PAPER

A robustness analysis method with fast estimation of dose uncertainty distributions for carbon-ion therapy treatment planning

IOP Publishing | Institute of Physics and Engineering in Medicine Physics in Medicine & Biology

Phys. Med. Biol. **61** (2016) 5818–5836 doi:10.1088/0031-9155/61/15/5818

A robustness analysis method with fast estimation of dose uncertainty distributions for carbon-ion therapy treatment planning

Makoto Sakama, Nobuyuki Kanematsu and Taku Inaniwa

Department of Accelerator and Medical Physics, National Institute of Radiological Sciences, Anagawa 4-9-1, Inage-ku, Chiba 263-8555, Japan

E-mail: sakama.makoto@qst.go.jp

Received 16 December 2015, revised 26 April 2016
Accepted for publication 31 May 2016
Published 18 July 2016

Abstract

A simple and efficient approach is needed for robustness evaluation and optimization of treatment planning in routine clinical particle therapy. Here we propose a robustness analysis method using dose standard deviation (SD) in possible scenarios such as the robustness indicator and a fast dose warping method, i.e. deformation of dose distributions, taking into account the setup and range errors in carbon-ion therapy. The dose warping method is based on the nominal dose distribution and the water-equivalent path length obtained from planning computed tomography data with a clinically commissioned treatment planning system (TPS). We compared, in a limited number of scenarios at the extreme boundaries of the assumed error, the dose SD distributions obtained by the warping method with those obtained using the TPS dose recalculations. The accuracy of the warping method was examined by the standard-deviation-volume histograms (SDVHs) for varying degrees of setup and range errors for three different tumor sites. Furthermore, the influence of dose fractionation on the combined dose uncertainty, taking into consideration the correlation of setup and range errors between fractions, was evaluated with simple equations using the SDVHs and the mean value of SDs in the defined volume of interest. The results of the proposed method agreed well with those obtained with the dose recalculations in these comparisons, and the effectiveness of dose SD evaluations at the extreme boundaries of given errors was confirmed from the responsivity and DVH analysis of relative SD values for each error. The combined dose uncertainties depended heavily on the number of fractions, assumed errors and tumor sites. The typical computation time of the warping method is approximately 60 times less than that of the

full dose calculation method using the TPS. The dose SD distributions and SDVHs with the fractionation effect will be useful indicators for robustness analysis in treatment planning, and the results of our comparative study show that the proposed analysis method would be beneficial in routine clinical use.

Keywords: carbon-ion therapy, robustness analysis method, fractionation effect, dose warping method

(Some figures may appear in colour only in the online journal)

1. Introduction

In radiation therapy the correspondence of a treatment plan to the actual treatment is subject to setup and range uncertainties, and the dose concentration in particle therapy is highly sensitive to these uncertainties because the scattering effect is small and the Bragg peak is narrow. Uncertainty analysis for the individual patient's treatment plans, however, is not currently included in clinical practice (Bortfeld 2006). While uncertainties are generally taken into account by using the margin concept for the tumor coverage (van Herk 2004), probabilistic approaches to treatment planning have been investigated (Baum *et al* 2006, Chan *et al* 2006, McShan *et al* 2006, Witte *et al* 2007, Unkelbach *et al* 2009, Sobotta *et al* 2010) and site-specific robustness-assessment protocols using retrospective analysis were recently proposed for proton therapy (McGowan *et al* 2015).

To evaluate the robustness of a treatment plan, one needs to calculate dose distributions for a great number of conditions or scenarios that simulate different possible errors. The probabilistic approach to the patient-specific evaluation, such as variance calculations based on dose blurring or sampling, has been proposed but those methods are computationally challenging or intrinsically limited (Ploquin *et al* 2006, Sobotta *et al* 2012, Trofimov *et al* 2012). In addition, it is difficult to make the robustness evaluation part of the clinical routine because of the high computational cost of calculating ion beam dose distributions under various conditions. Because the computational cost of dose calculation makes robustness evaluation a time-consuming task, Park *et al* (2012) instead used the dose-warping technique to quickly obtain perturbed dose distributions in proton therapy. In their implementation, the perturbed dose distribution can be created by simply shifting the pre-calculated dose profile along the beam direction by an amount based on an equivalent water-equivalent path length (WEPL). To obtain more accurate perturbed dose distributions, we use a similar approach with a different mapping process based on the nominal dose distribution.

The worst-case-scenario approach, which calculates a subset of these variations or the dose distributions at the extreme boundaries of the given setup or range errors, minimizes the number of dose calculations (Lomax *et al* 2004, Albertini *et al* 2011). However, this analysis, using the minimum and maximum worst dose distributions, does not evaluate the physically probable dose distributions and may overestimate the influence of these uncertainties on the planned dose distribution (Albertini *et al* 2011). Unkelbach *et al* (2009) introduced the dose standard-deviation-volume histograms (SDVHs) for the scenarios of assumed errors to perform a quantitative comparison of the plans in terms of dose uncertainty. For routine clinical use, however, no standard method or indicator for plan robustness has been established.

The dose uncertainties in an actual treatment situation are affected by the number of fractions in which the total dose is delivered and depend on the correlation of the errors. Analytical probabilistic modeling considering fractionation was proposed to quantify uncertainties in treatment plans with proton and photon beams (Bangert *et al* 2013), and the sampling of

geometric errors was performed in robustness analysis for fractionated radiotherapy (Tilly *et al* 2015). It is generally difficult, however, to estimate the effect of fractionation quickly in robustness analysis.

In this paper, we propose a robustness analysis method for routine use in carbon-ion therapy. Among the topics treated here are the following:

- Shortening the computation time of the perturbed dose distributions by using the fast dose warping method for arbitrary setup and range errors.
- Examination of the applicability of using the dose standard deviation (SD) distributions, SDVHs and DVHs as the indicator of dose uncertainty estimation in a limited number of scenarios.
- Fractionation effect for dose uncertainty estimation in consideration of the correlation between fractions with simple equations.

The proposed method is applied to three beams for typical cases of scanned carbon-ion therapy at the National Institute of Radiological Sciences (NIRS), and the warping method is compared with dose recalculation using the NIRS treatment planning system (TPS). First the perturbed dose distributions with the assumed errors are calculated using the fast dose warping method, and then the dose variations in the possible scenarios are evaluated quantitatively in several ways. Then the influence of fractionation on dose uncertainty is estimated by the relative combined dose uncertainty in consideration of the correlation of the setup and range errors between fractions.

2. Materials and methods

2.1. Dose warping method

In this work, we used the clinically commissioned TPS for scanning irradiations at the NIRS, which is implemented using the sub-pencil beam algorithm with a triple Gaussian beam model (Inaniwa *et al* 2014, 2015). The nominal dose distribution without any errors (D) was calculated as the reference distribution and the actual dose distributions under the influence of both setup and range errors (d) were derived. For convenience, the z-axis along the beam's axis was used as a beam's-eye view (BEV) coordinate. The nominal dose of a voxel at an arbitrary coordinate of (i, j, k) is expressed as $D(i, j, k)$, and the dose realized there with the setup error $(\Delta x, \Delta y)$ is expressed as $d_{\Delta x,\Delta y}(i, j, k)$. The parallel beam approximation was used for simple and fast calculations because of the small divergence effect. The WEPL of each voxel in the presence of the setup and range errors was compared with the WEPL of each voxel of the nominal plan in the overlapped area S_m shown in figure 1 with the voxel displaced by the setup error. The dose of each overlapped area D_m with the setup error is obtained from the corresponding dose in the nominal plan. The voxel dose with the setup error is calculated by the weighted mean dose using each overlapped area S_m and is obtained using equation (2) for linear interpolation of WEPL and equation (3) for area-weighted averaging.

$$\mathrm{WEPL}(i'_m, j'_m, k'_m - 1) < \mathrm{WEPL}(i, j, k) \leqslant \mathrm{WEPL}(i'_m, j'_m, k'_m) \tag{1}$$

$$\begin{aligned} D_m(i, j, k) = {} & \frac{D(i'_m, j'_m, k'_m) - D(i'_m, j'_m, k'_m - 1)}{\mathrm{WEPL}(i'_m, j'_m, k'_m) - \mathrm{WEPL}(i'_m, j'_m, k'_m - 1)} \\ & \times (\mathrm{WEPL}(i, j, k) - \mathrm{WEPL}(i'_m, j'_m, k'_m)) + D(i'_m, j'_m, k'_m) \end{aligned} \tag{2}$$

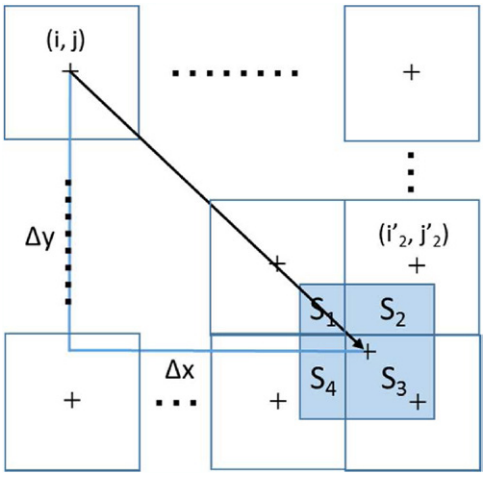

Figure 1. Schematic view of the voxel dose calculation in the setup error. The arrow indicates the direction of voxel displacement and the shaded region (S_1–S_4) shows the overlapped area with the voxels in the nominal plan. i'_2 and j'_2 are the voxel position in the presence of the overlapped part (2) in the nominal plan.

$$d_{\Delta x,\Delta y}(i,j,k) = \frac{\sum_m (S_m \cdot D_m(i,j,k))}{S} \tag{3}$$

S is the projected area of a voxel in the beam direction, S_m is each overlapped area of the voxel in the case of the setup error with the voxel in the nominal plan and m is the number of each area divided by the overlapped region with each voxel in the nominal plan. i'_m and j'_m are each voxel position in the presence of the overlapped part m in the nominal plan shown in figure 1 and k'_m is determined by the condition of equation (1). The voxel dose in the case of only the range error is obtained by equation (2) from the comparison of the WEPL in the errors arising from CT with that in the nominal plan.

2.2. Dose SD distribution, SDVH and DVH

All the individual dose distributions with the assumed errors were combined into a dose SD distribution in order to simply represent these distributions in the form of a single distribution for the nominal plan by using the following equation:

$$\sigma(i,j,k) = \sqrt{\frac{\sum_{\text{error}}^{N} (d_{\text{error}}(i,j,k) - {}_{\text{mean}}d(i,j,k))^2}{N-1}} \tag{4}$$

where N is the number of scenarios, and a 3D distribution of dose variation at each voxel in the nominal dose distribution is obtained by this calculation. SDVHs can also be calculated for each volume of interest in a way similar to that in which DVHs are calculated (Unkelbach *et al* 2009). The SDVHs and DVHs normalized by the prescribed target dose in the nominal plan can be used as helpful indicators to evaluate the plan robustness. We defined SD5% as the SD to the volume of 5% in SDVHs. This percentage in SDVHs was selected due to the estimation of influence of cold or hot spots and the characteristic of

comparatively stable indicators in order to compare the calculated values of the dose SD distributions. Minimum and maximum worst-dose distributions are given by the distributions of minimum and maximum dose for each voxel in all scenarios of the assumed errors defined as $d_{\max}(i,j,k) = \max_{\text{error}}(d_{\text{error}}(i,j,k))$ and $d_{\min}(i,j,k) = \min_{\text{error}}(d_{\text{error}}(i,j,k))$ (Lomax *et al* 2004). The mean dose distributions in various scenarios of the assumed error added or subtracted one SD for each voxel and were defined as $d_{+\sigma}(i,j,k) = {}_{\text{mean}}d(i,j,k) + \sigma(i,j,k)$ and $d_{-\sigma}(i,j,k) = {}_{\text{mean}}d(i,j,k) - \sigma(i,j,k)$.

2.3. Fractionation effect on dose uncertainty

Setup errors were assumed to be random errors with no correlation, and range errors were assumed to be systematic errors with perfect correlation between fractions. Therefore, under the assumption of no correlation between setup error and range error, the influence of fractionation on dose SD distribution was evaluated by relative combined dose uncertainties ${}_{\text{rel}}\sigma_{\text{c}}$ for these errors following equation (5) according to the propagation of uncertainties:

$$ {}_{\text{rel}}\sigma_{\text{c}}(i,j,k,n) = \frac{\sqrt{n\cdot\sigma^2_{\text{setup}}(i,j,k) + n^2\cdot\sigma^2_{\text{range}}(i,j,k)}}{n\cdot D_{\text{p}}} = \frac{\sqrt{\frac{1}{n}\cdot\sigma^2_{\text{setup}}(i,j,k) + \sigma^2_{\text{range}}(i,j,k)}}{D_{\text{p}}} \tag{5} $$

where D_{p} is the prescribed target dose and n is the number of fractions. In the case that all uncertainties are random, the relative combined dose uncertainty is given by equation (6), and in the case that all uncertainties are systematic, it is given by equation (7),

$$ {}_{\text{rel}}\sigma_{\text{r}}(i,j,k,n) = \frac{\sqrt{n\cdot\sigma^2_{\text{total}}(i,j,k)}}{n\cdot D_{\text{p}}} = \frac{\sigma_{\text{total}}(i,j,k)}{\sqrt{n}\,D_{\text{p}}} \tag{6} $$

$$ {}_{\text{rel}}\sigma_{\text{s}}(i,j,k,n) = \frac{\sqrt{n^2\cdot\sigma^2_{\text{total}}(i,j,k)}}{n\cdot D_{\text{p}}} = \frac{\sigma_{\text{total}}(i,j,k)}{D_{\text{p}}} \tag{7} $$

σ_{total} is the SD in the case of the combination of setup and range errors. The dependence of the mean value of SDs ${}_{\text{rel,mean}}\sigma'_{\text{c}}$ on the number of fractions in the target was calculated using the mean value of SDs of separately calculated setup and range errors following the simple equation (8) without the calculation of combined dose uncertainty distributions.

$$ {}_{\text{rel,mean}}\sigma'_{\text{c}}(n) = \frac{\sqrt{n\cdot {}_{\text{mean}}\sigma^2_{\text{setup}} + n^2\cdot {}_{\text{mean}}\sigma^2_{\text{range}}}}{n\cdot D_{\text{p}}} \tag{8} $$

2.4. Tested cases

In order to analyze the robustness of a treatment plan, we compared the dose SD distributions calculated using the warping method and the dose recalculation from the TPS for three different clinical cases. We validated the proposed method in lung, prostate and head & neck (HN) patient cases to show the applicability of the proposed method to these sites. All dose calculations were performed using the $2 \times 2 \times 2\,\text{mm}^3$voxel size. All the volumes of interest were contoured by a physician on the planning CT data. The nominal dose distribution and DVHs calculated using the planning CT data with no assumed setup and range errors by the spot scanning beam delivery method with single-field optimization are shown in figure 2. The dose

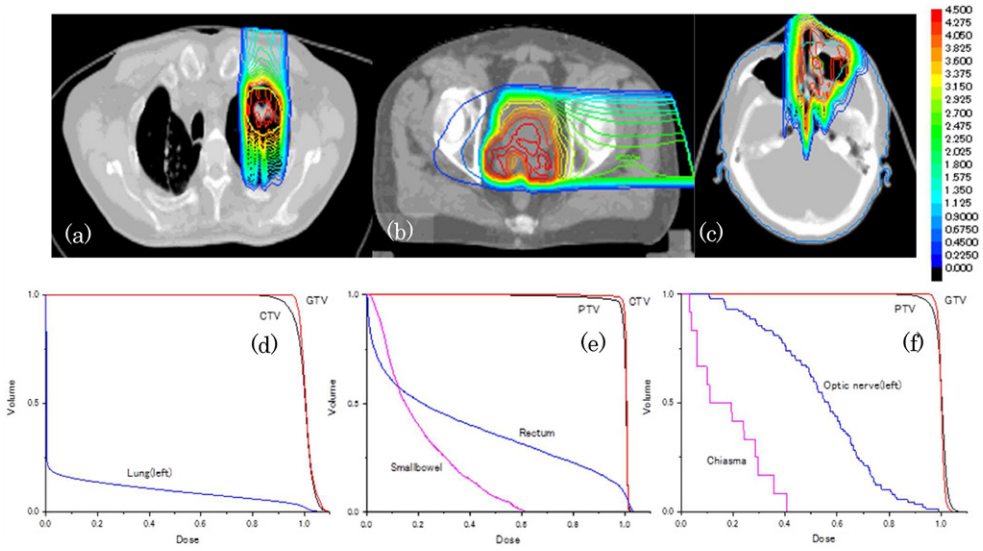

Figure 2. Nominal dose distributions and DVHs in the TPS for the lung (a) and (d), prostate (b) and (e) and head/neck (c) and (f) cases. The color scale for dose distributions is from 0 to 4.5 Gy (RBE).

SD distributions were calculated by the dose warping method based on these data. The prescribed doses with weighting of relative biological effectiveness (RBE) in the lung, prostate and HN cases were respectively 4, 4.3 and 4 Gy (RBE) (Inaniwa *et al* 2015). The overall accuracy of the proposed method was quantified using the dose SD distributions and the SDVHs. The warping method calculations were performed on a computer with an Intel® Xeon® CPU having a 3.1 GHz clock speed (E5-2687W v3), and the TPS calculations were performed on a computer with an Intel® Xeon® CPU with a 3.0 GHz clock speed (E5-2690 v2) with 4 threads in parallel computing. The computation time for the dose warping was compared with that for the full dose calculation using the TPS.

2.5. Evaluation procedure

We modeled the setup errors by recalculating the planned dose distribution in the condition of spatially shifted patient's CT. Bolsi *et al* (2008) reported that in proton therapy patients the size of systematic setup errors was below 0.6 mm and the size of random setup errors depended on the tumor location and the fixation technique. We assumed the influence of only random errors in a plane perpendicular to the beam axis. The impact of a patient shift along the beam axis is negligible due to a small change of WEPL by considering the additional air gap. Casiraghi *et al* (2013) reported the effectiveness of robustness evaluation using the worst dose distribution at the limited number of calculations in the extreme boundaries of given errors. Therefore we assumed that setup error is a uniform shift of ± 3 mm in both x and y directions in one instance, which comprised nine scenarios including the nominal dose distribution. The dose SD distribution in the setup error $\sigma_{\text{setup}}(i, j, k)$ was calculated using the individual perturbed dose distributions, and responsivity analysis was performed to confirm the applicability of the dose SDs at the extreme boundaries of assumed error in the target and the tendency of impact of the setup errors on the dose. In the responsivity analysis the dose SD distributions, SDVHs and DVHs in three tumor sites were calculated at setup errors ranging

from ±1 mm to ±7 mm in 1 mm steps and the mean value of SDs ${}_{\text{mean}}\sigma_{\text{setup}}$ in each volume of interest derived from the proposed method were compared with those obtained using the TPS recalculation.

Range uncertainties can usually be regarded as the critical source of error for particle therapy. When the dose distribution is typically calculated using a planning CT data, any errors arising from CT such as noise, artifacts and the conversion from CT Hounsfield units (HU) to relative stopping power are more likely to be systematic and propagate through the whole course in the treatment series (Schneider *et al* 1996, Schaffner and Pedroni 1998, Kanematsu *et al* 2003). Range errors were modeled by recalculating each plan with an error of ±3% for the nominal HU values in planning CT data according to the evaluation by Moyers *et al* (2010). Furthermore, the dose SD distribution in the range error $\sigma_{\text{range}}(i,j,k)$ was calculated using the perturbed dose distributions and the responsivity analysis was performed by calculations at range errors from ±1% to ±6% in 1% steps and comparisons similar to those in the case of the setup error were performed using the mean value of SDs ${}_{\text{mean}}\sigma_{\text{range}}$. In these cases, the DVHs of the worst dose distributions were calculated using the warping method and TPS dose recalculations. DVHs with one SD were also calculated from $d_{+\sigma}(i,j,k)$ and $d_{-\sigma}(i,j,k)$ in order to evaluate the applicability of the dose SD distributions. Considering the actual situation to have a combination of both setup and range errors, we evaluated dose SD by adding the range error of ±3% to the setup error of ±3 mm, which including the nominal dose distribution is 27 scenarios. The dose SD distribution and SDVH were calculated in the case of the combination of setup and range errors and combined dose uncertainty of separately calculated setup error and range error. The responsivity analysis of the combined dose uncertainties with setup and range errors was performed using the mean value of SDs in the target.

The percent SD difference distributions were calculated using equation (9) and the percent SD difference histograms for setup and range errors were obtained.

$$\text{diff}(i,j,k) = \frac{\sigma_{\text{warping}}(i,j,k) - \sigma_{\text{TP}}(i,j,k)}{D_{\text{p}}} \times 100 \tag{9}$$

The dose SD distribution and SDVH of the combination of setup and range errors were compared with those of combined dose uncertainty of separately calculated setup and range errors. The ${}_{\text{rel,mean}}\sigma'_{\text{c}}$ was compared with the mean value of SDs ${}_{\text{rel,mean}}\sigma_{\text{c}}$ in the target calculated by the combined dose uncertainty distributions ${}_{\text{rel}}\sigma_{\text{c}}$ of each number of fractions using the proposed method.

3. Results

3.1. Comparison of dose SD distributions

The dose SD distributions calculated by the dose warping method and the TPS recalculation with the setup error of ±3 mm for the lung, prostate and HN cases are shown in figure 3. The maximum value of the SD in these planes reached approximately 30% for the prescribed target dose. In the prostate case, the SD was larger than that at other sites in the edge of dose distribution. In the HN case the dose SD values in the warping method were slightly larger than those in the TPS recalculation shown in (d) of figure 3 and reached the deeper region behind the target due to the high heterogeneity, especially at the nasal cavity part. In the prostate case with a range error of ±3% (not shown here), the maximum value behind the target reached nearly 50%.

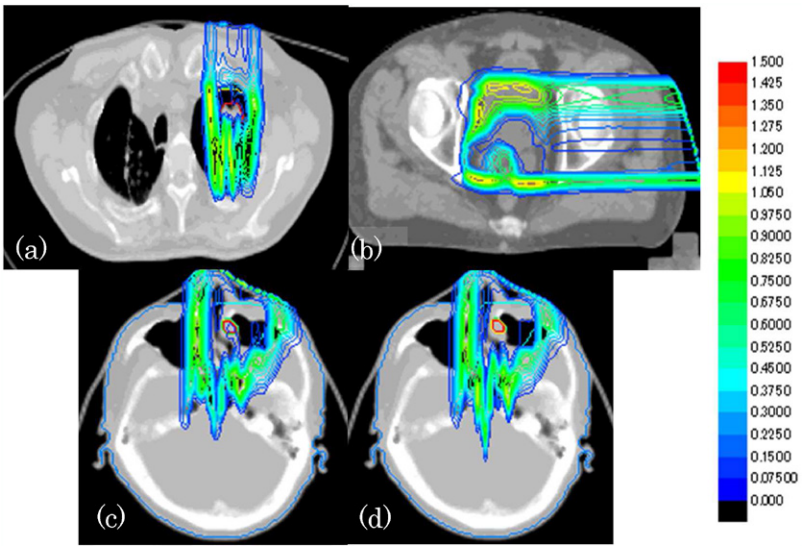

Figure 3. The dose SDs calculated by the dose warping method with the setup error of ±3 mm in the lung (a), prostate (b) and head/neck (d) cases. The SDs calculated using the TPS recalculation with the same setup error (c). The color scale runs from 0 to 1.5 Gy (RBE).

The percent SD difference distributions obtained using the TPS recalculation and the warping method are shown in figure 4. The differences between SDs calculated using the warping method and those calculated using the TPS recalculation were normalized by the prescribed target dose. The percent SD difference histograms for setup and range errors for lung, prostate and head/neck cases are shown. In this plane, the difference in the setup error was larger than that in the range error, especially in the HN case the maximum difference at the deeper region passing through the nasal cavity reached approximately 17%, which is overestimated by the warping method. In this plane of the prostate case, the difference in range and setup error was comparatively small and the maximum difference behind the target was about 6%. The dose SD distributions obtained by the warping method and recalculation of TPS agreed well generally. Table 1 shows the relative volumes in the irradiated region within ±0.5% and ±1% of percent SD difference histograms using the TPS recalculation and the warping method with each setup and range error. In the three tumor sites the relative volume within ±0.5% and ±1% was higher than or equal to 80% and 89%, respectively.

3.2. Volumetric analysis

The SDVHs for lung, prostate and head/neck cases are shown in figure 5. The histograms were calculated using the setup error of ±3 mm and the range error of ±3%. In all three tumor sites the SDVHs in the setup error ±3 mm were larger than those in the range error ±3%. The difference of SDVHs of organs at risk (OARs) calculated by the warping method and the TPS recalculation were comparatively large in the HN case due to the high heterogeneity and small volume of OARs. The SDVHs calculated using the warping method and those calculated using the TPS recalculation agreed well, particularly in the prostate case. Table 2 lists the SD values corresponding to the volume of 5% using the warping method and the TPS recalculation with the setup and range errors. In the prostate and HN cases, the SD values of

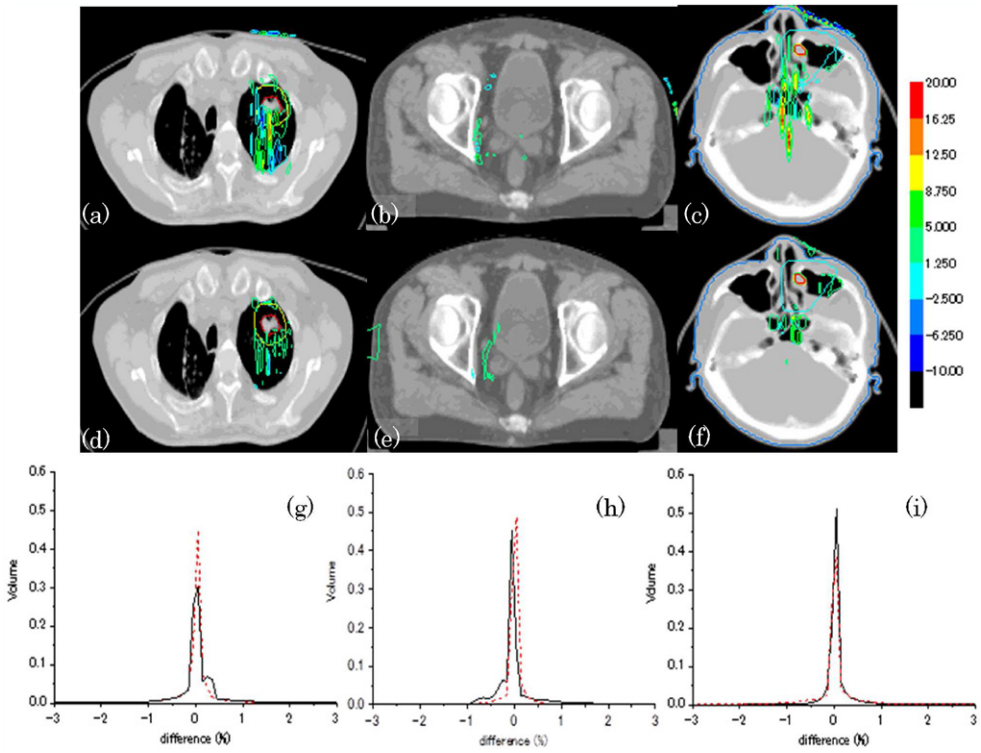

Figure 4. The percent SD difference distributions using the recalculation of TPS and the warping method with the setup error of ±3 mm (a)–(c) and the range error of ±3% (d)–(f) with the scale running from −10% to 20%. The differences between SDs calculated by the warping method and those obtained using the recalculation of TPS are relative to the prescribed dose of target. The percent SD difference histograms for the setup and range errors are indicated for lung (g), prostate (h) and head/neck cases (i). Red dashed lines and black solid lines were derived from a setup error of ±3 mm and a range error of ±3%, respectively.

Table 1. The results of percent SD difference histograms using the TPS recalculation and the dose warping method with each setup and range error.

	Relative volume (%) in the irradiated region									
	Lung				Prostate		HN			
Difference	Setup 3 ∗ 3 mm	Setup 6 ∗ 6 mm	Range 3%	Range 1%	Setup 3 ∗ 3 mm	Range 3%	Setup 3 ∗ 3 mm	Setup 2 ∗ 2 mm	Range 3%	Range 1%
±0.5%	83%	85%	85%	91%	95%	85%	80%	85%	91%	93%
±1.0%	89%	90%	92%	97%	99%	96%	89%	90%	96%	96%

Note. The relative volumes in the irradiated region within ±0.5% and ±1% of the SD difference are listed.

clinical target volume (CTV) and gross tumor volume (GTV) were generally small and those of OAR were generally large. The maximum difference of SD values calculated by the warping method and the TPS recalculation was 6% for the GTV in the lung case.

The DVHs and SDVHs for lung, prostate and head/neck cases are shown in figure 6. The DVHs were calculated in case of the combination of the setup error of ±3 mm and the range

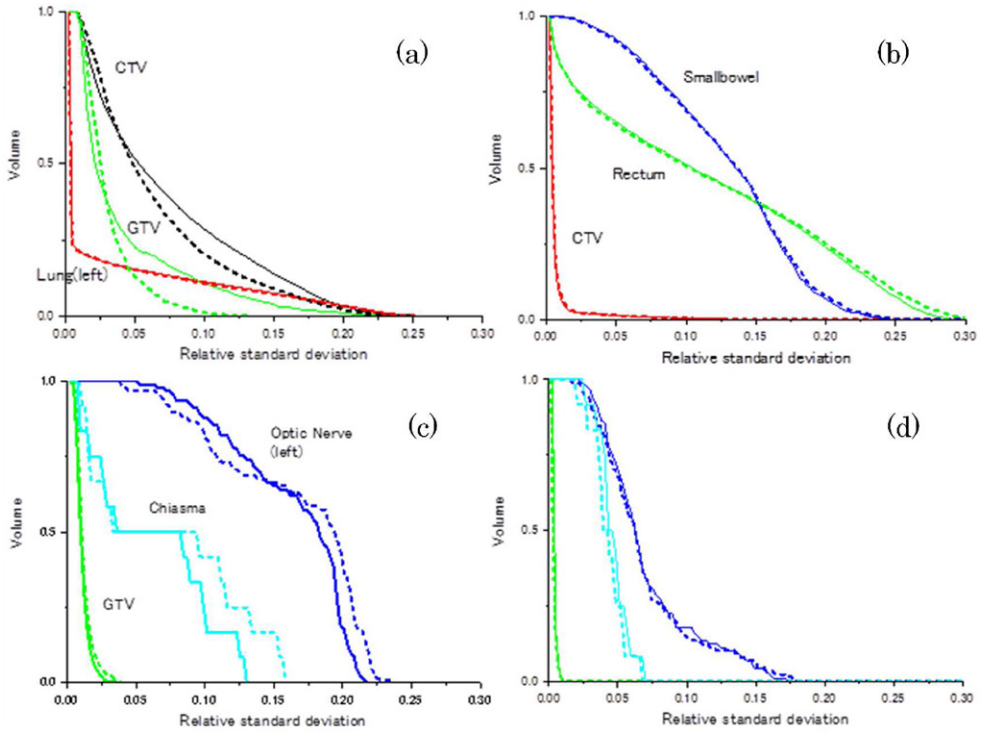

Figure 5. The SDVHs for lung (a), prostate (b) and head/neck (c) and (d) which are calculated using the setup error of ±3 mm ((a)–(c)) and the range error of ±3% (d). Dashed lines were derived from the TPS recalculation and solid lines were derived from the warping method.

Table 2. The relative SD values corresponding to volume of 5% using the dose warping method and TPS recalculation with the setup error, range error and the number of fractions.

		SD5% (warping method, TPS recalculation)					
		Lung		Prostate		HN	
		GTV	Lung (left)	CTV	Rectum	GTV	Optic nerve (left)
Setup 3 ∗ 3 mm	SD5%	13%, 7%	18%, 18%	1%, 1%	26%, 27%	2%, 2%	21%, 22%
Range 3%	SD5%	4%, 3%	11%, 11%	1%, 2%	14%, 14%	1%, 1%	15%, 15%
Setup 3 ∗ 3 mm + range 3%, fraction 1	SD5%	14%, 8%	22%, 22%	2%, 2%	28%, 29%	2%, 3%	23%, 25%
Setup 3 ∗ 3 mm + range 3%, fraction 4	SD5%	8%, 5%	14%, 14%	1%, 2%	17%, 19%	1%, 1%	17%, 18%

error of ±3%. DVHs with one SD in the target corresponded closely to the envelopes of DVH curves covering most DVH curves except for the low dose region. However DVHs of worst dose distributions in the target were excessive for the envelopes of DVH curves of each site. DVHs with one SD in the rectum and left optic nerve were not covering all DVH curves but will be useful as rough indicators of the DVH variations corresponding to the envelopes of

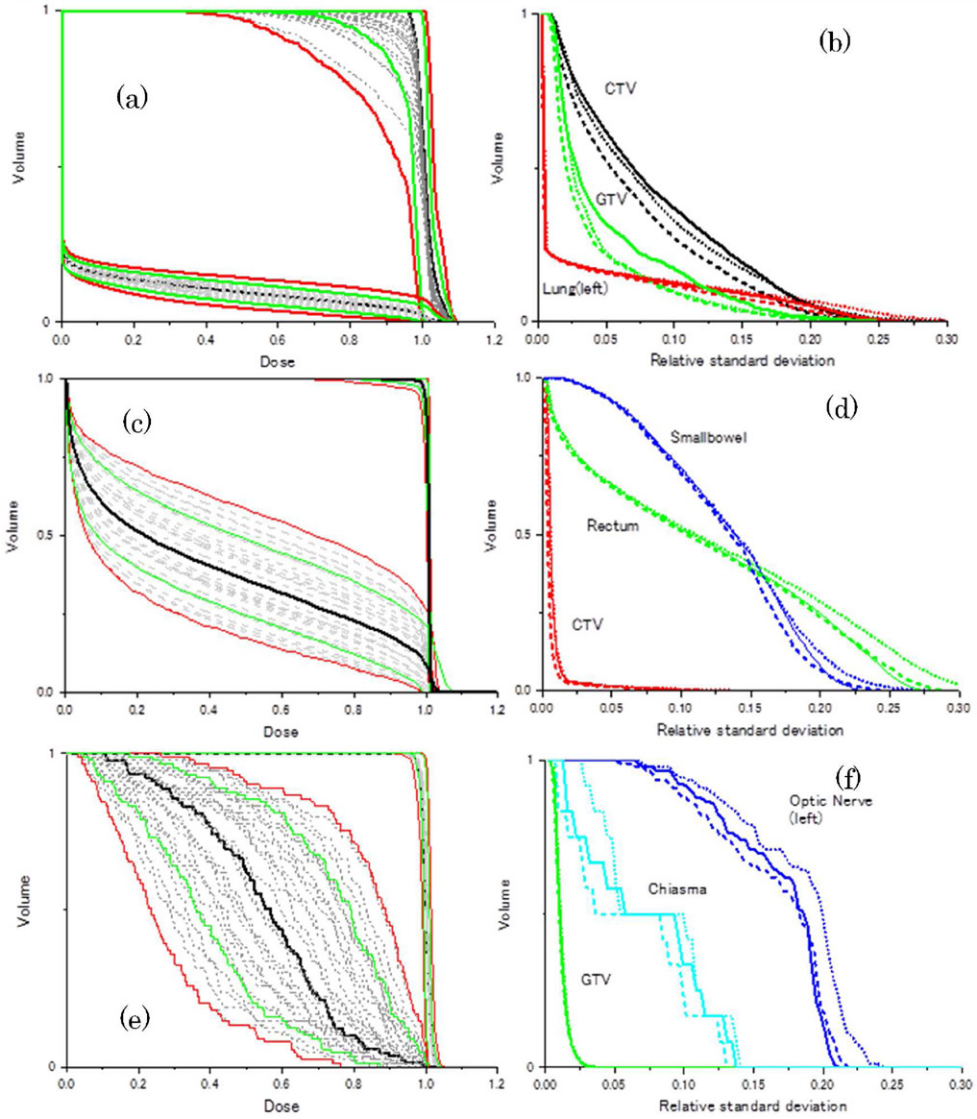

Figure 6. DVHs (left column) and SDVHs (right column) for lung (a) and (b), prostate (c) and (d) and head/neck (e) and (f). The DVHs were calculated in case of a combination of the setup error of ± 3 mm and the range error of $\pm 3\%$ relative to the mean dose of GTV (lung and head/neck) or CTV (prostate). In the DVHs red lines represent the DVHs of the minimum and maximum worst dose distribution. Green lines: the DVHs of the mean dose distribution added or subtracted one SD for each voxel. Gray lines: the DVHs of each scenario with the assumed errors. In the SDVHs solid lines correspond to a combination of both setup and range errors and dashed lines only indicate the case of the setup error. Dotted lines represent the combined SDs using the separately calculated setup and range errors.

DVH curves. The influence of range errors in limited places behind the target was smaller than that of the setup error in this assumed case. Therefore, SDVHs in a combination of both setup and range errors were nearly equivalent to SDVHs with only setup error. SDVHs obtained using the combined SDs ${}_{\mathrm{rel}}\sigma_{\mathrm{c}}$ of separately calculated setup error and range error agreed well

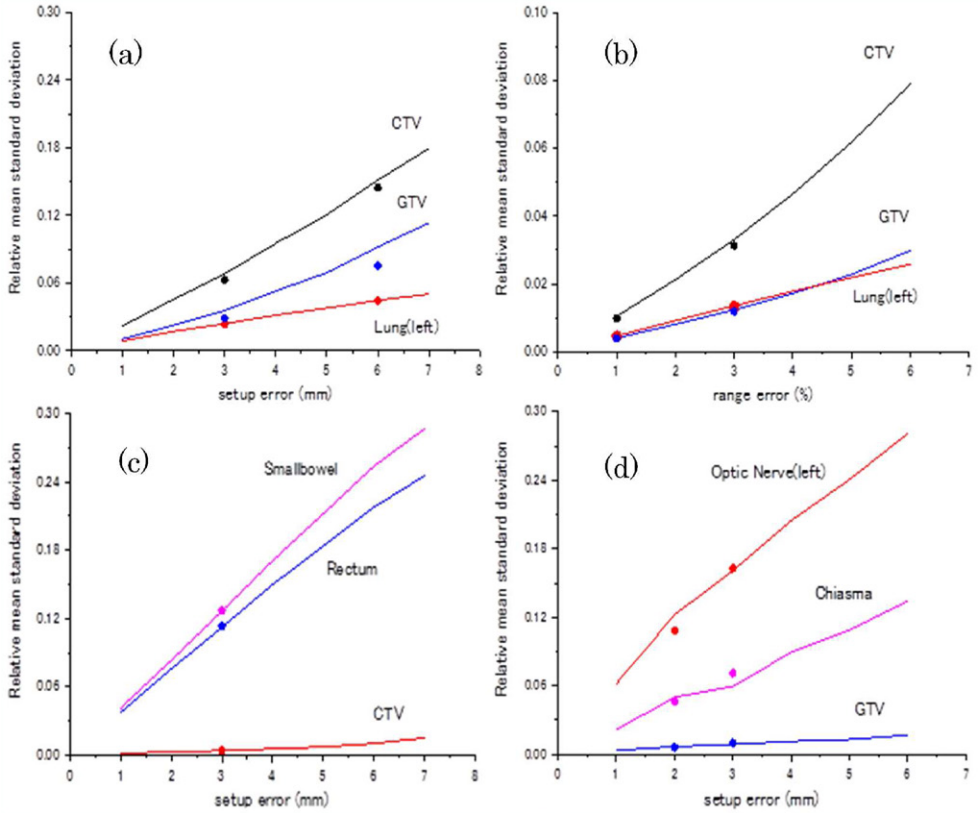

Figure 7. Mean value of relative SDs for lung (a) and (b), prostate (c) and head/neck (d) which were calculated depending on the setup error ((a), (c), (d)) and the range error (b). Solid lines were derived from the warping method and filled circles were obtained using the TPS recalculation.

with those obtained using the combination of setup and range errors. The dose SD distributions derived from the combined SD (equation (5)) were almost equal to those in the combination. However DVHs of the worst dose distribution and mean dose distribution with one SD in the combination of setup and range errors changed depending on the considered number of scenarios and conditions, but the change of the DVHs (red and green lines) in the OARs was comparatively small (figure 6).

Mean values of relative SDs in each volume of interest are shown in figure 7 for lung, prostate and head/neck cases. The SDs were calculated depending on the setup error and the range error. In these plans the SDs were monotonically increased with setup and range errors and the results of the warping method agreed well with TPS dose recalculation. The changes of mean values of SDs in the GTV or CTV were small with the increase of the setup error and range error due to the effect of the margin. In the lung case, the mean value of SDs of GTV rapidly increased with the change of setup and range errors.

3.3. Influence of fractionation

Relative dose SDs ${}_{\mathrm{rel}}\sigma_{\mathrm{c}}(i,j,k,n)$ combined with the setup error of ± 3 mm and range error of $\pm 3\%$ are shown in figure 8. The cases of the number of fractions 1 and 4 are shown in

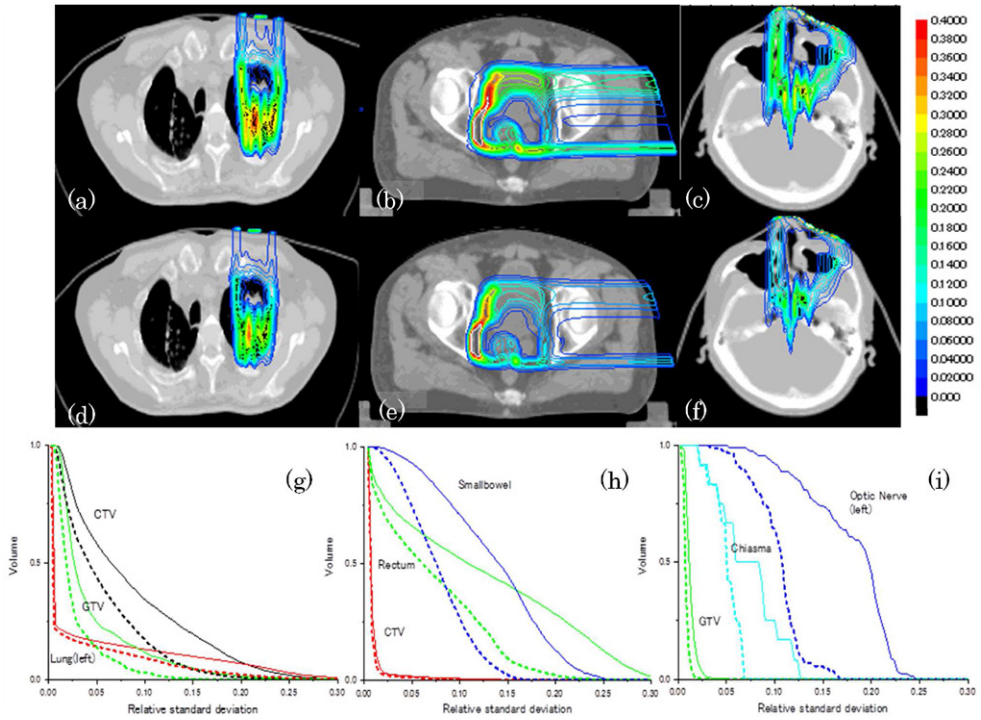

Figure 8. Relative combined SDs calculated by the proposed method with the setup error of ±3 mm and range error of ±3% with the scale running from 0 to 0.4. (a)–(c) are for when the number of fractions is 1, and (d)–(f) are for when the number of fractions is 4. The SDVHs for lung (g), prostate (h) and head/neck (i) correspond to the dose uncertainty distribution of the number of fractions (1 and 4) using the proposed method. Solid lines: number of fractions =1. Dashed lines: number of fractions =4.

(a)–(c) and (d)–(f), respectively. The SDVHs for lung, prostate and head/neck correspond to the dose SD distributions of the number of fractions (1 and 4) using the proposed method. The dose SD distributions decreased with the number of fractions, especially in OAR around the target due to averaging of the random setup error. In lung and HN cases, the SDVHs in GTV and CTV decreased slightly and the SDVHs in OAR changed largely. In rectum, SD5% decreased largely by about 10 pp with the change in number of fractions from 1 to 4 (table 2).

DVHs for lung calculated in the case of the setup error of ±4 mm and the range error of ±3% relative to the prescribed dose of target depending on the number of fractions and the fractionation effect are shown in figure 9. In the lung case, the combined dose uncertainty rapidly decreased with the change of the number of fractions from 1 to 4 and the DVHs with SD are changed with the number of fractions. In the case where all uncertainties are systematic, DVH with SD is largely changed from combined dose uncertainty due to the large contribution of setup error, while in the case that all uncertainties are random, it is almost equivalent to the DVH with combined dose uncertainty.

Figure 10 shows mean value of relative SDs for lung GTV, prostate CTV and head/neck GTV calculated depending on the number of fractions, setup error and range error. The setup error is fixed to ±3 mm in (b) and the range error is fixed to ±3% in (a), (c) and (d). In the lung case the combined SDs ${}_{\mathrm{rel,mean}}\sigma_{\mathrm{c}}'$ at the setup error ±3 mm reached the extent from 1 to 2.5% with the increase of the number of fractions depending on the range error and those at

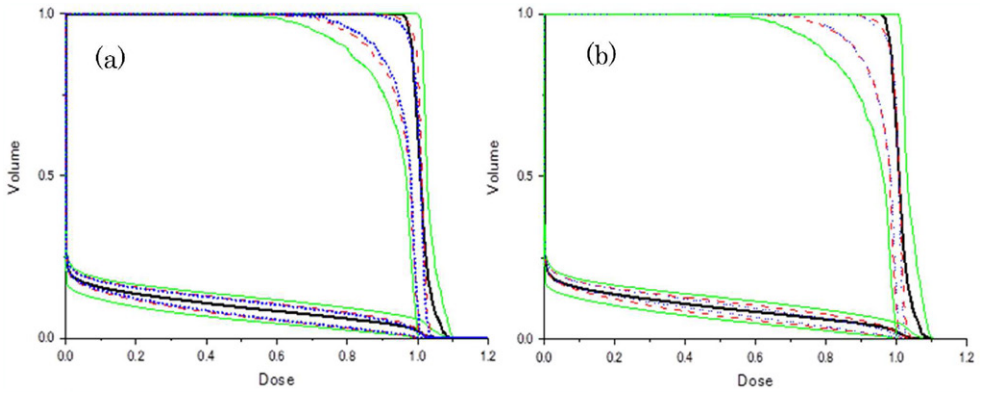

Figure 9. DVHs for lung calculated in case of the setup error of ±4 mm and the range error of ±3% relative to the prescribed target dose depending on the number of fractions (a) and the fractionation effect (b). In (a), red dashed lines represent the DVHs of the mean dose distribution added or subtracted one combined SD of the number of fractions 4. Green solid lines: number of fractions =1. Blue dotted lines: number of fractions =8. In (b), blue dotted lines represent the one combined SD of the number of fractions =8. Green solid lines: the number of fractions =8 and all uncertainties are systematic. Red dashed lines: the number of fractions =8 and all uncertainties are random.

the range error ±3% reached about 1.4% at minimum. The results calculated from mean value of SDs ${}_{\mathrm{rel,mean}}\sigma'_{\mathrm{c}}$ of the target agreed with ${}_{\mathrm{rel,mean}}\sigma_{\mathrm{c}}$ calculated from the combined dose uncertainty distributions using the proposed method. In the prostate the influence of the range error and setup error were similar and smaller than at other sites. In the lung they were considerably larger than at other sites and the influence of the range error was still large with the increase of the number of fractions from (b) of figure 10. In the influence of range error, the difference of combined dose uncertainty became larger with the increased the number of fractions, while in the influence of the setup error the combined dose uncertainty decreased rapidly and converged with the increased the number of fractions. When all uncertainties are systematic the relative combined dose uncertainty is constant, while when all uncertainties are random it decreases more rapidly with the increased number of fractions. In the prostate the difference between the random uncertainty case and combined dose uncertainty became larger with the increased the number of fractions due to the large contribution of the range error.

3.4. Computation time

The computation times of the dose warping method and TPS recalculation with the setup error and range error are listed in table 3. The computation time of the warping method was less than a few seconds for each scenario and that of the full dose calculation method using the TPS was a few minutes for each beam and depended on the tumor sites. The computation of the warping method for each scenario became faster with the increased number of scenarios and was approximately 60 times faster than that of the TPS in the setup error.

4. Discussion

The dose warping method does not consider the change of scattered dose components separately from the nominal dose distribution in the TPS and uses the parallel beam approximation,

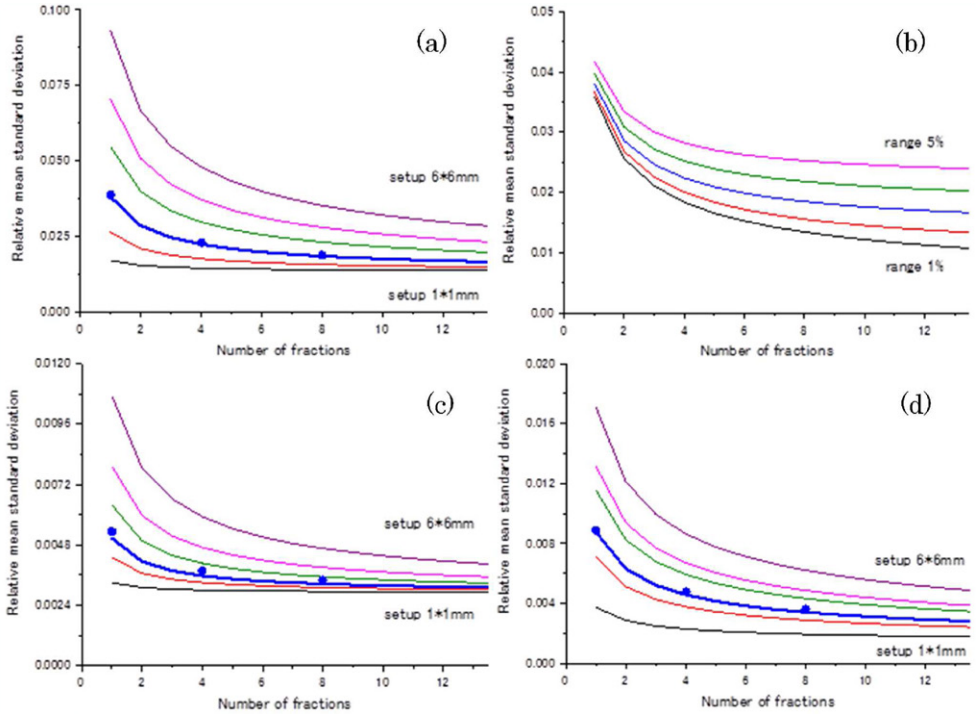

Figure 10. Mean value of relative SDs for lung GTV (a) and (b), prostate CTV (c) and head/neck GTV (d) calculated depending on the number of fractions, setup error and range error. Solid lines were derived from the proposed method using the mean value of SDs of separately calculated setup error and range error, and filled circles were derived from the proposed method using the combined dose uncertainty distributions in each number of fractions. The setup error is fixed to ± 3 mm in (b) and the range error is fixed to $\pm 3\%$ in (a), (c) and (d).

Table 3. Computation time of the dose warping method and TPS with the setup error and range error.

Tumor site	Number of scenarios	Volume of target (c.c.)	Computation time (s)	Computation time/scenario	Computation time/scenario (TPS)
Lung	8 (setup)	48.8	9.0	1.1	59
	26 (setup + range)		16.4	0.6	
Prostate	8 (setup)	197.3	14.7	1.8	187
	26 (setup + range)		19.8	0.8	
HN	8 (setup)	109.5	5.6	0.7	82
	26 (setup + range)		9.0	0.3	

which is limited with regard to the accuracy of dose uncertainty estimation. The limitations resulting from this approximation will become larger in the vicinity of regions with a large variation of tissue density lateral to the beam direction, and these can occur locally in the case of the small volume of OARs or near the beam penumbra and high dose gradient. The sub-pencil beam algorithm with a triple Gaussian beam model used in the TPS agreed well with the measured results in a heterogeneous head phantom (Inaniwa *et al* 2014) but is not ideal

in calculating the lateral scatter components under complex geometries, high heterogeneities or high dose gradients. Therefore a Monte Carlo dose calculation method will be suitable for more precise verification of the warping method.

The warping method is used to calculate the dose distributions under various conditions for evaluating the dose SD distributions, and for potential incorporation into a robust optimization process. Some researchers have reported that multiple dose distributions can be incorporated into treatment planning for robust plan optimization (Pflugfelder *et al* 2008, Unkelbach *et al* 2009, Fredriksson *et al* 2011, Inaniwa *et al* 2011, Liu *et al* 2012). The number of dose calculations needed to model the plan robustness will have to be balanced against the time needed to calculate the plans, especially in the case of iterative optimization algorithms. Conventional robust planning approaches may require hundreds of dose calculations in every iteration of the optimization process (Sobotta *et al* 2012). The approximated dose can be used for fast estimation of major dosimetric influence with various uncertainty sources or for intermediate dose calculations in plan optimization because highly accurate dose distributions are not needed. The fast computation time and adequate accuracy of the warping method will be useful for these evaluation and optimization processes.

The robustness analysis method is fast owing to approximate calculations in a small set of discrete calculations of the dose under the extreme boundaries of the assumed errors. This approach assumes that the dose distributions recalculated for the largest delivery errors represent the largest deviations of the delivered dose from the nominal dose, and these dose distributions can be used for the plan robustness evaluation in the case of practical delivery errors. It is not verified generally that the largest deviations of point dose from the nominal dose distribution are caused by the largest delivery errors. However, in a volumetric analysis the DVH obtained from the dose distributions at the extreme boundaries of assumed error represent the worst DVH (Casiraghi *et al* 2013). The minimum and maximum worst dose distributions and the DVHs may represent unphysical maximum deviation and an incorrect estimation of actual plan robustness. From the responsivity analysis the monotonic increase of mean value of SDs indicated that the evaluation at the extreme boundaries of assumed error is cautious for the dose uncertainty of the scenarios within the given error. These results showed the DVH curves with one SD in the target are covering most DVH curves in all scenarios within the extreme boundaries in the considered cases. Therefore the dose SD distributions will be used as practical indicators of dose uncertainty in a clinical situation. When the setup error is random and the range error is systematic for each field in multi-field irradiation of one field per day, which is generally used in particle therapy, the relative combined dose uncertainty is expressed by the following equation:

$$ {}^{m}_{\mathrm{rel}}\sigma_{\mathrm{c}}(i,j,k,n_f) = \frac{\sqrt{\sum_f n_f \cdot \sigma^2_{f,\mathrm{setup}}(i,j,k) + \left\{\sum_f n_f \cdot \sigma_{f,\mathrm{range}}(i,j,k)\right\}^2}}{\sum_f n_f \cdot D_{f,\mathrm{p}}} \tag{10} $$

where $D_{f,\mathrm{p}}$ is the prescribed target dose for each field and n_f is the number of fractions for each field.

In the analysis of the fractionation effect, it will be possible to estimate the influence of the number of fractions on the dose uncertainty by simply using the mean value of relative SDs ${}_{\mathrm{rel,mean}}\sigma'_{\mathrm{c}}$ shown in figure 10. The dose uncertainty distributions with consideration of fractionation in the simple geometry were reported using the analytical probabilistic modeling in proton and photon beams (Bangert *et al* 2013), and their similar tendency that the dose

uncertainties decrease with the number of fractions was confirmed. The method proposed in this paper is simpler and faster than the probabilistic modeling. When the setup error is completely random and comparatively large, the influence of the setup error decreases rapidly with the number of fractions. When the number of fractions is small, the influence of the setup error will be greater than that of the range error, and therefore accurate evaluation and suppression of the setup error are required. The comprehensive analysis of dose uncertainty will need to take into account the irradiation technique, the number of fractions, the immobilization technique, and inter and intra-fractional motion in order to optimize the particle therapy.

In clinical routines we propose the following procedures for robustness analysis. The combined dose uncertainty distribution is calculated for the condition of given errors and the number of fractions in order to check the large uncertainty region, and SDVH and DVH with SD are confirmed. Robustness indicators, such as SD5% or mean SD, can be used as a guide to improve treatment planning by comparing them to the criteria of the robustness protocols, which are determined by the retrospective analysis of treatment plans. However, detailed investigations of robustness indicators for each site or irradiation method will be needed. In next step the estimation of clinical effectiveness with dose uncertainty will be performed in connection with robustness analysis.

5. Conclusions

In this paper we have presented a method for evaluating the robustness of carbon-ion therapy treatment plans to spatial and range uncertainties by using the dose SD distributions and SDVHs. Irradiated areas that could be affected by these uncertainties can be readily visualized by using the dose SD distribution. The dose warping method used in the proposed robustness analysis method is based on the difference between WEPL under setup and range errors in order to derive a perturbed dose distribution from the planned dose distribution. We validated the warping method by comparing the dose SD distributions and SDVHs in a limited number of scenarios at the extreme boundaries of the assumed error for three different tumor sites. In these comparisons the results of the warping method agreed well with those of the TPS dose recalculation. The effectiveness of calculations at the extreme boundaries was confirmed by responsivity and DVH analysis. The typical computation of the warping method is approximately 60 times faster than that of the full dose calculation method using the NIRS TPS.

When considering a combination of both setup and range errors, dose SD distributions were evaluated by the SDVHs and DVHs and compared with the combined dose uncertainty of separately calculated setup and range errors. The uncertainty distributions and SDVHs calculated in both cases agreed well. The influence of fractionation on the combined SD $_{\mathrm{rel}}\sigma_{\mathrm{c}}$ was estimated using the SDVHs and mean value of SDs in defined volumes of interest with various setup and range errors. The combined SDs depended heavily on the number of fractions, assumed error and tumor site. The proposed robustness analysis method will be useful for routinely estimating the dose uncertainty associated with fractionation and for quickly comparing alternative plans.

Acknowledgments

The authors express the gratitude to the organizations and individuals who supported this study: W Furuichi of Accelerator Engineering Corporation for technical support in treatment planning, and Dr M Koto, Dr N Yamamoto, and Dr H Tsuji of the Radiation Oncology

Section of NIRS for the clinical cases. This study was supported in part by JSPS KAKENHI Grant Number 26860406.

References

Albertini F, Hug E B and Lomax A J 2011 Is it necessary to plan with safety margins for actively scanned proton therapy? *Phys. Med. Biol.* **56** 4399–413

Bangert M, Hennig P and Oelfke U 2013 Analytical probabilistic modeling for radiation therapy treatment planning *Phys. Med. Biol.* **58** 5401–19

Baum C, Alber M, Birkner M and Nusslin F 2006 Robust treatment planning for intensity modulated radiotherapy of prostate cancer based on coverage probabilities *Radiother. Oncol.* **78** 27–35

Bolsi A, Lomax A J, Pedroni E, Goitein G and Hug E 2008 Experiences at the Paul Scherrer Institute with a remote patient positioning procedure for high-throughput proton radiation therapy *Int. J. Radiat. Oncol. Biol. Phys.* **71** 1581–90

Bortfeld T 2006 IMRT: a review and preview *Phys. Med. Biol.* **51** R363

Casiraghi M, Albertini F and Lomax A J 2013 Advantages and limitations of the 'worst case scenario' approach in IMPT treatment planning *Phys. Med. Biol.* **58** 1323–39

Chan T C Y, Bortfeld T and Tsitsiklis J N 2006 A robust approach to IMRT optimization *Phys. Med. Biol.* **51** 2567–83

Fredriksson A, Forsgren A and Hardemark B 2011 Minimax optimization for handling range and set-up uncertainties in proton therapy *Med. Phys.* **38** 1672

Inaniwa T, Kanematsu N, Hara Y, Furukawa T, Fukahori M, Nakao M and Shirai T 2014 Implementation of a triple Gaussian beam model with subdivision and redefinition against density heterogeneities in treatment planning for scanned carbon-ion radiotherapy *Phys. Med. Biol.* **59** 5361–86

Inaniwa T, Kanematsu N, Matsufuji N, Kanai T, Shirai T, Noda K, Tsuji H, Kamada T and Tsujii H 2015 Reformulation of clinical-dose system for carbon-ion radiotherapy treatment planning at the Natural Institute of Radiological Sciences, Japan *Phys. Med. Biol.* **60** 3271–86

Inaniwa T, Kanematsu N, Furukawa T, and Hasegawa A 2011 A robust algorithm of intensity modulated proton therapy for critical tissue sparing and target coverage *Phys. Med. Biol.* **56** 4749–70

Kanematsu N, Matsufuji N, Kohno R, Minohara S and Kanai T 2003 A CT calibration method based on the polybinary tissue model for radiotherapy treatment planning *Phys. Med. Biol.* **48** 1053–64

Liu W, Zhang X, Li Y and Mohan R 2012 Robust optimization of intensity modulated proton therapy *Med. Phys.* **39** 1079–91

Lomax A J, Böhringer T, Bolsi A, Coray D, Emert F, Goitein G and Weber D C 2004 Treatment planning and verification of proton therapy using spot scanning: initial experiences *Med. Phys.* **31** 3150

McGowan S E, Albertini F, Thomas S J and Lomax A J 2015 Defining robustness protocols: a method to include and evaluate robustness in clinical plans *Phys. Med. Biol.* **60** 2671–84

McShan D L, Kessler M L, Vineberg K and Fraass B A 2006 Inverse plan optimization accounting for random geometric uncertainties with a multiple instance geometry approximation (MIGA) *Med. Phys.* **33** 1510–21

Moyers M F, Sardesai M, Sun S and Miller D W 2010 Ion stopping powers and CT numbers *Med. Dosim.* **35** 179–94

Park P C, Cheung J, Zhu X R, Sahoo N, Court L and Dong L 2012 Fast range-corrected proton dose approximation method using prior dose distribution *Phys. Med. Biol.* **57** 3555–69

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

Ploquin N, Kay I, Rangel-Baltazar A, Lau H and Dunscombe P 2006 A comparison of techniques for simulating set-up error and uncertainty in head and neck IMRT *Med. Phys.* **33** 3213

Schaffner B and Pedroni E 1998 The precision of proton range calculations in proton radiotherapy treatment planning: experimental verification of the relationship between CT-HU and proton stopping power *Phys. Med. Biol.* **43** 1579–92

Schneider U, Pedroni E and Lomax A J 1996 The calibration of CT Hounsfield units for radiotherapy treatment planning *Phys. Med. Biol.* **41** 111–24

Sobotta B, Sohn M and Alber M 2010 Robust optimization based upon statistical theory *Med. Phys.* **37** 4019–28

Sobotta B, Söhn M and Alber M 2012 Accelerated evaluation of the robustness of treatment plans against geometric uncertainties by Gaussian processes *Phys. Med. Biol.* **57** 8023

Tilly D and Ahnesjö A 2015 Fast dose algorithm for generation of dose coverage probability for robustness analysis of fractionated radiotherapy *Phys. Med. Biol.* **60** 5439–54

Trofimov A, Unkelbach J, DeLaney T F and Bortfeld T 2012 Visualization of a variety of possible dosimetric outcomes in radiation therapy using dose–volume histogram bands *Pract. Radiat. Oncol.* **2** 164–71

Unkelbach J, Bortfeld T, Martin B and Soukup M 2009 Reducing the sensitivity of IMPT treatment plans to setup errors and range uncertainties via probabilistic treatment planning *Med. Phys.* **36** 149–63

van Herk M 2004 Errors and margins in radiotherapy *Semin. Radiat. Oncol.* **14** 52–64

Witte M G, van der Geer J, Schneider C and Lebesque J V 2007 IMRT optimization including random and systematic geometric errors based on the expectation of TCP and NTCP *Med. Phys.* **34** 3544–55