Defining robustness protocols: a method to include and evaluate robustness in clinical plans

Defining robustness protocols: a method to include and evaluate robustness in clinical plans

S E McGowan[1], F Albertini[3], S J Thomas[2] and A J Lomax[3]

[1] Department of Oncology, University of Cambridge, Cambridge, UK
[2] Cambridge University Hospitals NHS Foundation Trust, Medical Physics and Clinical Engineering Cambridge, UK
[3] Paul Scherrer Institute, Centre for Proton Radiotherapy Villigen-PSI, Switzerland

E-mail: sem89@cam.ac.uk

Abstract

We aim to define a site-specific robustness protocol to be used during the clinical plan evaluation process. Plan robustness of 16 skull base IMPT plans to systematic range and random set-up errors have been retrospectively and systematically analysed. This was determined by calculating the error-bar dose distribution (ebDD) for all the plans and by defining some metrics used to define protocols aiding the plan assessment. Additionally, an example of how to clinically use the defined robustness database is given whereby a plan with sub-optimal brainstem robustness was identified. The advantage of using different beam arrangements to improve the plan robustness was analysed. Using the ebDD it was found range errors had a smaller effect on dose distribution than the corresponding set-up error in a single fraction, and that organs at risk were most robust to the range errors, whereas the target was more robust to set-up errors. A database was created to aid planners in terms of plan robustness aims in these volumes. This resulted in the definition of site-specific robustness protocols. The use of robustness constraints allowed for the identification of a specific patient that may have benefited from a treatment of greater individuality. A new beam arrangement showed to be preferential when balancing conformality and robustness for this case. The ebDD and error-bar volume histogram proved effective in analysing plan robustness. The process of retrospective analysis could be used to establish site-specific robustness planning protocols in proton therapy. These protocols

allow the planner to determine plans that, although delivering a dosimetrically adequate dose distribution, have resulted in sub-optimal robustness to these uncertainties. For these cases the use of different beam start conditions may improve the plan robustness to set-up and range uncertainties.

Keywords: radiotherapy, robustness, cancer, proton beam therapy, treatment planning

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

1. Introduction

Protons have a finite range, highly dependent on the electron density of the material they are traversing, resulting in a steep dose gradient at the distal edge of the Bragg peak. Due to these characteristics and the advancements in computation and technology it has led to the ability to produce and deliver treatments with greater conformality; sparing normal tissue through the use of intensity modulation, multiple fields and steep dose gradients. For these reasons proton therapy is considered to be advantageous in treating most childhood cancers and certain adult cancers; including those of the skull base and head and neck (Jones 2008, Ares *et al* 2009, Durante and Loeffler 2010, Allen *et al* 2012, De Ruysscher *et al* 2012). Besides meeting planning constraints the plan is also required to meet the aims of the planner at each and every fraction, therefore ensuring it is robust to uncertainties such as Hounsfield Unit (HU) error or patient motion. Due to the steep dose gradients and range sensitivity of proton therapy, robustness is a real concern (Lomax 2008b, Lassen-Ramshad *et al* 2011). Several factors will influence a plan's robustness including the delivery method, immobilization technique and beam orientation. Therefore, balancing the compromise between plan conformality and plan robustness through careful positioning of dose gradients is critical to successful planning and optimum treatment delivery (Lomax 2004, Albertini *et al* 2011, McGowan *et al* 2013).

At the fore-front of proton beam therapy is three-dimensional intensity modulated particle therapy (3D IMPT), the equivalent to intensity modulated radiation therapy in conventional x-ray radiotherapy, whereby Bragg peaks are optimally weighted and placed as 'spots' throughout the target volume (Lomax 1999). The spot weight is essentially its fluence (the number of protons in a given spot). Due to the number of spots required to fill the target an optimization algorithm is required to determine the optimal solution for their individual weights. Due to the highly degenerate nature of intensity modulation optimization (Webb 2003, Albertini *et al* 2011), there exist many solutions to the given problem of delivering a homogenous dose to target whilst ensuring the dose to certain organs at risk (OARs) are within predefined constraints. Despite there being many solutions that may satisfy the dosimetric constraints, some of these solutions may offer greater plan robustness and therefore a greater chance of cure whilst limiting the chance of normal tissue complications. Several authors have investigated this through including robustness in the optimization itself (Unkelbach *et al* 2007, Pflugfelder 2008, Chen 2012) or by altering the start conditions, such as number of beams and their orientation to obtain solutions, to satisfy both dosimetric and robustness requirements (Lomax 2008a, Albertini *et al* 2010).

Ultimately, the trade-off between plan conformality and plan robustness needs to be explored more thoroughly as well as the establishment of site-specific robustness thresholds. Increasing plan robustness may lead to a decrease in plan conformality. Nevertheless, clinical experience so far has been achieved without considering the robustness of the plan. Although

the importance of including robustness during the plan analysis is recognized, it is not yet clear how much the nominal dose distribution can be compromised without compromising the patient's treatment. The actual status is therefore that the robustness analysis is not yet fully included into the clinical plan evaluation phase. On one hand, no commercial planning system has a robustness evaluation module integrated. On the other hand, to the best of our knowledge, the research centres that have developed their own robustness analysis tool, are not yet being clinically using to make decisions for all treated patients. It is therefore of paramount importance to find a way to introduce the robustness analysis into the clinical plan evaluation process, without compromising the nominal dosimetric plan quality.

In this paper we suggest a way to introduce the robustness analysis gradually into the clinic process of plan evaluation. In particular we suggest a way to define site-specific, and also centre-specific, robustness database by retrospectively analysing patients already treated in the specific centre. This database can be used during the plan evaluation and during the plan calculation phase to aid the planner to select the optimal plan.

2. Methods and materials

We propose to define site-specific robustness protocols by retrospectively analysing the robustness, to both range and set-up errors, for a sample of Chordoma and Chondrosarcoma plans. All plans were clinically acceptable and had been delivered in the last 5 years at the Paul Scherrer Institute (PSI). All plans were calculated using the in-house treatment planning system at PSI. The treatment planning system uses a pencil beam model to calculate absolute dose delivered to the patient (Scheib and Pedroni 1992, Pedroni *et al* 2005) and uses a quasi-Newton optimization technique (Lomax 1999). The model parameters have been adjusted to fit the measured data, which takes into account range straggling, proton energy loss and the width of the initial energy spectrum. To take into account density heterogeneities in the patient, computed tomography (CT) data is converted into water equivalent depth using a ray-casting model (Schaffner *et al* 1999).

To determine plan robustness the set-up and range uncertainties were modelled using the method proposed by Albertini *et al* (2011) and later validated by Casiraghi *et al* (2013). This method will be briefly described here.

2.1. Set-up uncertainty

The effect of random set-up uncertainty on plan robustness was simulated by recalculating the nominal dose distribution on 14 isocentrically shifted CT data sets (Albertini *et al* 2011). The shifts used were applied along the major anatomical axes and their indices in both positive and negative directions (equivalent to the six faces and eight vertices of a cube) resulting in multiple spatially shifted does distributions associated with each shifted CT data set. Bolsi *et al* (2008) showed that the use of a remote patient positioning device reduced systematic set-up errors to negligible values (below 0.6 mm), whereas random set-up errors of over 2 mm can be observed, depending on tumour location and immobilization device used. Using these data a 3D shift vector was calculated from the standard deviation of each shift in left-right (LR), anterior-posterior (AP) and cranium-caudal (CC) directions which represents of the radius of the spherical error-space used in this analysis. Following the same method as Albertini *et al* (2011) a shift of magnitude equivalent to a confidence interval of 85% was applied. In particular shifts of ±3.2 mm and ±4.7 mm were used for bite block and for head mask fixation devices respectively. The nominal dose distribution and each recalculated dose distribution are then

combined into one 'error-bar' dose distribution (ebDD) by storing in each voxel the difference between the maximum and minimum value for that corresponding voxel from all the plans. This amplitude of dose errors, ΔD^i, in each voxel is calculated using formula (1).

$$\Delta D^i_{(random)} = \mathrm{Max}\left(D^i_{+LR}, D^i_{-LR}, ..., D^i_{-CC}, D^i_{nom}\right) - \mathrm{Min}\left(D^i_{+LR}, D^i_{-LR}, ..., D^i_{-CC}, D^i_{nom}\right), \quad (1)$$

The model then, assumes that the value in every voxel represents a dose-error-bar that brackets all the possible deviations from the nominal dose distribution that can be detected when patient shifts within an 85% confidence interval occur (Casiraghi *et al* 2013).

2.2. Range uncertainty

Radiotherapy plans are generally produced using a single CT data set of the patient and so any uncertainties in either the HUs, or their conversion to relative proton stopping power, will lead to a systematic uncertainty that will propagate throughout the treatment delivery (Lomax 2008b). The effect of systematic range uncertainty on plan robustness was simulated by recalculating the nominal dose distribution on CT data with HUs altered by ±3%, thus to simulate an overshoot or an undershoot scenario (Lomax 2008b). In this case the ebDD is calculated using formula (2).

$$\Delta D^i_{(systematic)} = \mathrm{Max}\left(D^i_{+HU}, D^i_{-HU}, D^i_{nom}\right) - \mathrm{Min}\left(D^i_{+HU}, D^i_{-HU}, D^i_{nom}\right), \quad (2)$$

The robustness of each volume has been analysed by extracting the error-bar volume histogram (ebVH). The error-bar volume histograms represent a useful and simple tool to evaluate the quality of the treatment: the closer the histogram is to the '0-error' line, the more robust the dose is to that specific error. In particular its similarity in presentation to the familiar dose volume histogram used conventionally in radiotherapy allows for easy integration into the planning process. Several metrics can be defined and analysed as for example the maximum error, the mean error and the volume receiving an error of 3, 5 and 10% ($V_{eb}3$, $V_{eb}5$, $V_{eb}10\%$).

There is a greater clinical relevance in displaying the potential over-dosage error (highest minus nominal value) for the organ at risk and the under-dosage error (nominal minus lowest value) for the target volume, instead of the more general error-bar dose distribution (ebDD) discussed above. In this study we have distinguished between under-dosages and over-dosages when creating the database and advise the same for clinical use.

2.3. Retrospective robustness analysis

Chordomas and chondrosarcomas are treated to 74 Gy(RBE) and 68 Gy(RBE) respectively. Each treatment consists of three series delivering a given proportion of the total prescribed dose to the target at each fraction. The first series is generally comprised of a three-field uniform dose distribution (single field uniform dose (SFUD)) plan used to deliver up to 34–40 Gy(RBE) of homogenous dose to the target volume. The only constraint within the SFUD plan optimization is to deliver a homogeneous dose to the target volume. As a consequence, the resulting target coverage will be robust to uncertainties, but at the compromise of OARs sparing (Albertini *et al* 2011). The following two series are comprised of 3D IMPT plans used to provide dose to the target whilst prioritizing OAR sparing. The second series then brings the dose to the primary target up to 54 Gy(RBE) sparing
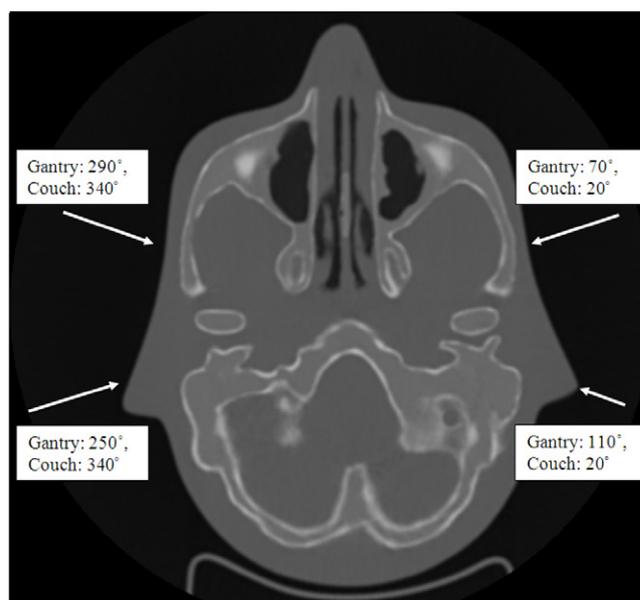
Figure 1. The A(4f) plan consists of four non-coplanar fields, two posterior obliques and two anterior lateral obliques. Clockwise from top left the beam angles used are (290°, 340°), (70°, 20°), (110°, 20°) and (250°, 340°), where the first angle is the gantry orientation and the second is the table orientation.

some critical structures. The third series (bringing the dose up to 68 or 74 Gy(RBE)) is delivered to boost a reduced target volume with the IMPT plan aimed to spare the critical structures. In all these cases, the planning target volume was isotropically grown from the clinical target volume (CTV) by 5 mm. The dose constraints are mainly defined for the brainstem (64 Gy(RBE) on the surface, 53 Gy(RBE) in the centre) and 60 Gy(RBE) in the optical structures (Ares *et al* 2009). In the IMPT plans, all the fluences of all the fields are optimized simultaneously. Unlike SFUD treatments, IMPT can deliver a number of non-uniform fields to produce the desired dose distribution.

The 3D-IMPT beam arrangement used clinically at PSI for the skull base is a four field, A(4f), beam arrangement illustrated in figure 1. This beam arrangement consists of four beams, two posterior obliques and two lateral obliques and was used to plan all IMPT plans in this study.

For each series, the plan robustness in the CTV for an underdose scenario was analysed for both types of uncertainty as well as the overdose scenario for the brainstem and for the optical structures. Due to their proximity with the target, the dose constraints defined for the brainstem and for the optical chiasm are driving the optimization outcome. The parameters measured include the mean and maximum dose uncertainties in the volume, the volume with an error of 3 and 5% ($V_{eb}3$ and $V_{eb}5\%$), and where these uncertainties occur in relation to anatomy and field direction. From these data a table of robustness parameters was created to aid planners in the selection between robustness and conformality.

2.4. Example case

To illustrate how the robustness database could be clinically used in practice, we identified one plan as an outlier. This plan was found to be un-robust in the brainstem compared to the

Table 1. Descriptions of each plan.

Plan	No. of fields	Beam arrangement
A(4f)	4	Two posterior obliques and two anterior lateral obliques
B(4f)	4	Two lateral and two posterior lateral obliques
A(6f)	6	Two posterior oblique, two anterior lateral obliques and two lateral fields
B(6f)	6	Two posterior obliques, two anterior lateral obliques and two posterior lateral obliques
C(4f)	4	Two lateral fields and two oblique fields all coplanar

sample. Due to the degeneracy associated with IMPT there exist many solutions to the given problem of ensuring target coverage whilst meeting dose and robustness constraints. Beyond changing the optimization itself (Unkelbach *et al* 2007, Pflugfelder 2008, Chen 2012), we can alter our starting conditions, such as number of beams and their orientation to obtain solutions to satisfy both dosimetric and robustness requirements. For this case, the plan was re-planned using four different beam arrangements.

Four additional start arrangements were investigated to determine which would be more optimal for this patient, including B(4f), B(6f), A(6f) and C(4f) shown in table 1 along with A(4f).

Each plan was created with the aim of keeping the dosimetric quality of the plan in the OARs and target as similar to those achieved in the A(4f) plan so that the degeneracy of robustness to start conditions could be investigated.

3. Results

3.1. Retrospective robustness analysis

Sixteen skull base 3D IMPT plans, series two and three, were retrospectively analysed in terms of robustness to systematic range and random set-up uncertainties. The volumes analysed were the brainstem, chiasm and the CTV. The results are shown in figure 2 where for each box plot the central mark is the median, the edges of the box are the 25th and 75th percentiles, the whiskers extend to the most extreme data points not considered outliers, and outliers are plotted individually. Outliers are considered to be points $\pm 2.7\sigma$. From the box plots in figure 2 it is seen that the impact of the random set-up error appears greater for an individual fraction than that resulting from the range error in all three volumes of interest (VOIs) for each parameter measured. Also that the mean and standard deviation in the range error was lower for the OARs than for the CTV; the opposite was true for the set-up error.

From our results an example of a robustness DATABASE for 3D IMPT skull-base plans has been generated (table 2). This table can be used to aid planners in visualizing the magnitude of effect that range and set-up errors can have on patient treatment in terms of percentage change from the nominal dose. The ranges represent the deviation from the nominal dose in that volume caused by the error for this patient sample. The values correspond to the 25th to 75th percentiles as seen in the box plots in figure 2 for an over-dosage scenario to OAR and under-dosage scenario to the CTV. The lower limits are included to allow the planner to see where a compromise between conformality and robustness can be made.
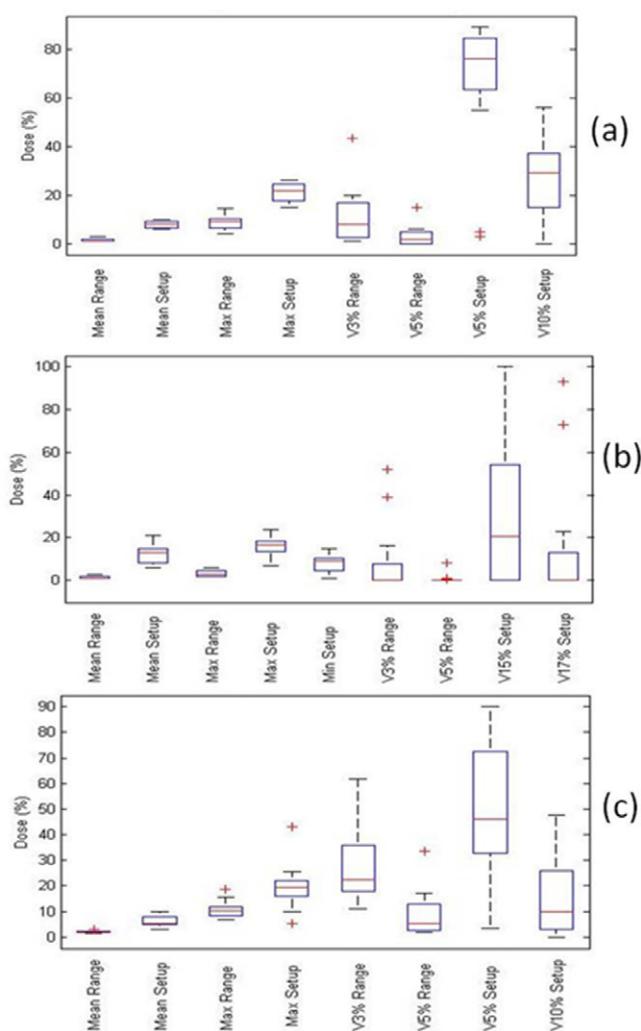
Figure 2. Robustness results for 16 skull base IMPT plans from eight patients to both the range and set-up uncertainties modelled in (*a*) in the brainstem, (*b*) in the chiasm and (*c*) in the CTV. On each box plot the central mark is the median, the edges of the box are the 25th and 75th percentiles, the whiskers extend to the most extreme data points not considered outliers, and outliers are plotted individually. Outliers are considered points $\pm 2.7\sigma$.

Table 2. An example of a robustness DATABASE defined for IMPT plans for skull-base cases.

VOI	Mean range (%)	Mean setup (%)	$V_{eb}3\%$ range (%)	$V_{eb}5\%$ range (%)	$V_{eb}5\%$ setup (%)	$V_{eb}10\%$ setup
Brainstem	1–2	6.5–9	2.5–17	5	63–85	15–37.5
Chiasm	1–1.5	8–15	<5	0	($V_{eb}15\%$) <50	($V_{eb}17\%$) <4
CTV	1.5–2.5	5–8	17.5–28	2.5–11.5	32.5–72.5	3–25.8

Table 3. Robustness values for each parameter set in the site specific ROBUSTNESS database in table 2 for plans A(4f) and B(4f). Bold values are those that are out of tolerance, and the italic values are those on the upper tolerance limit.

VOI	Mean range (%)	Mean setup (%)	$V_{eb}3\%$ range (%)	$V_{eb}5\%$ range (%)	$V_{eb}5\%$ setup (%)	$V_{eb}10\%$ range (%)
A(4f) Brainstem	***3.9***	9	43.5	15	84	38
B(4f) Brainstem	1.1	8.7	1.1	1.1	*85*	33.5
A(4f) Chiasm	1.4	6	2	***3.9***	1.8	0
B(4f) Chiasm	***2.4***	8.7	**23**	1.4	4	0
A(4f) CTV	2.1	5.2	16.5–28	3	36.5	7
B(4f) CTV	***2.8***	5.3	26	8.5	38	10

3.2. How to use the robustness database in practice: an example

To illustrate how the robustness database could be clinically used in practice, we will discuss here a clinical case. For patient X a nominal plan is generated and the robustness to set-up and range errors is calculated and compared to the robustness DATABASE table. If the robustness of plan X is within the defined values, the nominal plan is delivered. In contrast, if plan X has a robustness outside the DATABASE range for a given set of parameters for one or more volumes, the nominal plan has to be improved in terms of its robustness before delivering it to the patient. Or in the case where robustness is better than expected from the DATABASE it must be assured that plan conformality is not being compromised.

Of the plans retrospectively analysed, one plan showed to have a worse robustness in the brainstem compared to the sample. In particular, this plan resulted to have a potential mean range, $V_{eb}3\%$ and $V_{eb}5\%$ error in the brainstem of 3.9%, 43.5% and 15%, respectively, compared to 1–2%, 2.5–17% and <5% in the DATABASE in table 2. The possibility of improving the plan robustness, without using robust optimization, but by changing the start conditions of the plan has been investigated. While ensuring plan quality is maintained, four dosimetrically equivalent plans were created by changing the field setup used. For each new plan, the robustness and conformality of the CTV and each OAR considered during the optimization process, were analysed using the error bar dose histogram and by retrieving the robustness values for both the range and setup error from these (see tables 3 and 4).

The error bar dose histograms in the CTV for each plan created for this patient are shown in figure 3. The numbers in the plot key refer to the size of the area under each curve (the smaller the area, the greater the robustness). The B(4f) plan in figure 3(*a*) shows worse robustness in the CTV than the other plans. A comparison between the B(4f) and A(4f) nominal and ebDDs are shown in figure 3(*c*).

The B(4f) plan produces an underdose error to a larger volume of the target than the A(4f) plan. Table 3, shows that despite B(4f) having the worst plan robustness it only fails one robustness constraint, the mean range error which is 2.8% for the B(4f) plan compared to 2.1% for the A(4f) plan. The constraint set in table 2 was 1.5–2.5%. Figure 3(*b*) shows the volume histograms from the set-up error, A(4f) is the most robust plan, though plan B(4f) is within the robustness constraints in table 3. The C(4f) plan shows to be the most robust plan to range errors, though it is the least robust plan to set-up errors.

Figure 4(*a*) shows the volume histograms of error bars from range errors in the brainstem from all plans. The A(4f) plan is least robust to range errors, as discussed in the previous section of these results. Figure 4(*d*) shows the difference in ebDD between the B(4f) and A(4f) plans indicating the greater error in the brainstem on the A(4f) plan. Figure 4(*c*) is included to show how the plans react differently to random set-up errors, especially the C(4f) plan which shows a

Table 4. Robustness values for each parameter used the site specific ROBUSTNESS database in table 2 for the other OARs analysed for plans A(4f), B(4f), A(6f) and B(6f). Tolerances have not been set for these OARs currently and are simply used to further compare plans.

VOI beams	Mean range	Mean setup	$V_{eb}3\%$ range (%)	$V_{eb}5\%$ range (%)	$V_{eb}5\%$ setup (%)	$V_{eb}10\%$ setup (%)
Cochlea						
B(4f)	4	6.8	64.2	32.1	68.2	5.5
A(4f)	3.8	4	84.4	24.8	20.2	1.8
B(6f)	3.7	5.8	57.8	22.9	75.2	1.8
A(6f)	3.9	3.4	75.2	24.8	10.1	1.8
Left optic Nerve						
B(4f)	3.8	5.8	36	19	47	14.2
A(4f)	3.8	5.8	36	19	47	14.2
B(6f)	4.4	4.9	23.7	8.3	38.3	14.2
A(6f)	2.1	5.8	9.5	1.2	47.5	12.5
Left temporal lobe						
B(4f)	4.2	43	15.3	10.6	19.1	4.8
A(4f)	3	3.8	11.4	6.8	12.9	2
B(6f)	3.2	3.8	11.7	7.2	15.3	2.6
A(6f)	3.1	4	12.5	7.8	16.5	2.1

decrease in robustness when compared the other four beam arrangements for this type of error. As seen in the A(4f) plan there is a hot spot in the centre of the brainstem, whereas in the B(4f) the hot spot is shifted more anteriorly due to the direction of the lateral and posterior beams. Despite this improvement in robustness in the brainstem to range errors, the B(4f) plan is less robust in the chiasm (figure 4(*c*)) compared to the A(4f), whereas the B(6f) yields the greatest robustness to range errors in the chiasm of all the plans. In the case of the optic nerves again the B(4f) plan is less robust compared to the A(4f), but this time the A(6f) plan is most robust. These results can also be seen in table 3, the chiasm fails a robustness constraint with a $V_{eb}3\%$ range error of 23%, though this error in dose has fallen to 1.4% at $V_{eb}5\%$, matching that achieved by the A(4f) plan.

Table 4 has been included to show the results for four of the plans for other OARS, though constraints for these volumes have not been established currently.

4. Discussion

3D IMPT offers the ability to deliver highly conformal radiotherapy to the patient, reducing integral normal tissue dose (Chang *et al* 2006) and providing greater OAR sparing through the use of multiple inhomogeneous fields. However, without certainty in our ability to deliver these highly complex and conformal plans at each fraction caution must be heeded when choosing beam arrangements and positioning steep dose gradients near OARs (Lomax 2008a, Albertini *et al* 2010, McGowan *et al* 2013). As mentioned, any uncertainty might lead to severe target underdosage or OAR overdosage. Several authors are proposing to include the plan robustness as an extra parameter directly during the optimization process (Unkelbach *et al* 2007, Pflugfelder 2008, Fredriksson *et al* 2011, Chen *et al* 2012, Liu *et al* 2012). Unfortunately when introducing the robustness parameter into the optimization algorithm, the nominal plan conformality is often compromised (Unkelbach *et al* 2007, Fredriksson *et al* 2011). We believe
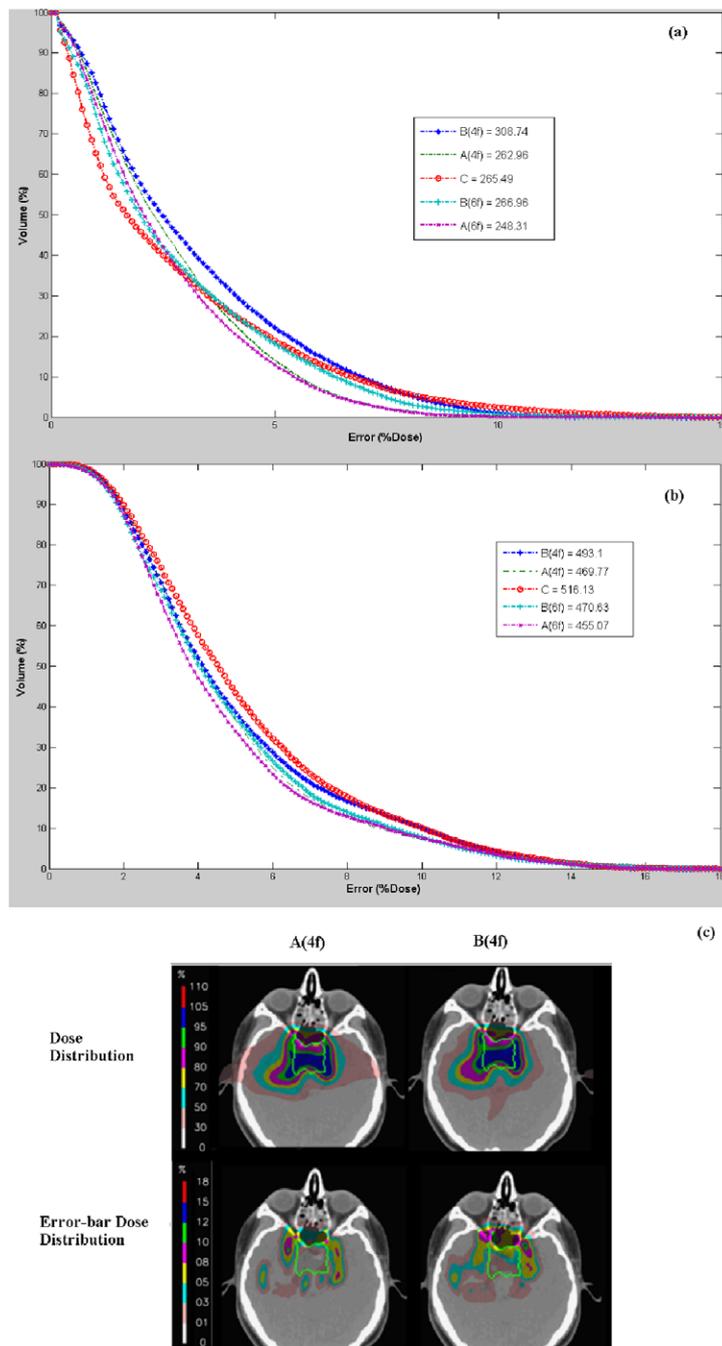
Figure 3. Volume histogram of (*a*) systematic range errors for an underdose scenario in the CTV and (*b*) random set-up errors for an underdose scenario in the CTV for five plans of different beam orientations, for five plans of different beam orientations. (*c*) Shows the dose distribution (top) in the CTV (green outline) for the A(4f) (top left) and B(4f) (top right) plans and ebDD of systematic range errors for the worst-case underdose scenario in the CTV (bottom) for the A(4f) (bottom left) and B(4f)(bottom right) plans.
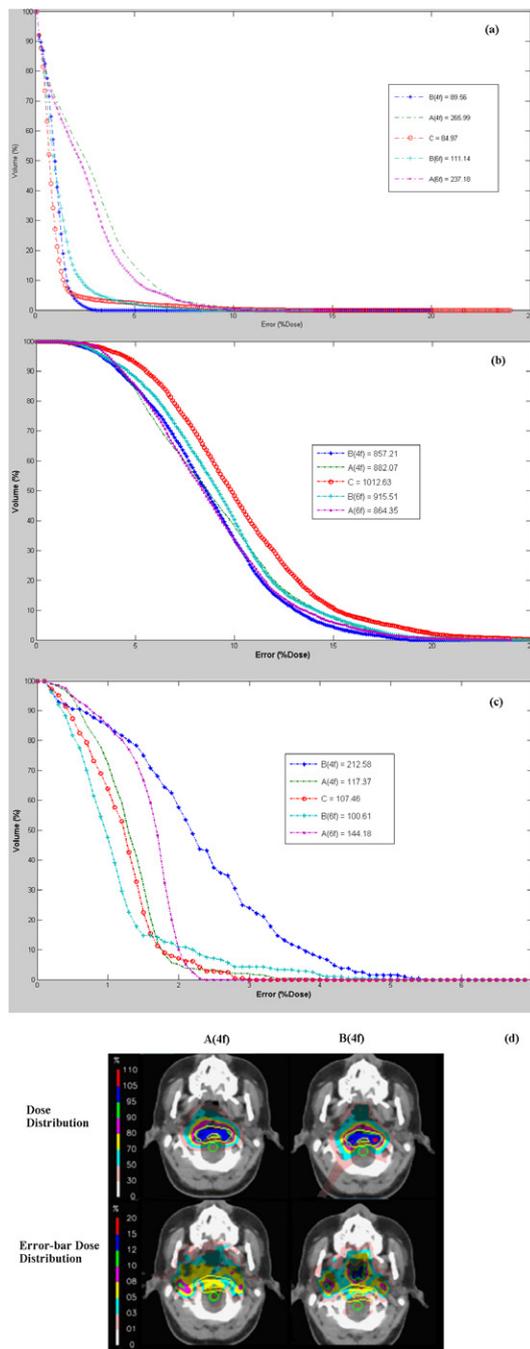
Figure 4. Volume histogram of (*a*) the systematic range errors, (*b*) the random set-up errors for an overdose scenario in the brainstem and (*c*) the random set-up errors for overdose scenario in the chiasm, for five plans of different beam orientations. (*d*) shows the dose distributions from the A(4f) (top left) and B(4f) (top right) plans and the corresponding ebDD from B(4f) (bottom left) and A(4f) (bottom right) treatment plans showing the brainstem in green and the target in yellow.

that the use of a robust plan is probably the correct way to deal with the problem of uncertainties for the IMPT plan. However, we also think that it is extremely important to define a threshold between robustness and plan conformality. In particular it is important to establish an adequate tolerance level of robustness for each volume. This can be used both as an input parameter during the robust optimization process, and also as control parameter during the plan evaluation phase. In this way the DATABASE is a compatible method for analysing plan robustness alongside robust optimization, as a method of providing feedback and of providing a threshold between plan robustness and plan conformality.

Additionally, in this paper it has been demonstrated how retrospective analysis of the robustness of clinical plans at PSI has furthered an understanding of how random set-up and systematic range errors effect the final dose distribution. It was seen that random set-up errors resulted in a greater difference in the resulting dose distribution than that from systematic range errors. However, this is valid when only one fraction is considered. When considering the full treatment, the random nature of set-up errors act to 'blur' the final dose distribution, therefore, the risk of reduced plan quality associated with these errors is less. In contrast range errors pose a greater concern due to their systematic origin, meaning they are present over the entire treatment leading to a cumulative 'shift' of the dose distribution, thereby reducing plan quality.

Through retrospectively analysing the robustness of each VOI a site-specific Robustness DATABASE was established (table 2). For the time being there are no established action levels, or methods for establishing action levels, for evaluating plan robustness. With the awareness that the clinical results achieved for patients treated intra-cranially for chordoma and chondrosarcomas at our institute were good (5 years of local control of 81% for chordomas and of 94% for chondrosarcomas patients, with only limited toxicities (Ares *et al* 2009)), we assumed that most of the plans clinically delivered in our centre were acceptable also in terms of robustness. Consequently, the proposed DATABASE can be used in the future to easily identify outliers. For these patients greater plan individualization may be required to improve the robustness to uncertainties. Although the DATABASE is specific to PSI, the idea behind can be used as an example for other centres to define their own robustness tolerance levels, based on experience and treatment technique. Hopefully, as more centres begin to introduce robustness evaluation into the clinical workflow we will in the future be able to define an adequate level of robustness for each volume of interest. In this way, robustness constraints that are accepted worldwide may be established in the same way as dose volume constraints to OAR and prescribed dose to target volumes are generally widely accepted and applied in the different protocols.

5. Conclusions

The error-bar-dose-distribution (ebDD) has been used to effectively analysis plan robustness to both systematic range and random set-up errors for 16 skull base IMPT plans previously treated at PSI. A site specific Robustness DATABASE was created as a simple solution to include robustness in plan analysis, as well as to aid planners produce plans to meet both dosimetric and robustness criteria by establishing site specific robustness thresholds. Using these methods further work can be carried out to fully explore how start conditions, and the optimization itself, affect both the plan robustness and conformality. This will further our ability to determine where a trade-off can be made between these two parameters to ensure the patient receives as optimal treatment as possible.

Acknowledgments

References

Albertini F, Hug E B and Lomax A J 2010 The influence of the optimization starting conditions on the robustness of intensity-modulated proton therapy plans *Phys. Med. Biol.* **55** 2863–78

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

Allen P *et al* 2012 An evidence based review of proton beam therapy: the report of ASTRO's emerging technology committee *Radiother. Oncol.* **103** 8–11

Ares C *et al* 2009 Effectiveness and safety of spot scanning proton radiation therapy for chordomas and chondrosarcomas of the skull base: first long-term report *Int. J. Radiat. Oncol. Biol. Phys.* **75** 1111–8

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

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

Chang J, Zhang X, Wang X, Kang Y, Riley B, Bilton S, Mohan R, Komaki R and Cox J 2006 Significant reduction of normal tissue dose by proton radiotherapy compared with 3D conformal or intensity-modulated radiation therapy in Stage I or Stage III non–small-cell lung cancer *Int. J. Radiat. Oncol. Biol. Phys.* **65** 1087–96

Chen W, Unkelbach J, Tromifov A, Madden T, Kooy H, Thomas B and Craft D 2012 Including robustness in multi-criteria optimization for intensity-modulated proton therapy *Phys. Med. Biol.* **57** 591

Cozzi L, Fogliata A, Lomax A and Bolsi A 2001 A treatment planning comparison of 3D conformal therapy, intensity modulated photon therapy and proton therapy for treatment of advanced head and neck tumours *Radiother. Oncol.* **61** 287–97

De Ruysscher D *et al* 2012 Charged particles in radiotherapy: a 5 year update of a systematic review *Radiother Oncol.* **103** 5–7

Durante M and Loeffler J S 2010 Charged particles in radiation oncology *Nature* **7** 37–43

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

Goitein M 1983 Nonstandard deviations *Med. Phys.* **10** 709–11

Jones B 2008 The potential clinical advantages of charged particle radiotherapy using protons or light ions *Clin. Oncol.* **20** 555–63

Lassen-Ramshad Y, Vestergaard A, Muren L P, Høyer M and Petersen J B B 2011 Plan robustness in proton beam therapy of a childhood brain tumour *Acta Oncol.* **50** 791–6

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 1999 Intensity modulation methods for proton radiotherapy *Phys. Med. Biol.* **44** 185–205

Lomax A J 2008a Intensity modulated proton therapy and its sensitivity to treatment uncertainties: II. The potential effects of inter-fraction and inter-field motions *Phys. Med. Biol.* **53** 1043–56

Lomax A J 2008b Intensity modulated proton therapy and its sensitivity to treatment uncertainties: I. The potential effects of calculation uncertainties *Phys. Med. Biol.* **53** 1027–42

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

Lomax A J *et al* 2001 Intensity modulated proton therapy: a clinical example *Med. Phys.* **28** 317–24

McGowan S E, Burnet N G and Lomax A J 2013 Treatment planning optimisation in proton therapy *Br. J. Radiol.* **86** 20120288

Pedroni E *et al* 2005 Experimental characterization and physical modelling of the dose distribution of scanned proton pencil beams *Phys. Med. Biol.* **50** 541

Pflugfelder D 2008 Worst case optimization: a method to account for uncertainties in the optimization of intensity *Cancer Res.* **53** 1689

Schaffner B, Pedroni E and Lomax A 1999 Dose calculation models for proton treatment planning using a dynamic beam delivery system: an attempt to include density heterogeneity effects in the analytical dose calculation *Phys. Med. Biol.* **44** 27–41

Scheib S and Pedroni E 1992 Dose calculation and optimization for 3D conformal voxel scanning *Radiat. Environ. Biophys.* **31** 251–6

Unkelbach J, Chan T C Y and Bortfeld T 2007 Accounting for range uncertainties in the optimization of intensity modulated proton therapy *Phys. Med. Biol.* **52** 2755–73

Webb S 2003 The physical basis of IMRT and inverse planning *Br. J. Radiol.* **76** 678–89