands

Abstract

Background and Purpose. Daily anatomical variations can cause considerable differences between delivered and planned dose. This study simulates and evaluates these effects in spot-scanning proton therapy for lung cancer patients.

Materials and methods. Robust intensity modulated treatment plans were designed on the mid-position CT scan for sixteen locally advanced lung cancer patients. To estimate dosimetric uncertainty, deformable registration was performed on their daily CBCTs to generate 4DCT equivalent scans for each fraction and to map recomputed dose to a common frame.

Results. Without adaptive planning, eight patients had an undercoverage of the targets of more than 2GyE (maximum of 14.1GyE) on the recalculated treatment dose from the daily anatomy variations including respiration. In organs at risk, a maximum increase of 4.7GyE in the D_1 was found in the mediastinal structures. The effect of respiratory motion alone is smaller: 1.4GyE undercoverage for targets and less than 1GyE for organs at risk.

Conclusions. Daily anatomical variations over the course of treatment can cause considerable dose differences in the robust planned dose distribution. An advanced planning strategy including knowledge of anatomical uncertainties would be recommended

to improve plan robustness against interfractional variations. For large anatomical changes, adaptive therapy is mandatory.

Keywords: IMPT, NSCLC, interfractional anatomical changes, respiratory motion, deformable image registration

---

Introduction

When treating locally advanced Non-Small Cell Lung Cancer (NSCLC) patients, proton therapy may spare Organs-At-Risk (OARs) significantly better than photon therapy [1, 2]. However, the sensitivity to changes in range [3] makes it challenging to deliver the planned dose. The classical Planning Target Volume (PTV) does not necessarily improve the robustness of a treatment plan against geometrical uncertainties, since dose errors can occur inside or beyond the target volume. Therefore alternative methods such as worst case optimization [4, 5], minmax optmization [6] and stochastic programming [7, 8] have been proposed. These methods consider range and rigid setup uncertainties in their robust optimization, but typically disregard inter- and intrafraction anatomical changes and respiratory motion.

Studies handling anatomical changes have mainly concentrated on the effects of the respiratory motion [9–13] on the dose delivery. A definition of Internal Gross Tumor Volume (IGTV) [14] has been suggested to mitigate these effects. Others proposed re-scanning and tracking.

While Intra-Thoracic Anatomical Changes (ITACs) were observed in 72% of lung cancer patients during the course of radiotherapy [15], the published number of studies on the effects of interfractional anatomical changes on the proton dose is small. The few that reported on the interfracional anatomical changes made use of limited repeat Computed Tomography (CT) scans [3, 16–18]. However, the effects of anatomical changes on the dose based on data representative of the whole course of treatment have not been reported.

The purpose of this study is to evaluate the separate and combined effects of respiration and interfractional anatomical changes during the course of treatment on the dose delivery in spot scanning proton therapy using daily Cone-Beam CT (CBCT) scans. These scans capture daily variations as well as progressive anatomical changes.

Method and materials

Patient selection

In our clinic, about 5% of NSCLC patients are treated to a lower dose than the standard 66Gy (24×2.75Gy) Intensity Modulated Radiotherapy (IMRT) prescription due to normal tissue constraints, mainly high mean lung dose. With proton therapy, these patients might be treated to the intended dose while meeting normal tissue constraints. Therefore we have selected retrospectively and consecutively sixteen of these stage IIIA/B patients (Table S1) with available daily motion compensated (MC)CBCTs, [19], treated in 2012-2013.

Treatment planning

Spot-scanning proton plans were created using the Pinnacle^3 research version 9.100 Treatment Planning System (TPS) with a prescription dose of 66GyE (relative biological effectiveness = 1.1 [20]). Plans were designed, using one oblique and one lateral beam if possible, on the mid-position (MidP) scan [21] derived from the 4D-planning-CT (pCT). To improve robustness against respiratory induced motion, a density override (1 g/cm^3) was applied to an IGTV [10] generated by expansion of the GTV using the motion trajectory of the primary tumor derived from local tumor registration on the 4DCT scan. As the lymph nodes are embedded in the mediastinum, a density override was not applied here. In our clinic, no margin for microscopic disease is applied, i.e., GTV = Clinical Target Volume (CTV). The generated IGTV was used only for the dose calculation density grid and not for margin purposes. Plans were optimized (1) without constraining beam uniformity (labeled Intensity Modulated Proton Therapy (IMPT)),

and (2) using a single field uniformity tolerance (amount of deviation from uniformity tolerated as a percentage of the target dose) of 3% (labeled Single Field Uniform Dose (SFUD)). For both, a weighted scenario-based robustness tool [7, 8] in the TPS was used for optimization (9 scenarios: nominal plan, ±3mm in three directions and ±3% range shift) to account for proton therapy specific uncertainties. Unlike the 66Gy IMRT plan, both 66GyE IMPT and SFUD plans were within clinical OAR constraints for all cases. More detailed information on the robust planning can be found in the supplementary material.

Deformable Image Registration

For dose comparison and accumulation over the varying patient anatomies, two Deformable Image Registration (DIR) techniques were applied using in-house software.

The first one creates a 4D-Deformation Vector Field (DVF) that maps each phase of the 4DCT to its mid-position (4D-DVF) to generate the MidP-pCT. This DIR is based on an Iterative Multiscale Motion Estimation (IMME) technique using image phase, and the validation was done by Wolthaus et al [21].

The second method uses the 3D-DVF to map the MidP-pCT to each 3D-MCCBCT (3D-DVF), as the 3D-MCCBCT represents the mid-position during treatment with strongly reduced respiratory induced blurring [19]. It uses a cubic b-spline algorithm as the representation of the DVF, driven by a correlation ratio [22]. Rigidity and volume constraints [23, 24] were used as regularization terms. A gradient descent based multi-resolution optimization [25] was performed with a final control point spacing of 1cm. The DIR was started after a local rigid body registration of the bony anatomy (vertebrae). The shift and rotation applied to the bony anatomy match was not included in the DVF, i.e. the obtained DVF contains only the anatomical variables. The rigid component was assumed to be minimized by a couch correction. The DIR precision was previously validated for CBCT-to-CT registration in lung cancer patients [26], where an accuracy about 1.5±1mm in vector length was found. The generation of the mod-

ified CT (mCT) uses the inverse of this DVF, and was validated by Veiga et al [27]. Note that this method maps the Hounsfield unit (HU) of pCT, i.e. the tissue densities are assumed to be stable.

Dose evaluation

For dose evaluation, biologically equivalent doses were used as in our daily clinical practice. We used the Linear Quadratic (LQ) model to convert the physical doses to the biologically equivalent doses as given in fractions of 2Gy (EQD_2) [28], using specific α/β ratios for the different tissues. The recalculated dose on every scan was converted to the biologically equivalent dose before summation.

To estimate the effects of realistic anatomical variations on the delivered dose, the accumulated dose distribution over the entire treatment course of five weeks was compared with the planned dose distribution for both IMPT and SFUD plans.

Target coverage was evaluated using the difference in the minimum dose to 99% of the volume ΔD_{99} (α/β = 10GyE) between the planned dose and recalculated dose. We have set a difference $\leq$2GyE to be acceptable. Although any threshold is somewhat arbitrary, 2GyE equals the standard dose of one fraction dose, and thus represents a clinically relevant unit.

OAR doses were evaluated for each patient individually, including the maximum dose to 1% of the volume D_1 and mean dose D_{mean} for the heart (α/β = 3GyE) and mediastinal structures [29] (α/β = 3GyE), D_{mean} for the lungs-GTV (α/β = 3GyE [30]), D_1 for the spinal cord (α/β = 2GyE) and the percentage volume receiving 50Gy (V_{50}) for the esophagus (α/β = 10GyE).

The time between the pCT and the first treatment day can be up to two weeks. To distinguish early and late changes, and to evaluate the effects of shorter treatment schedules, we accumulated the recalculated dose also over the first week and first two weeks of the treatment.

Three dose evaluation methods were analyzed: (1) only respiratory motion, (2)

only day-to-day anatomy variations and (3) day-to-day anatomy variation combined with respiratory motion (summarized in Figure A2). Note that only regular respiration is evaluated given the amplitudes reported in Table S1.

1. Respiratory motion

The 4D-pCT, consisting of the ten phases of the respiratory cycle, represents the geometrical variations due to respiratory motion. To evaluate the effect of the respiratory motion, dose was recalculated on these ten phases. To compare with the planned dose, the dose distribution of each phase was deformed back to the MidP-pCT with the 4D-DVF, and accumulated.

2. Daily anatomy variations

Daily 3D-MCCBCT scans were acquired to capture day-to-day variations (e.g. posture changes and baseline shifts) and progressive anatomical changes, without respiratory motion [19] and without setup errors. Daily setup errors were corrected by imaging and patient re-alignment. Due to limitations with the CBCT field of view and lack of HU calibration, a mCT was generated with the anatomy of the 3D-MCCBCT using the 3D-DVF and the MidP-pCT. The effect of interfractional variations during the course of treatment is evaluated by recalculating the dose on the mCT. As an example, the pCT and a mCT with dose distributions are displayed in Figure A3. Subsequently, the dose distributions were deformed back to the MidP-pCT by applying the inverse of the 3D-DVF. The accumulation of these daily fraction doses represents an estimation of the total treatment dose, corrected for the number of treatment days without available CBCT images present.

3. Daily anatomy variations including respiratory motion

Our plans were optimized to be robust against respiratory motion and day-to-day variation. An overestimation of the robustness of the plans is likely to happen by evaluating these separately. Here we have combined these two components. It has been

found previously that the respiratory motion is more irregular intra- than interfractionally [31] and that the tumor trajectory shape is very stable [32]. Therefore, assuming a stable interfraction respiratory pattern is a reasonable first order approximation. For the construction of the 4D-mCT, the 4D-DVF of each phase was applied to the daily mCT. Next, the dose was recalculated on the daily 4D-mCT. For each phase, the dose was mapped back to the daily mCT, and accumulated. These daily fraction dose distributions were thereafter deformed back to the MidP-pCT for further accumulation.

To establish the dose accuracy of this methodology we performed a limited validation comparing dose recalculated on a CBCT and a repeat CT acquired on the same day (see supplementary materials). An average dose difference of -0.1 with range [−0.8–0.3] GyE was found between these two scans.

Results

Data from sixteen patients (Table S1) was used. Patient 1 was treated for 2 weeks only, thus end of treatment was after two weeks in this case. Patient 3 was withdrawn due to failure of deformable registration. Patient 15 had no SFUD plan as a result of limitation of the beam size. No significant difference ($p > 0.05$, t-test) between IMPT and SFUD plans (Figures 1 and 3) was found in terms of robustness against anatomical variations.

Target coverage

The median of the ΔD_{99} was close to zero in all three cases (Figure 1). The size of the 25-75 percentile box and min-max whiskers for evaluation 1, i.e. only respiratory motion, are considerably smaller than for the other two evaluations. Eight patients had a ΔD_{99} of more than 2GyE in at least one of the targets at end of treatment, with a maximum of 12GyE (Figure 2). The ΔD_{99} of the dose accumulated over the first week and the first two weeks are closer to each other than to that of the dose accumulated over the whole course of treatment.

Organs-at-risk

Compared to the effects on the targets, the differences in the accumulated dose and planned dose were small in the OARs (Figure3). The largest difference was found in the mediastinal structures, where an increase of 4.7Gy in the D_1 was observed for an IMPT plan in evaluation 2. The respiratory motion has only a small influence on the dose distribution, as is the case for the target coverage.

Discussion

Daily anatomical variations, occurring over the course of treatment, can cause considerable dose differences in the planned dose distribution. We built an infrastructure that allows evaluation of the effect of daily anatomical variations and the accumulated effects of respiratory motion on spot scanning proton delivery. To our knowledge, this is the first study published on this topic using 4DCT and daily CBCTs.

During the course of treatment, anatomical changes should not lead to unacceptable underdosage of the targets or an unacceptable overdosage of the OARs. The IGTV concept applied here leads to plans that are fairly robust against respiratory motion alone. We observed a much larger effect on the accumulated dose due to day-to-day variations, as the anatomy of lung cancer patients changes significantly over a time span of five weeks. According to Kwint et al [15], 55% of the ITACs occurred in the first week. Our results show that eight out of fifteen patients have ITACs in the first week, which was scored during their IMRT treatment according to our clinical protocol for photons [15], see Table S1. Although half of the patients had code red (dangerous) or orange (risky), none had an adaptive plan. The scoring mainly concentrates on the influence of the ITACs on the GTV position, therefore it is not always predictive to our outcome. For example, patient 8 had code green (safe), but the GTV had an underdosage of about 5GyE due to changing bone position. Another frequently occurring ITAC is atelectasis. Since protons are more sensitive to density variation than photons,

dose variations in the targets will be larger when atelectasis changes along the path of the proton beams. Considering that robust planning will not be able to cope with such large density variations, adaptive planning is required in these cases.

The patients selected for this study were treated with IMRT to a lower dose than our standard of 66Gy due to OAR constraints. Proton plans of 66GyE on the other hand, met all OAR dose constraints but were susceptible to anatomical changes. As IMRT plans are typically more robust against anatomical changes, it is currently unclear which patients would have benefited from proton therapy over photon therapy. Our results show that adaptive radiotherapy and/or advanced planning strategies capable of producing plans that are robust against anatomical changes will be needed to provide the full benefit of scanning beam proton therapy for advanced lung cancer patients.

Other studies reported similar observations of effects of interfractional anatomical changes on proton delivery for lung cancer patients, mainly with passive scattering, using only one repeat CT or weekly CTs, and excluding respiratory motion [3, 16, 18]. Here, they all disregarded the possibility of underdosage of the targets over the whole course of treatment. One case reported on interfractional variations using daily CTs and dose accumulation over the whole course of treatment [33], but using the less complicated prostate patient group, concluding that their IMPT plans are robust to interfractional variations.

We disregarded the interplay effect between interfraction motion and spot delivery time [12, 13] for simplicity. The use of repainting strategies [34] can reduce this effect. Besides, this effect was reported to be much smaller than the effect of anatomical variations [35], and by using IGTV, the targets were adequately covered by the proton beam [36]. The boxplots in our results indeed show that the effect of the respiratory motion on the dose distribution is minimal. However, we did not take into account the irregularity of breathing and the density changes in lung, which could make the IGTV

concept less robust [37]. While we expect that these two effects will have little influence on the dose compared to the very large interfractional variation, further studies should be performed including the interplay effect and irregular respiratory motion to confirm these expectations.

In the planning strategy, plans were generated by robust optimization instead using PTV-based plans, as Liu et al. [5] found that robustness and OARs sparing improved by using robust optimization. The results could differ quantitatively when using PTV-based plans, but qualitatively anatomical changes are still expected to affect the dose distribution unacceptably. Note that the plans were robust against setup error (3mm) which was eliminated for using a simulated online correction protocol. So this study essentially tested if plans designed to be robust against setup and range uncertainties were also robust against anatomical changes, and the results suggest that this is not the case.

We attempted to create robust IMPT and SFUD plans using the same number of beams (two), and the same IMPT and SFUD beam setup for a patient to make comparison easier. For the robust optimization an additional target margin was used, as the target coverage objectives were not reached when using GTV alone (see supplementary material). Using more beams could result in less dose in normal tissue, while using a different beam setup could also create a more robust plan. The same applies for the used SFUD parameters which were not rigorously optimized. We also did not push the objectives to their limits for the optimization. So despite the fact that we used robust optimization, we could very likely have achieved a ”better” plan.

Concerning DIR, deforming CTs to CBCT anatomy is only an approximation of the anatomy of the patient. Unfortunately, no CBCT and CT were acquired simultaneously to test the accuracy of our entire methodology. However, an accuracy of about 1.5mm±1mm in vector length was found by Abdoli et al. [26], which is small relative to the day-to-day geometric uncertainties. They also concluded that this method is less

accurate in patients with large deformations, especially in case of atelectasis. The DIR has difficulties in creating/removing a large volume of tissue as required when atelectasis occurs/disappears, and the tissue-to-tissue correspondence will not be possible. Due to this, the dose differences in patients with large differences in the amount of atelectasis might be underestimated, since the difference between the mCT and the pCT will be smaller than it actually is. In the extreme case, such as for Patient 3, the atelectasis was totally resolved from the first CBCT, while it was largely present in the pCT. The mCTs deviate visibly from the CBCTs and were therefore not representative of the patient's anatomy on the treatment days. The dosimetric impact of the registration error was estimated on one patient where a repeat CT was available. Another limitation of this method is the HU accuracy. It uses the HU of the pCT and does not anticipate changes in densities [38], which is reasonable for most soft tissues, but not fully correct for lung or tumor tissues over the course of treatment. As a result, the deformed CT will slightly underestimate anatomical changes. The same applies for the 4D-mCT, where the lung density changes during breathing. On the other hand, by using the pCT to create the mCTs, we excluded potential effects of contrast in the analyses.

Another assumption is that the tumor shrinkage is elastic [39], which is not the case for all tumor types.

Although the method based on daily CBCT and deformable registration has its limitations, the results do give an approximation of the effect of interfractional and respiratory variations on the dose distributions for a group of potential proton therapy patients. Further studies should be carried out to support our preliminary findings. Using the available tools in the current planning strategy, anatomical changes can cause considerable dose differences between the planned dose and the delivered dose. As setup errors can be minimized by image guidance, the focus of these robustness tools should be broadened towards robustness against respiratory motion, and especially anatomical variations. For some cases, adaptive radiotherapy is inevitable, but an advanced

planning strategy including the knowledge of the anatomical uncertainties will also be needed to improve plan robustness against interfractional variations.

Acknowledgment

The authors acknowledge Angela Tijhuis for help with treatment planning and Philips Healthcare for technical support.

[1] Nichols RC, Huh SH, Hoppe BS, et al. Protons safely allow coverage of high-risk nodes for patients with regionally advanced non-small-cell lung cancer. Technol Cancer Res Treat 2011;10:317–22.

[2] Nichols RC, Huh SN, Henderson RH, et al. Proton Radiation Therapy Offers Reduced Normal Lung and Bone Marrow Exposure for Patients Receiving Dose-Escalated Radiation Therapy for Unresectable Stage III Non-Small-Cell Lung Cancer: A Dosimetric Study. Clin Lung Cancer 2011;12:252–257.

[3] Hui Z, Zhang X, Starkschall G, et al. Effects of interfractional motion and anatomic changes on proton therapy dose distribution in lung cancer. Int J Radiat Oncol Biol Phys 2008;72:1385–95.

[4] Pflugfelder D, Wilkens JJ, Oelfke U. Worst case optimization: a method to account for uncertainties in the optimization of intensity modulated proton therapy. Phys Med Biol 2008;53:1689–700.

[5] Liu W, Zhang X, Li Y, Mohan R. Robust optimization of intensity modulated proton therapy. Med Phys 2012;39:1079–91.

[6] Fredriksson A, Forsgren A, Hårdemark B. Minimax optimization for handling range and setup uncertainties in proton therapy. Med Phys 2011;38:1672.

[7] Unkelbach J, Chan TCY, Bortfeld T. Accounting for range uncertainties in the optimization of intensity modulated proton therapy. Phys Med Biol 2007;52:2755–73.

[8] Unkelbach J, Bortfeld T, Martin BC, Soukup M. Reducing the sensitivity of IMPT treatment plans to setup errors and range uncertainties via probabilistic treatment planning. Med Phys 2009;36:149.

[9] Engelsman M, Rietzel E, Kooy HM. Four-dimensional proton treatment planning for lung tumors. Int J Radiat Oncol Biol Phys 2006;64:1589–95.

[10] Kang Y, Zhang X, Chang JY, et al. 4D Proton treatment planning strategy for mobile lung tumors. Int J Radiat Oncol Biol Phys 2007;67:906–14.

[11] Knopf AC, Hong TS, Lomax A. 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 2011;56:7257–71.

[12] Munck af Rosenschöld P, Aznar MC, Nygaard DE, et al. A treatment planning study of the potential of geometrical tracking for intensity modulated proton therapy of lung cancer. Acta Oncol 2010;49:1141–8.

[13] Grassberger C, Dowdell S, Lomax A, et al. Motion Interplay as a Function of Patient Parameters and Spot Size in Spot Scanning Proton Therapy for Lung Cancer. Int J Radiat Oncol 2013;86:380–386.

[14] Chang JY, Dong L, Starkschall G, et al. Image-Guided Radiation Therapy for Non-small Cell Lung Cancer. J Thorac Oncol 2008;3:177–186.

[15] Kwint M, Conijn S, Schaake E, et al. Intra thoracic anatomical changes in lung cancer patients during the course of radiotherapy. Radiother Oncol 2014;113:392–397.

[16] Shi W, Nichols Jr RC, Flampouri S, et al. CLINICIAN ' S PERSPECTIVE Tumour Shrinkage during Proton-based Chemoradiation for Non – small-cell Lung Cancer May Necessitate Adaptive Replanning during Treatment. Hong Kong J Radiol 2011;14:190–4.

[17] Shi W, Nichols Jr RC, Flampouri S, et al. Proton-based chemoradiation for synchronous bilateral non-small-cell lung cancers: A case report. Thorac Cancer 2013;4:198–202.

[18] Koay EJ, Lege D, Mohan R, Komaki R, Cox JD, Chang JY. Adaptive/nonadaptive proton radiation planning and outcomes in a phase II trial for locally advanced non-small cell lung cancer. Int J Radiat Oncol Biol Phys 2012;84:1093–100.

[19] Rit S, Wolthaus JWH, van Herk M, Sonke JJ. On-the-fly motion-compensated cone-beam CT using an a priori model of the respiratory motion. Med Phys 2009;36:2283.

[20] Paganetti H, Niemierko A, Ancukiewicz M, et al. Relative biological effectiveness (RBE) values for proton beam therapy. Int J Radiat Oncol Biol Phys 2002;53:407–21.

[21] Wolthaus JWH, Sonke JJ, van Herk M, Damen EMF. Reconstruction of a time-averaged midposition CT scan for radiotherapy planning of lung cancer patients using deformable registration. Med Phys 2008;35:3998.

[22] van Kranen S, Mencarelli A, van Beek S, Rasch C, van Herk M, Sonke JJ. Adaptive radiotherapy with an average anatomy model: evaluation and quantification of residual deformations in head and neck cancer patients. Radiother Oncol 2013;109:463–8.

[23] Staring M, Klein S, Pluim JPW. A rigidity penalty term for nonrigid registration. Med Phys 2007;34:4098.

[24] Loeckx D, Slagmolen P, Maes F, Vandermeulen D, Suetens P. Nonrigid Image Registration Using Conditional Mutual Information. Med Imaging, IEEE Trans 2010;29:19–29.

[25] Mattes D, Haynor DR, Vesselle H, Lewellen TK, Eubank W. PET-CT image registration in the chest using free-form deformations. IEEE Trans Med Imaging 2003;22:120–8.

[26] Abdoli M, van Kranen SR, Sonke JJ. Validation of a Deformable Image Registration for Adaptive Radiotherapy of Lung Cancer. In: Radiother. Oncol. 3rd ESTRO FORUM. Elsevier Inc; 2015, p. 5491.

[27] Veiga C, Lourenço AM, Mouinuddin S, et al. Toward adaptive radiotherapy for head and neck patients: Uncertainties in dose warping due to the choice of deformable registration algorithm. Med Phys 2015;42:760–9.

[28] Joiner M, Bentzen MS. Fractionation: The Linear Quadratic Approach. In: Joiner M, Van der Kogel A, editors. Basic Clin. Radiobiol.; chap. 8; 4th ed. London: Hodder Arnold; 2009, p. 102–119.

[29] van Elmpt W, De Ruysscher D, van der Salm A, et al. The PET-boost randomised phase II dose-escalation trial in non-small cell lung cancer. Radiother Oncol 2012;104:67–71.

[30] Borst GR, Ishikawa M, Nijkamp J, et al. Radiation pneumonitis after hypofractionated radiotherapy: evaluation of the LQ(L) model and different dose parameters. Int J Radiat Oncol Biol Phys 2010;77:1596–603.

[31] Rit S, van Herk M, Zijp L, Sonke JJ. Quantification of the variability of diaphragm motion and implications for treatment margin construction. Int J Radiat Oncol Biol Phys 2012;82:399–407.

[32] Sonke JJ, Lebesque J, van Herk M. Variability of four-dimensional computed tomography patient models. Int J Radiat Oncol Biol Phys 2008;70:590–8.

[33] Wang Y, Efstathiou Ja, Sharp GC, Lu HM, Frank Ciernik I, Trofimov AV. Evaluation of the dosimetric impact of interfractional anatomical variations on prostate proton therapy using daily in-room CT images. Med Phys 2011;38:4623.

[34] Zenklusen SM, Pedroni E, Meer D. 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 2010;55:5103–5121.

[35] Li Y, Kardar L, Li X, et al. On the interplay effects with proton scanning beams in stage III lung cancer. Med Phys 2014;41:021721.

[36] Chang JY, Zhang X, Wang X, et al. Significant reduction of normal tissue dose by proton radiotherapy compared with three-dimensional conformal or intensity-modulated radiation therapy in Stage I or Stage III non-small-cell lung cancer. Int J Radiat Oncol Biol Phys 2006;65:1087–96.

[37] Koybasi O, Mishra P, St James S, Lewis JH, Seco J. Simulation of dosimetric consequences of 4D-CT-based motion margin estimation for proton radiotherapy using patient tumor motion data. Phys Med Biol 2014;59:853–67.

[38] Bertelsen A, Schytte T, Bentzen SrM, Hansen O, Nielsen M, Brink C. Radiation dose response of normal lung assessed by Cone Beam CT - a potential tool for biologically adaptive radiation therapy. Radiother Oncol 2011;100:351–5.

[39] Sonke JJ, Belderbos J. Adaptive radiotherapy for lung cancer. Semin Radiat Oncol 2010;20:94–106.

[40] Feuvret L, Noël G, Mazeron JJ, Bey P. Conformity index: a review. Int J Radiat Oncol Biol Phys 2006;64:333–42.

[41] Wang X, Zhang X, Dong L, et al. Effectiveness of noncoplanar IMRT planning using a parallelized multiresolution beam angle optimization method for paranasal sinus carcinoma. Int J Radiat Oncol Biol Phys 2005;63:594–601.

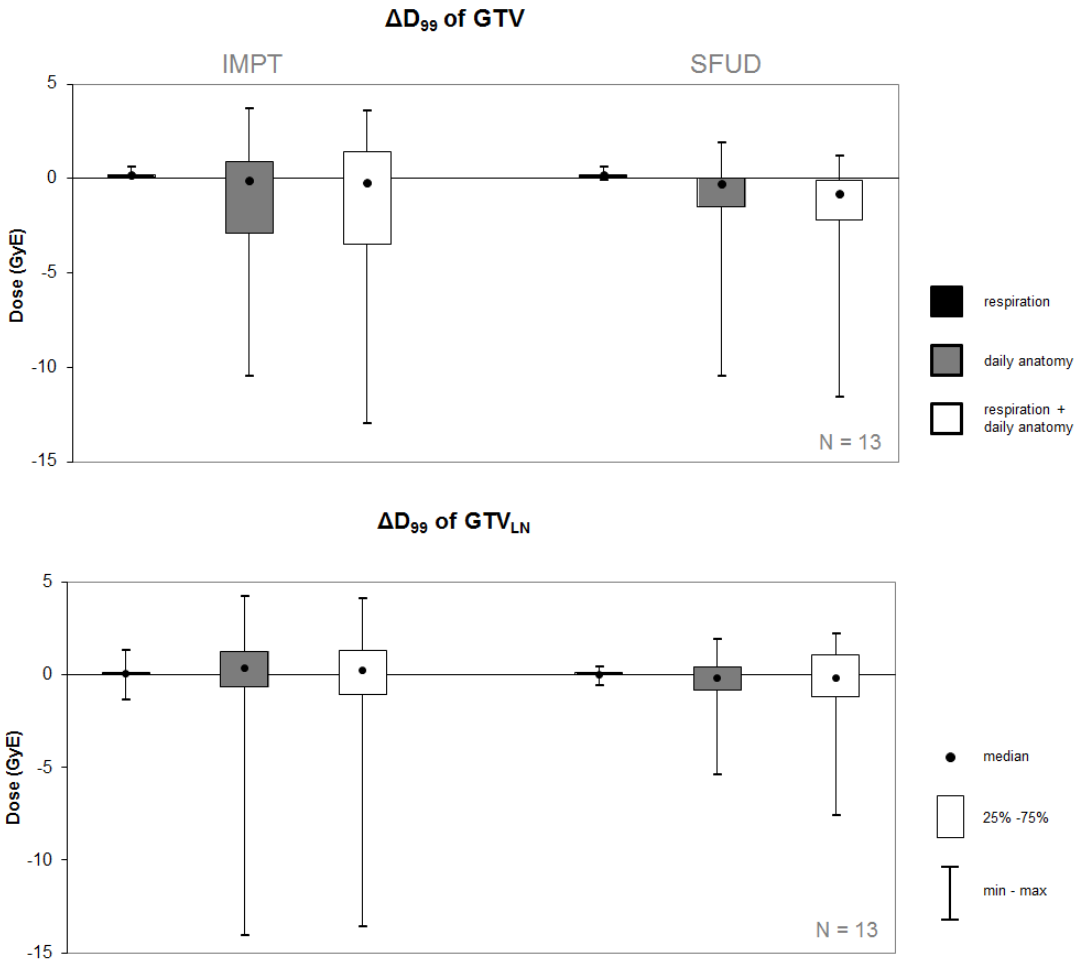

Figure 1: Boxplots of the difference in minimum dose ΔD_{99} between the planned dose (IMPT and SFUD) and the three different recalculated treatment doses (respiratory motion, daily anatomy variations and daily anatomy variations including respiratory motion) for both the primary tumor (GTV, top) and involved lymph nodes (GTV_{LN}, bottom). Patient 3 and Patient 15 were withdrawn from both plots. Patient 5 had no GTV and not included in the GTV analyses, while Patient 16 had no GTV_{LN} and thus excluded in the GTV_{LN} analyses.

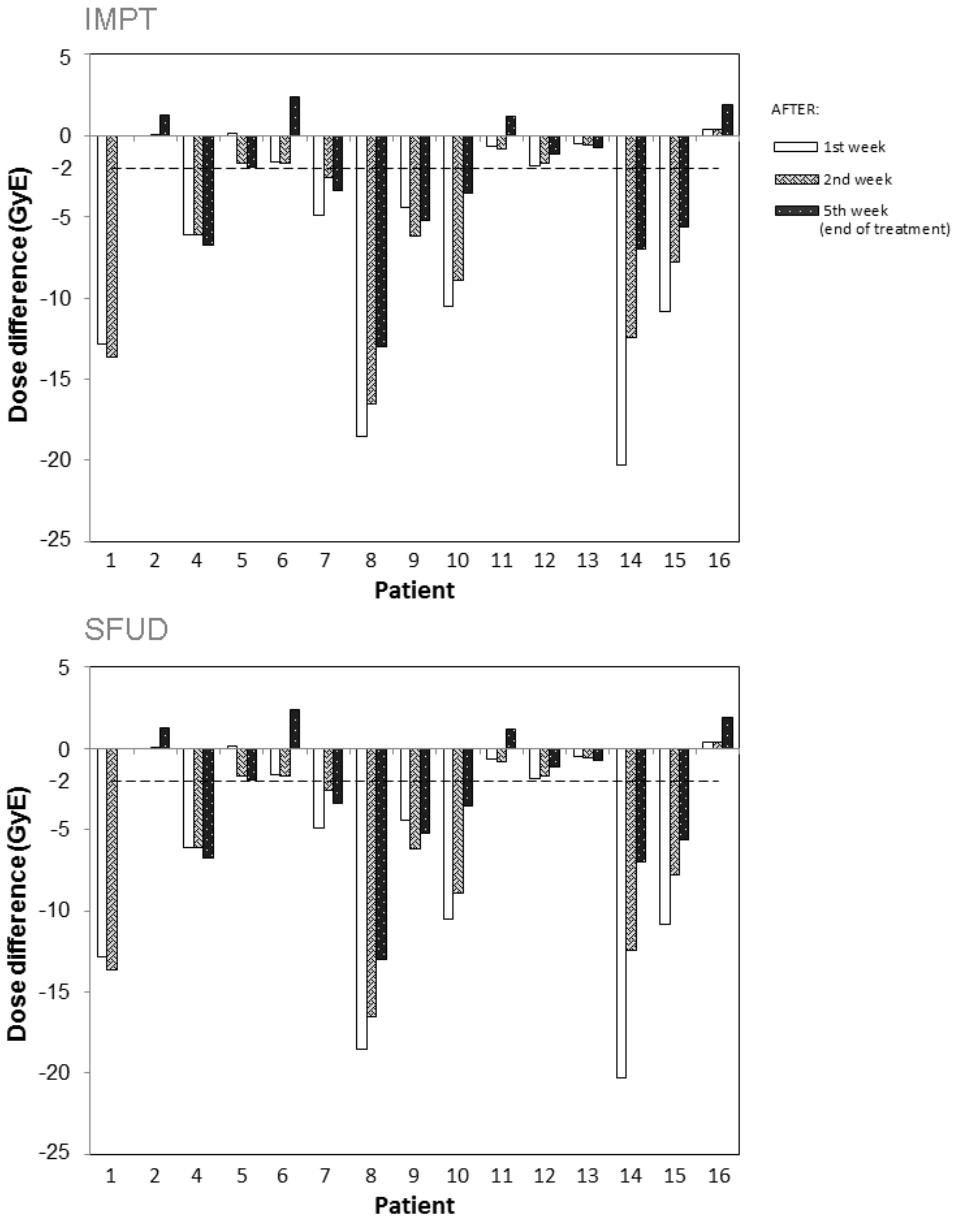

Figure 2: Barplots of the difference in minumum dose ΔD_{99} between the planned dose and the recalculated treatment dose accumulated over the first week, first two weeks and the whole course of treatment, for the day-to-day variations including respiratory motion in the worst target (primary tumor (GTV) or one of the involved lymph nodes (GTV_{LN})) of 15 patients for both IMPT and SFUD plans. Patient 1 had a treatment of only two weeks. Patient 3 was withdrawn from the analyses. Patient 15 only has an IMPT plan. In the supplementary material, figures A4 and A5 show respectively the barplots of GTV and worst GTV_{LN}.

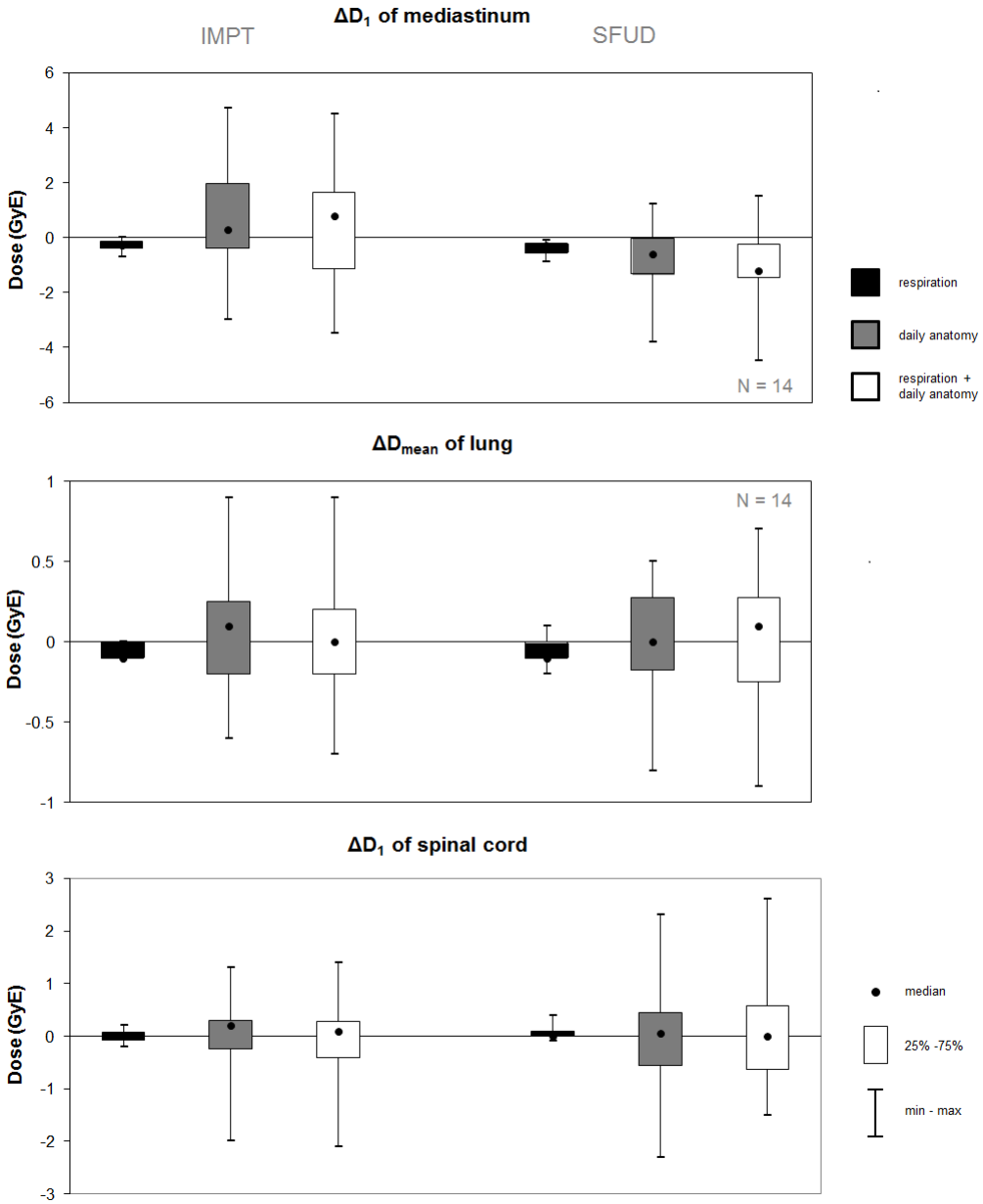

Figure 3: Boxplots of the difference in the mean lung dose D_{mean} and the D_1 of the spinal cord and the mediastinum between the planned dose and the three recalculated treatment doses (respiratory motion, daily anatomy variations and daily anatomy variations including respiratory motion) for IMPT and SFUD plans of the 14 patients. Patient 3 and Patient 15 were withdrawn from the analyses.

Supplementary material

Patient characteristics for the selected patients in this study can be found in table S1. A schematic summary of the three different evaluation methods used in this study is shown in figure A2. Figure A3 displays two examples of the planned dose on the pCT and the recalculated dose on the mCT. Figure 2 can be split in two: (1) GTV (Figure A4) and (2) worst GTV_{LN} (Figure A5).

Dosimetric accuracy of modified CT

CBCT to CT registration was used to generage a modified CT with a reported accuracy of about 1.5±1mm [26]. To estimate the dosimetric accuracy of this approach, we performed a side by side comparison between a 4D MC-CBCT and a 4D repeat CT (rCT), both acquired on the same day. As none of the patients in our study received a rCT, we selected a NSCLC patient (Patient A) treated in 2015 with IMRT. A large tumor shift was observed during the course of treatment, thus a rCT was acquired to create an adaptive plan.

An IMPT plan was made on its initial MidP pCT using the same method as described in this study. We used the 3D-DVF to map the MidP pCT to the 3D MC-CBCT to generate the mCT, and the 4D-DVF to construct the 4D mCT. Unfortunately , the field of view of this acquired CBCT does not cover the external of the patient. Therefore, the repeat CT was used to cover the missing parts of the mCT, as dosimetric effects due to anatomical changes outside the field of view are out of the scope of this paper. Figure A1 shows the scans of patient A.

Dose was recalculated on each phase of the 4D mCT using the IMPT plan and deformed back to the mid-position before dose accumulation. The same dose recalculation and dose accumulation steps were performed for the 4D rCT. For comparison purposes, the accumulated dose distributions were deformed to the pCT, and evaluated with the same dose parameters as reported in this study using the pCT contours. The results are shown in Table S2. The average dose difference was -0.1 with range [-0.8

- 0.3] GyE. The differences are induced by the limitations of our simulation model as well as anatomical differences between the CBCT and the repeat CT (Figure A1I)

Note that there are some limitations in this procedure. Although the repeat 4DCT and the CBCT are performed on the same day, there is an hour gap between the two scans at different scanners. Therefore the 4DCT and the CBCT are similar to each other, but are not identical (Figure A1(H)). The field of view of the CBCT in this case is more centered in the patient than normally. In most cases, the field of view of the CBCT is more centered in the ipsilateral lung, therefore, unlike for Patient A, the external of that side is included in the scan. Further studies with more data are needed to provide a more accurate estimation of the dosimetric accuracy of the error due to the registration error.

Robust treatment planning

The TPS utilized a weighted scenario-based robustness tool as described earlier. This robust optimization was performed using default settings of the TPS, where the scenarios were weighted as follows: 25% for range, 25% for setup in 3 directions and 50% for nominal. The weight is a multiplier of the objective function of each scenario and the optimizer minimizes the weighted sum of the composite objective functions. As this tool is based on the weighted sum of costs from the different scenarios, full coverage of all scenarios could not be attained by optimizing on the GTVs without an additional help structure. To reach the target coverage objectives, we performed the optimization on the Planning Target Volumes (PTVs) that were already clinically used for IMRT planning. The PTV of the primary tumor is defined as the GTV plus a margin (0.25 $\times$ peak-to-peak amplitude + 12mm), whereas the lymph node (GTV_{LN}) was expanded by 12mm to PTV_{LN}. Note that the PTV is only used as a supporting structure to ensure sufficient GTV coverage and no other attempt was made to find a different supporting structure since the GTV was covered. To evaluate the robustness of each plan, we used the additional 8 scenarios utilized by the robust optimization,

i.e., 6 plans where the isocenter of the beams are shifted in three directions (±anterior-posterior, ±left-right and ±cranial-caudal) and 2 plans where a range correction of ±3% is applied. The D_{99} of the targets, i.e., GTVs and GTV_{LN}, should be at least 95% of the prescribed dose to be clinically acceptable. We increased the shifts of the isocenter by 1mm from 3mm until we found $D_{99} < 62.7$GyE in at least one of the targets in one direction or range correction, as this would not be clinically acceptable anymore. Note that we did not increase the range correction. Figure A6 illustrates the robustness of the different plans using this approach. For example, the SFUD plan of Patient 10 is robust up to 8mm shift, while the SFUD plan of Patient 13 is robust up to 4mm. However, Figure 2 shows that Patient 10 has more target undercoverage than Patient 13. Therefore, the plans do not appear to be overly robust despite using the PTV to optimize. Between the IMPT and SFUD plans in terms of plan robustness, no statistically significant difference ($p = 0.57$, t-test) was found.

Due to the help structure used in plan optimization, it is difficult to define a proper conformity index [40] to analyze the dose conformity. Instead we evaluated the volume receiving more than 95% of the prescribed dose. To compare IMPT and SFUD plans, we calculated the ratio of this volume of these plans (Table S3):

$$HDR = \frac{V_{\text{IMPT}}}{V_{\text{SFUD}}} \tag{1}$$

This ratio indicates which plan achieved a robust plan with less high dose volume using the same objectives and constraints. Although for 10 patients out of 16, the SFUD plans needed a larger high dose volume than the IMPT plans (HDR < 1) to achieve a robust plan, no statistically significant difference (p=0.096, Wilcoxon signed ranks test) was found.

To compare the beam homogeneity between the IMPT and SFUD plans, we have calculated a heterogeneity index (HI) [41] per beam (see Table S3). The HI is defined as follows:

$$HI = \frac{|D_5 - D_{95}|}{D_{\text{mean}}} \tag{2}$$

where D_5 is the minimum dose to 5% of the target volume, D_{95} is the minimum dose to 95% of the target volume and D_{mean} is the mean dose. The smaller the HI, the more homogeneous the beam is. The SFUD beams yielded significantly more homogeneous dose than the IMPT beams, with $p < 0.001$ (t-test).

Table S1: Patient characteristics, including stage, volume of the primary tumor, volume of the total involved lymph nodes, number of involved lymph nodes (LN), peak-to-peak amplitude of the primary tumor, days between the CT scan (t_{CT}) and first treatment day (t_1), number of available CBCTs of each patient, and the intra-thoracic anatomical changes (ITACs) over course of treatment [15]. The ITACs are classified in four color codes: red (dangerous), orange (risky), yellow (cautious) and green (safe). Note that this classification was developed for photons and would be different in the context of protons. The primary tumor of Patient 5 was not visible anymore after induction chemotherapy and was therefore not treated. In this particular case, the peak-to-peak amplitude was derived from two lymph nodes. Patient A is included for the dose error estimation due to the registration error. The values in the parentheses are the values found for the repeat CT.

patient	stage	GTV vol (cc)	GTV_{LN} vol (cc)	# LN	peak-peak ampl (cm) LR	CC	AP	$t_1 - t_{CT}$ (days)	# CBCT	intra-thoracic anatomical changes type	code
1	IIIB	127.7	32.4	6	0.1	1.2	0.5	11	7	atelectasis development	orange
2	IIIA	72.7	42.8	2	0.1	1.0	0.4	11	23	atelectasis development & changing bone position	green
3	IIIB	309.7	22.3	2	0.1	0.8	0.1	8	23	atelectasis totally resolved	orange
4	IIIB	254.3	82.9	4	0.1	1.0	0	6	17	noticable tumor regression	orange
5	IIIB	-	44.2	8	0.1	1.1	0.2	13	18	very small anatomy changes	green
6	IIIB	12.3	233.1	3	0.1	0.1	0.2	5	23	tumor regression	orange
7	IIIB	18.8	89.1	11	0.3	1.8	0.3	12	22	shift mediastinum	green
8	IIIB	88.1	4.9	1	1.2	1.5	0.5	8	21	atelectasis development	orange
9	IIIB/IV	23.9	67.9	8	0.1	0.7	0.1	15	21	changing bone position	green
10	IIIA	60.4	21.7	3	0.1	0.7	0.2	13	17	changing body contour	orange
11	IIIA	129.0	13.5	2	0.3	0.6	0.2	9	17	very small anatomy changes	green
12	IIIA	42.1	29.3	3	0.2	1.7	0.3	11	17	very small anatomy changes	green
13	IIIB	4.2	98.4	6	0.1	0.2	0.3	13	17	minor tumor regression	green
14	IIIB	145.0	69.1	4	0.1	0.9	0.2	7	16	changing body contour and bone position	yellow
15	IIIB	500.7	65.7	9	0.2	2.0	0.4	12	19	moving atelectasis and tumor regression	orange
16	IIIA	212.8	-	0	0.3	0.1	0.2	5	17	disappearance of atlectasis	red
A	IIA	13.9 (15.1)	10.9 (9.1)	1	0.2 (0.2)	0.5 (1.0)	0.3 (0.6)	11	24	tumor shift	orange

Table S2: Dose evaluation of the planned dose, the recalculated dose on the 4D mCT and the recalculated dose on the 4D rCT for Patient A

	GTV	GTV_{LN}	Lungs-GTV	Heart		Spinal cord	Esophagus
	D_{99} (GyE)	D_{99} (GyE)	D_{mean} (GyE)	D_1 (GyE)	D_{mean} (GyE)	D_1 (GyE)	V_{50} (%)
pCT	64.6	64.3	6	13.5	0.4	0.1	0
4D mCT	39.5	63.3	5.7	22.3	0.7	0.2	0
4D rCT	39.4	63.5	5.4	23.1	0.7	0.3	0

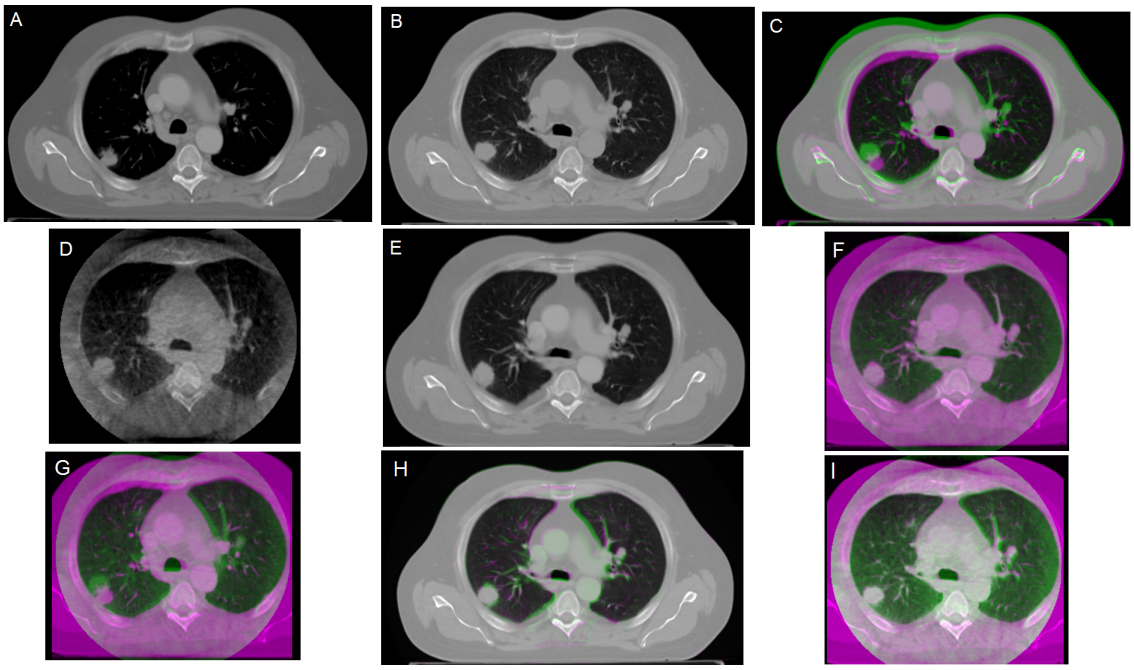

Figure A1: The four scans of Patient A: planning CT (A), repeat CT (B), CBCT (D) and mCT (E). And overlays of scans for better visualization: planning CT (purple) and repeat CT (green) (C), mCT (purple) and CBCT (green) (F), planning CT (purple) and CBCT (green) (G), repeat CT (purple) and mCT (green) (H), and repeat CT (purple) and CBCT (green) (I)

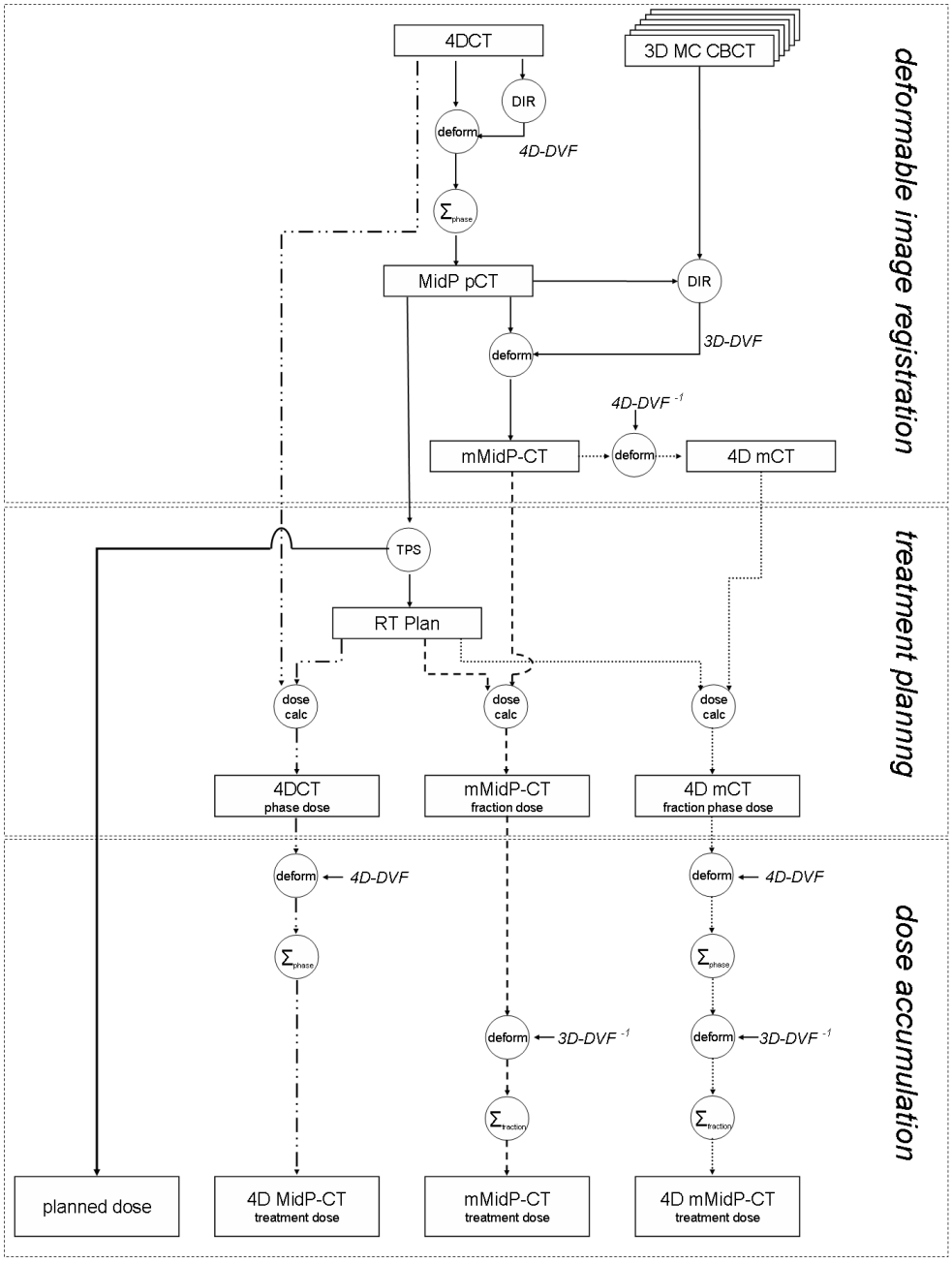

Figure A2: Summary of the three different evaluation methods to evaluate the effects of respiratory and anatomical variations: only respiratory motion (4D MidP-CT), daily anatomy variations without respiratory motion (mMidP-CT) and daily anatomy variations including respiratory motion (4D mMidP-CT). 4D-DVF represents the deformation vector field mapping each phase of the 4D CT to its mid-position (MidP). 3D-DVF deforms the planning CT to the daily anatomy of the CBCT. TPS is the optimization in the treatment planning system, whereas "dose calc" refers to dose recalculation in the same system.

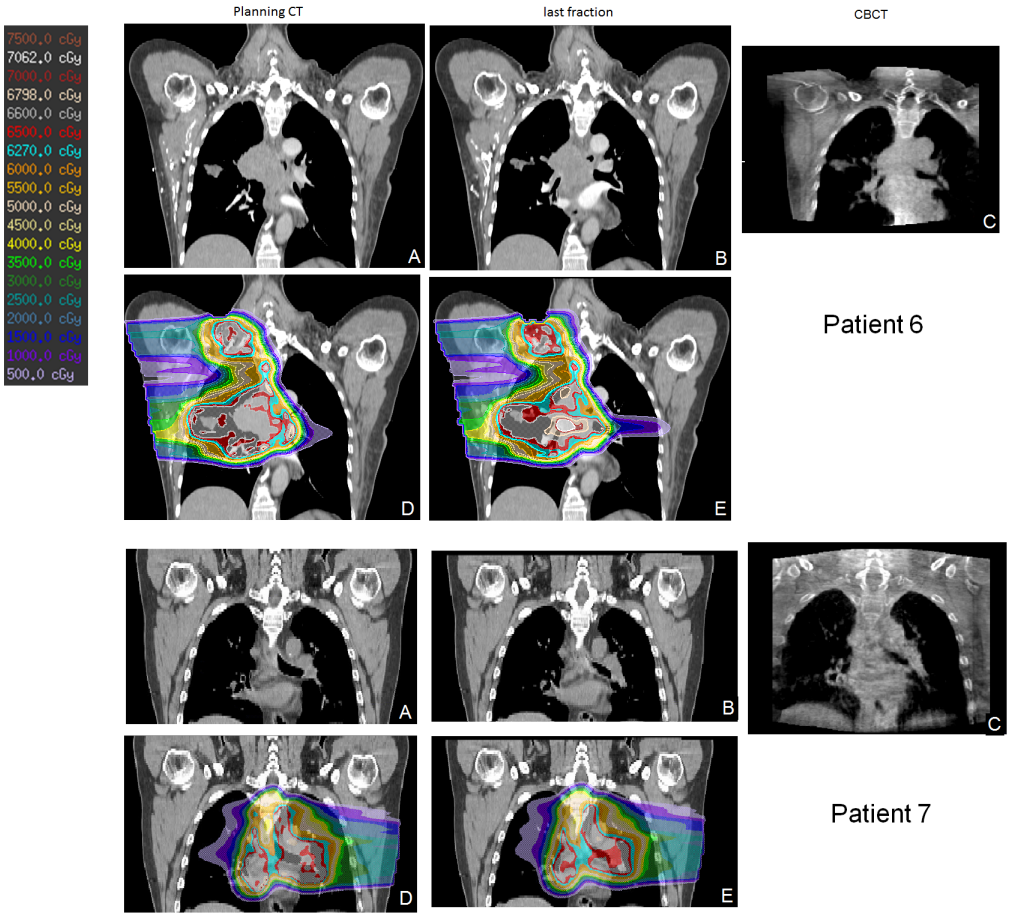

Figure A3: Planning CT (A) and the modified CT (B) of the last fraction of Patient 6 (top) and Patient 7 (bottom) with the planned dose (D) and the recalculated dose (E) on the modified CT. Daily CBCT (C) is used to deform the planning CT to the modified CT.

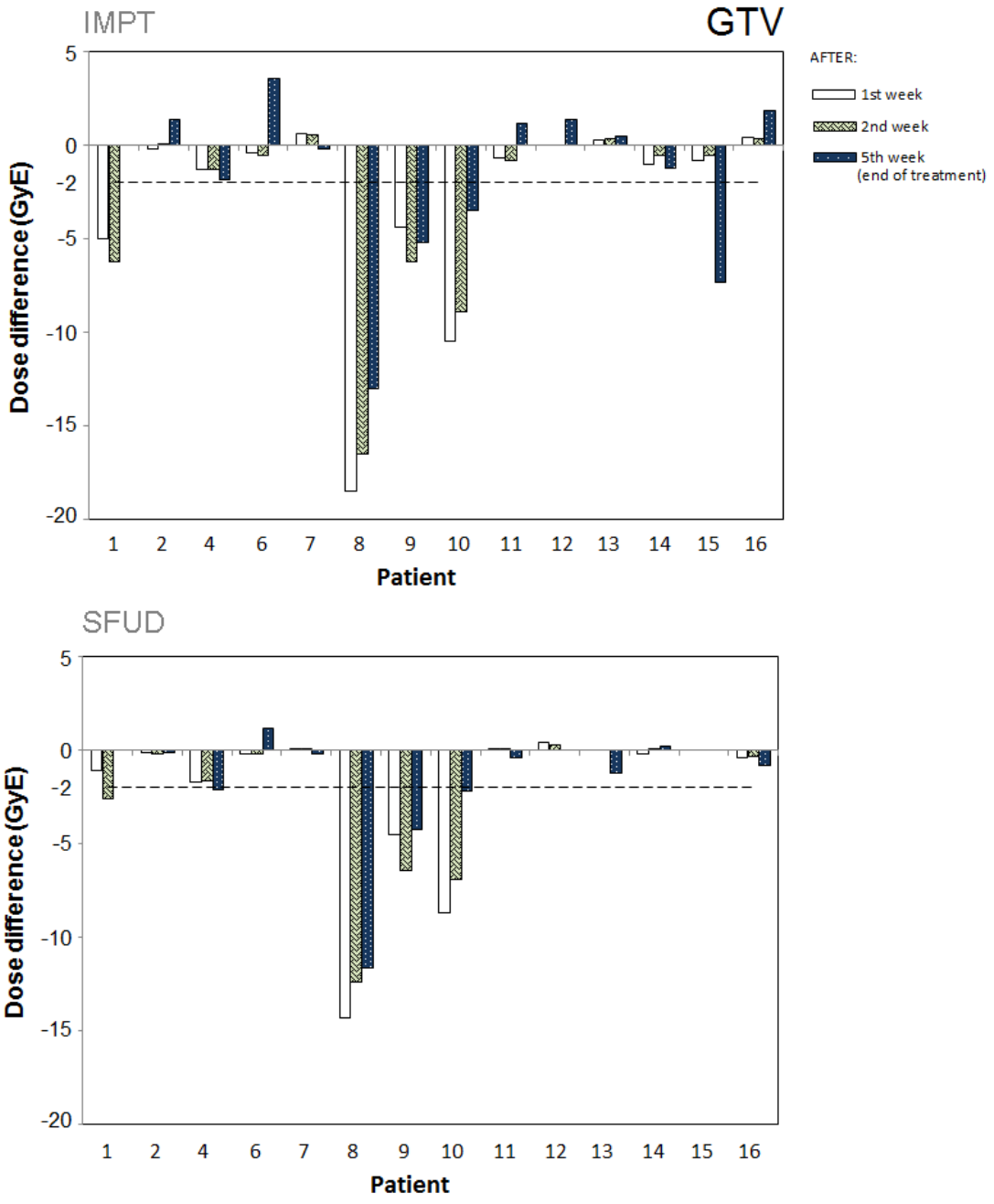

Figure A4: Barplots of the difference in minumum dose ΔD_{99} in the GTV between the planned dose and the recalculated treatment dose accumulated over the first week, first two weeks and the whole course of treatment, for the day-to-day variations including respiratory motion of 15 patients for both IMPT and SFUD plans. Patient 1 was treated only for two weeks. Patient 15 only had an IMPT plan.

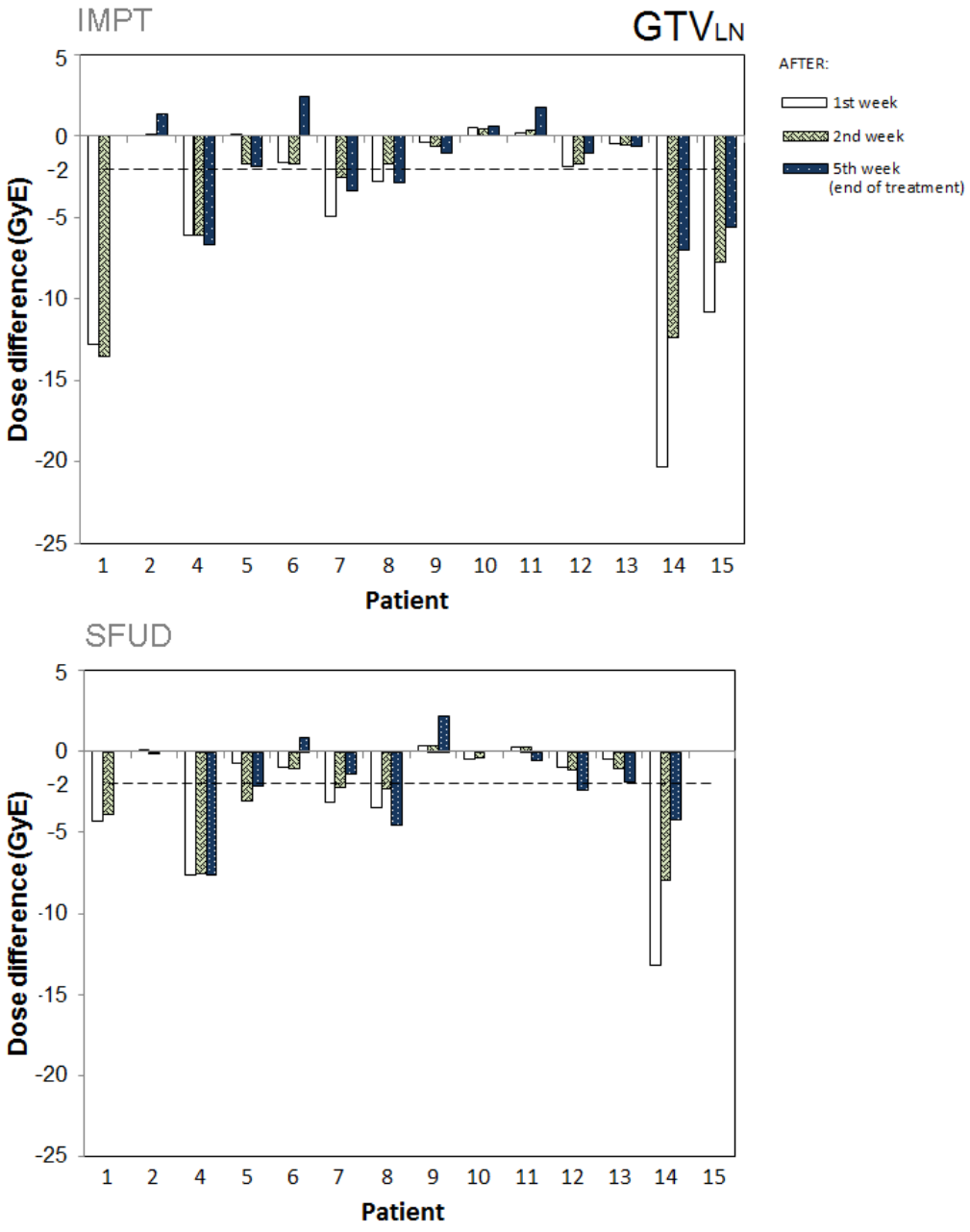

Figure A5: Barplots of the difference in minumum dose ΔD_{99} in the worst GTV_{LN} between the planned dose and the recalculated treatment dose accumulated over the first week, first two weeks and the whole course of treatment, for the day-to-day variations including respiratory motion of 15 patients for both IMPT and SFUD plans. Patient 1 was treated only for two weeks. Patient 15 only had an IMPT plan.

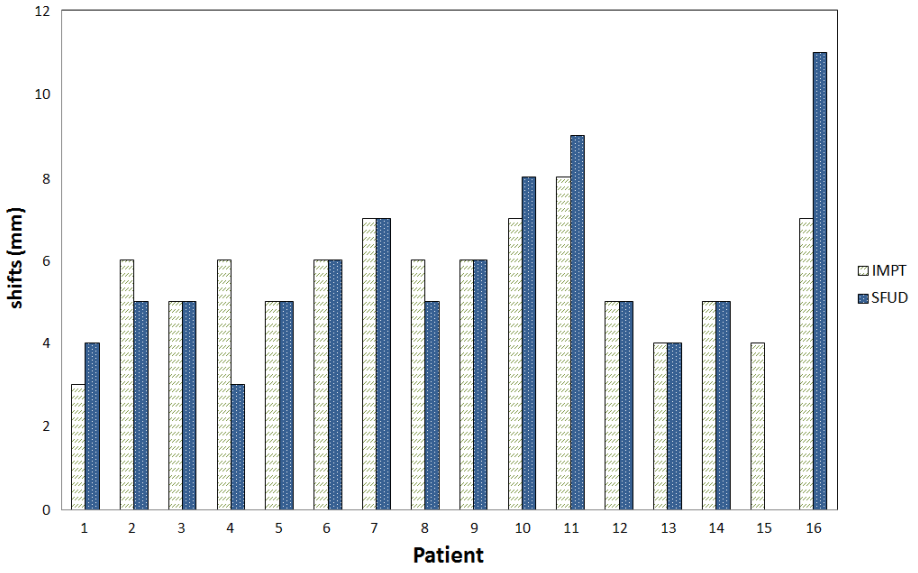

Figure A6: Robustness evaluation of the IMPT and SFUD plans for 16 patients, showing the shift that is needed to have an undercoverage ($D_{99} < 62,7$GyE) in a target. The isocenter of the beams were shifted from the nominal plan in three directions or a range correction was applied. For each plan of a patient, only one case is shown, i.e. worst target (either GTV or GTV_{LN}) and worst direction (a shift or range correction). The isocenter was shifted from 3mm to 11mm, while only a range correction of ±3% was applied.

Table S3: High Dose Ratio (HDR) and beam Heterogeneity Index (HI) of the different plans. Patient 15 is disregarded, because of the absence of the SFUD plan

		HI			
		IMPT		SFUD	
patient	HDR	beam1	beam2	beam1	beam2
1	1.03	1.32 (0.28)	1.35 (0.37)	1.04 (0.01)	1.05 (0.02)
2	0.97	1.21 (0.18)	1.24 (0.20)	1.05 (0.01)	1.07 (0.01)
3	1.00	1.55 (0.75)	1.35 (0.41)	1.08 (0.07)	1.05 (0.04)
4	1.09	2.07 (0.83)	1.70 (0.47)	1.06 (0.01)	1.04 (0.01)
5	0.92	1.25 (0.17)	1.32 (0.25)	1.04 (0.02)	1.06 (0.02)
6	0.97	1.45 (0.21)	1.57 (0.24)	1.05 (0.02)	1.05 (0.02)
7	0.97	1.29 (0.10)	1.40 (0.32)	1.05 (0.03)	1.08 (0.02)
8	0.98	3.57 (3.32)	1.57 (0.63)	1.07 (0.03)	1.07 (0.02)
9	0.93	1.82 (0.77)	1.23 (0.21)	1.07 (0.03)	1.03 (0.01)
10	0.98	1.29 (0.15)	1.28 (0.17)	1.04 (0.02)	1.05 (0.02)
11	0.97	1.37 (0.35)	1.30 (0.31)	1.06 (0.01)	1.05 (0.02)
12	1.02	1.90 (0.82)	1.38 (0.25)	1.06 (0.01)	1.05 (0.01)
13	0.91	1.22 (0.08)	1.23 (0.09)	1.05 (0.02)	1.06 (0.02)
14	1.00	2.26 (0.81)	2.37 (0.43)	1.07 (0.04)	1.07 (0.01)
16	0.99	1.59 (-)	1.87 (-)	1.06 (-)	1.07 (-)
mean		1.676	1.477	1.056	1.056
sd		0.619	0.309	0.003	0.003