International Journal of Radiation Oncology biology • physics

www.redjournal.org

Physics Contribution

Dose Uncertainties in IMPT for Oropharyngeal Cancer in the Presence of Anatomical, Range, and Setup Errors

Aafke C. Kraan, PhD,* Steven van de Water, MSc,* David N. Teguh, MD, PhD,* Abrahim Al-Mamgani, MD, PhD,* Tom Madden, PhD,† Hanne M. Kooy, PhD,† Ben J.M. Heijmen, PhD,* and Mischa S. Hoogeman, PhD*

**Erasmus MC Daniel den Hoed Cancer Center, Rotterdam, The Netherlands; and †Massachusetts General Hospital and Harvard Medical School, Boston, Massachusetts*

Received Apr 22, 2013, and in revised form Sep 6, 2013. Accepted for publication Sep 8, 2013.

Summary

We quantified the impact of treatment-related uncertainties on delivered dose in intensity modulated proton therapy for oropharyngeal cancer patients by simulating approximately 3700 treatments of 10 patients. Large dose deviations can occur especially when various errors are combined. We quantified the effect of adaptive planning on improvement of treatment quality and investigated how plan robustness changes when more beam directions are included.

Purpose: Setup, range, and anatomical uncertainties influence the dose delivered with intensity modulated proton therapy (IMPT), but clinical quantification of these errors for oropharyngeal cancer is lacking. We quantified these factors and investigated treatment fidelity, that is, robustness, as influenced by adaptive planning and by applying more beam directions.

Methods and Materials: We used an in-house treatment planning system with multicriteria optimization of pencil beam energies, directions, and weights to create treatment plans for 3-, 5-, and 7-beam directions for 10 oropharyngeal cancer patients. The dose prescription was a simultaneously integrated boost scheme, prescribing 66 Gy to primary tumor and positive neck levels (clinical target volume-66 Gy; CTV-66 Gy) and 54 Gy to elective neck levels (CTV-54 Gy). Doses were recalculated in 3700 simulations of setup, range, and anatomical uncertainties. Repeat computed tomography (CT) scans were used to evaluate an adaptive planning strategy using nonrigid registration for dose accumulation.

Results: For the recalculated 3-beam plans including all treatment uncertainty sources, only 69% (CTV-66 Gy) and 88% (CTV-54 Gy) of the simulations had a dose received by 98% of the target volume (D98%) >95% of the prescription dose. Doses to organs at risk (OARs) showed considerable spread around planned values. Causes for major deviations were mixed. Adaptive planning based on repeat imaging positively affected dose delivery accuracy: in the presence of the other errors, percentages of treatments with D98% >95% increased to 96% (CTV-66 Gy) and 100% (CTV-54 Gy). Plans with more beam directions were not more robust.

Conclusions: For oropharyngeal cancer patients, treatment uncertainties can result in significant differences between planned and delivered IMPT doses. Given the mixed causes for major deviations, we advise repeat diagnostic CT scans during treatment, recalculation of the dose, and if required, adaptive planning to improve adequate IMPT dose delivery.

Reprint requests to: Aafke C. Kraan, PhD, Department of Radiation Oncology, Erasmus MC-Daniel den Hoed Cancer Center, Groene Hilledijk 301, 3075 EA Rotterdam, The Netherlands. Tel: (31) 10 704 17 48; E-mail: aafke.kraan@pi.infn.it

A.C. Kraan is currently at Istituto Nazionale di Fisica Nucleare Sezione di Pisa, Pisa, Italy.

Conflict of interest: none.

Supplementary material for this article can be found at www.redjournal.org.

Int J Radiation Oncol Biol Phys, Vol. 87, No. 5, pp. 888–896, 2013
0360-3016/$ - see front matter
http://dx.doi.org/10.1016/j.ijrobp.2013.09.014

Introduction

Intensity modulated proton therapy (IMPT) uses proton pencil beams with varying position and energy, whose intensities are optimized individually. Oropharyngeal tumors form an attractive site for IMPT because of their complex shape and proximity to organs at risk (OARs). Improved OAR sparing is expected compared to intensity modulated radiation therapy (IMRT) (1, 2), potentially resulting in reduced negative side effects. In IMPT, however, uncertainties may exist that cause the delivered dose to seriously deviate from planning and thus reduce the benefit of IMPT. These uncertainties include anatomical changes (eg, organ motion, changes in air cavities, tumor regression, weight loss), patient setup errors and range uncertainties from uncertainties in computed tomography (CT) Hounsfield units (HU), conversion of HUs into stopping power; and reconstruction artifacts (3-8).

Various works discuss IMPT treatment uncertainties (2-4, 6, 8) and mitigation by robust optimization (4, 9-16). Many studies focusing on dose impacts of treatment uncertainties (3, 4, 6, 8) show one or few patient cases, do not include oropharyngeal patients, or do not include combined effects of treatment uncertainties. Robust optimization studies (4, 9-16) typically present results obtained with robust treatment plans (ie, to confirm the effectiveness of a particularly robust algorithm. In this context, for head-and-neck cancer patients, a recent article by Liu et al (9) focused on differences between various robust optimization strategies to compensate range and patient setup errors. However, the impact of anatomical changes and realistic simulation of all treatment-related uncertainties was not presented. Simone et al (2) compared adaptive IMRT with IMPT for head-and-neck cancer patients but did not investigate the impact of range and setup errors. Despite the concerns for robustness of IMPT for head-and-neck cancer and the interest to treat head-and-neck cancer with IMPT, a realistic and comprehensive overview of the effects of treatment uncertainties on the delivered dose is still lacking. Furthermore, such studies are needed, a priori, to assess the required level of robustness mitigation in clinical practice.

This study presents a realistic and accurate (3700 simulations from 10 oropharyngeal cancer patients) analysis of the effects of treatment uncertainties on delivered dose. First, we studied the impact of anatomy, range, and setup errors, separately and in combination with simulating realistic treatments. Moreover, we evaluated the effect of range and setup errors of different magnitudes. Results for target and OARs are readily interpreted in terms of dose−volume histogram (DVH) parameters. Second, we investigated 2 approaches, not yet reported, to making treatments robust and applicable independent of robust optimization by posing the questions: (*1*) how is treatment precision affected by adaptive planning ie, replanning based on repeated imaging, as recommended in IMRT (17)?; and (*2*) does plan robustness improve by choosing more beam directions, as postulated by Unkelbach et al (4)?

Methods and Materials

Patient group

We included 10 prospectively selected oropharyngeal cancer patients treated at Erasmus Medical Center in 2004 and 2005, aged 48 to 83 years old (mean, 60 years old) (Table 1). Each patient had a planning and a repeat CT scan, the latter at 46 Gy. Scans were enhanced by administration of intravenous contrast (100 mL of Omnipaque, 647 mg of iohexol per mL [see Discussion]). The primary tumor was delineated by a radiation oncologist on each CT. Contouring of OARs was assisted by atlas-based autosegmentation (18) and edited if needed. Artifact reduction software (19) was used for 5 patients with metal teeth implants. The dose prescription was a simultaneously integrated boost scheme, prescribing 66 Gy to the primary tumor and positive neck levels (CTV-66 Gy) and 54 Gy to elective neck levels (CTV-54 Gy) in 30 IMPT fractions. CTV volumes are shown in Table 1. We defined the planning target volumes (PTV)-66 Gy and PTV-54 Gy as the CTV-66 Gy and CTV-54 Gy with a 5-mm margins. Being aware that margin recipes are not fully valid for IMPT, we recall that this would be sufficient in IMRT to account for setup errors of 1.5 mm standard deviation on systematic and random components (1, 20). We quantified the amount of anatomical change between the 2 CTs as the average deformation of salivary glands and primary tumor measured in 6 directions, as suggested by Vasquez-Osorio et al (21) (Table 1, last column).

Proton delivery system

We used the pencil beam scanning techniques from Hong et al (22) and Kooy et al (23). The proton energy ranged from 45 MeV to 230 MeV. Range shifters were applied when needed for lower energies. The distance between energy layers was 2 mm in water. The full width at half the maximum (FWHM) of the pencil beam ranged from 9 mm at 230 MeV to 21 mm at 45 MeV (numbers for 5 cm of air excluding range shifters [22, 23]).

Planning technique

Treatment planning used an in-house-developed IMPT treatment planning system. The proton dose calculation algorithm from Hong et al (22) and Kooy et al (23) was integrated in an automated plan generation platform called iCycle (24, 25). It performed multicriteria optimization, which optimized objectives sequentially according to a list of prioritized objectives and hard constraints, identical for all-patients. The minimal PTV dose was a hard constraint, while prioritized objectives included (in order of priority) minimizing maximum PTV dose, target conformality, and minimization of dose to all OARs and unspecified tissues (see Supplementary Fig. E1 for details). We applied a newly developed method of iterative resampling of randomly placed pencil beams (26). The dose grid resolution was $3 \times 3 \times 3$ mm^3.

For each patient, treatment plans with 3, 5, and 7 beam directions were created for planning and repeat CT.

The 3-beam plan is the basis against which dose deviations and adaptive planning strategy were quantified and has coplanar beams at $-50°$, $50°$, and $180°$ as published previously (1). We generated plans with 5 and 7 beam directions with uniformly spaced angles, starting at $0°$, to investigate robustness versus more beams. Examples of dose distributions are shown in Figure 1.

Simulation and quantification of dose deviations

Anatomical uncertainties

These were modeled by calculating the treatment plan dose for fractions 1 to 15 on the planning CT and for fractions 16 to 30 on the repeat CT. Dose accumulation for OARs and target, both subject to deformation, of the first and second part of the treatment was performed with an in-house-developed nonrigid registration method (21).

Table 1 Summary of patient characteristics

Patient	Site	TNM staging	Planning CT: Volume CTV-66 Gy (cm^3)	Planning CT: Volume CTV-54 Gy (cm^3)	Repeat CT: Volume CTV-66 Gy (cm^3)	Repeat CT: Volume CTV-54 Gy (cm^3)	Average deformation* (mm)
1	Base of tongue	T1N2c	106	199	82	176	4.2
2	Base of tongue	T3N2a	99	313	73	268	3.3
3	Tonsil	T2N1	43	165	37	156	2.5
4	Tonsil	T2N0	11	77	9	75	3.1
5	Soft palate	T2N0	14	72	10	67	2.4
6	Base of tongue	T3N2a	68	221	47	191	6.6
7	Tonsil	T2N0	5	67	5	72	2.3
8	Tonsil	T1N1	41	95	35	87	2.5
9	Base of tongue	T3N3	178	343	132	296	6.6
10	Base of tongue	T1N2c	70	294	63	252	4.5
Patient average			63	185	49	164	3.8

Abbreviations: CT = computed tomography; CTV-66 Gy = Clinical Target Volume receiving 66 Gy (primary tumor and positive neck levels); CTV-54 Gy: Clinical Target Volume receiving 54 Gy (elective neck levels).

* Average deformation was measured in 6 directions as suggested by Vasquez Osorio et al (21).

Range uncertainties

These were simulated by recalculating the treatment plans on a CT scan with all CT numbers systematically increased or decreased by a given percentage of its value, as proposed by Lomax et al (3). We investigated shifts in CT numbers ranging from −5% to +5% with steps of 1%. Negative (positive) values represent an overshoot (undershoot) of the Bragg peak.

Setup uncertainties

We simulated an online patient setup correction protocol with bony anatomy as reference. Ideally, setup errors are zero after applying online corrections. Residual errors, however, remain from intrafractional motion or uncorrected deformations of the bony structures (7). These were simulated by applying rigid isocenter shifts in anteroposterior (AP), superior-inferior (SI), and

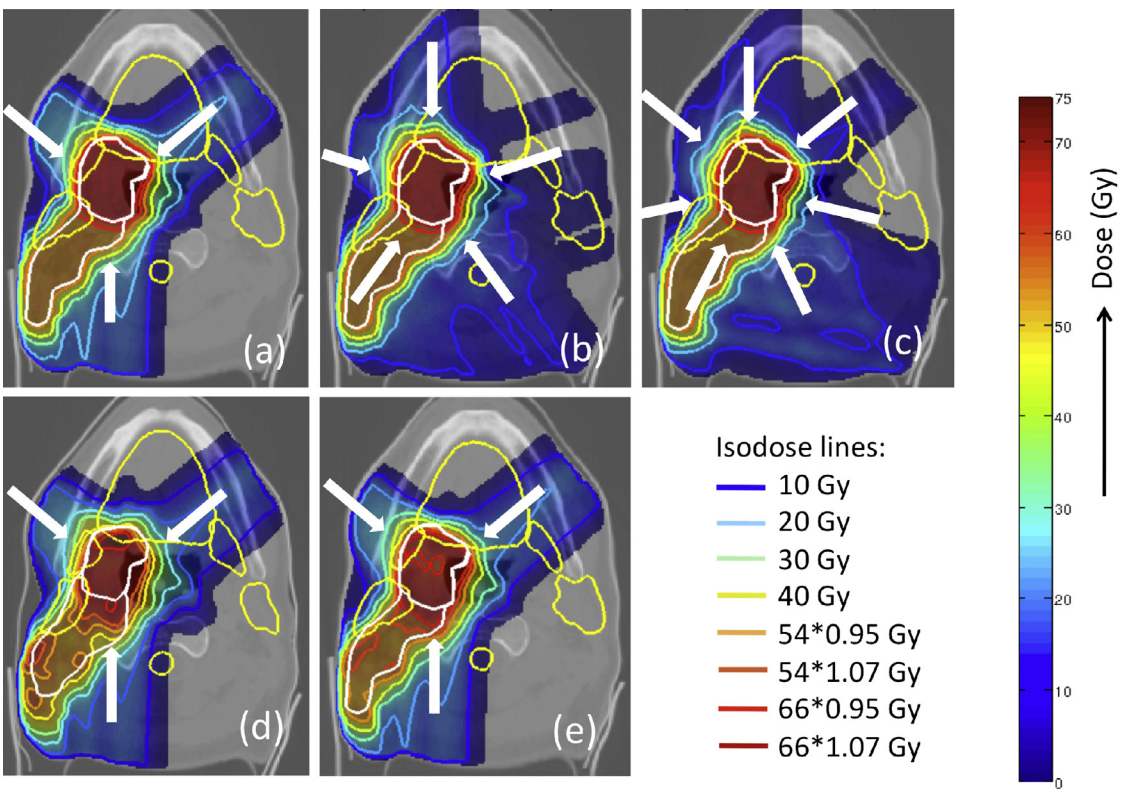

Fig. 1. Dose distributions for patient 4. (Top) planned dose for 3 (a), 5 (b), and 7 (c) beam directions. (Bottom, d) calculated dose delivered to repeat CT when treated with original 3-beam plan, and (e) delivered dose when applying a 4-mm posterior isocenter shift to the original CT. Target is shown in white; organs at risk (cord, parotids, oral cavity) are in yellow.

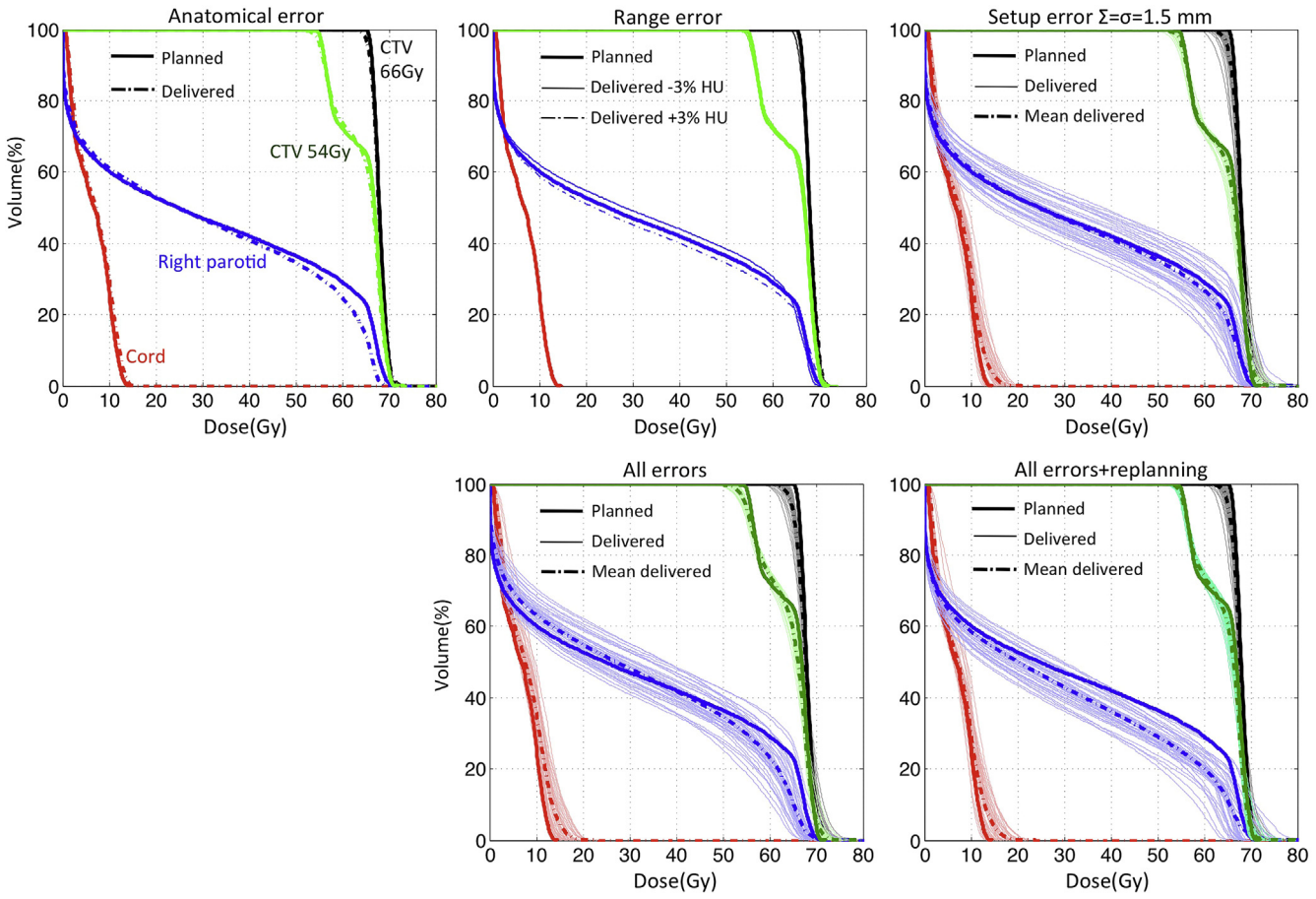

Fig. 2. Example of dose−volume histograms of patient 1 with various errors included for 5 sets of simulations. For visualization, we show only the target regions and 2 organs at risk.

right-left (RL) lateral directions. For each patient, we simulated 40 treatments of 30 fractions each. For each simulated treatment, we did the following:

- A systematic shift ($\delta_{AP},\delta_{SI},\delta_{RL}$) was applied, fixed for all fractions. The size of each δ (AP, SI, or RL) was determined randomly and independently from a normal distribution with standard deviation Σ (20).
- A random shift ($\Delta_{AP},\Delta_{SI},\Delta_{RL}$) was applied, varying for each fraction. The size of each Δ was determined randomly from a normal distribution with standard deviation σ (20). We assumed that $\sigma = \Sigma$ (20).
- The total isocenter shift for each fraction was determined, which was the vector sum of the above-described 2 contributions; and a dose matrix was calculated for each fraction.
- The resulting 30 dose matrices of all fractions was added to obtain the total dose for 1 treatment.

This was repeated 40 times, yielding 40 dose matrices per patient, representing 40 treatments. We repeated this procedure for $\Sigma = \sigma =$ 1.0, 1.5, 2.0, 2.5, and 3.0 mm.

All uncertainties combined

We investigated the combined effect of all errors by including them simultaneously. We included a range shift of −3% (3) and setup errors $\Sigma = \sigma =$ 1.5 mm (both following the procedure above for both planning and repeat CT), corresponding to the average residual setup error from bony anatomy deformation (7). We applied the treatment plan to the planning CT for fractions 1 to 15 and to the repeat CT for fractions 16 to 30. Nonrigid registration was used for dose accumulation between the CTs. We simulated 40 treatments of 30 fractions, yielding 40 dose distributions per patient.

Improving treatment quality

Adaptive planning

We investigated the effect of adaptive planning (adaptive planning based on repeat imaging) on treatment quality by repeating the simulations and including all errors as described above, but now applying for fractions 16 to 30 a newly created plan based on the repeat CT. Forty treatments were simulated per patient.

Increased beam directions

Analogue to the 3-beam plans, we simulated 40 treatments that included all errors for 5 and 7 beam directions (see section "All uncertainties combined", above above) to investigate whether plan robustness improved when applying more beam directions.

In summary, for each patient we performed 291 treatment simulations for 3-beam directions (nominal situation: 1 simulation; range error: 9 simulations; anatomical uncertainty: 1 simulation, setup error: 40 × 5 = 200 simulations; combined errors: 40 simulations; adaptive plan: 40 simulations), 40 simulations for 5-beams, and 40 simulations for 7-beams. For 10 patients this resulted in a total of 3710 treatment simulations.

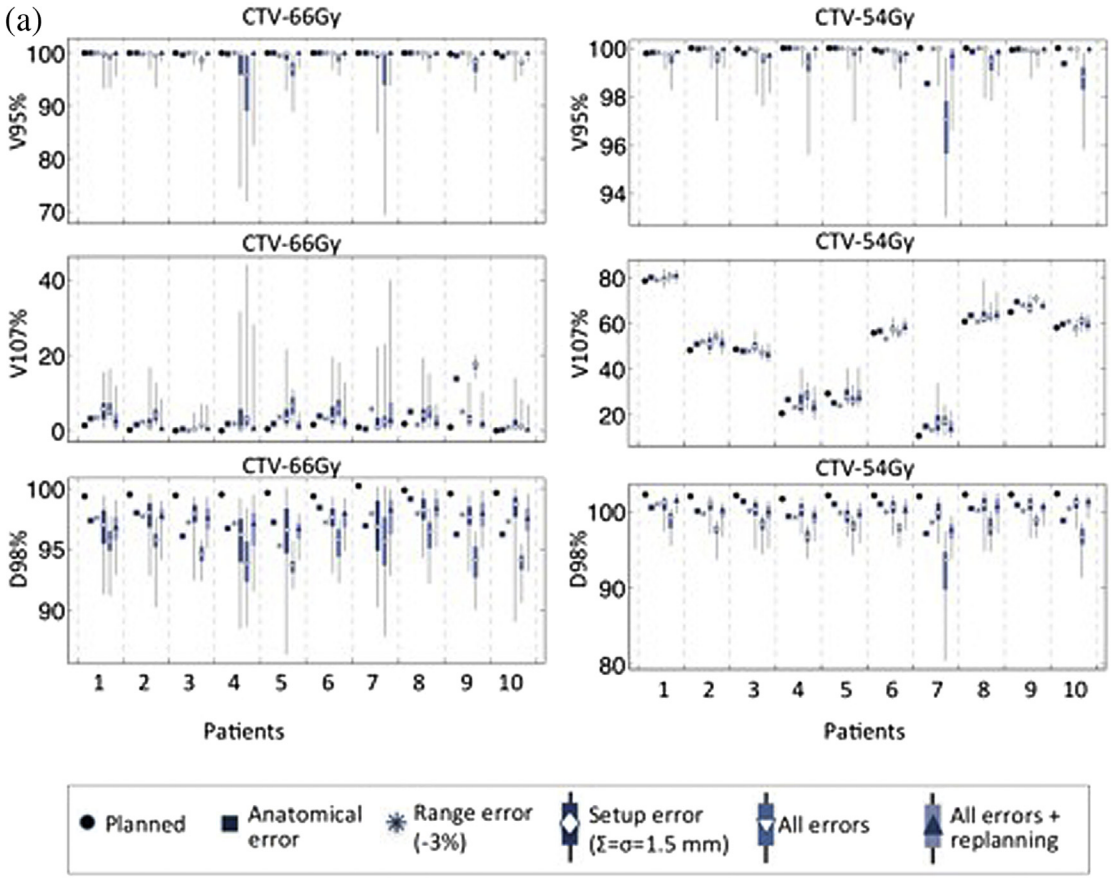

Fig. 3. (a) For the 10 patients, impact of various uncertainties on target dose are shown. Solid markers represent medians, boxes are spread between the 25th and 75th percentiles, and lines extend to the minimum and maximum of the 40 simulations. Because CTV-54 Gy contains CTV-66 Gy region, the V107% and D98% of CTV-54 Gy are naturally high for all patients. (b) For the 10 patients, impact of various uncertainties on organs at risk are shown. (figure 3b is displayed on the following page) Markers are as shown in panel a.

Evaluation

Plan evaluation was based on DVH following ICRU recommendations (27) (Fig. 2). For CTV-66 Gy and CTV-54 Gy, we considered the percentages of the target volumes receiving at least 95% and 107% of the prescribed dose (V95% and V107%, respectively) (27) and dose received by 98% of the target volume (D98%) (27). For cord and brainstem, we report dose to 2% of the volume (D2%) for near maximum dose (27), while for other OARs, we report mean doses (27). Statistical analysis of the differences among the plans for 3, 5, and 7 beam directions used the ANOVA test (a P value $<.05$ was considered significant).

Results

Below we report the dose deviations in our simulated treatments and corroborate our findings with the treatment intent that 90% of the patient population should have D98% $\geq$95% (20). We also highlight a few large dose deviations.

Quantification of dose deviations

Anatomical errors

Figure 1d shows an example of the dose degradation that resulted from applying the treatment plan from the planning CT to the repeat CT. Figure 3 summarizes dose effects for all simulations. For the target (Fig. 3a, "anatomical error"), the effect of anatomical errors was clearly visible in the D98%. The increase in target hot spots was small, except for patient 9 (V107% = 14% [see Supplementary Fig. E2]). Anatomical errors had a negative impact, but D98% for all patients was >95%. For CTV-66 Gy and CTV-54 Gy, the average D98% reduction was −1.9% and −2.1%, respectively. For CTV-66 Gy, the V107% increase due to anatomical errors was correlated with its size (Pearson correlation, 0.76; P = .011) and a trend toward a moderate correlation was observed for the average deformation (Pearson correlation, 0.64; P = .058). The biggest increase in OAR dose (Fig. 3b, "anatomical error") was in patient 6, with a 9.2Gy increase in the brainstem (see Supplementary Fig. E3).

Range error

Figure 4 (left panels) shows the impact of range errors. For CTV66-Gy, we note that for range errors within 4%, more than 90% of the patient population has D98% $\geq$95%, whereas for CTV54-Gy, even larger range errors do not cause a violation. Figure 3 shows the impact of a −3% range error on target (Fig. 3a) and OARs (Fig. 3b). For CTV-66 Gy and CTV-54 Gy (spaces don't look good) the average D98% reductions were −2.0% and −2.2%, respectively. The average increase for the 10 patients in D2% (for cord, brainstem) and mean dose (for other OARs) was less than 1 Gy for all OARs.

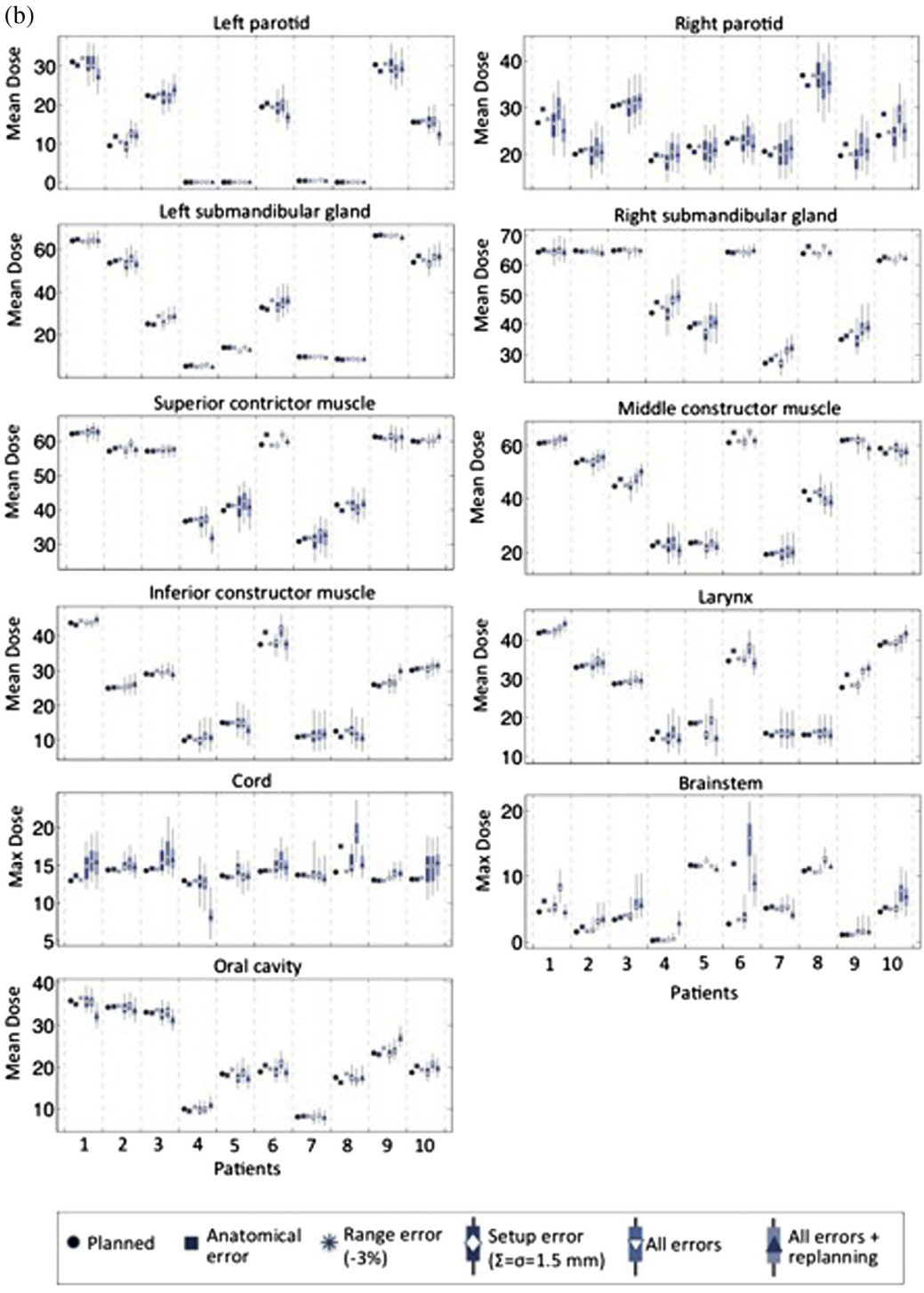

Fig. 3. (*continued*)

Setup error

Figures 1a and 1e demonstrate the impact of a 4-mm rigid isocenter shift on delivered dose. Figure 4 (right panels) shows the effects of setup errors. All parameters show an increase in target dose deterioration with increasing setup errors. At $\Sigma = \sigma = 2$ mm (CTV-66 Gy), the treatment intent that 90% of the population have D98% >95%, was not satisfied anymore. Figure 3a ("setup error") summarizes the dose effects for $\Sigma = \sigma = 1.5$ mm for target. For both CTVs, the median reduction in D98% was small, although with a large spread around the median. The largest D98% spread in CTV-66 Gy was seen in patient 4 with a relatively small tumor. For OARs (Fig. 3b), we quantified the effect of the

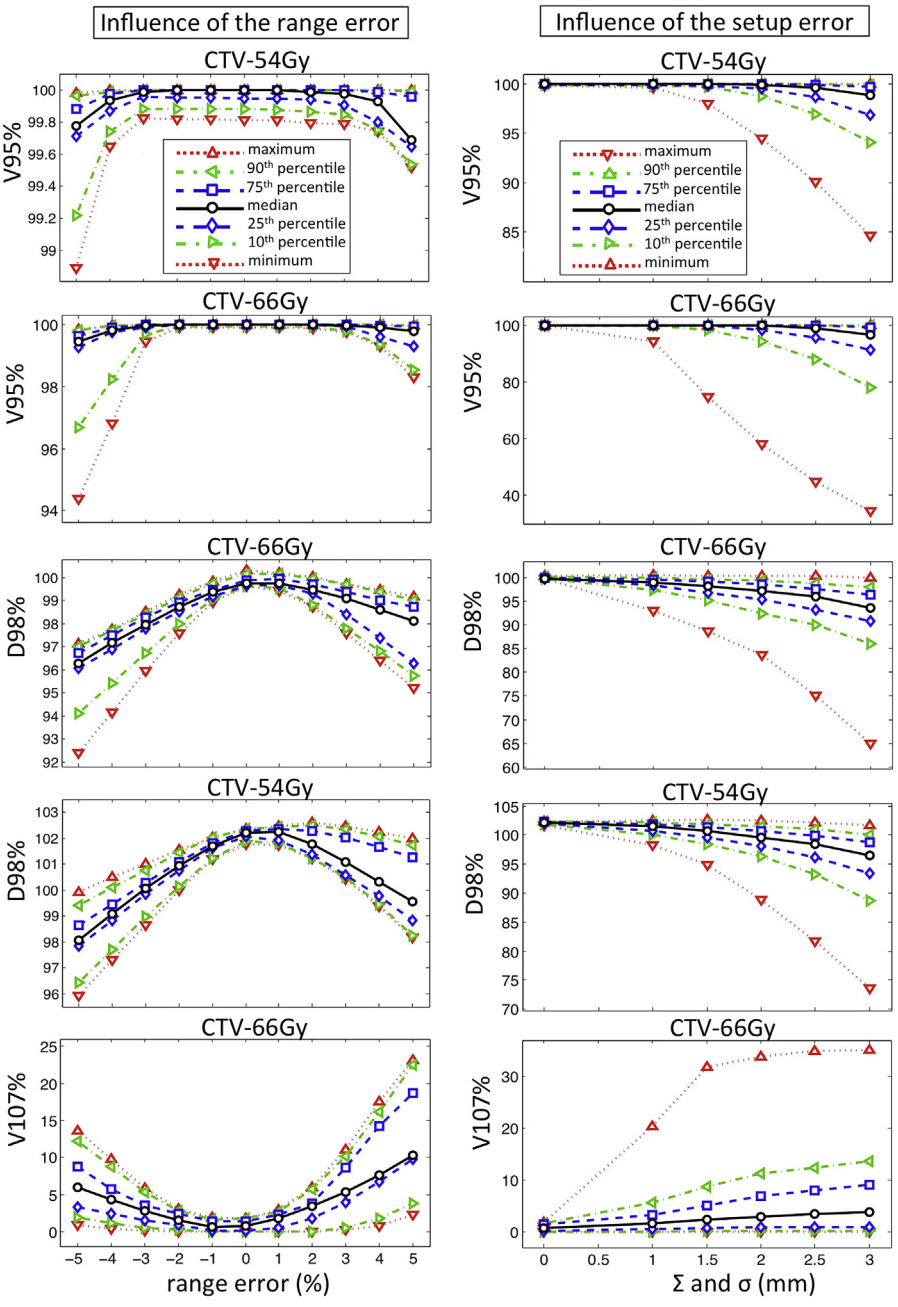

Fig. 4. Influence of range (left) and setup (right) uncertainties on target coverage. For the range error, maximum, minimum, and percentiles refer to the statistics of the group of 10 patients. For the setup error, they refer to 400 simulations (40 simulations × 10 patients). Planned values correspond to an error of zero.

setup error $\Sigma = \sigma = 1.5$ mm by calculating the average, standard deviation, and maximum values across all simulations of the delivered minus the planned maximum (cord, brainstem) or mean (other OARs) dose. The average dose increase across all simulations was less than 1 Gy for each OAR, although some large variations between individual treatments were observed (Fig. 3a).

All errors combined

Figures 3a and 3b show that the combination of errors can lead to inferior treatments. For CTV-66 Gy and CTV-54 Gy, percentages of simulations with D98% >95% were only 69% and 88%, respectively (Table 2), which is clinically unsatisfying. Hot spots increase: the average increase in V107% (ie, delivered minus planned V107%) across all simulations was 5.1% in V107% (CTV-66 Gy) and 2.9% (CTV-54 Gy). For OARs, we found average dose increases below 1 Gy but again large variations among treatments were present. The largest variation was in the right parotid, with a standard deviation in mean dose difference of 3.6 Gy and a maximum difference of 11.3 Gy (see Supplementary Fig. E4). The spinal cord showed a standard deviation in the maximum dose difference of 2.2 Gy and a maximum difference of 9.5 Gy (see Supplementary Fig. E5).

Table 2 Percentage of simulations satisfying D98% ≥95% (or V95% ≥98%)

	All errors		Replanning	
Patient	CTV-66 Gy	CTV-54 Gy	CTV-66 Gy	CTV-54 Gy
1	88%	100%	92%	100%
2	90%	95%	100%	100%
3	75%	98%	100%	100%
4	40%	90%	85%	100%
5	28%	95%	98%	100%
6	82%	100%	100%	100%
7	68%	20%	90%	98%
8	95%	98%	100%	100%
9	60%	100%	100%	100%
10	62%	82%	100%	100%
All patients	69%	88%	96%	100%

Treatment quality improvement

Adaptive planning

Figure 3 and Table 2 demonstrate how adaptive planning positively affects dose delivery. Target coverage showed a clear increase. For CTV-66 Gy and CTV-54 Gy, 96% and 100%, respectively, of simulations had a D98%>95%, now safely complying with the treatment intent that 90% of the population has D98% >95% (Table 2). Target hot spots decreased compared to the scenario were no adaptive planning was applied: the delivered minus planned V107% was now only 1.5% (CTV-66 Gy) and 2.0% (CTV-54 Gy). For OARs, the dose spread was reduced. For patient 4, the spinal cord dose was reduced greatly due to the lower cord dose in the repeat CT plan.

More beam directions

Supplementary Figure E6 shows dose effects for plans with 3, 5, and 7 beam directions. As shown in Supplementary Figure E6, plans with 5 or 7 beam directions were not significantly more robust than plans with 3 beam directions.

Discussion

An important observation is that the errors individually do not lead to serious target underdosage (D98%<95%) but that combined effects of the errors generate serious problems and cause the treatment intent (D98% ≥95% for at least 90% of the patient population [20]) to not be fulfilled. We recommend for IMPT studies that simulations combine all uncertainties to study site-specific clinical robustness requirements.

Concerning anatomical deformations, 1 of the 10 patients had a significant increase in hot spots. This patient had substantial target volume shrinkage (178 to 132 ml for CTV-66 Gy) (Table 1), deformation (6.6 mm [Table 1]), and weight loss, resulting in incorrect spatial pencil beam matching in the target (see Supplementary Fig. E2 for details). Patient 6, with similar deformation, did not show serious target hot spots but instead showed a large dose increase to the brainstem (see Supplementary Fig. E3). Small changes in V107% were seen in patients 3, 5, 7, and 8, with small CTV volumes and small deformations (2.3-2.5 mm). Factors such as initial tumor size can help to estimate which patients risk large dose deviations, as shown by the significant correlation between the increase in V107% and the tumor size. On the other hand, only a moderate correlation was found between increase in V107% and deformation. Much higher patient statistics and a thorough classification of anatomy changes (tumor shrinkage, weight loss, swelling, and others) are needed to identify patients at risk. In the absence of accurate prognostic factors, repeat imaging, dose recalculation, and, if required, adaptive planning are recommended to ensure sufficient target coverage and to avoid unwanted exposure of OARs. The timing and frequency of repeated imaging are subjects for future research.

We found no dramatic dose modifications when studying range errors separately. This is a significant clinical observation, as range errors are considered one of the primary concerns in robust optimization algorithms. A different spot placement technique could alter this, as shown by Lomax et al (3).

In the simulations that include setup errors, median dose deviations are small, but the dose spread is large both for OARs and target. Dose deteriorations in the target are caused by incorrect spatial matching of pencil beams, resulting in target hot spots, and by target misalignment (Figs. 1e and 3a). The impact on V95% is seen to be largest for small tumors (Table 1, patients 4 and 7), which is explained by the larger relative impact of a misalignment on these small volumes. The 5-mm CTV margin helps to account for target misalignment but cannot fully avoid dose impacts. Robust optimization could reduce the contribution of this error source (9). In addition, based on Figure 4, it is highly recommended that the systematic and random setup errors (Σ and σ, respectively) are kept below 2 mm, for instance by the use of online image guidance. This would also work to deal with problems such as a different position of the neck at planning and during treatment.

Concerning combined errors, the simulations (Fig. 3) showed that the combined occurrence of errors can amplify negative dose effects. We saw for patient 6 a large dose increase in the brainstem (see Supplementary Fig. E3), where the combined effect of the errors was larger than the sum of the effects separately (Fig. 3b). This was also observed for the cord of patient 8 (see Supplementary Fig. E5). Given the mixed causes for deviations, dose distortions are very difficult to predict without recalculating the dose with uncertainties included.

Treatment accuracy can be substantially improved by applying adaptive planning, increasing the percentage of treatments with D98% >95% safely to above 90%. Because adaptive planning increases staff workload and costs, the optimal adaptive strategy would have to be defined. Treatment plans with more beam directions were instead not more robust, in contrast to what was suggested by by Unkelbach et al (4). We suggest that the use of more beam directions does not necessarily solve dose deteriorations resulting from pencil beam mismatching or anatomical changes. We are aware that the choice of uniformly spaced beam angles may not be optimal and that beam angle optimization (28) may result in a more optimal choice, which could alter our observation. Other ways to improve treatment robustness not investigated here are to use wider pencil beams (however, this is expected to worsen OAR sparing [29]) or to apply robust optimization to dampen the effect of range and patient setup errors (9).

A limitation of this study is that intrafractional setup errors (ie, changes in patient position during the fraction) were not modeled separately and dose calculation errors were not included. Both are expected to have only minor dose impact (3, 30). We verified that the contrast fluid in our CTs, which led to a slight increase in inhomogeneities and thereby possibly influenced the simulations, did not alter our results.

Conclusions

Based on 3700 simulated treatments for oropharyngeal cancer patients, we quantified the dose differences between planned and delivered IMPT doses in targets and OARs. Without action against treatment uncertainties, the treatment intent (D98% $\geq$95% for at least 90% of the patient population) was not fulfilled. Given the mixed causes for major deviations observed, we advise acquisition of repeat CT scans and dose recalculation to properly assess delivered dose. If required, adaptive planning is effective for mitigating the effect of treatment-related uncertainties. Applying more than 3 beam directions did not increase plan robustness.

References

1. van de Water TA, Lomax AJ, Bijl HP, et al. Potential benefits of scanned intensity-modulated proton therapy versus advanced photon therapy with regard to sparing of the salivary glands in oropharyngeal cancer. *Int J Radiat Oncol Biol Phys* 2011;79:1216-1224.
2. Simone CB II, Ly D, Dan TD, et al. Comparison of intensity-modulated radiotherapy, adaptive radiotherapy, proton radiotherapy, and adaptive proton radiotherapy for treatment of locally advanced head-and-neck cancer. *Radiother Oncol* 2011;101:376-382.
3. Lomax AJ. Intensity-modulated proton therapy and its sensitivity to treatment uncertainties 1: the potential effects of calculational uncertainties. *Phys Med Biol* 2008;53:1027-1042.
4. Unkelbach J, Martin B, Soukup M, et al. Reducing the sensitivity of IMPT treatment plans to setup errors and range uncertainties via probabilistic treatment planning. *Med Phys* 2009;36:149-163.
5. Barker JL Jr., Garden AS, Ang KK, et al. Quantification of volumetric and geometric changes occurring during fractionated radiotherapy for head-and-neck cancer using an integrated CT/linear accelerator system. *Int J Radiat Oncol Biol Phys* 2004;59:960-970.
6. Lomax AJ. Intensity modulated proton therapy and its sensitivity to treatment uncertainties 2: the potential effects of inter-fraction and inter-field motions. *Phys Med Biol* 2008;53:1043-1056.
7. van Kranen S, van Beek S, Rasch C, et al. Setup uncertainties of anatomical sub-regions in head-and-neck cancer patients after offline CBCT guidance. *Int J Radiat Oncol Biol Phys* 2009;73:1566-1573.
8. Albertini F, Bolsi A, Lomax AJ, et al. Sensitivity of intensity-modulated proton therapy plans to changes in patient weight. *Radiother Oncol* 2008;86:187-194.
9. Liu W, Frank SJ, Li X, et al. Effectiveness of robust optimization in intensity modulated proton therapy planning for head and neck cancers. *Med Phys* 2013;40:051711.
10. Liu W, Zhang X, Li Y, et al. Robust optimization of intensity modulated proton therapy. *Med Phys* 2012;39:1079-1091.
11. Liu W, Frank SJ, Li X, et al. PTV-based IMPT optimization incorporating planning risk volumes. *Med Phys* 2013;40:021709.
12. Chen W, Unkelbach J, Trofimov A, et al. Including robustness in multi-criteria optimization for intensity-modulated proton therapy. *Phys Med Biol* 2012;57:91.
13. Pflugfelder D, Wilkens JJ, Oelfke U. Worst case optimization: a method to account for uncertainties in the optimization of intensity modulated proton therapy. *Med Phys Biol* 2008;53:1689-1700.
14. Inaniwa T, Kanematsu N, Furukawa T, et al. A robust algorithm of intensity modulated proton therapy for critical tissue sparing and target coverage. *Med Phys Biol* 2011;56:4749-4770.
15. Fredriksson A, Forsgren A, Hårdemark B. Minimax optimization for handling range and setup uncertainties in proton therapy. *Med Phys* 2011;38:1672-1684.
16. Fredriksson A. A characterization of robust radiation therapy planning methods: from expected value to worst case optimization. *Med Phys* 2012;39:5169-5181.
17. Castadot P, Lee JA, Geets X, et al. Adaptive radiotherapy of head and neck cancer. *Semin Radiat Oncol* 2010;20:84-93.
18. Teguh DN, Levendag PC, Voet PW, et al. Clinical validation of atlas-based auto-segmentation of multiple target volumes and normal tissue (swallowing/mastication) structures in the head-and-neck. *Int J Radiat Oncol Biol Phys* 2011;81:950-957.
19. Revision Radiology. CT metal artifact reduction. Available at: http://www.revisionrads.com/. Accessed Month Day, Year.
20. van Herk M, Remeijer P, Rasch C, et al. The probability of correct target dosage: dose-population histograms for deriving treatment margins in radiotherapy. *Int J Radiat Oncol Biol Phys* 2000;47:1121-1135.
21. Vásquez Osorio EM, Hoogeman MS, Al-Mamgani A, et al. Local anatomic changes in parotid and submandibular glands during radiotherapy for oropharynx cancer and correlation with dose, studied in detail with nonrigid registration. *Int J Radiat Oncol Biol Phys* 2008;70:875-882.
22. Hong LM, Goitein M, Bucciolini, et al. A pencil beam algorithm for proton dose calculations. *Phys Med Biol* 1996;4:1305-1330.
23. Kooy HM, Clasie BM, Lu HM, et al. A case study in proton pencil-beam scanning delivery. *Int J Radiat Oncol Biol Phys* 2010;76:624-630.
24. Breedveld S, Storchi PR, Voet PW, et al. iCycle: Integrated, multicriterial beam angle, and profile optimization for generation of coplanar and noncoplanar IMRT plans. *Med Phys* 2012;39:951-963.
25. Voet P, Breedveld S, Dirkx MLP, et al. Integrated multicriterial optimization of beam angles and intensity profiles for coplanar and noncoplanar head and neck IMRT and implications for VMAT. *Med Phys* 2012;39:4858-4865.
26. van de Water S, Kraan AC, Breedveld S. Improved efficiency of multi-criteria IMPT treatment planning using iterative resampling of randomly placed pencil beams. *Phys Med Biol* 2013;58:6969-6983.
27. ICRU Report 83. Prescribing, recording, and reporting photon-beam intensity-modulated radiation therapy (IMRT), vol. 10. Oxford, UK: Oxford University Press; 2010.
28. Cao W, Lim GJ, Lee A, et al. Uncertainty incorporated beam angle optimization for IMPT treatment planning. *Med Phys* 2012;39:5248-5256.
29. van de Water TA, Lomax AJ, Bijl HP, et al. Using a reduced spot size for intensity-modulated proton therapy potentially improves salivary gland-sparing in oropharyngeal cancer. *Int J Radiat Oncol Biol Phys* 2012;82:313-319.
30. Hoogeman MS, Nuyttens JJ, Levendag PC, et al. Time dependence of intrafraction patient motion assessed by repeat stereoscopic imaging. *Int J Radiat Oncol Biol Phys* 2008;70:609-618.