DISS. ETH $N_O$:19525

# Planning and Optimizing Treatment Plans for Actively Scanned Proton Therapy: evaluating and estimating the effect of uncertainties

A dissertation submitted to
ETH ZURICH

for the degree of
Doctor of Sciences

presented by
FRANCESCA ALBERTINI
Dip. Phys., Università degli Studi di Milano, Italy

born on February 28, 1978
citizen of Italy

accepted on the recommendation of

Prof Dr. G. Dissertori, examiner
Prof Dr A.J. Lomax, co-examiner
PD Dr U. Schneider, co-examiner

2011

A Claudio, per esserci. SEMPRE.

A Gaia, la mia stellina
e al piccolo Mattia

# Acknowledgment

I wish here to thank a lot of people that have been fundamental for this thesis.
First of all I have to thank my direct supervisor, **Prof Dr. Antony Lomax**. I have learned much more from him than from anybody else. He has always supported me, both academically and personally. It is a real pleasure working with him, as he is always motivating you by enhancing the positive aspects of results that, at first glance, might not look that nice! Moreover, working with him is a great honor, as he is one of the worldwide experts for the treatment planning of proton therapy.
I would also like to thank **Prof Dr. G Dissertori** who supervised this work from the ETH side and **PD Dr. Uwe Schneider** for accepting to be my co-supervisor.
A special thanks goes to **Dr Gudrun Goitein**. After completing my specialization in medical physics, she convinced me to start a PhD, rather than directly accepting a job as a full-time clinical medical physicist. Following this recommendation was one of the best choices of my life. I have really enjoyed this last four years and I have had the chance to deepen different aspects of my medical physics knowledge, especially on the treatment planning side. A personal thanks here also goes to **Prof Dr. Eugen Hug**, the director of the Center for Proton Therapy, for his scientific support and careful reviewing of all my papers

As during my PhD work, I have been also been working part-time as a medical physicist, I want to thank different colleagues who have supported me and shared with me the clinical duties. In particular, I would like to thanks **Juergen Salk** (for IT support and useful discussions), **Matthias Hartman**, **Martijin Engelsman**, **Kathy Haller** and **Carlo Algranati** (in particular for allowing me to make some phantom measurements just before the weekly patient QA!). An extra special thanks goes to **Alessandra Bolsi**. She is a colleague, and most importantly a friend! As a colleague, I have always appreciated her help when discussing different subjects, and especially in sharing of our clinical duties. As a friend, she has been always present, and we have shared all the joys and the troubles of the last 6 years! A personal thanks also goes to **Jorn Vervey**. I always enjoyed discussing with him and appreciated his direct and honest opinions.
Big thanks to all the operators: **Hansueli Stauble**, **Benno Rohrer**, **Daniel Lempen** and **Martin Heller**. They have all helped me a lot during this work, and we have shared many lunches too. In particular, they have all supported me in performing the phantom measurements and have advised me both on the mechanical aspects of the gear construction and on how to electronically steer it.
As a medical physicist you can never forget about our team of MTRA's. Here at PSI we have one of the best groups of MTRAs that I have ever met. They are all motivated and always helpful (and I clearly had to ask many times for their help!). I would also like to acknowledge in

particular **Lydia Leder**, **Alexander Lehde**, **Sandra Hergsperger**, **Jeanette Fricker** and **Petra Thoma.**
A big thank you also to **Stefano Lorentini** and **Margherita Casiraghi** for helping me in the acquisition and analysis of the GafCromic films of the Charlie phantom. I am grateful also to **Dr Med Carmen Ares** and **Dr Med Barbara Rombi** for helping me in defining the anatomical area and forms of a typical titanium implant and for drawing the volumes in the phantom CT. A special thanks goes also to **Dirk Boye** (for IT support and help in software development) and **Maria Luisa Belli** (for the analysis of the stopping power values of Charlie).
I have to admit that sharing an office with **Dr Antje Knopf** has been more than a pleasure! I like to see her smile in the morning and chat about the previous evening. Sunday brunches at her place have been always delicious. Moreover, she was always happy to listen, discuss and advise me on many issues. A particular thank goes to **Lamberto Widesott**, **Dr Sairos Safai** and **Dr Silvan Zenkluesen** for our common lunches together, their friendship, but most of all for their willingness to discuss with me many details of my research. A big thanks also goes to **Dr Andrea Carminati**. He is a good friend since the time of University, and helped enormously in putting together the theoretical background of the optimization algorithm. A final special thank goes also to **Alexander Tourovsky:** he is an infinite source of knowledge on all aspects of physics - I have used and 'abused' his patience more than once.

Finally, I wish here to thank my two families.
Firstly, my parents and my sister for their constant support in the last few years.
Finally **Claudio**, my love, and **Gaia** and **Mattia**, my little-big jewels to whom this thesis is dedicated. For their endless love and their immense patience during the writing of this thesis.

# Contents

# Abstract

Great advances have been made in the delivery of external beam radiotherapy with photons, through the development of intensity modulated radiotherapy (IMRT). Similar developments are being pursued in proton therapy, work that is being pioneered at the Paul Scherrer Institute (PSI) through the development of active scanning and intensity modulated proton therapy (IMPT). With IMPT, highly conformal dose distributions can be delivered through the application of multiple, angularly spaced fields, each applying an optimized pattern of spatially varying particle fluences. In other words, Bragg peaks are spatially distributed in three dimensions and are simultaneously optimized for all field directions. Due to the large number of optimization parameters and the relatively basic goals of radiotherapy planning, it has been found for IMRT, and it is even more valid for protons, that the problem of optimization is highly degenerate: that is, there are many fluence profiles that meet the main planning aims. Thus, the result of the optimization will generally depend on the starting conditions. In this work it has been studied how the manipulation of the starting conditions influences the optimized IMPT plan. In particular, how results can be 'steered' by the user such as to provide plans which are, for example, more robust to potential delivery errors, and consequently safer to be delivered.

Furthermore, to properly perform the analysis of plan robustness a new method to evaluate and display plan robustness have been introduced, namely the max-to-min error bar distribution. With this model all the distributions of uncertainties are collapsed into a single distribution, for more simple analysis. Although this concept has been applied to proton plans, it can be easily extended to all kind of radiation therapy plans (e.g photons, protons, ions).

Finally, clinically similar plans but with very different Bragg peak distributions have been dosimetrically verified to assess the accuracy of their delivery both in nominal and under different error conditions. An anthropomorphic phantom has been customized to acquire these measurements as closely as possible to the clinical situation. Furthermore, these measurements have been used to validate the robustness analysis we have developed.

# Riassunto

La radioterapia coi fotoni ha compiuto negli ultimi anni degli enormi progressi con l'avvento della radioterpia ad intensità modulata (IMRT). Un simile progresso è altresi' avvenuto nel campo della protonterapia, sviluppo che ha avuto come pioniere il Paul Scherrer Institute (PSI) attraverso l'implementazione dello scanning attivo dei protoni (*active scanning*) e della protonterapia ad intensità modulata (IMPT). I piani IMPT consentono di rilasciare al tumore una distribuzione di dose altamente conformazionale. Questo si ottiene mediante la sovrapposizione di piu' fasci di protoni, ciascuno dei quali emette delle particelle distribuite nello spazio con un profilo di fluenza ottimizzato e piuttosto complesso. In altre parole, i picchi di Bragg sono spazialmente distribuiti in 3D e sono ottimizzati contemporaneamente per tutte le direzioni dei campi. Grazie sia al numero di gradi di libertà a disposizione dell'algoritmo di ottimizzazione sia ai realtivamente semplici obiettivi definiti per la pianificazione del trattamento di radioterapia, il problema da risolvere durante il processo di ottimizzazione è altamente degenerativo. Questo significa che esistono più soluzioni di profili di fluenza che soddisfano gli stessi obiettivi di pianificazione (risultato in precedenza dimostrato per l'IMRT ed è ancora più valido per i protoni).Quindi, il risultato dell'ottimizzazione dipende, tra gli altri fattori, anche dalle condizioni iniziali.

In questo lavoro abbiamo studiato come la scelte di differenti condizioni iniziali possa influenzare l'ottimizzazione dei piani IMPT. In particolare abbiamo verificato come le soluzioni possano essere 'guidati' dall'utente in modo da ottenere piani con caratteristiche diverse, come, ad esempio, il fatto di essere più robusti (cioè, meno dipendenti) per eventuali incertezze (i.e. errori nel set-up del paziente, nella calibrazione della CT, etc..) che possono capitare durante la terapia, e di conseguenza essere più sicuri. Inoltre, abbiamo introdotto un nuovo metodo per poter valutare la robustezza dei piani in funzione degli errori e per visualizzare le aree piu' soggette all'effetto delle incertezze. Con questo metodo si riducono tutte le distribuzioni di dose derivanti dalla varie incertezze in un'unica distribuzione, che pertanto risulta più facile da analizzare. Questo concetto è stato sviluppato ed applicato ai piani di protoni, ma può essere facilmente esteso a tutti i tipi di piani di radioterapia (ad esempio fotoni e ioni).

Infine, piani clinicamente molto simili tra loro (cioè con la stessa distribuzione di dose nel volume bersaglio e che rispettano gli stessi limiti di dose agli organi a rischio) ma con diverse distribuzioni di picchi di Bragg sono stati valutati dosimetricamente per stimare l'accuratezza della loro distribuzione di dose, sia in condizioni nominali che in presenza di errori. Un fantoccio antropomorfo è stato personalizzato per acquisire tali misurazioni in condizioni il piu' possibile simile ad ad una situazione clinica. Queste misure sono altresi state utilizzate per la validazione del metodo di analisi della robustezza qui introdotto.

# 1 Introduction

## 1.1 Proton therapy

There are many sophisticated and advanced methods currently being proposed and introduced into radiotherapy. One of the most promising is proton therapy. When quasi-mono-energetic proton beams are applied to a medium, they display the rather attractive property of having a well defined range, beyond which no energy is deposited, and a concentrated and focused high dose volume (dose = deposited energy) in a narrow region known as the 'Bragg peak'. This property was recognised by the physicist Robert Wilson in 1946 (Wilson 1946) as being almost ideal for radiotherapy, in that, by exploiting the three-dimensional localisation of dose in the Bragg peak, increased dose can potentially be applied to the tumour whilst simultaneously sparing surrounding normal tissues. It took approximately 8 years however before the potential of proton therapy could be demonstrated on patients, with the first being treated at the Lawrence Berkeley Laboratory in 1954.

One of the challenges facing the early pioneers of proton therapy was how to modulate the typical proton beams that are produced by particle accelerators such as to produce a clinically useful field size that can treat arbitrarily sized tumours. For instance, although the Bragg peak has many desirable characteristics, it is typically only a few mm's wide in depth. In addition, proton beams produced by modern accelerators have a rather narrow cross section, again typically of only a few millimetres. Thus, in order to irradiate anything but the smallest tumours, such proton beams must be broadened (or 'spread') both along, and orthogonal to, the beam direction.

For the first proton treatments this was done using what has subsequently been called the passive scattering technique (Koehler et al 1975, Koehler et al 1977). In depth, the field size (i.e. the volume over which the dose can be considered homogenous) can be extended by the simple technique of delivering a number of range adjusted and appropriately weighted individual Bragg

peaks (achieved through a rapidly rotating, varying thickness wheel). When the weights (or more correctly stated, the relative fluences) of the individual Bragg peaks are correctly calculated, then a uniform dose profile along the beam direction can be produced, known as the 'Spread-Out-Bragg-Peak' (SOBP).

Orthogonal to the beam direction, the narrow pencil beam emitted from the accelerator nozzle is broadened through the insertion of 'scattering-foils' into the beam. As protons are scattered predominantly by electro-magnetic interactions with the atomic nuclei as they pass through any medium, and the amount of scatter is dependent on the atomic number of the material, a narrow proton beam can be effectively spread by the insertion of high-Z foils directly in the beam. Unfortunately, the resulting profile is Gaussian in form rather than uniform. Nevertheless, by working in the highest portion of this Gaussian profile, a lateral uniformity of ± 2-5%, depending on the required field size, can be achieved. Finally, the resulting delivered fields are shaped to better conform the delivered dose to the tumour through the use of patient and field specific collimators (to conform the cross-sectional shape of the delivered field to the projected cross-sectional shape of the target volume) and compensators (to match the distal end of the SOBP to the distal end of the tumour).

Although a relatively simple approach, passive scattering has been a remarkably successful delivery technique for proton therapy, with more than 66,000 patients having been treated worldwide with this technology. Indeed, even today, most new proton facilities being planned or built will still be based on the passive scattering approach. Nevertheless, passive scattering must slowly be considered to be somewhat limited in its capabilities, and new delivery methods for proton therapy must also be developed.

## 1.2 Proton therapy using active scanning

One such approach is active scanning. This was first suggested by Kanai et al (1980), and has subsequently been implemented in our clinical proton therapy facility at the Paul Scherrer Institute (PSI) in Switzerland, where the world's first spot scanning gantry was developed in the early 1990's (Pedroni et al 1995), and where the first patient was treated in 1996.

The principle of active scanning is deceptively simple, and is based on the fact that protons, being charged particles, are subject to Lorentz forces. That is, when subjected to an electric field, they are accelerated and when subjected to a magnetic field, they are deflected. It is this second characteristic that is exploited in active scanning to spread the beam laterally. Thus, in this approach, the narrow proton 'pencil beam' is no longer broadened through scattering, but is 'scanned' across the tumour using a scanning magnet. In depth, the Bragg peaks are stacked by

altering the initial proton energy at the accelerator source. Through this combination of scanning and energy variation, the Bragg peak can be effectively placed anywhere in three dimensions within the tumour. Dose uniformity is then achieved through a mathematical optimisation of the individual fluences of each pencil beam (see Chapter 2).

Active scanning has a number of advantages over the more traditional passive scattering approach. First, it can be fully automated and computer controlled, meaning that, as only Bragg peaks that stop in the tumour volume are delivered, it is not necessary to use collimators and compensators to achieve dose conformation. Second, it is more efficient than passive scattering in that less protons need to be delivered in order to achieve a given total dose to the tumour and therefore less dose is delivered to the normal tissues. Last, active scanning is an inherently more flexible technique than passive scattering.

Although active scanning has many advantages, it is a rather new delivery technique, and currently there are only three facilities world wide which are using active scanned protons in clinical practice, with an additional centre using a similar approach with carbon-ions. Of the proton centres, the Paul Scherrer Institute has so far the largest experience, with more than 600 patients having been treated, all using active scanning. Initial clinical results are promising, and based on this success, most new proton therapy centres being currently built or planned will have active scanning at least as an option, if not the only delivery approach.

## 1.3 Treatment planning

Given the sophistication of the delivery methods available to the contemporary radiotherapist, and the precision with which these methods can conform the high dose to the tumour volume, it has become increasingly important that accurate predictions of the form of the dose delivered to the patient are available. Thus, treatment planning - the process of designing and calculating the likely three-dimensional dose distribution in the patient – remains one of the critical parts of the radiotherapy chain.

The basis of all treatment planning systems, be they for photon, electron or proton therapies, is the dose calculation. For example, in proton therapy, such models are usually based on the Bethe-Block equation for the calculation of energy loss as a function of depth, on collision cross-sections for the calculation of fluence loss and nuclear interactions and Fermi-Eyges theory of scattering to predict the angular and spatial distribution of the beam as a result of multiple Coulomb scattering(see.e.g Scheib and Pedroni et al 1992, Pedroni et a 2005 ). Such analytical dose calculations can reach high accuracy (agreement between measurement and calculation to a

few percent is not uncommon) in homogenous mediums such as water, in which most routine dosimetry and quality assurance checks are made.

However, no matter how good the dose calculation is, it can only be as accurate as the model, or description, of the patient in which it is calculated. Although many imaging modalities can be used for helping the medical staff delineate the tumour (e.g. Magnetic resonance imaging or Positron Emission tomography), X-ray computer tomographic (CT) data is the only imaging modality that provides the data necessary for computing accurate *in-vivo* dose distributions. CT data provides a three-dimensional model of the relative attenuation of the patient's anatomy to low energy X-rays (typically of about 120KeV), which can be translated into either electron densities for photon dose calculations or proton stopping powers for proton therapy (Schneider et al 1996). Given the extreme difficulty of performing direct dose measurements *in-vivo*, treatment planning, based on high resolution CT data of the patient and accurate dose calculation models, remains the only method to estimate with any accuracy the distribution of dose delivered to the patient.

Thus, the process of treatment planning is fundamental to the success of radiotherapy. Despite the delivery approach chosen to treat the patient, the process of treatment planning is the same and can be summarized as follows (Goitein 2008):

1. Evaluate the patient using all relevant diagnostic tools, and decide whether to employ proton radiation therapy as part of the patient's treatment.
2. Obtain and inter-register appropriate imaging studies. This includes always the acquisition of a planning computer tomography (CT) study that is taken with the patient lying in the position and in the immobilization device that will be used for treatment, together with complementary imaging studies such as MRI, PET etc.
3. Delineate on the planning CT the target volumes and all organs-at-risk (OARs) whose proximity to the target volume makes them of particular interest.
4. Establish the planning aims for the treatment, which includes the intended dose to be delivered to the target volume (e.g. 74 $Gy_{RBE,}$ with a proton radio biological effectiveness of 1.1), the limit doses to the OARs and the *fractionation scheme* (e.g. $2Gy_{RBE}$ per day, 5 days a week).
5. Design and calculate the dose distribution for one or more plans -i.e., sets of beams each of which fulfill as much as possible the requirements of the planning aims.
6. Evaluate these plan(s) and select one of them for use in treatment. In case the treatment requirements cannot be met, then revise the planning aims and return to step 5.
7. Finalize the prescription, which includes all the information specifying how the patient is to be treated and how the treatment is to be delivered

8. Deliver the treatment, and verify daily that the delivery is correct, i.e. by allowing the treatment only if the machine settings match the prescribed parameters within a defined tolerance.
9. Re-evaluate the patient during the course of treatment to ensure that the plan remains appropriate (e.g., check that weight loss or tumor regression have not affected the treatment geometry) and, if that is not the case, return to step 5, or even 2, to re-plan the remainder of the treatment
10. Document and archive the final treatment plan.

Despite the critical importance of the treatment planning process to radiotherapy, it is important to bear in mind that the resultant dose distribution ('plan') is an image on a screen, and as such can never be any more than an estimate of the actually delivered dose to the patient. Thus, during the plan analysis step, in addition to the careful evaluation of the 3-dimensional dose distribution, it is extremely important to identify all the (major) source of errors and evaluate their impact on the treatment dose distribution. Clearly, for the same dose distribution delivered to the target volume and critical structures (i.e. *plan degeneracy*, see following section), the treatment plan that will be less affected by the delivery errors is the more desirable as it can be more safely delivered. That is, the plan is *more robust* to delivery uncertainties. This concept is further discussed in section 1.6.

## 1.4 SFUD and IMPT

As implemented at PSI, active scanning with protons can be applied in two 'flavours' (Lomax 2008c). The first essentially 'mimics' passive scattering delivery, such that each individually delivered field applies a more or less homogenous dose across the selected target volume, much as the scatterer and SOBP delivery of passive scattering applies a uniform dose across the target. However, the active scanning approach adds some flexibility, and can be considered more efficient in that only Bragg peaks delivered within (or close to) the target are delivered. One can think of this as the delivery of a number SOBP shaped beamlets, but where the extent of each SOBP is matched to the thickness of the target along the beam direction. This is, however, somewhat of an over simplification. In practice, it is necessary to individually optimise the weight (fluence) of each delivered beamlet in order to truly deliver a homogenous dose to arbitrarily complex shaped target volumes. A full treatment plan then consists of the simple linear addition of such individually optimised fields. For this reason, we call this method Single Field, Uniform Dose (SFUD) delivery, which is the direct active scanning equivalent of open field photon plans or passive scattered proton treatments. Indeed, the SFUD approach was the method of choice for the first patients treated at PSI, and is still used in about 60% of the currently applied plans at

PSI.

The flexibility of the spot scanning approach has been recently extended to deliver so called 'Intensity Modulated Proton Therapy' (IMPT) (Lomax 1999, Lomax et al 2001). IMPT is the direct equivalent of IMRT (Intensity Modulated Radiotherapy) with photons, in that each individual field delivers an optimized, and potentially highly irregular, fluence pattern, and the desired dose distribution in the patient is then only achieved when all such fields are combined. The major difference in this approach is that the 'uniform, single field' constraint is removed, and the optimisation is given full reign to weight Bragg peaks regardless of the final form of the individual field dose distributions, as long as the total dose (the addition of all the individual field dose distributions) combine to give the desired result. Although the difference to SFUD planning is rather subtle, the IMPT approach best exploits the full potential of scanned proton therapy, in that it provides even more flexibility in tailoring the dose distribution to the target - and in selectively avoiding critical structures -than the SFUD.

Consider for instance Figure 1.1. A 3 field SFUD plan applied to the defined target volume would look like the dose distribution in Figure 1.1a. It would achieve a good dose homogeneity over the target volume (green line in the figure), but inevitably, the organ at risk (red contour) close to the target would receive some dose, mainly due to the plateau dose being delivered from the two cranial beams (F1,F2). By relaxing the constraint that each field must uniformly irradiate the whole target volume however, we could irradiate the target in the way indicated in Figure 1.1b. Here, the beamlets passing through the organ at risk (OAR) from the cranial beams are switched off, thus sparing this organ. By properly modulating the beamlet fluences also for the lateral field (F0) then a homogenous dose can still be delivered to the target while sparing the OAR. For scanning, such a field and dose arrangement would be achieved through IMPT, or the simultaneous optimisation of all beamlets from the three field directions, in that, the missing dose from one of the field (i.e. the cranial field for which the beamlets have been reduced in weight as they pass through the OAR) are compensated for by contributions from the other fields. For the optimisation to achieve this, then it is obvious that it must know the positions and spatial whereabouts of all beamlets of all fields.

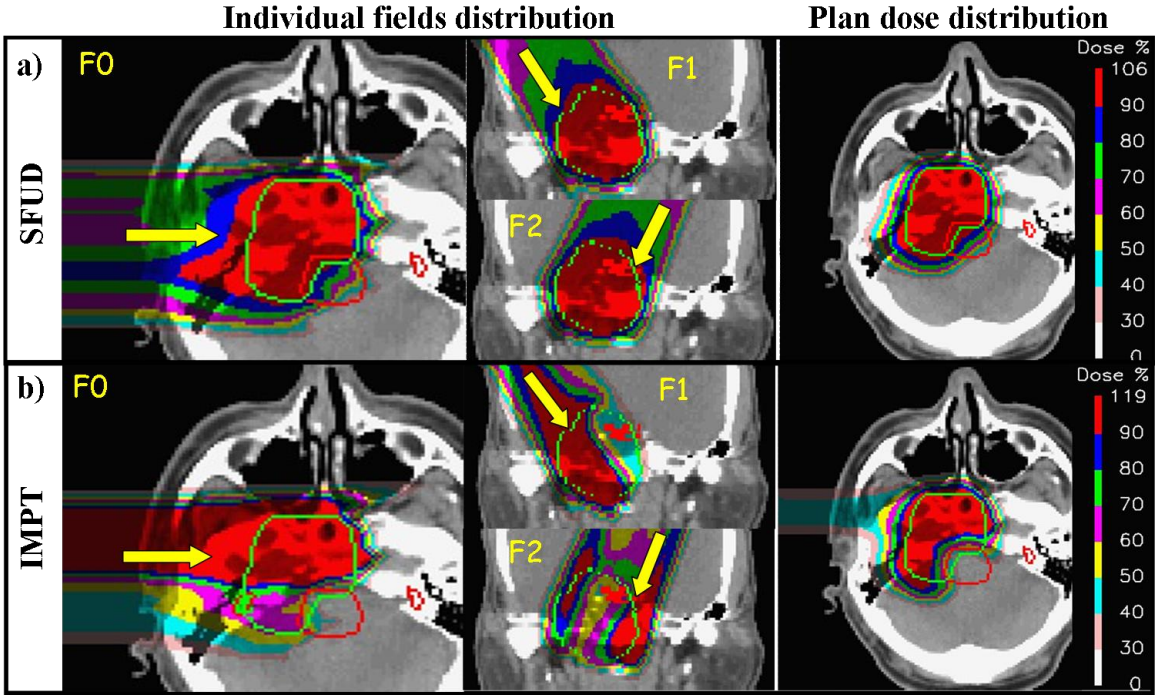


**Figure 1.1.** (a) Example of a SFUD (Single Field, Uniform Dose) plan. Each of the three fields (indicated by the arrows) delivers a homogenous dose to the whole target volume.(b) Example of an IMPT plan. All fields are highly modulated to avoid the direct passage through red critical structure (i.e. the brainstem). The combination of all fields together still provides a homogeneous dose through the target volume. The advantage of IMPT lies in its ability to individually modulate fields in order to make up for under- or overdosing from the other fields due to their sparing of neighboring critical structures.

In summary, regardless of whether one uses SFUD or IMPT planning, an optimisation algorithm for the calculation of a set of fluences which provides the desired dose distribution is a pre-requisite for active scanned proton therapy. The optimisation approach used for such calculations at our institute will be outlined in Chapter 2.

## 1.5 IMPT & Degeneracy

In a typical SFUD or IMPT plan, there are many thousands of applied beamlets, each of which is individually weighted by the optimisation algorithm. As the number of beamlets that can be varied determines the degrees of freedom (DOF) available to the optimisation process, in comparison to other optimisation problems in radiotherapy, the DOF's available when optimising SFUD, and in particular IMPT, plans are enormous. The main reason for this is the fact that in scanned proton therapy, beamlets are not just distributed in a plane orthogonal to the field direction, but also *in depth* along the beam direction. It is this shifting (and weighting) of Bragg peaks which allows for construction of the SOBP in passive scattering, and this same technique can be exploited in SFUD and IMPT optimisation to individually modify the depth dose curve along the beam

direction, such as to deliver *different* effective depth dose curves through different parts of the target. However, even in the optimisation of IMRT with photons (where the modulation matrix per field is 'only' two dimensional), the optimisation problem is rather *degenerate* (Alber et al 2002, Llacer et al 2003, Webb 2003, Llacer et al 2004). That is, there are a large number of different beamlet fluence combinations for each field which result in very similar calculated dose distributions. Due to the addition of a third dimension to the problem in proton therapy, it is perhaps not surprising that *degeneracy* of the solution for SFUD and IMPT is even larger. It is our belief that degeneracy can be utilised in a number of ways to drive the plans to more desirable results based on non-dosimetric criteria, and investigating the potential of this is one of the main aims of this work.

## 1.6 The problem of uncertainties

As mentioned in section 1.3, the evaluation of a treatment plan is the primary method for a clinician to estimate the quality of a planned or delivered treatment. However, one must always be aware that a calculated dose distribution, even if calculated in three dimensions using the most sophisticated dose calculation methods, is still nothing more than an image on the screen, and at best only an approximate estimation of the doses actually delivered to the patient. This is particularly true given the nature of radiotherapy, where the full treatment is given in small doses per day (usually of about $2Gy_{RBE}$), over many weeks. This is known as *fractionation*. Thus, although a calculated, 'one-off' dose calculation may meet all clinical needs, this is not necessarily the true dose delivered to the patient, as there will inevitably be uncertainties during the treatment process. Some of these will vary day-by-day in a more or less random fashion and can then be considered to smooth-out over the full course of the fractionated delivery (e.g. variations in the daily positioning of the patient). On the other hand, there can also be systematic errors which could well be the same every day of treatment, and which are therefore potentially much more serious.

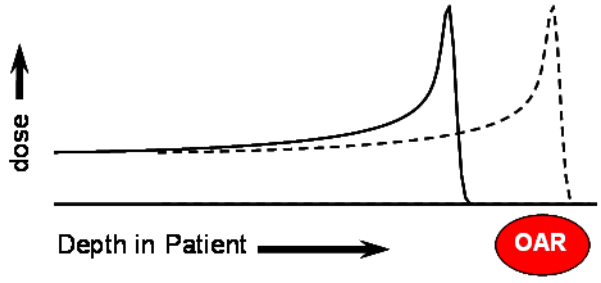


**Fig 1.2** Schematic representation of the effect of range error: the nominal Bragg-peak (solid line) is calculated with a range that spares completely the organ at risk (OAR); a shift (dashed line) could instead bring the Bragg peak in the middle of the OAR.

This is true for radiotherapy in general and possibly even more for proton therapy. The advantage of proton therapy, as mentioned, is related to the spatial localization of the dose in the Bragg peak, and consequently to the presence of a finite range (i.e a well defined depth) in matter. This allows the delivery of highly conformal dose distributions to the target volume, while sparing the surrounding critical structures. However, due to the presence of the steep dose gradient in the depth direction, any shift of the Bragg peak in depth could potentially increase the dose to the organ at risk (Figure 1.2). Thus, the estimation of the exact range in the patient is of fundamental importance for proton therapy. Unfortunately, there are a number of different sources affecting the accuracy of the calculation of range in the patient, such as uncertainties in the derived Hounsfield Unit (HU) number of the planning computed tomography (CT) study; reconstruction artefacts in the CT, uncertainties in the beam delivery system or changes in patient anatomy during therapy. Indeed, these range uncertainties are almost certainly mainly systematic in that they will propagate throughout the course of the treatment (Lomax 2004), and therefore should be carefully considered during the process of planning.

### *1.6.1 Dealing with uncertainties*

As discussed above, both random and systematic uncertainties are inherent to any radiotherapy process and methods should be employed to deal with them. There are a number of different strategies for dealing with such uncertainties. Firstly, regular imaging can be used, either directly before each treatment session (fraction) or even within the time it takes to deliver each fraction. For instance, on-line imaging of PET or prompt-gamma activation has been proposed for on-line range verification (Parodi and Enghardt 2000, Polf et al 2009) as has soft tissue imaging using MRI (Gensheimer et al 2010), or the direct measurement of proton range via the use of a range-probe (Mumot et al 2010) or proton-radiography (Schneider et al 1995), Alternatively, range can also be derived indirectly through daily imaging of the patients positioning using daily x-rays and/or CT imaging or surface imaging (e.g. using the VisionRT system©). Nevertheless, although such methods can certainly help reducing spatial and range uncertainties, they can never completely eliminate them. Thus, it would seem to make sense that when assessing the 'quality' of a treatment (de-facto the 'quality' of a treatment *plan*), an assessment of the '*robustness*' of the plan to these potential uncertainties should be performed and, ideally, such a concept should be incorporated into the planning process directly.

Consequently, in order to make treatment planning more robust to delivery uncertainties, two criteria should be included into the planning process:

i. Tools for robust planning to calculate a plan that is as insensitive as possible to the estimated delivery uncertainties for any given case. This could be done, for example, either by incorporating potential errors directly in the optimization algorithm, as proposed by Unkelbach et al (2007,2009) and Pflugfelder et al (2008), or by manually changing the starting conditions (as, for example, the initial field arrangements and the initial pencil beam weight) to steer the optimization result to solutions more robust to uncertainties, as proposed by Lomax (2004, 2008a).

ii. Tools for the evaluation of plan robustness. Despite the early work of Goitein to systematically calculate error bars for treatment plans taking different delivery uncertainties into account (Goitein et al 1985), such tools are still sadly absent from most treatment planning systems. Nevertheless, simple tools for assessing the sensitivity of a plan to delivery uncertainties could be invaluable for generally improving the quality of radiotherapy treatments, particularly for proton therapy where accurate estimates of the proton range are a pre-requisite for accurate treatments.

## 1.7 Aims and structure of the thesis.

Based on these criteria for robust treatment planning, it has been the aim of this work to investigate to what extent degeneracy can be exploited to produce more robust treatments and to develop tools whereby plan robustness can be assessed at the treatment planning stage.

In the first part of this work, the potential of manipulating the starting conditions of the optimization as a tool for 'steering' the IMPT optimization procedure to more robust solutions has been investigated. In the second part, a new method for evaluating and displaying plan robustness has been developed, with the aim of providing a succinct and easy to interpret tool for treatment plan analysis. Finally, through the development of an anthropomorphic phantom, the concepts developed in parts 1 and 2 of the thesis have been validated experimentally under conditions that are as close as possible to those observed clinically at the Paul Scherrer Institute.

As such, the thesis is structured in the following way. In *Chapter 2* we provide the theoretical background to the dose calculation and optimization algorithms used for planning IMPT treatments at PSI. In *Chapter 3* we then discuss in more detail the concept of degeneracy for the IMPT optimization problem, and discuss how degeneracy can be used to 'steer' the solution to results based on criteria other than simple dose-based constraints. This concept is further expanded in *Chapter 4*, in which the concept of 'starting condition' driven optimization is developed and applied to a number of clinical examples. *Chapter 5* then introduces the concept of the 'error-bar dose distribution' as a tool for comparing and assessing the robustness of different plans to spatial

and range uncertainties. In *Chapter 6*, the ideas of chapters 4 and 5 have been tested experimentally through the development of the anthropomorphic phantom and the use of two dimensional film based dosimetry in the clinical proton beam of PSI. Finally *Chapters 7* and *8* propose some new ideas and areas of future investigation that have arisen as a consequence of this work and a summary of the whole thesis.

---

Note: Parts of this Chapter has been published in the following paper: Albertini F, Gaignat S, Bossherdt M, Lomax AJ 2009 *Planning and optimizing Treatment Plans for actively Scanned Proton Therapy* in Biomedical Mathematics: Promising Directions in Imaging, therapy Planning, and Inverse Problems (Y Censor, M Jiang, G Wang Editors) (Albertini et al 2009)

# 2 Theory

## 2.1 The PSI proton pencil beam model

In this section we describe the basic components of the physical model used to calculate the dose delivered to the patient (Scheib et al 1996, Pedroni et al 2005).

The 3D dose distribution for a proton pencil beam at a given position $d(x,y,z)$ in water is given by

$$d(x,y,z) = \frac{T(z)}{2\pi\sigma_{x(z)}\sigma_{y(z)}} e^{-\frac{x^2}{2\sigma_x^2(z)}} e^{-\frac{y^2}{2\sigma_y^2(z)}} \qquad \text{Eq. 2.1.}$$

*T(z)* is the depth-dose curve (expressed in Gy cm$^2$) integrated over the whole plane perpendicular to the beam at the depth *z*. The two exponential terms represent the spatial distribution of the single pencil beam projected on the x-y plane (orthogonal to the proton direction *z*). $\sigma_x(z)$ and $\sigma_y(z)$ are the width of the distribution (standard deviation) projected respectively on the *x*-*z* plane and *y*-*z* plane as a function of the depth *z*. They include all contributions to the beam width (i.e. initial angular-spatial distribution (*phase space*), beam broadening due to multiple Coulomb scattering when protons pass through range shifter plates, air, patients). The beam width due to multiple Coulomb scattering is modeled based on the work by Øverås (1960), in which the magnitude of the scattering is integrated as a function of depth. Contributions from the width of the incident beam and the combined effect of range shifter plates and the air gap to the patient are added quadratically. For a pencil beam centered in $x=x_0$ and $y=y_0$ the argument *x* and *y* in the above expression should be substitute with $x-x_0$ and $y-y_0$.

The depth-dose curve *T(z)* is described by a physical model, which takes into account the width of the initial energy spectrum (so called *momentum band*[1]), the proton energy loss derived by the Bethe-Bloch equation (Bischel 1972), the *range straggling*[2] in the patient and the attenuation with penetration of the primary photon flux due to inelastic nuclear interaction. *T(z)* has been measured experimentally with a large plane-parallel ionization chamber inserted in a water tank. The free parameters of the model have been adjusted to fit the shape of the measured depth-dose curve. *T(z)* is stored as a normalized dose per incident proton in a Look Up Table as well as $\sigma_x(z)$ and $\sigma_y(z)$.

The dose applied at a given point consists in the superposition of the dose distribution of a large number of single pencil beams (typically 8000 spots in one liter volume) individually calculated in 3D. Therefore the overall dose $D_{tot}(x, y, z)$ is

$$D_{tot}(x, y, z) = \sum_{j=1}^{M} \omega_j d_j (x - x_j y - y_j, z) \qquad \text{Eq 2.2}$$

where this summation is taken over all pencil beams (spots) *j* applied for a given field and M is the number of pencil beams. $d_j(x - x_j y - y_j, z)$ should be calculated for each pencil beam *j* centered at $(x_j,y_j)$. $\omega_j$ is the weight (fluence) of the pencil beam *j*, corresponding to the number of protons per pencil beam, and it should be optimized to achieve the desired prescribed dose (see next section).

Protons are normalized to monitor units of the ionization chamber used to measure and control the proton beam flux via Faraday cup measurements (Coray et al 2002) The Faraday cup measures the total charge deposited in 'the cup' when protons stops entirely within it, and by knowing the proton charge, the number of protons that stopped in the material is directly calculated.

In the PSI treatment planning package the proton dose is calculated taking into account variations of the density distribution of the patient body on the basis of calibrated computer tomography (CT) data (Schneider et al 1996). Therefore all dose grid points, which are regularly spaced in the CT, are transformed into corresponding water equivalent depth. To properly include the effect of density heterogeneity a ray casting model has been introduced by Schaffner et al

[1] The *momentum band* ($\delta p/p$) describes the initial spread of the beam at the exit of the nozzle when no RS are inserted in the beam path. This is a free parameter of the model and at PSI the values used are around ±1%.

[2] Range straggling describes the fluctuations of energy losses while the beam is slowing down in matter. Uncertainty in the range *R* is described by the *range straggling* parameter $\sigma_R/R$, a typical value is about 1% (Goitein 2008).

(1999), in which the depth-dose distribution is scaled by the projected water-equivalent depth of each dose grid spot.

## 2.2 The PSI Optimization algorithm

As described above, active scanning consists of the delivery of a three dimensional pattern of individually weighted proton pencil beams, where in this context a pencil beam's weight can be considered to be it's fluence (i.e. the number of protons delivered). However, assuming that pencil beams are delivered with a spacing of 5mm in all directions (both laterally and in between successive Bragg peaks in depth), then many thousands of individual pencil beams are required to fully cover the tumour volume. Thus, the only tractable solution to find the 'optimal' set of relative fluences for all these pencil beams is to apply an optimisation algorithm. For the system at PSI, we have adopted the following optimisation approach.

Our planning system uses a quasi-Newton technique, similar to that described by Bortfeld et al (1990) and Wu and Mohan (2000) for photon therapy and already described by elsewhere (Scheib and Pedroni 1992, Lomax et al 1996, Lomax 1999) for protons, and re-formalized here. The cost function $F(\boldsymbol{\omega})$ , to be minimized, is defined as

$$F(\boldsymbol{\omega}) = \chi^2 = \sum_i^N g_i^2 (P_i - D_i)^2 \qquad \text{Eq. 2.3}$$

where $\boldsymbol{\omega} = (\omega_1, \omega_2 ..., \omega_M)$ are pencil beam fluences, $N$ is the total number of grid points, $D_i$ is the calculated dose to the grid point $i$, $P_i$ [3] is the prescribed dose at that point and $g_i$ [4] is the associated importance factor, based on planner's prescriptions.

When the pencil beam weights $\boldsymbol{\omega} = (\omega_1, \omega_2 ..., \omega_M)$ are updated during the iterative process, then the dose $\mathbf{D} = (D_1, D_2, ... D_N)$ can be calculated through

$$D_i = \sum_{j=1}^{M} d_{ij} \omega_j \ , \qquad \text{Eq. 2.4}$$

where $d_{ij}$ [5] is the un-weighted dose contribution of pencil beam $j$ to the grid point $i$, $M$ is the number of pencil beams.

---

[3] The prescription dose is assumed to be homogeneous within the volume apart from close to the surface, where it has a Gaussian fall-off calculated in such a way that the dose to the target volume's surface is 90% of the maximum prescribed dose to the target (this is done to provide a more realistic template that approximately matches with the lateral fall-off of the individual proton pencil beam)

[4] For the weighting matrix (**g**) all dose grid points within the target volume are assigned a fixed weight of 1

[5] As an initial step outside of the optimization loop, the treatment planning pre-calculates the dose, $d_{ij}$ (with equation 2.1), for each pencil beam $j$, for every depth (i.e. for every Energy and every range shifters inserted), to every voxel $i$ within a distance in the x-y plane of half lateral distribution standard deviation ( $\sigma_{x,y}$ )from the center of the pencil beam.

The inverse problem to be solved here is to find the set of pencil beam fluences (**ω**) which minimizes the cost function $F(\boldsymbol{\omega})$. As stated by Bortfeld (Bortfeld et al 1990), there exists a huge variety of iterative algorithms for the solution of such problems which are similar to the Newton iteration:

$$\omega(k+1) = \omega(k) - \alpha_k (\nabla^2 F(\omega(k)))^{-1} \nabla F(\omega(k)), \qquad \text{Eq. 2.5}$$

The main problem to be solved is how to invert the Hessian matrix $\nabla^2 F(\omega(k))$ which is difficult due to the huge size of **D**; therefore the matrix has to be approximated to make the problem tractable. A common approximation of this matrix is by a diagonal matrix **S** whose elements are those of the Hessian matrix: $S_{j,j} = \sum_i^N D^2{}_{ij}$ [6] (see e.g. Bortfeld et al 1990).

In one dimension, for a single pencil beam *j*, the iterative update of the beam weight $\omega_j$ from iteration (k) to iteration (k+1) then becomes

$$\omega_j(k+1) = \omega_j(k) - \alpha_k \left(\frac{\frac{dF}{d\omega_j}}{\frac{d^2F}{d\omega^2{}_j}}\right) = \omega_j(k) + \alpha_k \left(\frac{\sum_i^N g_i^2 (P_i - D_i) d_{ij}}{\sum_i^N g_i^2 d^2{}_{ij}}\right). \qquad \text{Eq. 2.6}$$

where $\alpha_k$ is a damping factor to ensure convergence[7] for the iterative process.

All $\omega_j(k)$ should fulfill physical constraints which require that only positive fluencies can be delivered. Often this request is satisfied by using a positive operator which turns to zero all negative constraints, $\boldsymbol{\omega} = \begin{cases} \boldsymbol{\omega} & if \quad \omega \geq 0 \\ 0 & if \quad \omega < 0 \end{cases}$ (see e.g (Bortfeld et al 1990)).

However, as mentioned by Pflugfelder (Pflugfelder et al 2008a,b), the use of the positive operator can slow down the algorithm since it can repeatedly provide the same solution and can result in solutions which are worse than the previous iteration.

---

[6] The use of the diagonal matrix **S** to approximate the Hessian matrix implies that there are only a few pencil beams delivering dose to the same voxel (Pflugfelder et al 2008a,b). Off-diagonal elements of the Hessian matrix are non-zero if two beamlets deliver dose to the same voxel. By setting these elements to zero every beamlet is treated independently during the optimization process. As this assumption is patently incorrect for particle therapy (as there is large overlap between beamlets laterally, and particularly along the beam directions where many plateau doses overlap), convergence can only be achieved by introducing a damping function.

[7] Often $\alpha_k$ is chosen to satisfy the **Wolfe conditions.** The Wolfe conditions are a set of inequalities that guaranties convergence in (unconstrained) optimization especially in quasi-Newton methods for performing inexact linesearches. Inexact linesearches provide an efficient way of computing an acceptable step length that reduces the cost 'sufficiently', rather than minimizing the cost function exactly.

Our planning system uses a damping factor introduced by Lomax et al (1996) with the following form:

$$\alpha_{ij}(k) = {}^{\omega_j(k)d_{ij}}\!\!\diagup\!\!_{D_i(k)}\ , \qquad \text{Eq. 2.7}$$

which represents the fraction of the total calculated dose at the kth-iteration contributed by spot *j* to grid point *i*. Eq. 2.6 then becomes

$$\omega_j(k+1) = \omega_j(k) + \alpha_k\left(\frac{\sum_i^N g_i^2(P_i - D_i)d_{ij}}{\sum_i^N g_i^2 d^2{}_{ij}}\right) = \omega_j(k) + \omega_j(k)\left(\frac{\sum_i^N g_i^2 \frac{d_{ij}}{D_i}(P_i - D_i)d_{ij}}{\sum_i^N g_i^2 d^2{}_{ij}}\right) =$$

$$= \omega_j(k)\left[1 + \left(\frac{\sum_i^N g_i^2 \frac{d_{ij}{}^2}{D_i}P_i}{\sum_i^N g_i^2 d^2{}_{ij}} - \frac{\sum_i^N g_i^2 \frac{d_{ij}{}^2}{D_i}D_i}{\sum_i^N g_i^2 d^2{}_{ij}}\right)\right] = \omega_j(k)\left[1 + \frac{\sum_i^N g_i^2 \frac{d_{ij}{}^2}{D_i}P_i}{\sum_i^N g_i^2 d^2{}_{ij}} - 1\right] =$$

$$= \omega_j(k)\frac{\sum_i^N g_i^2 d_{ij}{}^2 \frac{P_i}{D_i}}{\sum_i^N g_i^2 d^2{}_{ij}}. \qquad \text{Eq. 2.8}$$

With the use of the damping function described in Eq 2.7, this final optimization algorithm clearly fulfills the physical request of positive fluences without adding additional constraints to the problem[8], as long as both the initial fluences and the doses are always positive.

Both the target dependent prescription dose $P_i$ and the weighting function $g_i$ can be modified by the presence of critical structures. All critical organ VOIs can be assigned individual prescription doses and weights, depending on their relative importance in the planning process. The weighting function for critical structures has been defined in the same manner as described by Bortfeld *et al (1990)*. So to express it explicitly, the objective function $F(\boldsymbol{\omega})$ maybe be rewritten as

[8] The damping function not only guarantees convergence to the iterative process but also elegantly transform the constrained problem to an un-constrained one. The absence of additional constraints is probably the reason for the rapid convergence observed with our optimization algorithm. Matsinos E. et al (2007) have in fact shown for a single-case study that PSI optimization algorithm produces results comparable to conjugate gradient (CG) and simulated annealing (SA) optimization methods, but converging to a lower value of the objective function and, most importantly, in a faster way ($t_{CG}$=~1.5 $t_{PSI}$ and $t_{SA}$=~2.2 $t_{PSI}$).

$$F(\boldsymbol{\omega}) = \sum_{i}^{N_T} g_i^2 (P_i^T - D_i)^2 + \sum_{i}^{N_{OAR}} g_i^2 (P_i^{OAR} - D_i)^2 H(P_i^{OAR} - D_i) \qquad \text{Eq. 2.9}$$

where $N_T$ and $N_{OAR}$ are respectively the number of voxels in a target volume and in an organ at risk $H(P_i^{OAR} - D_i)$ is the step function defined as

$$H(P_i^{OAR} - D_i) = \begin{cases} 1 & D_i > P_i^{OAR} \\ 0 & D_i \leq P_i^{OAR} \end{cases}. \qquad \text{Eq. 2.10}$$

In other words, only the voxels in the OAR with dose greater than the tolerance dose will contribute to its objective function components. This concept has proven to result in a fast and stable optimization (Bortfeld et al 1990, Thiecke et al 2003).

As a final comment, the widely used quasi-Newton method here discussed is a gradient based optimization algorithm. Hence, it finds the local minimum closest to the defined starting condition. By manipulating the starting conditions it is therefore possible to 'steer' the optimization outcomes to results defined by the user. This leads to the concept of 'planner driven' optimization, a characteristic that is expanded and exploited in the following chapters of this work.

---

Note: Section 2.2 has been published as part of the following paper: Albertini F, Gaignat S, Bossherdt M, Lomax AJ 2009 *Planning and optimizing Treatment Plans for actively Scanned Proton Therapy* in Biomedical Mathematics: Promising Directions in Imaging, therapy Planning, and Inverse Problems (Y Censor, M Jiang, G Wang Editors) (Albertini et al 2009)

# 3 IMPT & Degeneracy

## 3.1 Introduction

Given the number of variables available to the optimisation engine, and the relatively basic goals of radiotherapy planning (a homogenous dose to the target, no more than a certain defined dose to one or more critical structures), it has been found, both practically and theoretically, that the problem of IMRT optimisation is highly degenerate (Alber et al 2002, Llacer et al 2003, Webb 2003, Llacer et al 2004). That is, there are many, sometimes very different fluence profiles that meet the primary planning constraints of a given dose to the target volume whilst keeping dose to all critical structures to below pre-determined values. Thus, without additional guidance, the result of the optimisation will generally depend on the starting conditions. Although it could be argued that when the resulting dose distribution is acceptable, degeneracy in the problem doesn't matter, degeneracy can also be exploited in order to find solutions to the problem based on alternative, non-clinical criteria. For instance, Hou et al (2003) have suggested exploiting plan degeneracy in IMRT such as to force the result towards smoother fluence patterns, which should be easier and safer to deliver.

Intensity Modulated Proton Therapy (IMPT) can be considered to be the proton equivalent of IMRT with photons. It consists of the simultaneous optimization of all Bragg peaks from all field directions, with or without additional dose constraints to neighboring critical structures. However, given that the Bragg peak characteristic of the proton depth-dose curve is well localized in space in three dimensions (unlike the photon depth dose curve, which falls off slowly and more or less exponentially), with IMPT we literally have a third dimension which can be modulated. As discussed in Chapter 2, the optimization problem consists of determining, for each individual field, the optimum weights of a set of energy (and therefore depth) modulated

Bragg peaks, with each energy layer consisting of a two dimensional distribution of proton pencil beams.

Consequently, given the greater degrees of freedom available to the optimiser in IMPT, it is likely that IMPT optimisation will be even more degenerate than photon IMRT (Lomax 2004). Several possible IMPT approaches have been already described and categorised (Lomax 1999), ranging from full 3D-IMPT, in which Bragg peaks are distributed throughout the target volume in three-dimensions from every incident beam direction, to Distal Edge Tracking (DET) (Deasy et al 1997), in which for each field, Bragg peaks are only deposited on the distal edge of the target volume. The near equivalence of these approaches from the point of view of sparing structures and of target dose homogeneity has also been documented by several authors (Lomax 1999, Oelfke and Bortfeld 2000, Lomax et al 2004, Lomax 2008a-b). However, it has recently been shown that two dosimetrically equivalent solutions will not necessarily behave similarly to delivery uncertainties (Lomax 2008a-b).

We believe that the problem of degeneracy is rather complicated: it can indeed have both a negative and a positive face. From one side, as the problem is generally under-defined (i.e. there are many more variables available than constraints defined) the outcome of the optimization engine can result in a sub-optimal plan for instance for the normal tissues for which no constraints are defined . On the other side, we believe that degeneracy can also be utilised in a number of ways to drive the plans to more desirable results based on non-dosimetric criteria that are, for example, easier or safer to deliver. These two opposite faces of the problem of degeneracy are here illustrated with two examples. In the first example we show how the use of different starting conditions, although delivering the same dose distribution to the target volume (plan degeneracy), leads to a difference in the dose to the normal tissue, which increases the potential risk of developing secondary neoplasm. In the second example we illustrate how degeneracy can be positively exploited to construct plans that are safer to deliver. In this example, a number of plans are calculated with different starting conditions (in this case, field arrangements), and the robustness of each plan assessed *post-priori*.

## 3.2 A negative example of degeneracy: normal tissue integral dose

The first example of degeneracy illustrates how, if the problem is too ill-defined, the optimization can result in a sub-optimal plan if the starting conditions are not well defined. In particular it shows how the choice of the starting conditions (in this case, the initial beamlet fluences) affects the normal tissue integral dose. As any extra dose delivered to normal tissue

increases the risk of developing a radiation-induced secondary neoplasm, it is important to reduce as much as possible the normal tissue integral dose.

To illustrate this, we have calculated 2-field IMPT plans optimized with the single, and only constraint of delivering a homogeneous dose to the target volume. For the two plans, different starting conditions have been used:

(i) **Forward wedge:** uses Bragg peaks equally weighted in depth for each field direction, which leads to an initial dose distribution along the beam direction with a gradient from the proximal edge(maximal dose) to the distal (minimal dose) part of the target (see Figure 3.1 (a));

(ii) **Flat SOBP:** uses a pre-weighted set of Bragg peaks, in which the weights are reduced in depth from distally to proximally such as to deliver a flat Spread-Out-Bragg-Peak (SOBP) type profile along the beam direction (see Figure 3.1 (b)).

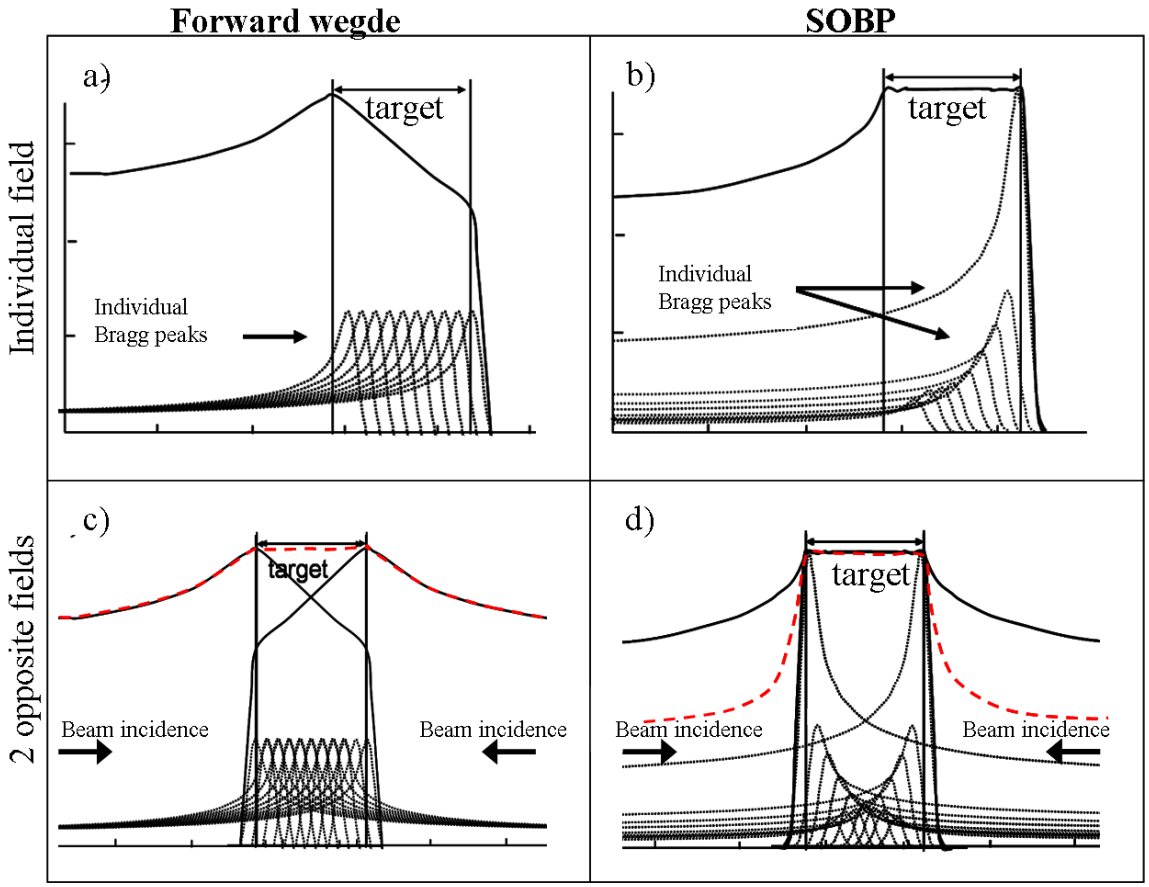


**Figure 3.1** a)-b)Schematic representation of the single field dose distribution for both individual Bragg peaks with identical weighting in depth resulting in a '**forward wedge**' and individual Bragg peaks with reduced weighting from distally to proximally resulting in a '**flat SOBP**' and reduced cumulative entrance dose. c)-d)schematic representation of the composite dose distribution (red dashed line) for 2 opposite fields, for the two approaches delivering the same homogeneous dose to the target volume. Note the reduction of the entrance dose when the SOBP approach is used.

As schematically shown in Figure 3.1(c)-(d), when two opposite fields are combined together, although the coverage and the dose to the target volume will be the same (plan degeneracy), in the absence of additional normal tissue constraints, the plan achieved with the **flat SOBP** approach significantly reduces the dose to the normal tissue.

To analyze this effect for a real case, we have calculated for a prostate patient, 2-field IMPT plans with the aim of delivering homogeneously 79.2 $Gy_{RBE}$ (with a radio biological effectiveness of 1.1) to the target volume, while keeping the dose to the rectum below 75 $Gy_{RBE}$. The dose distribution for a lateral field of each approach is shown in Figures 3.2.

Figure 3.2a shows that, with the '**forward wedge**' approach, a significant amount of the dose is delivered in the proximal portion of the target volume. In contrast, Figure 3.2b shows that, with the '**flat-SOBP**' approach, a more uniform dose is delivered throughout the target volume and moreover, that the dose delivered proximal to the target volume is substantially reduced. This last result is confirmed in the full plans shown respectively in Figures 3.2(c) and (d). Indeed, the volume of the femoral head receiving a dose above 40 $Gy_{RBE}$ ($V40_{GyRBE}$) is reduced from 89% to 24% by starting the optimization process with pre-weighted Bragg peaks (i.e. with the 'flat-SOBP' approach). In addition, a general dose reduction is observed between the plans for all tissues in the entrance paths of the two fields. In particular, this results in half of the normal tissue volume receiving an integral dose above 30 $Gy_{RBE}$ as compared with the '**forward wedge**' approach.

In conclusion, although both starting condition fully satisfy the treatment objectives of designing an homogeneous target coverage while reducing the dose to the rectum (plan degeneracy), starting the optimization process with the '**forward wedge**' approach does not provide an optimal plan from the point of view of dose delivered proximal to the target volume, and in particular to the normal tissue. In particular, this is the approach investigated at a different proton center (see e.g. Trofimov et al 2007), whereas the '**flat-SOBP**' is the routine approach for delivering 3D-IMPT treatments used at PSI (Lomax 1999).

Finally, we would like to mention that the dose to the proximal normal tissues can be even more reduced using a distal-proximal wedge (maximum dose at the distal end) or, at the extreme, by delivering only the most distal beamlet (i.e. applying the DET technique) as the starting point. The impact of using these starting conditions is further pursued in the following chapter.

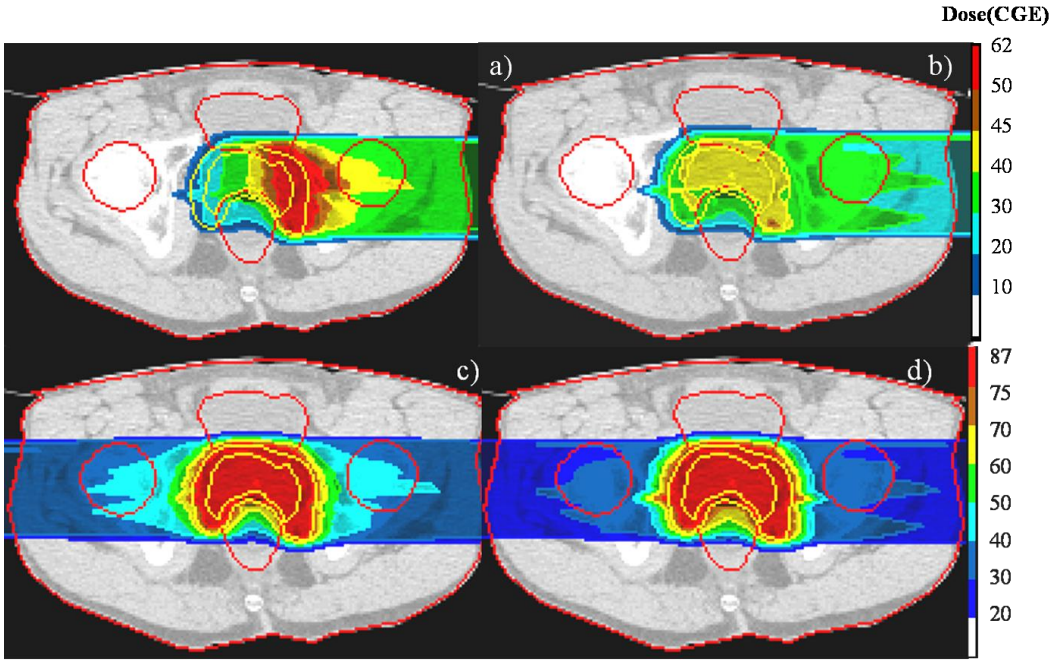


**Figure 3.2** a)-c) Single field and full plan dose distribution achieved with the '**forward wedge**' and b)-d) with the '**flat SOBP**' approach respectively. Targets (PTV and CTV) are outlined in yellow; organs at risk are in red. Both single fields and plans are normalized to the mean dose to the target. Both single fields and plans are normalized to the total dose to the target.

## 3.3 A positive example of degeneracy: field numbers and orientations in IMPT

As a second example of degeneracy, we present a contrasting study looking into the effects of field directions and number of fields on IMPT plans calculated for a set of skull-base chordoma cases. These are radiation resistant tumours originating in the bony structures of the skull base. Although relatively slow growing, these tumours are particularly problematic due to their radio-resistance and very close proximity to dose limiting critical structures such as the brainstem and optic structures (optic nerves and chiasm). Although relatively rare tumours, these have become a standard indication for proton therapy, which allows for high levels of conformation allowing for dose escalation to the tumour bed, whilst simultaneously sparing the neighbouring critical structures. Figure 3.3 shows a typical case, with a central and somewhat anteriorly positioned target volume, slightly wrapping around the brain stem.

The optic structures are not seen in this figure, as they are positioned just cranially of the target volume. The aim of this study was to look into a set of different field arrangements for treating such cases, to investigate if a 'standard' set of beam arrangements could be determined. The field arrangements investigated are shown in Figure 3.3. Five arrangements were investigated:

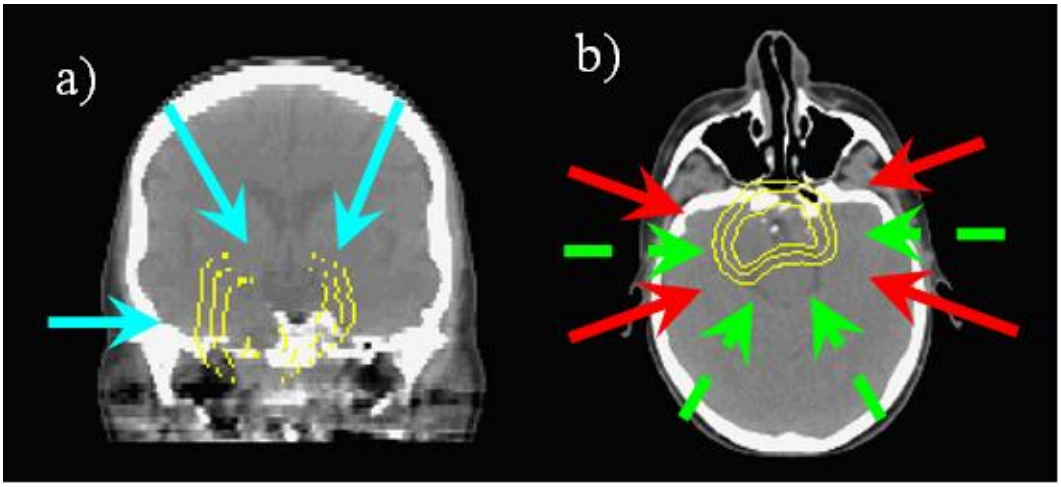


**Figure 3.3.** An example of a skull base chordoma used for analyzing the effect of field directions and number of fields on IMPT treatment plans. The inner contour is the gross tumor volume (GTV), the middle the clinical target volume (CTV), and the outer contour the planning target volume (PTV).The aim is to cover the PTV contour with as homogenous dose as possible, while limiting the dose to the defined critical structures. The different field directions are as shown in the two figures. (a) The field configuration of plan A.(b) The other field directions used to construct plans B–E. The red fields show the four fields of the 4F-star arrangement and the green fields, the four fields of the 4F-Pi arrangement. The other two arrangements are combinations of these red and green fields.

A) Three fields consisting of one lateral field and two non-coplanar, superior-lateral obliques (3F);
B) Four field star arrangement consisting of two posterior lateral obliques and two anterior lateral obliques (4F-star);
C) Four field 'PI' arrangement consisting of two lateral fields and two posterior-lateral obliques, forming a 'PI' arrangement (4F-PI);
D) Six field star, combining the four fields of the 4F-star arrangement plus an additional two lateral fields;
E) Six field 'PI" arrangement, consisting of the four fields of the 4F-star plan, plus the two posterior fields of the PI arrangement.

For each field arrangement, IMPT plans were calculated assuming a target dose of 74 $Gy_{RBE}$ and with exactly the same dose constraints on the neighbouring critical structures – maximum doses of 64 $Gy_{RBE}$ to the brainstem, 53 $Gy_{RBE}$ to the centre of the brainstem and no more than 60 $Gy_{RBE}$ to the optical structures. Priority was given to the critical structure constraints, and the plans were evaluated looking at the various dose based parameters representing the dose coverage and dose homogeneity to the various target volumes.

Figure 3.4 shows the results of this analysis. For each target volume defined (namely the PTV, the CTV and the GTV), the $V_{95\%}$ and $D_{98\%}$ for each plan are plotted. PTV, CTV and GTV are the name respectively used for the planning target volume, the clinical target volume and the gross tumour volume (ICRU 50). The $V_{95\%}$ is simply the percentage of the target volume that receives 95% of the prescription dose or more, whereas $D_{98\%}$ is the dose (as percentage of the prescription dose) that 98% of the target volume receives. Both parameters are a measure of the minimum dose coverage of the target volume, which is, for such tumours, considered to be a strong indicator for tumour control (Qu et al 2002). In addition, also in Figure 3.4, the absolute maximum dose in each plan is plotted.

The results presented in Figure 3.4 are a clear example of the degeneracy of the optimisation problem in IMPT. Despite a wide range of different field arrangements (and different numbers of fields), all the measured parameters are more or less identical, indicating that, regardless of if we take a simple 3 field approach, or a more complicated 6 field arrangement, the target coverage (for the same maximum doses in the neighbouring critical structures) will be the same. Clearly then, in the absence of any other factors, one would tend towards the simplest three field approach, as this would be the easiest and fastest to deliver, and increasing the number of fields (or changing their arrangement) brings little or no advantage.

However, for this study, we were also interested in investigating if the problem of degeneracy can be resolved through the introduction of other, non dose-based characteristic, as for example, whether the plan is more robust to possible delivery uncertainties.

As mentioned in section 1.6, one of the most important uncertainties in proton therapy is the range. It is perhaps a truism to state that, the advantage of protons is that they stop whereas the disadvantage of protons is that we don't always know where. Nevertheless, potential uncertainty in the range must be taken seriously, particularly as this error could very well be systematic – that is, it is the same for every fraction. Thus all plans have been re-calculated assuming a +/-3% systematic change to the CT values of the planning CT. In this case, reducing CT values by 3% will result in a systematic *overshoot*, whereas increasing them by 3% results in a systematic *undershoot* of delivered proton pencil beams.

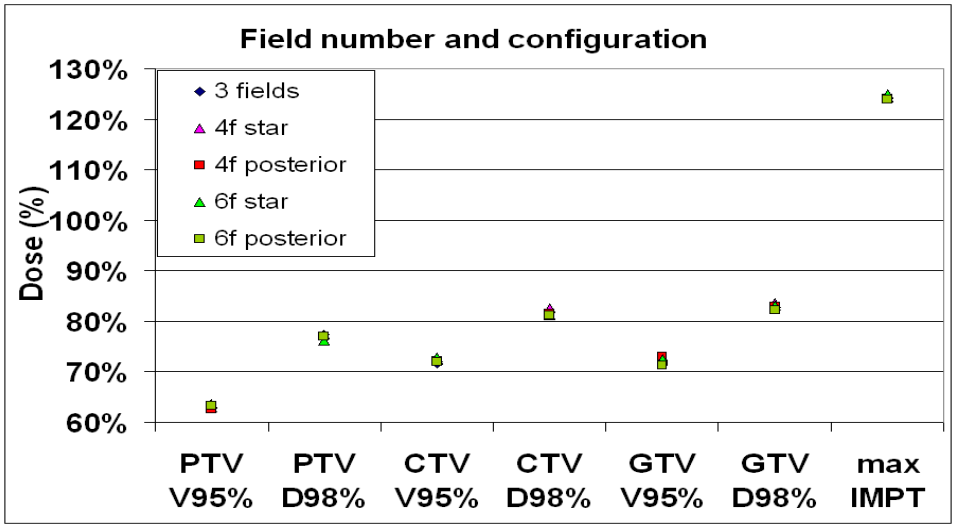


**Figure 3.4** Various plan evaluation parameters calculated for the five field arrangements (A–E).V95 is the volume that receives 95% of the percentage dose. D98 is the dose that 98% of the volume receives, and max is the maximum dose anywhere in each plan. Despite the various field arrangements and different numbers of fields (between 3–6), all parameters are almost identical across all plans (including the maximum dose)

Figure 3.5 shows for a cohort of five patients, an analysis of the robustness for field arrangements B and C (4 field star and Pi arrangement). The figure 3.5 shows the maximum doses of the nominal plans and the maximum doses when both plans are recalculated with the range errors included. As would be expected from the discussion above, the maximum doses for the nominal plans are essentially identical, as are the maximum doses for the condition of systematic undershoot (+3% HU). However, when simulating an overshoot, the star arrangement is substantially more robust, with the maximum dose increasing only 1-2% in comparison with the nominal plan. For the PI arrangement on the other hand, the maximum dose increases substantially by 9-10%, an alarming amount given the relatively modest range uncertainty modelled. Clearly, when using the absolute maximum dose in the plan as a criterion, the 4F-Pi arrangement is much more sensitive to 3% overshoot than the 4F-star arrangement.

With this example we have begun to study how the problem of degeneracy is resolved when other non dose-based characteristic are introduced. In particular how degeneracy can be used to design plans that are more robustness to delivery uncertainties. Based on the results of the basic robustness analysis here presented, we have further followed this analysis in a more extended and formal way in the next chapter.

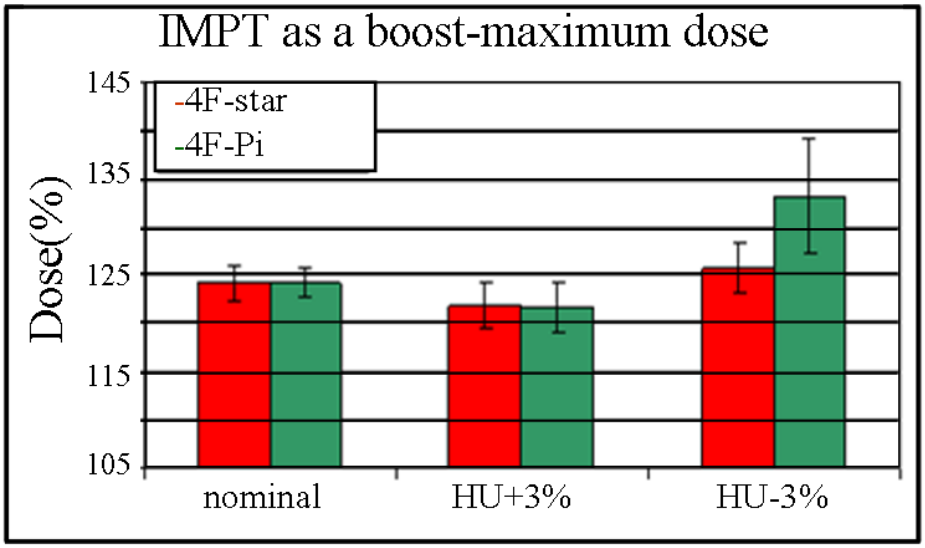


**Figure 3.5**. Changes in the maximum dose in the 4F-star and 4F-Pi field arrangements when recalculated on CT datasets with all Hounsfield units (HU) altered by either +3% or –3%.These simulate approximately 3% undershoot (+3% HU) or 3% overshoot (–3%HU), respectively. As seen in Figure 3.4, there is no difference in maximum dose for the nominal plans, and only a very minimal difference for the undershoot plan (+3% HU). However, for the –3% plans (overshoot), there is a large and significant difference in the maximum dose between to two field arrangements.

## 3.4 Summary

From both studies reported above, it can be seen that optimisation for proton therapy is a highly degenerate problem. Moreover, as discussed in Chapter 2, it is a characteristic of the optimisation algorithm used here, that the resulting dose distribution is heavily dependent on the starting conditions. These aspects have to be seriously considered when planning, as, with the 'wrong' starting conditions, the optimization engine can result in a sub-optimal plans.

However, degeneracy can also be utilised to design plans that fulfil other, non-dose based criteria. These include plan robustness constraints. Anyhow, this is only one of the possibilities, and many other plan 'characteristics' could perhaps be studied and incorporated into the planning process. This could be utilised by incorporating successively more 'constraints' into the optimisation process. On the other hand, and this is our philosophy, one can give control of the outcome of the planning system to the planner. For example, due to the strong dependence of the resulting dose distribution on the starting conditions, if these can be stipulated in a straight forward and intuitive way, then this leads to the concept of 'planner driven' optimisation. In other word, the user drives the optimisation towards plans with desirable characteristics by steering it through the

manipulation of these starting conditions. This concept will be further pursued in the following chapter.

---

Notes:
Section 3.2 has been published as part of the following paper: Albertini F, Lomax AJ and Hug EB 2007 *In regard to Trofimov et al: Radiotherapy treatment of early-stage prostate cancer with IMRT and protons:a treatment planning comparison (Int J Radiat Oncol Biol Physics 2007;69:444-453* (Albertini et al 2007)
Section 3.3 has been published as part of the following paper: Albertini F, Gaignat S, Bossherdt M, Lomax AJ 2009 *Planning and optimizing Treatment Plans for actively Scanned Proton Therapy* in Biomedical Mathematics: Promising Directions in Imaging, therapy Planning, and Inverse Problems (Y Censor, M Jiang, G Wang Editors) (Albertini et al 2009)

# 4 Dealing with uncertainties: Robust planning through manipulation of the starting conditions

## 4.1. Introduction

As already discussed in Chapter 3, the problem of Intensity Modulated Proton Therapy (IMPT) optimisation is highly degenerate. In other words, there are many, sometimes very different, fluence profiles that meet the primary planning constraints of delivering a given dose to the target volume whilst limiting doses to the critical structures. Thus, without additional guidance, the result of the optimisation will generally depend on the starting conditions. We have shown in Chapter 3 an example of how degeneracy can be exploited in order to find solutions to the problem based on alternative, non-clinical criteria, as for example being more robust to delivery uncertainties. Here we want to pursue this concept in more detail.

The problem of assessing the robustness of a plan in the case of uncertainties is of major importance for radiotherapy with protons in general, due to the spatially well defined Bragg peak curve, and even more for intensity modulated therapy, given the presence of steep dose gradients used to shape the dose. In particular any delivery error has the potential to shift the Bragg peak curve and thereby strongly affect the dose distribution. This could potentially result in over-dosage of a critical structure.

To solve the problem, some authors have proposed methods to integrate range and set-up uncertainties directly in the optimization algorithm (Unkelbach 2007 and 2009 and Pflugfelder 2008) however, these improvements have the disadvantage of significantly increasing the computational time.

In this Chapter we follow-up on the recent publication by our group (Lomax 2008a) in which the effect of systematic range uncertainties have been analysed for a single case (a skull-

base tumour) by comparing 2 different IMPT plans (3D- and the DET-based plans). In particular we have applied four 'flavours' of IMPT, which differ in their initial set of beamlet weights, to three different types of clinical situations. For each of these, we have studied in detail the individual field distributions to better understand the characteristics of the different IMPT plans and their influence on robustness to range errors. Our aims in doing this are to 1) investigate how the choice of starting conditions alone can be used to drive treatment plans to more robust solutions and 2) to better understand the plan and individual field characteristics that effect robustness. By manipulating starting conditions, we aim to find 'case solutions' for different plan types that are both robust and meet clinical constraints, whilst also providing insights into the details of individual field dose distributions resulting from IMPT which could be useful for subsequent research into robust optimisation techniques.

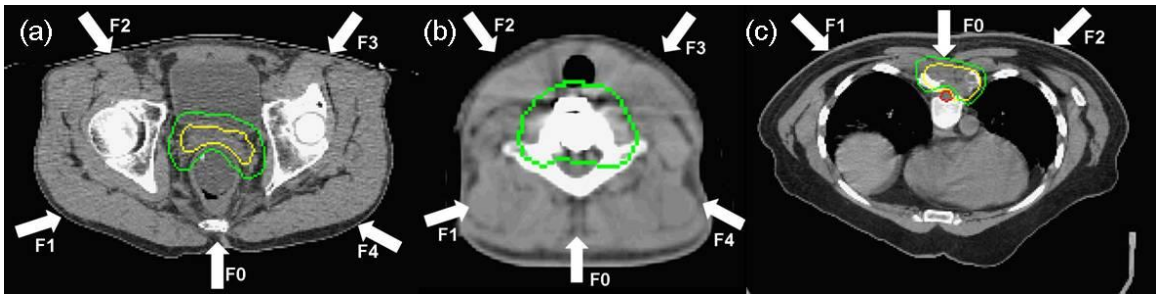


**Figure 4.1** Three clinical cases selected for the study and their corresponding field directions (white arrows): (a) Prostate carcinoma, (b) Chondrosarcoma of the C-spine, (c) Chndrosarcoma of the thoracic spine. Target volumes are indicated by green contours in all cases.

# 4. 2. Materials and methods

## *4.2.1 Patient data*

Three different patients have been included in this study:

### *4.2.1.1. Case A. Prostate carcinoma*

A patient with prostate cancer (figure 4.1a) has been selected for this study, as it represents a medium-large target volume (~210 $cm^3$) and requires only a rather "simple" IMPT plan, as no (or only minor) compromises to the target coverage were necessary in order to meet the organ-at-risk (OAR) constraints. In fact, as both the prescribed dose to the PTV and the maximum tolerance dose to the rectum, were set to 74 $Gy_{RBE,}$ (assuming a radio biological effectiveness (RBE) of 1.1), then the clinical requirements were satisfied with relatively relaxed planning constraints. For this case, five evenly distributed fields, separated by 72° starting from the posterior direction have been used for all plans.

#### *4.2.1.2. Case-B. Cervical chondrosarcoma*

This is a cervical chondrosarcoma patient (figure 4.1b) with a relatively small target volume (~70 $cm^3$) circumferentially encompassing the spinal cord. For this case we simulated our clinical procedure by splitting the treatment in two series (Rutz et al 2008): an initial Single Field, Uniform Dose (SFUD) plan delivering uniformly a total dose of 44 $Gy_{RBE}$ to the target volume, followed by an IMPT plan delivering the remaining 28 $Gy_{RBE}$ to reach a total dose of 72 $Gy_{RBE}$. For the second series IMPT plans, relatively severe constraints had to be applied to the main critical structure (in this case the spinal cord) in order to stay within the maximum dose constraint (63 $Gy_{RBE}$ to surface and 54 $Gy_{RBE}$ to the spinal cord center). In our approach we had to force the individual fields to have much greater in-field dose gradients compared to case-A. The field configuration was the same as for case A.

#### *4.2.1.3. Case-C. Thoracic chondrosarcoma*

The last case, a thoracic chondrosarcoma (figure 4.1c), was included as it was the first patient clinically treated with IMPT in 1999 (Lomax et al 2001). In addition, a similar geometry has been recently used as a model to study the inclusion of delivery uncertainties in the optimization algorithm from different centres (Unkelbach et al 2007, 2009 and Pflugfelder et al 2008). In 1999, this patient was treated with a "manual" 3-field IMPT plan, in the sense that field-specific target volumes were designed to avoid beamlets directly stopping in front of the spinal cord (Lomax et al 2001). Using this target configuration, the starting conditions were essentially 'manually' adjusted to drive the optimization to produce a plan that was significantly more robust to range errors. In this work, IMPT plans have been calculated with the same beam arrangement as that used by Lomax (Lomax et al 2001), but without the use of field specific targets. That is, three coplanar beams equally spaced by 45°, entering from the posterior aspect of the patient. In addition, we assumed to deliver a total dose of 72 $Gy_{RBE}$ and a maximum tolerance dose of 60 $Gy_{RBE}$ to the surface and 54 $Gy_{RBE}$ to the center of the spinal cord (SC),

### *4.2.2 Treatment planning*

IMPT plans were calculated using the PSI treatment planning system, assuming an in-air pencil beam width of 3mm ($\sigma$) at the entrance point (Lomax et al 2004). However, as the beam energy is varied at the PSI Gantry by the insertion of range shifters plates (RS) in front of the patient, the actual size of the pencil beam at the entrance point is somewhat larger, and strongly depends on the number of RS plates inserted and on the air-gap between the last plate and the patient surface (Pedroni et al 2004). An average value for $\sigma$ can be considered to be ~8mm

(Lomax et al 2004). The dose has been calculated analytically using the ray casting pencil beam algorithm described in Chapter 2.1 (Schaffner et al 1999).

We exploited the full 3D-IMPT method as described by Lomax (Lomax 1999), for which individual Bragg peaks are distributed, for every beam, throughout the target volume in three-dimensions. Spot positions are defined as the position of the maximal dose for a given Bragg peak, and are calculated taking into account density heterogeneities derived from the planning CT. Spots are placed on a regular, rectangular grid, which is in line with the proton beam, spaced 4x5 mm in the lateral direction and 4.6 mm in depth (water equivalent), the latter value being determined by the thickness of the range shifter plates used by the PSI gantry for varying energy.

A Bragg peak for a given lateral position and a given depth is referred to as a beamlet and is assigned a fluence weight during the optimization. The initial set of beamlets available to the optimization engine includes all those positioned within the defined target volume plus up to 5 mm outside the target volume. It has been previously shown that the selection of additional spots a small distance outside the target volume improves target coverage (Lomax 1999).

### *4.2.3 IMPT starting condition: initial beamlet fluences*

For each case, IMPT plans have been calculated using four different initial sets of beamlet fluences (figure 4.2), i.e. the set of fluences applied to all pencil beams as a starting condition for the subsequent optimisation. These were as follows:

a) **Forward wedge**: All beamlet weights set to the same weight, which leads to an initial dose distribution with a gradient from the proximal edge(maximal dose) to the distal (minimal dose) part of the target (figure 4.2a);
b) **SOBP**: a pre-weighted set of beamlets, in which the weights are reduced from distally to proximally such as to deliver a flat 'Spread-Out-Bragg-Peak' throughout the target (the default for IMPT optimisation at PSI);
c) **Inverse wedge**: a pre-weighted set of beamlets, whereby the weights are reduced from distally to proximally resulting in a gradient from the distal edge (maximal dose) to the proximal (minimal dose) aspect of the target;
d) **DET**: initial beamlet weights set to 0 universally, except for the most distal ones for each given lateral direction. The selected beamlets are then assigned the same initial weight (Deasy et al 1997, Lomax 1999 ).

These four starting conditions have been selected in order to cover a wide spectrum ranging from the proximal to distal decreasing wedge ('**forward wedge**') to the other extreme of **DET**, which has been proposed by various authors as a means to reduce integral dose and decrease delivery time (Oelfke and Bortfeld 2000, Nill et al 2004).

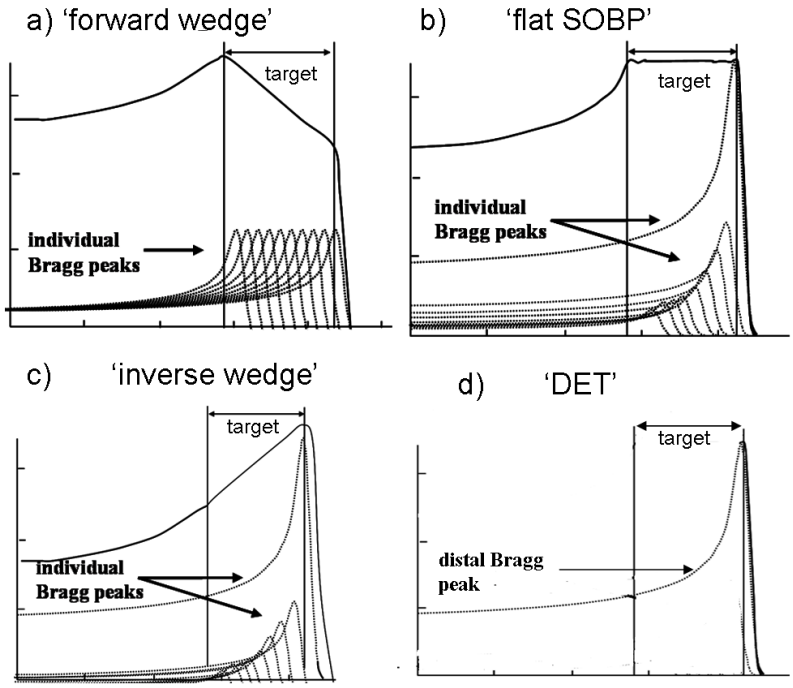


**Figure 4.2** Schematic representation of the initial beamlet weights along the field direction for a single field approaching the target from left to right: (a) individual Bragg peaks with identical weighting in depth resulting in an initial 'forward wedge' dose distribution; (b) individual Bragg peaks with reduced weighting from distally to proximally resulting in a 'flat SOBP' depth dose distribution; (c) individual Bragg peaks with reduced weighting from distally to proximally such as to deliver an initial 'inverse wedge' dose distribution; (d) selection of the most distal Bragg peak only (DET approach)

### *4.2.4 Plan analysis*

Plans were analyzed using the cumulative dose–volume histograms (DVHs) calculated for each plan and for the target and main dose limiting OAR's. The DVH is a simple tool to reduce the 3-dimensional dose distribution to a single plot (Drzymala et al 1991, Goitein 2008). In particular, it summarizes the distribution of dose within a particular volume of interest (VOI) by means of a cumulative frequency distribution of the dose within the VOI. Each point on the curve gives the information of the total volume of the VOI receiving a dose greater or equal to the dose indicated by the abscissa (see figure 4.3). Ideally, the DVH of the target volume will look like a square box, with the 100% of the volume covered with 100% of the dose. In contrast, for an

OAR, the closest the DVH is to the bottom-left corner, the better it is, as it means that the critical structure is receiving less dose.

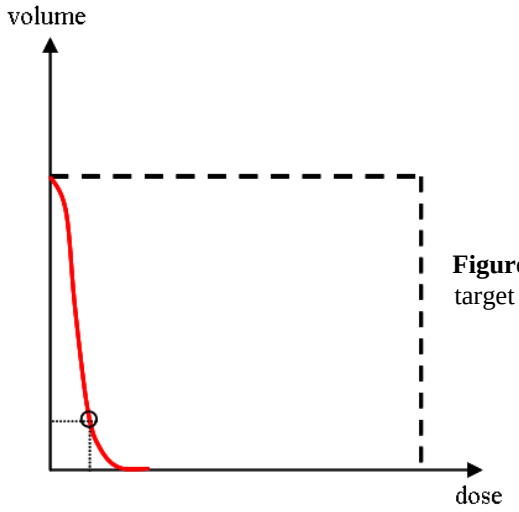


**Figure 4.3** Illustrations of the ideal cumulative DVHs for both target volume (black dashed line) and OAR (red solid line).

All dose distributions were normalized to the mean dose delivered to the target volume excluding the voxels of OARs when overlapping with the target. Individual field dose distributions were systematically analyzed to better understand the detailed characteristics of the different plans.

### *4.2.5 Range uncertainty*

Our main goal was to evaluate the robustness of different IMPT plans as a function of changing starting conditions. To that effect, we have only considered the effects of range uncertainties as they will propagate through the entire treatment course. Consequently, all IMPT plans have been recalculated following the approaches already proposed by Lomax et at (Lomax 2001) and before by Goitein et al (Goitein 1985) assuming a general error of ±3% of all Hounsfield values in patients' CT, a value generally being accepted as realistic for current dose calculations and CT calibration procedures.

In addition to DVH analysis, plan robustness has been analysed with the help of the generalised Equivalent Uniform Dose (gEUD) equation (Niemierko A 1997, Niemierko A 1999):

$$gEUD = \left( \frac{1}{N} \sum_i D_i^a \right)^{\frac{1}{a}} \qquad \text{Eq 4.1}$$

gEUD is a single parameter function, in which the parameter a is varied in order to mimic either parallel structures organs ($a$ =1) or serial organs ($a$ >> 1). The gEUD can conveniently assess both, target tissues (in which case the $a$-parameter is always negative) or critical tissues ($a$ >= 1). For each nominal IMPT plan, and for the corresponding range error plans (±3% HU),

curves of gEUD as a function of the $a$-parameter for both, target and critical structures were calculated. For target volumes, values of $a \in (-1;-20)$ were used (where for large negative $a$ -values the minimum value is approached and $a = -10$ may corresponds to a typical value to evaluate cold spots (Bortfeld et al 2008)) and for critical structures, values of $a \in (1;20)$. In order to highlight differences, the results were presented as the differences in gEUD between the $EUD_a$ curves of the range error plans and the nominal ones.

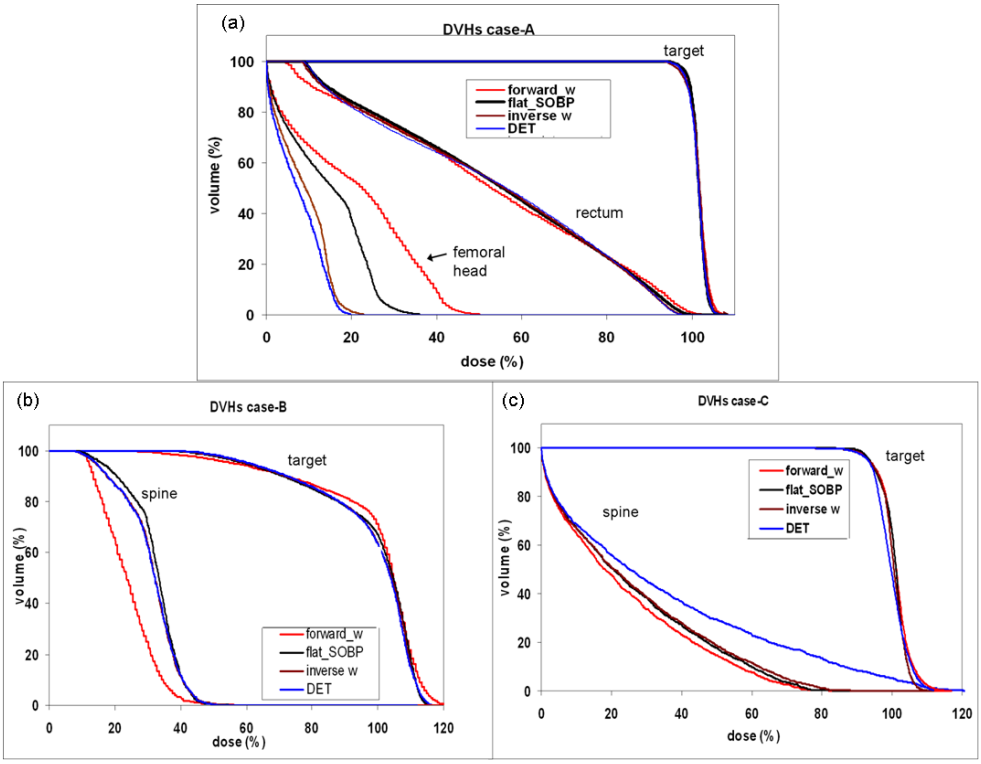


**Figure 4.4** DVHs obtained with the four initial starting conditions are here reported: (a) case-A (prostate carcinoma):DVHs for target, rectum and femoral head;(b) case-B (cervical chondrosarcoma): DVHs for target and for the spine; (c) case-C (thoracic chondrosarcoma): DVHs for target and for the spine

# 4.3. Results

## *4.3.1 Case-A. Prostate carcinoma*

### *4.3.1.1 Comparison of nominal plans.*

The four IMPT plans resulting from the different sets of starting conditions produced similar high dose conformation, clinical target volume (CTV) coverage and rectum sparing (figure 4.4(a), (figure 4.5 (o)-(r)). Hoverer, for the femoral heads the **forward wedge** approach delivered a significantly higher dose (figure 4.4(a)). This is consistent with the results presented

in Chapter 3.2 regarding IMPT planning of prostate cancer with parallel opposed fields. Also in keeping with other work the dose is minimized to the femoral heads using the **DET** approach (Oelfke and Bortfeld 2000-2001) and to a lesser extent by the **inverse wedge** approach.

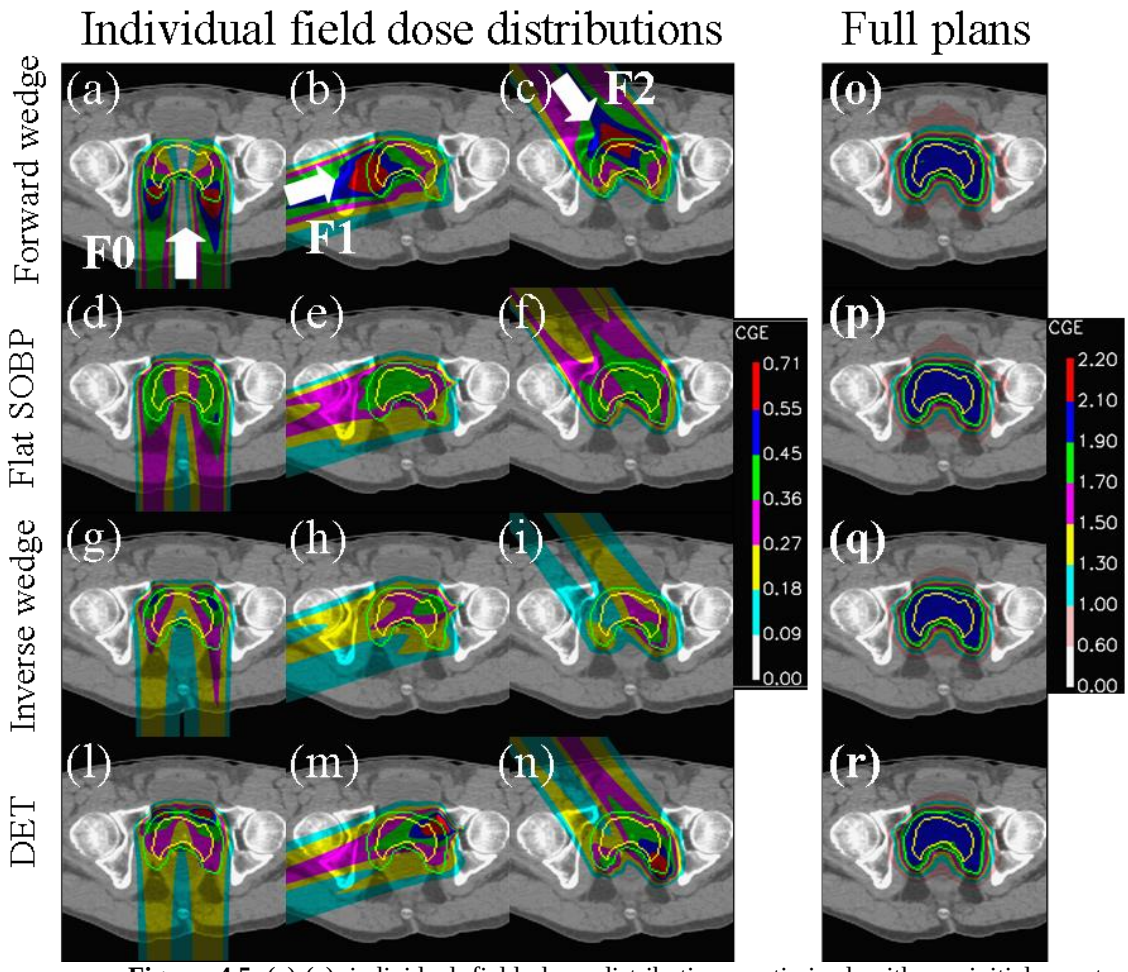


**Figure 4.5** (a)-(c) individual field dose distributions optimized with an initial constant beamlets fluence matrix and (o) resulting plan; (d)-(f) single field dose distributions optimized with an initial pre-weighted set of Bragg peaks such to deliver a flat SOBP and (p)resulting plan; (g)-(i) single field dose distributions optimized with an initial pre-weighted set of Bragg peaks such to deliver a gradient (in the text: inverse wedge) and (q) resulting plan;(l)-(n) single field dose distributions optimized starting from the DET and (r) resulting plan. Note the different dose scale between the single field and the plan dose distributions.

#### *4.3.1.2 Individual field distributions.*

Figure 4.5(a)-(n) depicts results for each starting condition for 3 of 5 of the individual field dose distributions (due to symmetry of the field arrangement and target volume, the missing fields show similar characteristics). The effects of the different starting conditions are now clear with the form of the individual fields showing quite different characteristics depending on the starting conditions, even though the resultant full dose distributions across the target are almost identical (see figures 4.5 (o),(p),(q) and (r)). In particular, for the **forward wedge** approach, the

two lateral-anterior fields (see e.g. figure 4.5(c)) clearly placed highly weighted Bragg peaks against the rectum (note high doses at the end of the field stopping against the rectum). In contrast, for the **SOBP** approach, target coverage was achieved by adding together quasi-homogeneous in-field dose distributions (figure 4.4(d)-(f)). Finally, for the **DET** and **inverse wedge** approaches, the single field dose distributions had the highest doses positioned at the distal end of the target (figure 4.5(g)-(n)) with dose homogeneity across the target being achieved by patching the proximal and lateral edges of each field.

#### *4.3.1.3 Sensitivity to range uncertainties.*

Figure 4.6 shows the different gEUD curves between the plans re-calculated with range errors and the nominal plans. In order to only consider the worse case scenarios (under dosage in the CTV and over dosage in the rectum) only positive differences for the rectum and negative differences for the target have been considered. In addition, only the curves for the range error (+ or – 3%) giving the worst results are shown (i.e. the most negative values for CTV and most positive for rectum). Note, as these are plots of the *difference* between the range error and nominal plans, the most robust plans are those closest to the zero line (no change), whilst those further away are less robust.

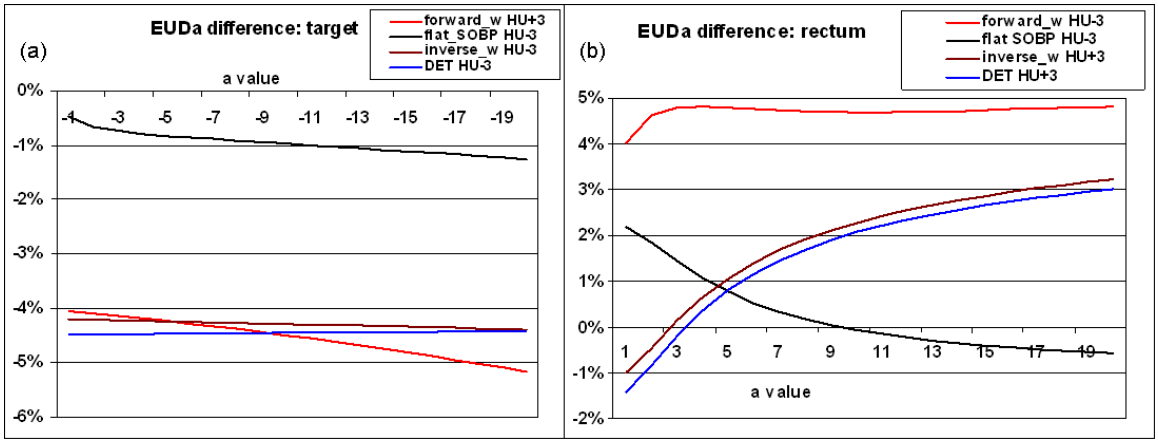


**Figure 4.6** Differences of g $EUD_a$ curves as a function of the a-parameter and depending on the four starting conditions under conditions of range uncertainty for a) target and b) rectum for case-A

It is evident from figure 4.6 that, of all the methods, the **SOBP** approach provided the more robust plan to range errors, as the difference $EUD_a$ curves, for both the CTV and the rectum, were closest to the zero line (little change). In contrast, the **forward wedge** approach was the least robust with $EUD_a$ values for the rectum increasing by up to 0.1 $Gy_{RBE}$ per fraction (assuming to deliver 2$Gy_{RBE}$ per fraction), which corresponded to a general increase of 3.7 $Gy_{RBE}$

for the total treatment (i.e. from 54.4 $Gy_{RBE}$ to 58.1 $Gy_{RBE}$ when considering a value of $a$ equal to 6 as proposed by Wu for the prostate region (Wu et al 2002)). The curves for **DET** and **inverse wedge** approaches were in between the other two plans. Interestingly, for the **forward wedge** approach, the generalised under-dosage of the CTV resulted from a 3% *undershoot* (+3% in HU) of all Bragg peaks, whereas for the **DET** and **inverse wedge** approaches, the under-dosage of the CTV was caused by an *overshoot* (-3% in HU values). The reverse was true for both cases when considering the doses to rectum (increased dose to the rectum for *overshoot* and *undershoot* for the **forward wedge** and **DET/inverse wedge** approaches, respectively).

### *4.3.2 Case-B. Cervical chondrosarcoma.*

#### *4.3.2.1 Analysis of full plans.*

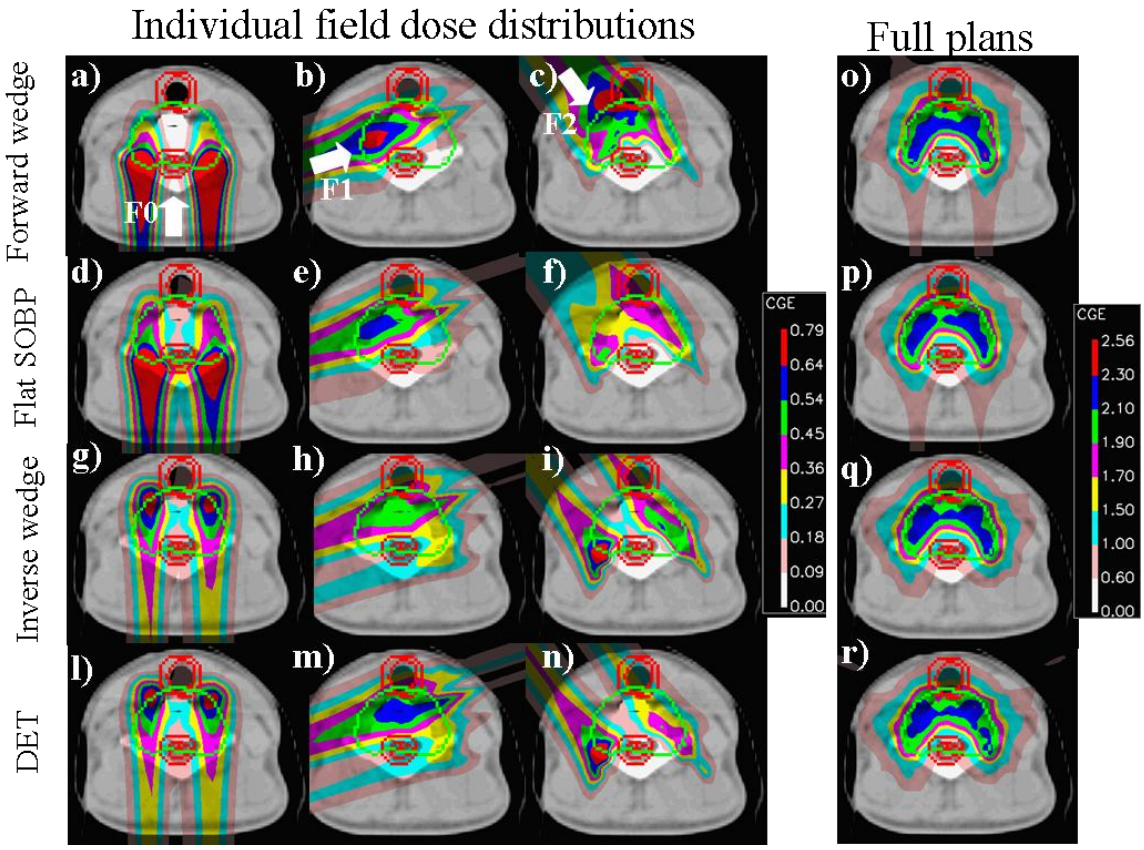


**Figure 4.7** (a)-(c) single field dose distributions optimized with an initial constant beamlets fluence matrix and (o) resulting plan; (d)-(f) single field dose distributions optimized with an initial pre-weighted set of Bragg peaks in order to deliver a flat SOBP and (p)reaulting plan; (g)-(i) single field dose distributions optimized with an initial pre-weighted set of Bragg peaks to deliver a gradient (in the text: inverse wedge approach) and (q) resulting plan; (l)-(n) single field dose distributions optimized starting from the DET and (r) resulting plan

The four plans with representative individual field dose distributions for this case are shown in figure 4.7. Corresponding DVH's are displayed in figure 4.4(b). For all plans the DVH's for the CTV were compromised due to the low constraints assigned to an expansion of the spinal cord (5mm), which overlapped with the planning target volume (PTV).

Several differences can already be detected when analysing the full plans: The **forward wedge** approach reduced the dose to the spinal cord (SC) slightly more (1-3%) compared to the other approaches (Figure 4.4(b)) due to a steeper gradient in front of the spine caused by a 'patch' of the individual fields (see next section). In addition, the percentage of the target volume receiving at least 95% of the prescribed dose (i.e. $V_{95\%}$ ) increased by up to 4% compared to the other plans; the $D_{98\%}$ (our surrogate for minimum dose) was reduced by 11%. This reduction was related to the low dose delivered to the SC which partially overlapped with the PTV volume.

***4.3.2.2. Analysis of the individual field distributions*.**

The individual dose distributions for 3 of the 5 fields of the plans are shown in figure 4.7. Analogous to the first case, the **forward wedge** approach resulted in single fields with a high entrance dose (Figure 4.7(a)-(c)) and high in-field dose gradients around the spinal cord in order to reduce the dose to the SC. The distal fall-offs of the two anterior fields (F2 and F3) were essentially 'patched' to the lateral fall-off's of the remaining three fields. Although for the **SOBP** approach a similar patching effect was observed (Figure 4.7(d)-(f)), the resulting distal fall-offs of the two anterior-lateral fields were smoother compared to the equivalent **forward wedge** approach. In particular, the anterior-lateral fields deposited most of the high weighted Bragg peaks laterally to the spine, rather than directly in front of it. The **inverse gradient** and **DET** approaches provided very similar solutions (figure 4.7 (g)-(n)). For both plans, the spine was spared by patching together the lateral fall-off of the anterior-lateral and posterior fields (F2,F3,F0) with the *proximal-edges* of the two lateral fields (F1,F4). No fields deposited highly weighted Bragg peaks in front of the spinal cord.

***4. 3.2.3. Range uncertainties***

Figure 4.8 shows the difference g $EUD_a$ curves for this case. Since individual fields were patched together around the spine through lateral fall-off and proximal edges, the **DET** or **inverse wedge** approach resulted in IMPT plans that were the *most* robust to range uncertainties both for the CTV and for the SC (difference g $EUD_a$ curves closest to zero in figure 4.8 (a)-(b)). For both approaches, the curves reported here were calculated simulating the worst case scenario (i.e. HU decreased by 3%). Even under those circumstances only a minor under-dosage of the CTV and a

slight increase of the $EUD_a$ value of the SC (up to 0.02Gy$_{RBE}$ per fraction) were detected. On the other hand, the **forward wedge** approach provided a substantially less robust plan, with the g $EUD_a$ value for the SC increasing by about 0.1Gy$_{RBE}$ per fraction (5%), for an $a = 20$.

As would be expected from the individual field discussion in the previous section, the **SOBP** approach produced intermediate results, being generally more robust than the **forward wedge** approach, but less robust than the **DET**/**inverse wedge**. This held true for both the CTV and the spinal cord.

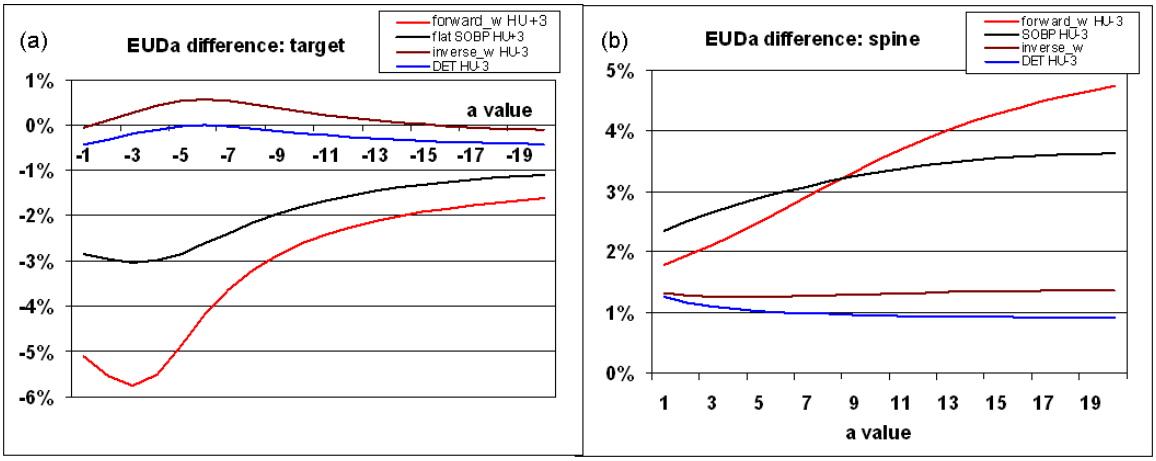


**Figure 4.8**.Differences of g $EUD_a$ curves as a function of the a-parameter and depending on the four starting conditions under conditions of range uncertainty for a) target and b) spine for case-B

### *4.3.3 Case-C. Thoracic chondrosarcoma*

#### *4.3.3.1 Analysis of full plans.*

All plans for case C are shown in figure 4.8, together with the resulting DVH's in figure 4.4(c). Due to the limited angular approach used for this case, the **DET** approach showed slightly compromised PTV coverage compared to the other approaches. It also proved difficult to reduce the dose to the spinal cord (figure 4.4(c)), indicating that this approach, at least when using only 3 fields as in the case here, is probably not a practical method for such treatment. The other approaches resulted in similar coverage and spinal cord sparing.

#### *4.3.3.2 Analysis of the individual field distributions.*

Figures 4.9 (a)-(n) show the individual dose distributions obtained for each field of each approach. Surprisingly, with both, the **forward wedge** and the **SOBP** approaches, no high weighted Bragg peaks were positioned against the SC, with the dose 'hole' around the organ at risk being achieved mainly by the lateral fall-off of the pencil beams. Conversely, for the **DET** and **inverse wedge** approaches, for the more posterior field, the dose distributions had part of the

high doses positioned on the distal end of the target, just in front of the spinal cord (Figure 4.9 (h)-(m)).

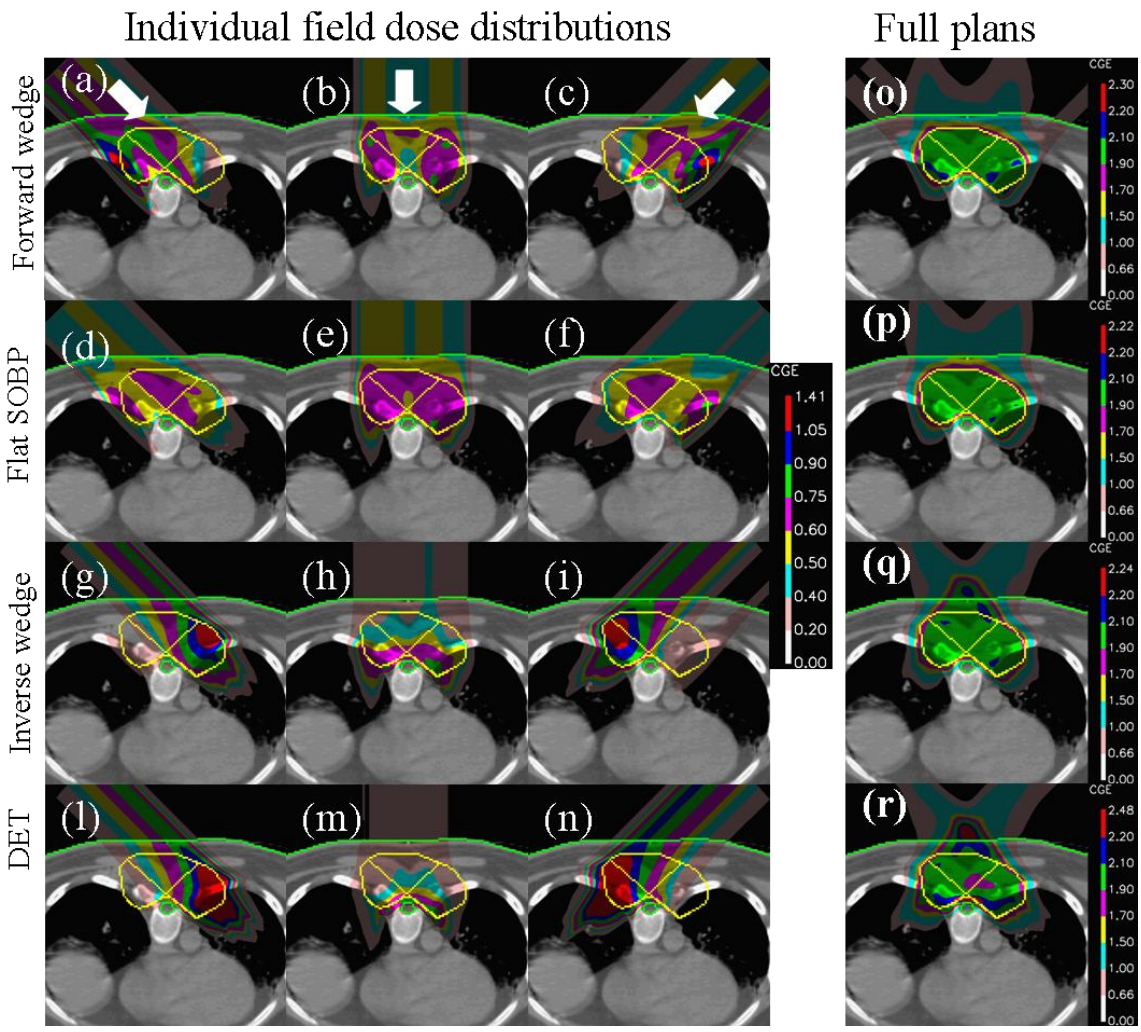


**Figure 4.9** (a)-(c) individual field dose distributions optimized with an initial constant beamlet fluence matrix (in the text forward wedge) and (o) resulting plan; (d)-(f) single field dose distributions optimized with an initial pre-weighted set of Bragg peaks to deliver a flat SOBP and (p) resulting plan; (g)-(i) single field dose distributions optimized with an initial pre-weighted set of Bragg peaks to deliver a gradient (in the text: inverse wedge) and (q) resulting plan; (l)-(n) single field dose distributions optimized starting from the DET and (r) resulting plan. Note the different dose scale between the single field and the plan dose distributions.

#### *4.3.3.3 Range uncertainties.*

Differences in gEUD curves for this case are depicted in figure 4.10. Similar to the other cases, for the spinal cord only positive differences were considered, whereas for the target volume only negative differences are considered.

All fields entered from the posterior aspect of the patient, and the spine was positioned at the distal end of the target. Hence, a range *overshoot* (-3% in the HU) resulted in an increased

dose to the spinal cord for all plans. Conversely, under-dosage of the target resulted from a 3% *undershoot* (+3% in HU) of all Bragg peaks for all the starting conditions, except for the **DET** approach. Using the DET method an *undershoot* of all Bragg peaks caused a small 'contraction' of all Bragg peaks towards the center of the target. This contraction increased the Bragg peaks concentration within the target volume, resulting in an increase in overall dose.

It is evident from figures 4.10 (a)-(b) that the DET approach provided overall the generally more robust plan, as the difference in $EUD_a$ curves, for both, target and spine, were the closest to the zero line (no change). However the nominal plan was considerably worse by use of DET than for the other approaches. On the other hand, the inverse wedge approach also provided a rather robust plan without major compromises to the target volume (see figure 4.10). As seen for other cases, the forward wedge approach was probably the least robust, as, for the target volume, the $EUD_a$ value decreased with respect to the nominal plan. Finally, the SOBP approach provided again results that are somewhere in between.

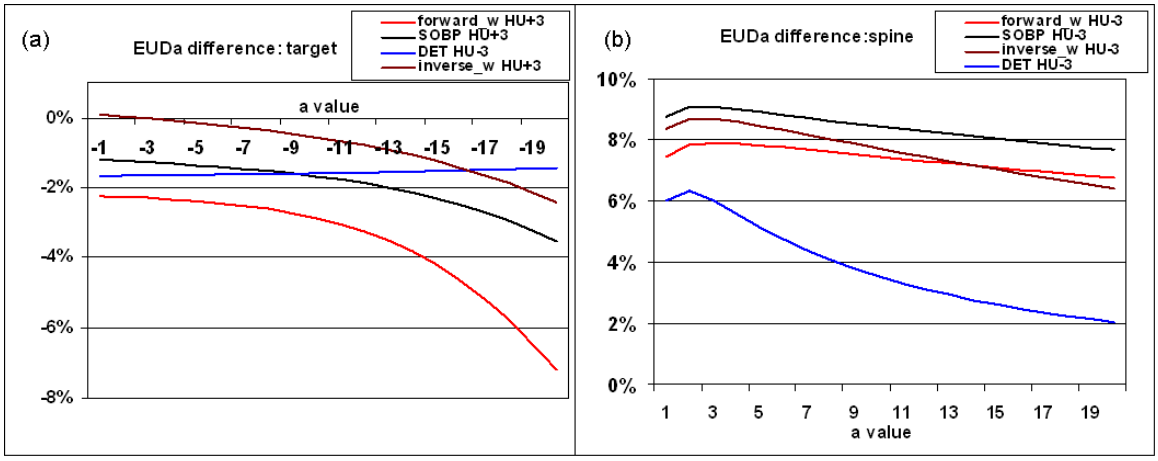


**Figure 4.10.** Differences of g $EUD_a$ curves as a function of the a-parameter and depending on the four starting conditions under conditions of range uncertainty for a) target and b) spine for case-C

## 4.4. Discussion

In this Chapter we analysed how the selection of initial beamlet weights as input into the optimisation algorithm can affect IMPT plans. Endpoints were plan quality and plan robustness to systematic range uncertainties. It has been shown that IMPT plans calculated for different 'flavours' of starting conditions, despite similar target coverage and sparing of critical structures, exhibited important differences when the effect of range uncertainties was considered. In the case of no- or low-dose constraints to neighboring critical structures (e.g. case A), the **SOBP** approach provided the most robust plan (see figure 4.6). These results are in agreement with the single case

study previously published by our group (Lomax 2008a). In the present work we have looked in more detail at the individual fields. Using the **SOBP** approach, the optimized single field distributions are quasi-homogeneous and never place high weighted distal edges against the critical structure (figure 4.5 (d)-(f)). On the other hand, for the more complex situations involving a dose limiting critical structure, partially overlapping with the target volume (case-B), either the **DET** or **inverse wedge** approach provided the most robust plans (see figure 4.8) (this is somewhat in contrast to what has been previously published (Lomax 2008a)). Indeed, as the dose hole around the spinal cord was achieved for these plans by patching together single field lateral fall-off and proximal edges (where the gradients are shallower than for the distal fall-off), range errors resulted only in a minor effect on dose distribution (figure 4.7 (g)-(n)). In contrast, for the **forward wedge** approach, as the dose around the spinal cord is predominantly shaped through the distal fall-off of highly weighted Bragg peaks stopping directly on this organ, any changes in range calculation will significantly affect the plan directly and could lead to severe organ-at-risk over-dosage. However in the nominal plan, by using the Bragg peak's distal fall-off to spare the spine, this organ was better spared with the **forward wedge** approach than by the other approaches (see figure 4.4(b)). Therefore, even in the case of range uncertainties, perhaps the absolute dose would only tend to the dose received in the nominal cases of the other plans. This indicates the inherent necessity of solving the conflict between the achievement of a robust plan (i.e. a plan –more or less- constant even in case of uncertainties) and of a good nominal plan.

In contrast, of the 3-field IMPT plans calculated to the thoracic chordrosarcoma (case-C) which provided clinically acceptable nominal plans, the **inverse wedge** approach provided the most robust plans to range uncertainties. As such, the **inverse wedge** approach could represent a good solution to treat such a case.

As has been previously reported (see e.g. Oelfke and Bortfeld 2000,2001, Lomax 2008a), one of the advantages of the **DET** approach is to reduce the total dose delivered to the patient (see figure 4.4(a)). In addition, we have shown that using an initial **inverse wedge** starting condition is also an appropriate method to reduce the integral dose. On the other hand, as previously reported (Chapter 3.1, Albertini et al 2007), starting the optimization algorithm with the **forward wedge** approach, increases the dose to the non-target normal tissue significantly. Without additional constraints, this approach generated a sub-optimal plan in all cases considered here.

An advantage of the **DET** approach is the reduction in delivery time by reducing the applied Bragg peaks. For example, for the prostate case (case-A) the **SOBP** approach delivered ~3000 Bragg peaks per field, whereas the **DET** approach required only ~300 Bragg peaks per field. However, the **DET** approach (using a limited number of angles) did not provide in all

instances acceptable target coverage while sparing a neighbouring critical structure (case-C). Although the addition of fields to the plan improved the situation it increased both the total number of Bragg peaks and the time necessary to rotate the gantry between fields and consequently the delivery time.

On the other hand, the **inverse wedge** approach provided for all the cases studied a good balance between target coverage, critical structure sparing and robustness to range errors.

For all plans the initial sets of spots available to the optimization algorithm were those within the selected target volume and extending up to 5mm outside of the target surface. Therefore, for cases where the organ at risk was positioned at the distal part of the target in relation to the beam angles, without any constraints, both the **DET** and the **inverse wedge** approaches placed high weighted Bragg peaks directly in the OAR. However, when OAR dose constraints were assigned, these Bragg peaks were reduced in weight or even switched-off completely during the optimization process. This relative simple measure increases the plan robustness, without the need to include range uncertainty into the optimisation.

For this work the dose has been calculated using the analytical approach described in Chapter 2.1. Although it will be interesting to repeat the analysis using a Monte Carlo algorithm, we do not expect significant differences in the results here reported. Previous studies (Tourovsky *et al* 2005 and Lomax 2008a) have indeed confirmed a good agreement (>90% volume points with a dose difference within 5%) between our analytical and the Monte Carlo calculation, even when the DET approach was considered.

It should be noted that due to the presence of a 'damping function' used in the optimization (see Chapter 2.2), the optimization algorithm implemented in our treatment planning was potentially more dependent than others on the choice of the starting condition. As this damping function means that, from one iteration to the next, the degree of change of the weight of a pencil beam is somewhat constrained, initially low pencil beam weights will rarely be increased during the optimization process. We believe that this can be an advantage, as it means that we can 'steer' the results in a transparent manner by manipulating the starting conditions. Indeed, similar results could be found also with other dose-gradient based optimization algorithms. For instance, in a recent paper by Soukup *et al* (2009) similar results to those presented here were found, in that for a single prostate case an **SOBP** approach provided a more robust IMPT plan to organ motion than a **forward-wedge** like approach.

Finally, as already mentioned, the problem of assessing plan robustness in case of uncertainties is of major importance when treating with protons. Consequently, some authors have suggested to deal with the problem by incorporating range uncertainties directly at the level

of the optimization algorithm (see Unkelbach 2007 and 2009 and Pflugfelder 2008). This approach can perhaps elegantly solve the problem, given that no unexpected uncertainties happen during the treatment course (e.g. change in weight bigger than expected). However, currently, one of the biggest limits of this approach is the significant increase of the computational time. In this work we have tried to find a more pragmatic approach to solve the problem, whereby after having defined a "class solution" (i.e. a set of starting conditions consisting of a given beam configuration and initial beamlet weights) that is specific for different tumour geometries, results could be steered to a more robust plan. Nevertheless, the two approaches are no doubt complementary to each other and we also believe that the results presented here could be helpful in understanding the problem of plan robustness such as help the optimisation process in robust treatment planning.

## 4.5. Summary

In this Chapter, we have studied in detail the characteristics of IMPT plans calculated using four different starting conditions, and have looked at the relative robustness of the resulting dose distributions to possible range uncertainties. In the case of non-overlapping, dose limiting OAR's, the **SOBP** approach provides the most robust plan, as it delivers quasi-homogeneous single field dose distributions. In contrast, for the case of overlapping targets and critical structures, then both the **DET** and **inverse wedge** approaches provide the most robust plans. However, as the **DET** approach does not always provide a clinically deliverable solution, we suggest that the **inverse wedge** approach represents a good solution, for both plan robustness and quality. Finally, we believe that the **forward wedge** approach (i.e. all Bragg peaks set to the same initial weights) should be avoided.

However, it is hopefully clear from the analysis in this chapter that robustness could be an important characteristic of IMPT treatments, and should be evaluated when comparing different potential treatment plans for delivery to the patient. Hence, it would be helpful to have a tool to properly analyse and easily evaluate plan robustness. Such an apporach will be presented in the next chapter.

---

Note: This Chapter has been published as part of the following paper: Albertini F, Hug EB and Lomax AJ 2010 *The influence of the optimization starting conditions on the robustness of intensity-modulated proton therapy plans* in Phys. Med. Biol (55) 2863-2878 (Albertini et al 2010)

# 5 Evaluating uncertainties: the use of Error-bar distributions to assess plan robustness

## 5.1. Introduction

Radiation treatment planning is considered to be robust if the calculated and the delivered dose are in good agreement even in the case of different uncertainties. In order to accomplish that, the current practice in photon radiotherapy is to use safety margins to expand the clinical target volume (CTV) sufficiently enough to account for treatment uncertainties. This practice, however, may not necessarily be the ideal solution when treating with active protons and in particular with Intensity Modulated Proton Therapy (IMPT). IMPT plans can deliver highly conformal dose distributions to the target volume, while sparing embedded critical structures. This is achieved by patching together complex, individually in-homogeneous dose distribution which, in case of errors can lead to degradations in the planned dose not only at the target boundary (which can generally be dealt with through the use of a safety margin, van Herk et al 2000) but also in the middle of the target as a result of misalignments of these high in-field dose gradients. Indeed, the precision of proton radiotherapy strongly depends on both spatial misalignments of the patient and the accuracy of the position of the Bragg peak. Although range uncertainties likely represent the most critical source of errors in case of proton therapy - as these are generally systematic and therefore will tend to propagate through the whole course of the treatment (Lomax 2008a) – spatial errors can also not be ignored, especially when the IMPT fields are highly modulated (i.e. have large in-field gradients).

It was suggested more than 20 years ago (Goitein 1985) that error-bars should be associated to any calculated dose distribution as "*at every point within the patient there is in fact a range of possible doses that may be delivered*". Nevertheless, although some work has been done to evaluate the effect of range and set-up errors on the proton dose distribution and in

particular on IMPT plans (see for example Lomax et al 2001, Unkelbach et al 2007, Ellerbrock et al 2007, Pflugfelder et al 2008, Unkelbach et al 2009, Lomax 2008a-b, Albertini et al 2008, Albertini et al 2010), treatment plans are still routinely evaluated generally assuming no-uncertainties in the delivery. Different methods have been proposed so far, such as the worst case analysis introduced by Goitein (1985). However, possibly due to the problems of interpreting unphysical dose distributions, their use has not yet been widely spread.

In this Chapter, following the work of Lomax (2004) and Negreanu et al (2008), we propose a simple way to analyze and easily display dose variations due to different delivery errors. With this tool, comparisons of plan robustness have been assessed for different SFUD and IMPT plans to both assess the usefulness of the PTV concept for these delivery methods, and to evaluate the general robustness of such plans. Finally, we show that plan robustness could be predicted by the simple expedient of calculating, combining and directly displaying dose gradients of the individual fields of IMPT plans.

## 5.2. Material and Methods

Three clinical cases have each been planned using SFUD (Single Field Uniform Dose) and IMPT (Intensity Modulated Proton Therapy) and their robustness to set-up and range errors analysed. The cases include tumors in the skull-base, along the spinal axis and in the head and neck (H&N) region. The cases have been selected in order to 1) introduce methods for the analysis of plan robustness and then to use these approaches to 2) to assess the advantage (or otherwise) of the use of Planning Target Volume (PTV) concepts for SFUD and IMPT plans and 3) to evaluate the potential variation of dose in organs at risk due to set-up and range errors.

As already discussed in Chapter 1, with SFUD plans we refer to plans which are constructed by adding together fields which have been individually optimized to each deliver a homogeneous dose through the target volume (Figure 5.1a). In contrast, IMPT plans are calculated through the simultaneous optimization of all Bragg peaks of two or more field, which only when combined provide a homogeneous dose through the target and, additionally, allow for the selective sparing of specific organs-at-risk (OAR's) (Lomax 1999, Lomax 2008c). If the IMPT plan is optimized with stringent dose-volume constraints, the individual fields can exhibit an in-homogeneous dose distribution resulting in high in-field dose gradients (Figure 5.1b). In this work we refer to such plans as **High gradient IMPT** plans. On the other hand, when no (or relaxed) planning dose constraints are used the resulting IMPT plan is composed by single fields with a quasi-homogeneous dose distribution (**Low gradient IMPT**).

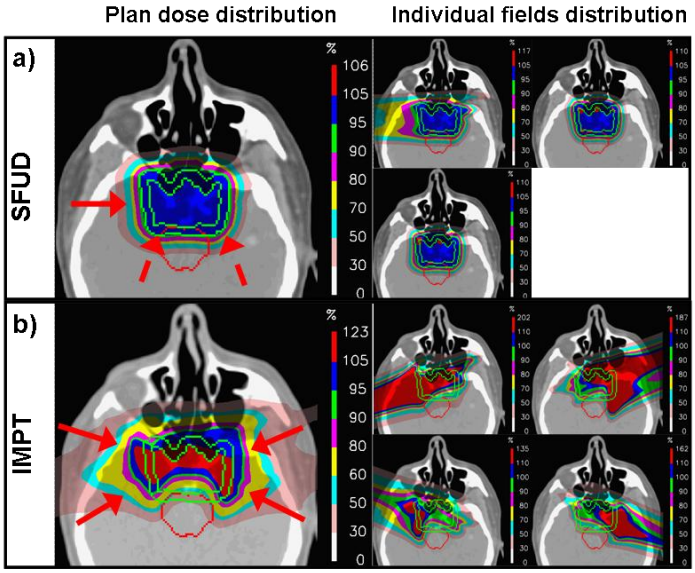


**Figure 5.1**: (a)Dose distribution for a 3-field SFUD plan and relatively homogeneous individual fields dose distribution; (b) dose distribution for a 4-fields IMPT plan and relatively highly in-homogeneous individual field dose distribution.Green- lines represent the CTV (innermost) and the PTV (outermost). Red arrows represent the beam arrangements

All plans have been investigated from the point of view of the efficacy of safety margins for planning, and to also assess the robustness of OAR doses for the different plan types to both set-up and range errors. Plans have been calculated using the PSI treatment planning system and our clinically used optimization algorithm (for details, see Chapter 2). Margins used in this study were those clinically applied: 5mm for the skull base and head & neck region and 7-8mm extra-cranially. All plans have been normalized to the mean dose delivered to the planning target volume (PTV), minus dose limiting organ-at-risk (OARs) volumes when these overlap with the target.

### *5.2.1 Modeling set-up errors*

Set-up errors have been modeled by simply recalculating the planned dose distribution in a number of spatially shifted versions of the patient's CT. Typically, these consist of scalar shifts of the CT in relation to the treatment isocenter along each of the major anatomical axes (i.e. the left-right (LR), anterior-posterior (AP) and cranio-caudal (CC) directions), in both the positive and negative directions. As a result of these calculations, a number of spatially shifted 3D dose distributions are generated for each plan. The magnitudes of the applied shifts have been calculated based on positioning data reported by Bolsi (Bolsi et al 2008) for our institute. In that work, systematic set-up errors were found to be below 0.6mm for all patients, whereas random errors depended on the tumor location (i.e. skull-base, H&N and extra-cranial cases) and on the

| | $\sigma_{LR*}$ | $\sigma_{AP*}$ | $\sigma_{CC*}$ | $\sigma_{tot}$ | $\sigma_{85\%}$ |
|---|---|---|---|---|---|
| Head Bite-Block | 1.3mm | 1.2mm | 1.3mm | 2.2mm | 3.2mm |
| H&N Mask | 2.0mm | 1.8mm | 2.0mm | 3.3mm | 4.8mm |
| T8-S5 mask | 2.1mm | 2.4mm | 2.0mm | 3.8mm | 5.5mm |

Abbreviations: LR, left-right; AP, anterior-posterior; CC, cranio-caudal.

**Table 5.1**: residual set-up errors (i.e. random errors) after the daily patient positioning correction for different fixation devices and region of treatment, along the three axes, $\sigma_{LR*,AP*,CC*}$ (Bolsi et al 2008), $\sigma_{tot}$ is the combined error and $\sigma_{85\%}$ is the corresponding 85% confidence limit.

fixation device (e.g. bite block or thermoplastic mask). For the set-up analysis, we have therefore neglected the systematic error and have considered only the random errors along each of the cardinal axes of the patient ($\sigma_{LR}$, $\sigma_{AP}$, $\sigma_{CC}$). The magnitudes of these errors are reported in Table 5.1. With the assumption that residual rotational errors are negligible, the combined effect (if the errors are uncorrelated) is described by a sphere of radius $\sigma_{tot}$ (see Table 5.1), where $\sigma_{tot}$ is the sum of the individual axial errors summed in quadrature, for each fixation device and tumour location. From this value, different iso-probability shifts can be calculated based on the confidence level desired.

In accordance with the recommendations by Goitein (1983), the actual shifts applied to the CT data sets have been calculated to satisfy the 85% confidence level (C.I.) based on the sigma values reported in table 5.1 ($\sigma_{85\%}$). This confidence limit corresponds to spatial shifts of almost 1.5 the combined sigma of the positioning error, which, as argued by Goitein (1983), is a reasonable parameter to use for clinical purposes. For each tumour location and fixation device a sample (n=14) of all possible shifts with an 85% C.I. along the 3 major axes have been calculated and, correspondingly, fourteen spatially shifted 3D dose distributions have been calculated (Table 5.2) per plan.

| Head Bite-Block | Head & Neck Mask | Extra-cranial mask |
|---|---|---|
| $\begin{pmatrix} \pm 3.2;0;0 \\ 0;\pm 3.2;0 \\ 0;0;\pm 3.2 \end{pmatrix} + (\pm 1.8;\pm 1.8;\pm 1.8)$ | $\begin{pmatrix} \pm 4.8;0;0 \\ 0;\pm 4.8;0 \\ 0;0;\pm 4.8 \end{pmatrix} + (\pm 2.8;\pm 2.8;\pm 2.8)$ | $\begin{pmatrix} \pm 5.5;0;0 \\ 0;\pm 5.5;0 \\ 0;0;\pm 5.5 \end{pmatrix} + (\pm 3.2;\pm 3.2;\pm 3.2)$ |

**Table 5.2**: Shifts (in mm) along the 3 major axes satisfying the 85% C.I. applied for the different fixation devices and tumour location

#### *5.2.1.1 Error-Bar dose distribution/ Error-Bar volume histograms*

In order to succinctly represent this data in the form of a single distribution, for each nominal plan, all the individual spatial 3D 'error' dose distributions can be combined into one unique 'error-bar' distribution representing the spread of dose values at each point at the 85% confidence limit thus:

$$\Delta D^i{}_{random-error} = Max\left(D^i_{+x}, D^i_{-x},..., D^i_{-z}\right) - Min\left(D^i_{+x}, D^i_{-x},..., D^i_{-z}\right),$$

where $\Delta D^i{}_{random-error}$ is the value at the dose grid point i in the resulting error distribution and $D^i_{+x}$,is the dose at point i for the recalculated plan shifted in the positive x- direction etc. Given that the result of this calculation is a three dimensional distribution of dose variation at each dose calculation point in the nominal dose distribution, cumulative Error-Bar Volume Histograms (**EVH**) can also be extracted for each volume of interest in an analogous way to the calculation of dose volume histograms (DVH's). These histograms represent a useful and easy tool to evaluate the quality of the treatment: the closer the histogram is to the '0-error' line, the more robust (i.e. insensitive) is the dose in the selected volume to set-up errors. All values in such plots are represented as percentage differences relative to the prescribed dose of the nominal plan (e.g. 5% corresponds to a difference of 0.1 $Gy_{RBE,}$ (with a radio biological effectiveness (RBE) of 1.1) for a plan delivering 2 $Gy_{RBE}$ /fraction).

### *5.2.2 Range errors*

Range uncertainties can be considered the most important source of error for proton therapy. As the dose is usually calculated based on a single planning-CT ('snap-shot'), any errors in the CT data itself, or in the transformation from CT Hounsfield units (HU) to relative proton stopping power (see e.g. Schneider et al 1996, Schaffner et al 1998), are likely to be systematic and propagate through the whole course of the treatment.

As previously suggested (e.g. Lomax et al 2001) range errors can be modeled by the simple expedient of recalculating each plan with an error of ±3% on the nominal HU values in patients' CT. Due to the lack of data supporting a defined statistical value for range errors, the corresponding **Range Error-Bar dose distribution** is therefore calculated as

$$\Delta D^i{}_{systematic\_error} = Max\left(D^i_{+HU}, D^i_{-HU}\right) - Min\left(D^i_{+HU}, D^i_{-HU}\right)$$

and does not have a corresponding statistical interval of validity. Nevertheless, regarding all the possible source of range uncertainties, an uncertainty of at least ±3% error in the HU values can be considered a reasonable value (see e.g. Moyers et al 2010).

### *5.2.3 Composite error-bar dose distribution*

From the two approaches outlined in sections 2.2 and 2.3, for each plan, there will be two distinct error-bar distributions, one for the (random) set-up errors, and one for the (systematic) range errors. For convenience, it is useful to collapse these two distributions into a single, composite error-bar distribution. Here we have adapted the 'margin recipe formula' suggested by van Herk at al (2000) in the following way:

$$\Delta D^{i}_{composite-error} = 2.5\Delta D^{i}_{systematic-error} + 0.7\Delta D^{1}_{random-error}$$

This equation has been originally derived to calculate the PTV margin necessary to ensure a 95% dose coverage of the CTV volume, with a confidence level of 90% and assuming a standard deviation of 3.2mm for the beam lateral penumbra. We have here revisited it to combine together random and systematic errors.

### *5.2.4 Gradient dose distribution*

As the calculation of the errors distributions described above involve a number of dose calculations under various error conditions, and each of these calculations can potentially be rather time consuming, we have also investigated whether a simple surrogate can be defined that predicts such error-bar distributions without performing a full dose calculation-based error analysis. As we believe that the error-bars resulting from set-up and range errors are likely to be correlated to dose gradients within the individual fields of the plan, we have looked into the application of a Sobel filter (Gonzales and Woods 1992) to the individual field dose distributions of a plan. The results of this filtering process for each field are then added together to make a composite 'gradient' image of the plan. The Sobel operator is well known in image processing applications as a simple and fast edge detector which essentially calculates the local gradient of an image or distribution at each point and along each axis.

## 5.3. Results

### *5.3.1 Robustness of the target volume: effect of planning with safety margins*

In order to investigate the effect of using safety margins (PTV based planning) on SFUD and IMPT plans, we have calculated for a standard skull-base case (case-A), 4 different plans (Figure 5.2(a)-(d)):

(i) a 3-field SFUD plan delivering a homogeneous dose to the PTV (**SFUD**);

(ii) a 3-field IMPT plan delivering a homogeneous dose to the PTV (**low gradient IMPT**);

(iii) a 4-field IMPT plan, optimized to satisfy stringent dose constraints on the brainstem (BS) and on the optical structures (**high gradient IMPT**);

(iv) the same 4-field high gradient IMPT as in (iii) but calculated without safety margins, i.e. directly on the CTV volume (**high gradient IMPT-CTV)**.

***5.3.1.1 Set-up errors***

Figure 5.2(e)-(h) show the Error-Bar dose distributions resulting from set-up uncertainties (with a C.I. of 85%), as described in section 5.2.1. From these figures, it is clear that for both the **SFUD** plan and for the **low gradient IMPT** plans, the use of a safety margin is adequate to deal with set-up errors, as the dose inside the CTV volume is relatively robust: only 2% of the volume has an error bar of 10% or more (roughly ±5% of the nominal dose) and of 16% or more (roughly ±8%) respectively for these plans. Indeed, it is clear that the largest error-bars are generally in the CTV-PTV margin, thus indicating that the use of a PTV helps guarantee appropriate target coverage.

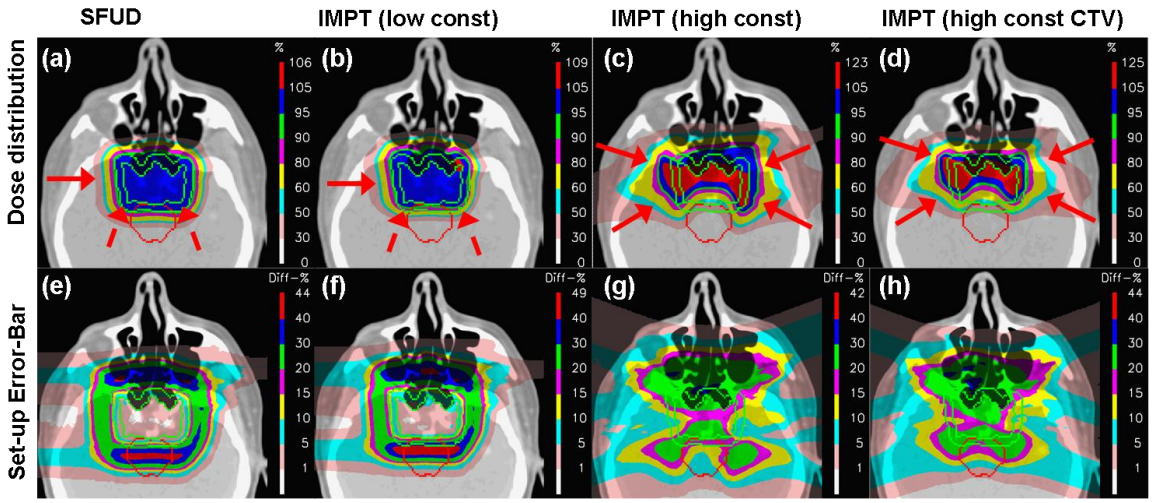


**Figure 5.2**: (a)-(d) Dose distributions respectively for a 3-field SFUD plan; a 3-field IMPT plan with no constraints on organs-at-risk; a 4-field IMPT plan with high constraints on the brainstem (red-line) and optics structure; a 4-field IMPT plan with high constraints on the OARs optimized on the CTV; (e)-(h) corresponding error bar for set-up uncertainties with an 85% confidence limit. Green- lines represent the CTV (innermost) and the PTV (outermost). Red arrows represent the beam arrangements

In contrast, for the **high gradient IMPT** plan, no difference in robustness is discernable when planning on the PTV or directly on the CTV. Indeed, as is clear in figures 5.2(g) and (h), the regions with the highest uncertainty (pink and green in the figure) are situated within the target volume and not necessarily at the target edges. Thus, although the use of a safety margin somewhat mitigates positional uncertainties, there still remain significant error-bars in the middle

of the target volume (including the GTV), for which the PTV expansion has absolutely no effect. Thus, simple positioning errors alone can lead to possible dose in-homogeneities within the irradiated volume, and not just to shifts of the whole dose distribution in relation to the target volume.

This is even more graphically seen when analyzing the Error Bar Volume Histograms (EVH) of the PTV and CTV in Figure 5.3(a), which shows the EVH's for all plans and for both volumes. For both the **SFUD** plan and **low gradient IMPT** plans, the dose to the CTV (broken lines) is considerably more robust than those to the PTV (solid line), as the EVH's are shifted towards the lower left corner of the plot, indicating that the use of a PTV margin is generally sufficient to ensure good coverage of the CTV under all simulated positional errors. In contrast, the histograms of the PTV and of the CTV are almost identical for the **high gradient IMPT** plan (pink lines in figure 5.3a), indicating that little or no advantage results from the use of a PTV. Moreover, there is also little difference when comparing the EVH of the CTV of the **low gradient IMPT** plan calculated on the PTV with that of the **high gradient IMPT** plan calculated directly on the CTV volume (Figure 5.4a).

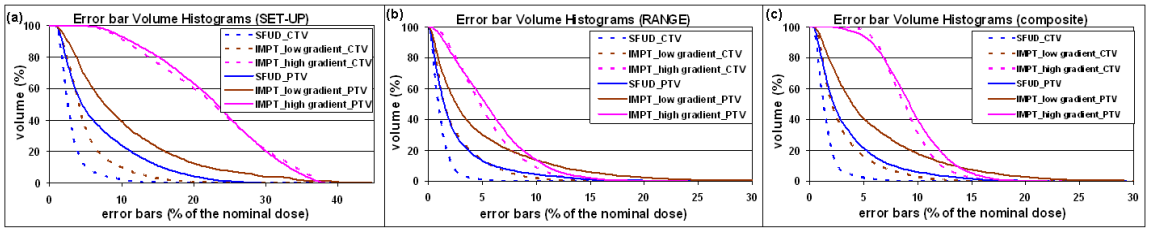


**Figure 5.3**: Error-bars-Volume-Histograms for both the PTV (solid lines) and the CTV (broken lines) for (a) set-up uncertainties and (b) range uncertainties alone and (c) for the combination of the two uncertainties

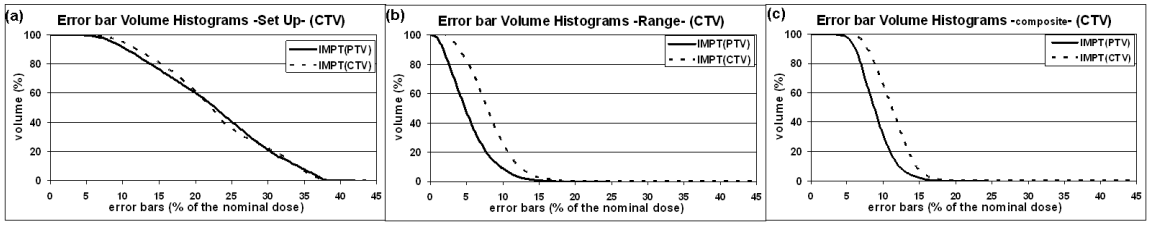


**Figure 5.4:** Error-bars-Volume-Histograms for the CTV volume for the 4-field **IMPT high gradient** plans optimized to the PTV (solid line) and directly to the CTV (dashed line) for (a) set-up uncertainties and (b) range uncertainties alone and (c) for the combination of the two uncertainties

### *5.3.1.2 Range errors*

Figure 5.5(a)-(d) and Figure 5.3(b) show the corresponding Error-Bar distributions and the corresponding histograms for all plans as a result of range uncertainties alone. As seen for positional errors, both the **SFUD** and **low gradient IMPT** plans are robust to range errors, with

less than 2% of the CTV volume having error bars of 4% and 10% or more for the two plans respectively (equivalent to about ±2% and ± 5% of the prescribed dose) (Figure 5.3(b)).

Moreover, from Figure 5.5(a)-(b) and Figure 5.3(b) it is evident that the use of a PTV margin is again adequate to deal with the range error, as the error bar values are largest at the target edges. In contrast, as the **high gradient IMPT** plan (Figure 5.5(c)) is constructed through the patching of fields with steep dose gradients in the middle of the target, the use of a safety margin only marginally improves plan robustness, as is clear also from the EVHs for the PTV and CTV shown in Figure 5.3(b). However, when the IMPT plan is calculated without the use of safety margins (Figure 5.5(d) and Figure 5.4(b)) the areas of larger dose variations increase and collapse towards the middle of the target. Thus, the EVH of the CTV obtained when planning using safety margins indicates a more robust plan (solid line) than that achieved when planning without safety margins (broken line).

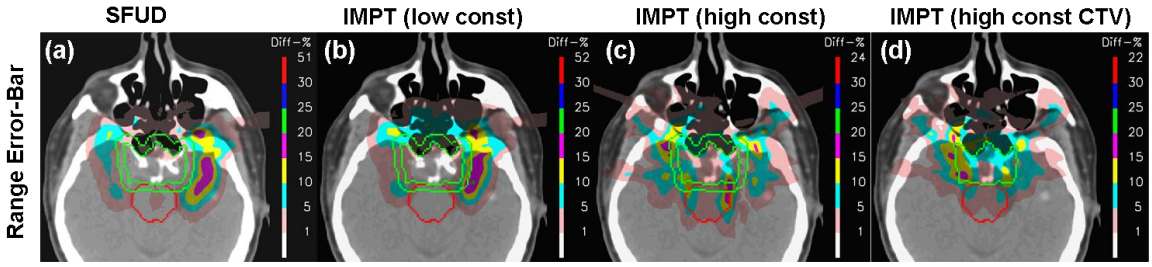


**Figure 5.5**:(a)-(d) Error bar distributions for range uncertainties; Green- lines represent the CTV (innermost) and the PTV (outermost). Red arrows represent the beam arrangements

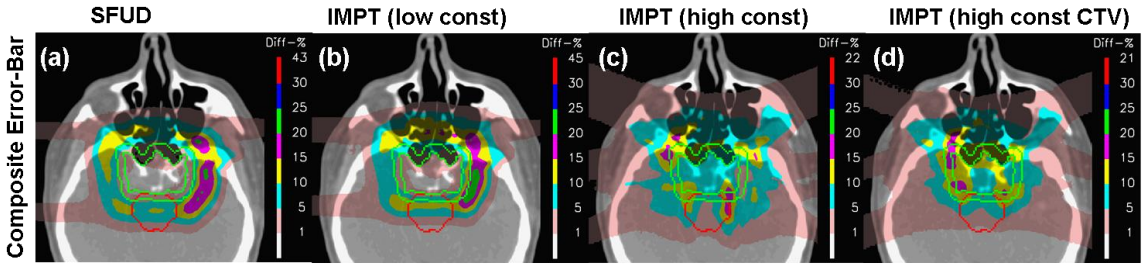


**Figure 5.6**: (a)-(b) Composite error-bar distributions for the 4 plans; Green- lines represent the CTV (innermost) and the PTV (outermost). Red arrows represent the beam arrangements

### *5.3.1.3 Composite errors*

Finally, figures 5.6(a)-(d) and Figure 5.3(c) show the composite error distributions and the histograms for all plans, calculated as described in section 2.3. The combination of set-up (i.e. random component) and range (systematic) errors with an appropriate weighting factor provides a succinct and informative method for visualising the robustness of plans. As expected, it is clear that set-up errors play a less important role in the composite error distribution than range uncertainties (see the similarity of the error distributions shape of Figures 5.6 and 5.5). In the composite calculation, for the **SFUD** plan and the **low gradient IMPT** plans, less than 5% of the

CTV volume has error bars of more than 4% and 9% respectively, and of more than 13% and 15% for the **high gradient IMPT** plans, respectively with and without a PTV (Figure 5.3(c)), again indicating substantial differences in plan robustness between the different plans.

### *5.3.2 Organ-at-risk analysis*

Now we concentrate our attention on the analysis of the robustness of dose in organs at risk that have been spared in the optimization process. In particular, the results for the following cases will be analysed:

(i) case-A: the skull-base cases discussed in section 5.3.1 for which we analyze an **SFUD** plan, a **low gradient IMPT** and a **high gradient IMPT** plan;

(ii) case-B: a skull-base case for which we will compare the robustness for the optical chiasm of two **high gradient IMPT** plans calculated using different beam angles approach;

(iii) case-C: a patient planned with an **high gradient IMPT** plan delivering a 'dose hole' around the cervical spine, similar to the case discussed by Lomax et al (2001) and more recently by authors evaluating the robustness of IMPT plans in case of both spatial and range uncertainties (Unkelbach et al 2007, Pflugfelder et al 2008, Unkelbach et al 2009, Albertini et al 2010);

#### *5.3.2.1 Case-A: SFUD vs IMPT plans*

Figure 5.2, Figure 5.5 and Figure 5.6 show the robustness to set-up and range errors of all plans for the relevant OARs. Due to the geometry of the target and the beam arrangement selected, both the **SFUD** plan and the **low gradient IMPT** are very sensitive to any anterior-posterior misalignment of the patient, as seen in the brainstem volume (red contour in Figure 5.2), which has an error bar above 40%. However, it should be said, that both plans were calculated with the goal of delivering a uniform dose to the PTV only which encompasses part of the brainstem. Indeed, the high gradient penumbra around the PTV is directly in the middle of this structure, leading to the large dose variance in this region. However, both plans are quite robust in the case of range errors, as less than 5% of the brainstem volume has an error bar of 10% or more. In contrast, the dose to the brainstem is more robust for the **high gradient IMPT** plan, as only 5% of the volume has an error of more than 25% when spatial errors are considered alone and of more than 13% when both uncertainties are combined. This indicates that critical structures spared during the optimization process are rather robust to uncertainties.

### *5.3.2.2 Case-B: effect of beam angle selection*

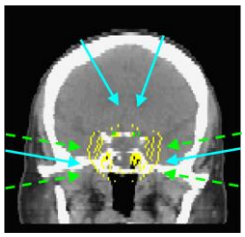

**Figure 5.7:** An example of the skull base chordoma used to analyse the effect of field arrangments. The yellow contours represent the target volumes, the green contour represents the optical chiasm. The different beam arrangments are as shown in the figure: the green fields are used to generate the 4 fields quasi-parallel to the plane including the optical chiasm; 2 (out of 4) of the blue fields are crossing the chiasm.

Figure 5.8 shows the dose distribution of two, 4-field **high gradient IMPT** plans optimized to spare the optical chiasm and calculated with different beam configurations. In particular, the first plan is calculated with all fields being quasi co-planar, and parallel to the chiasm volume. In contrast, the second plan has two (out of 4) fields coming from the cranial aspect of the patient (see Figure 5.7). Error bar distributions for both spatial and range uncertainties are presented (Figure 5.8(c)-(h)).

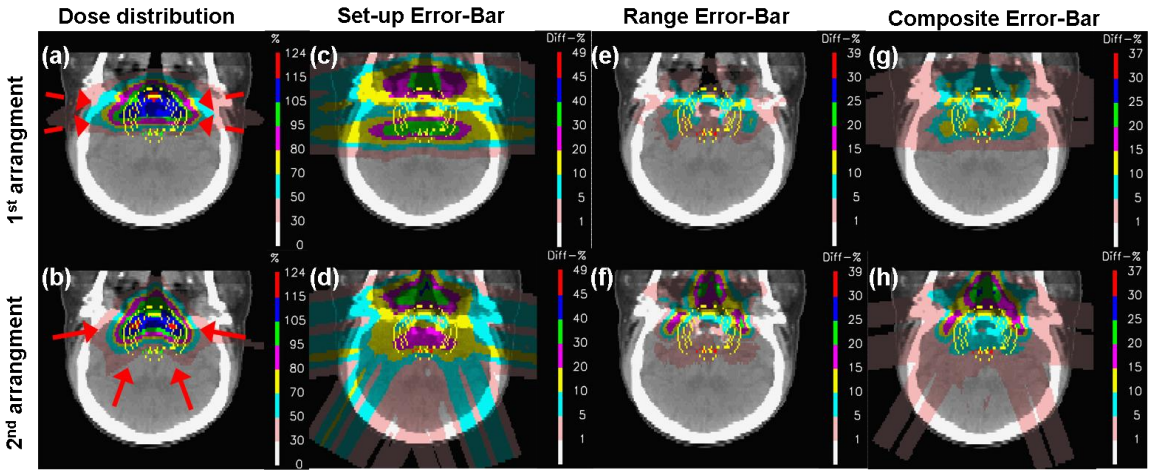


**Figure 5.8:** (a)-(b) Dose distributions for the two IMPT high gradient plan with the different beam arrangements, as shown by red arrows; (c)-(d) corresponding error bar for set-up uncertainties with an 85% confidence limit; (e)-(f) corresponding error bar for range uncertainties; (g)-(h) corresponding composite error bar distribution

When the fields are all in a plane parallel to the structure that is to be spared (e.g. chiasm), then the corresponding IMPT plan is highly sensitive (i.e. not robust) to patient misalignments. The reason for this great uncertainty lies in the very low tolerance dose assigned to the chiasm. This organ totally overlaps with the CTV volume in the most cranial three slices. To reduce the dose to this organ, a large gradient along the *z*-axis (the cranio-caudal axis of the patient) must be generated within the CTV. As all fields are incident in the plane orthogonal to this gradient, the dominant error vectors for *all* fields are in this *z*-direction. That is, a shift of the

patient along this axis can be extremely problematic. In contrast, for the second beam arrangement, when 2 out of 4 fields cross the chiasm volume, the dose gradient in the *z*-direction is reduced and consequently the plan robustness to spatial misalignment is improved (Figure 5.6(d)). Thus the volume of the optical chiasm with an error bar above 5% is reduced from 33% to 16%. Further, when the positional and range errors are combined, the advantage of using the two vertical field approaches is still evident, as the percentage of the chiasm volume with a composite error of more than 5% is reduced from 11% to only 5%. However, as with the non-coplanar arrangement the dose around the critical structure is shaped by combining proximal with lateral fall-offs, the quality of the plan is slightly compromised, thus the $V_{95\%}$ dose to the CTV volume is reduced by 6%.

#### *5.3.2.3 Case-C: robustness of the 'dose hole'*

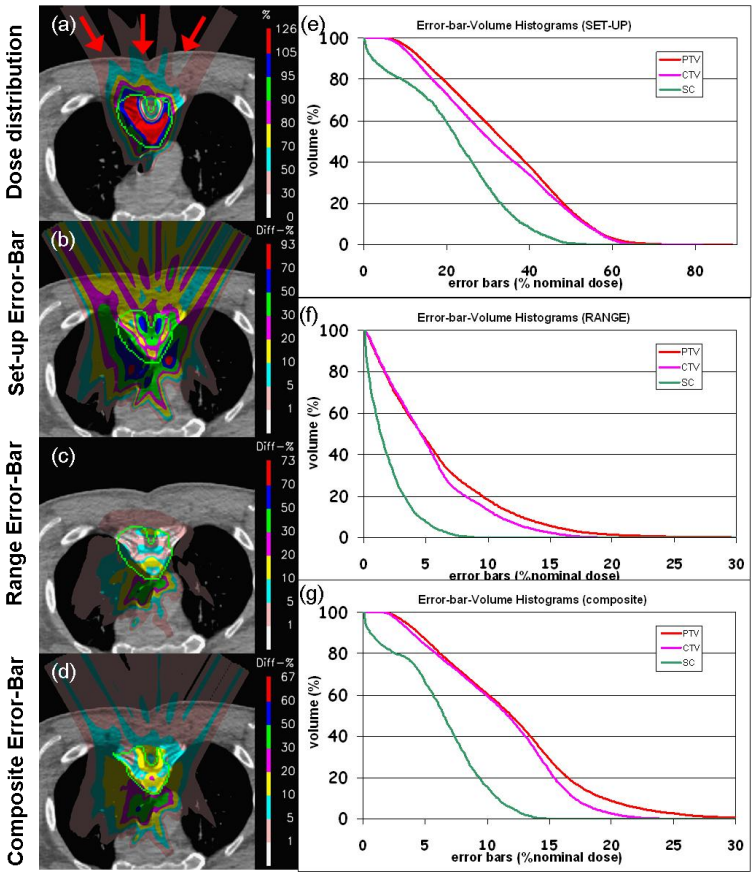


**Figure 5.9:**case-C. Dose distributions for a 3-field **high gradient IMPT** plan (a) and corresponding error bar for set-up (b) for range uncertainties (c) and for the composite uncertainties.Green- line represents the PTV, and red arrows show the beam arrangements. (e)-(f) EVHs for the CTV, PTV and for the spinal cord (SC) volumes in case of set-up and range uncertainties respectively and (g) for the composite uncertainties.

Figure 5.9 shows the dose distribution for a 3-field IMPT plan (**high gradient IMPT**) optimized with a rather low dose constraint to the spinal cord (SC), together with error-bar distributions for set-up, range and combined errors (Figure 5.9(b)-(d)). Due to the low doses prescribed to the spinal cord, the resultant plan has high in-field dose gradients between the fields within the target volume. Therefore, similarly to the **high gradient IMPT** plans previously discussed (case-A and case-B), the EVHs for the PTV and CTV for this case overlap (red and pink line in Figure 5.9(e)-(f)) indicating that there may be no advantage to planning using safety margins. Surprisingly however, the dose to the spinal cord, which was optimized to receive at most 50% of the prescribed dose, is extremely robust to range errors, as only 2% of the volume has an error bar of 7% or more. For set-up uncertainties, errors bars are larger (Figure 5.6(e)), but still lower than those for the PTV and CTV, thus indicating that critical structure spared in the optimization process are also rather robust in the case of patient misalignments. These conclusions are also valid for the composite error distribution (85% of the SC volume has an error bar below 10%).

### *5.3.3 Effect of dose gradients*

Finally, we present the results regarding the effect of using a surrogate to represent plan robustness. For simplicity, results are shown only for case-A. Figure 5.10(a)-(d) show for each plan the gradient distribution, achieved by applying a Sobel filter to each field as described in section 2.4. From these figures, this method seems to be a valid surrogate to quickly represent the likely error bar distribution for both the set-up and the range uncertainties. Comparing Figure 5.10(a)-(d) with Figures 5.2(e)-(h), Figures 5.5(a)-(d) and Figures 5.6(a)-(d), there is a similarity in the shape of the distribution, thus indicating, as expected, that regions with the steepest gradients are generally more susceptible to errors. Moreover, as the calculation of a dose-gradient via the application of a Sobel filter is a very fast operation, these results suggest that such a simple tool could be a useful surrogate for comparing potential plan robustness without the need to recalculate many dose distributions on a number of modified versions of the patient's CT. Indeed, it could be interesting to see if such a method could be used at the optimization stage in order to provide a fast surrogate for robustness.

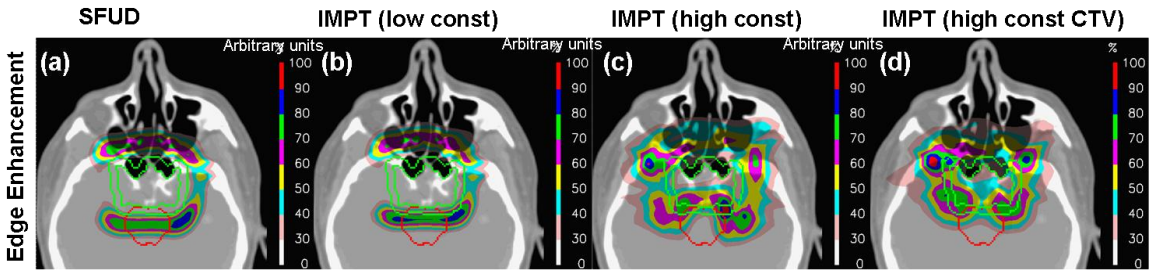


**Figure 5.10** : (a)-(d) dose distributions, based on the Sobel filter technique for the 4 plans as discussed for case-A; Green- lines represent the CTV (innermost) and the PTV (outermost). Red arrows represent the beam arrangements.

## 5.4. Discussions

In this Chapter, we have presented a method to evaluate the robustness of treatment plans to spatial and range uncertainties. By using the concept of the 'Error Bar Dose Distribution', it is possible to detect the irradiated areas that could be affected by uncertainties and to quickly compare two plans in terms of their robustness. We have applied this tool to different active scanning protons plans to (i) assess the benefit of planning using safety margins to deal with uncertainties and to (ii) gauge the robustness to uncertainties of both IMPT and SFUD plans. We have here reported the analysis only for 3 selected cases, but the results are confirmed for other 10 patients additionally studied.

Based on our analysis, it appears that the use of a safety margin (i.e. the PTV-approach), although still useful to improve the target coverage in the case of geometric and range uncertainties for SFUD and moderately modulated IMPT plans, may not be as helpful for highly modulated IMPT plans. As for the SFUD plan, each individual field delivers a homogeneous dose through the target, in-field dose gradients can only occur at the edge of the target itself. Therefore, misalignments and range errors can consequently only affect the edge of the PTV, without modifying the CTV coverage significantly. The same is true for the IMPT plan with low in-field gradient, as again the individual fields are delivering a quasi-homogeneous dose trough the target volume.

In contrast, for the IMPT high-gradient plans, individual fields are forced to have much greater in-field dose gradients. In this case, large dose uncertainties can also appear in the middle of the target volume and the use of a PTV has been shown to not necessarily improve the robustness of the plan in the CTV region. This is graphically demonstrated in the Error-Bar Volume Histograms (EVH's). Whereas for both the SFUD plan and for the moderately modulated IMPT plans EVH's for the CTV are improved in contrast to the corresponding EVH's for the PTV (Figure 5.3), for the highly modulated IMPT plans, the EVH's are almost overlapping.

Additionally, the results presented here show that even when complex IMPT plans are calculated without the use of a safety margins (i.e. directly on the CTV volume itself), there appears to be no deterioration in the target coverage in the case of patient misalignments (Figure 5.4(a)). In the case of range uncertainties alone or in combination with spatial uncertainties, only a marginal improvement in the CTV robustness is detectable with the use of a PTV (Figure 5.4(b),(c)).

Thus, rather than the use of safety margins, it would appear that more sophisticated methods to take into account uncertainties may be required for highly modulated IMPT plans. For example, either by incorporating errors directly in the optimization algorithm, as proposed by Unkelbach et al (2007,2009) and Pflugfelder et al (2008), or by adjusting the starting conditions (as, for example, the initial field arrangements and the initial pencil beam weight) to steer the optimization result to a solution more robust to uncertainties, as discussed in Chapter 4 and already proposed by Lomax (2004, 2008a). Indeed, for all patients analyzed, the SFUD plans were found to be slightly more robust to misalignments and range errors, at least from the point of view of target coverage, than IMPT plans (as is shown in Figures 5.3). Surprisingly however, IMPT plans appear to be quite robust for organs at risk that were spared during the optimization process (e.g. brainstem in Figures 5.2, Figures 5.5 and Figure 5.6, chiasm in Figures 5.8 and spine in Figures 5.9).

With Case-B we have also shown how the Error-Bar distributions can be used to directly compare the robustness to spatial and range uncertainties of two dosimetrically similar plans (Figure 5.7). We have seen that just by changing beam configurations it is possible to improve the robustness of a plan significantly, as 5% of the optical chiasm has a reduction of the error bar from 33% to 16% in case of set-up errors, when planning with two fields coming from the cranial aspect of the patient.

The approach we have adopted for our robustness analysis is simple and fast to calculate. It is based on a relatively small set of discrete calculations of the dose under well defined error conditions. As doses are only calculated for the nominal and some set of bounding conditions based the magnitudes of the assumed errors, we are not attempting to fully characterize the Probability Distribution Function (PDF) of the resulting spectrum of doses as others have done (see e.g. Unkelbach et al). In the latter case, although there is potentially more information, as the resulting spectrum of doses will generally not be (indeed will very rarely be) normally distributed, it is extremely difficult to then represent this spectrum in a succinct manner. Generally, this is done by reporting the standard deviation of the dose, but due to the complex and non-normal shape of the dose PDF's, this has little direct statistical meaning. In contrast, our approach does not attempt to make any assumption about the shape of the dose PDF other than

that all possible dose values within the defined magnitude of errors used in the error calculations are bounded by the minimum and maximum values of the doses calculated for the nominal situation and all error bounds. If this assumption is correct, then we can say that (for the magnitude of errors used in this work) we have an 85% confidence limit that the dose delivered to any point is within the error-bounds calculated for our error-bar distributions. On the other hand, we cannot say with any accuracy what is the chance that the dose within a given voxel is within (for example) +/- 3% of the nominal dose. For this, one requires an accurate representation of the actual dose PDF, and one that is more detailed than a simple calculation of the standard deviation of this distribution.

As mentioned above, we assume that all possible dose values of the PDF are bounded by the minimum and maximum calculated doses. We believe that for the magnitude of spatial and range errors used here this is a valid assumption. The spatial errors used here (of the order of a few millimeters) are substantially less than the sigma of the pencil beams being used for delivery, which for our facility have an average value of 8mm at the Bragg peak (Lomax et al 2001). Thus, even for our largest spatial errors (as used for example for case C) the magnitude of shift is still less than the average beam sigma (5.5mm vs 8mm). Thus, we believe it is extremely unlikely that dose values outside of the minimum or maximum calculated values will occur. Nevertheless, showing that this is the case is the subject of ongoing work in our group.

In addition, it should be noted that all error bar distributions reported here for set-up uncertainties are applicable only to the error that could occur on a *single* fraction. Consequently, even in the case of rather high uncertainty values, the combined treatment error-bars will be significantly reduced when fractionation is taken into account (assuming that set-up errors are indeed random). In contrast, as range errors are likely to be systematic, and will therefore propagate throughout the whole treatment course, the error distributions estimated here could be relevant for the whole treatment. Consequently, we have combined random and systematic errors into a composite error distribution using the "margin recipe formula" described by van Herk et al (2000) in order to provide a single distribution which reflects the different natures of systematic and random errors. The resulting composite error distributions are, however, just one example of how the systematic and random errors could be combined. Other approaches include those already suggested by different authors for dose error displays in photon therapy (see e.g. Hosang et al (2007), Flynn et al (2009) or Maleike et al (2006)). An alternative approach could also be to combine the error distributions as a function of the number of fractions. Clearly, there is only a difference between random and systematic errors for an individual patient if the full treatment

consists of multiple fractions. For single fraction, radiosurgery type treatments, the error distributions calculated here for set-up errors would also become potentially systematic for any given patient and could then be directly combined with the range errors. Thus, an alternative approach could be to calculate composite error-bar distributions by weighting the set-up error contribution by the inverse of the square-root of the number of fractions. In this case, the reduced weighting factor used in this work for random errors (based on the van Herk formula) would correspond roughly to a 13 fraction treatment ($0.7/2.5 = 0.28$, $1/\sqrt{13} = 0.277$). Either way, more work is required in order to determine the best weighting between the range and set-up error-bar distributions, such that one single error distribution can succinctly and meaningfully represent clinically relevant scenarios.

Finally, as we have found that in-field dose gradients are generally responsible for dose variations *within* the target volume, we have suggested that the application of a simple Sobel (edge enhancing) filter to the individual field dose distributions could provide a simple surrogate for predicting potential areas of uncertainty without having to perform a full uncertainty analysis for the plan, and that possibly, such an approach could have some applications in the optimisation process as an additional constraint in addition to purely dose based constraints.

## 5.5. Summary

We have presented a method for calculating, presenting and evaluating distributions of uncertainties for treatment plans in a form that is easy to calculate and can be collapsed into a single distribution for analysis. Although here applied to proton plans, this concept can be easily applied to all kind of radiation therapy plans (i.e. photons, protons, ions).

Using this approach for a number of difference cases, it has been found that, in general, SFUD plans are more robust than IMPT plans for both set-up and range errors. In addition, it has also been shown that IMPT plans which are optimised to spare neighboring critical structures seem to generally preserve this sparing effect even under conditions of positioning and range uncertainties. However, although planning with the "PTV approach" (i.e. the use of safety margin) improves the plan robustness to set-up and range errors for SFUD plans and for IMPT low gradient plans, in contrast, for IMPT high gradient plans, the "PTV approach" seems to have at best a marginal effect for dealing with such uncertainties. This result points to the fact that the simple PTV margin approach is not necessarily sufficient for complex IMPT plans.

Finally, a simple surrogate for identifying potential regions of high uncertainty has been proposed through the application of an edge enhancing filter to the nominal dose distribution

which we believe could provide a quick first check of the robustness of a treatment plan, and may have applications in the optimization process to aid robust optimization.

In the next Chapter we will present experimental verification of IMPT plans to assess the accuracy of their delivery both in nominal and error conditions. Furthermore, these measurements have been used to validate the robustness analysis model here discussed.

---

Note: This Chapter has been published as part of the following paper: Albertini F, Hug EB and Lomax AJ *Is it necessary to plan with safety margins for proton scanning plans?* in Phys. Med. Biol (56) 4399-4413 (Albertini et al 2011a)

# 6 Experimental verification of IMPT treatment plans and robustness

## 6.1 Introduction

Proton therapy provides a tool for delivering highly conformal dose distributions. The first proton treatments started in 1954 at the Lawrence Berkley Laboratory using what has subsequently been called the passive scattering technique (Koehler et al 1975, Koehler et al 1977). With this approach, each individual field delivers a homogeneous dose through the target volume, with the distal end of the spread out Bragg Peaks (SOBP) matching the distal end of the tumour.

In 1996 the first patient has been treated at the Paul Scherrer Institute with the active scanning technique (Pedroni et al 1995). With this technique, suggested in 1980 by Kanai (Kanai et al 1980), protons are deflected through the application of magnetic fields in the direction orthogonal to the beam direction. And in depth, the position of the individual Bragg peaks is varied by changing the energy at the accelerator source. With this delivery technique, the Bragg peak can be placed everywhere in three dimensions in the tumour volume. The total dose distribution is then achieved through a mathematical optimisation of the individual fluences of each Bragg peak.

The first patients were treated using the active scanning technique by essentially mimicking the passive scattering technique. That is, each individually delivered field applied a more or less homogeneous dose throughout the target volume. This approach is what has been called Single Field, Uniform Dose (SFUD) planning (Lomax 2008c). A full treatment plan then consists of the linear addition of each individually homogeneous field.

Ten years ago the flexibility of the spot scanning technique was extended to deliver the so called ‘Intensity Modulated Proton Therapy’ (IMPT) (Lomax 1999, Lomax et al 2001). With this method, individually highly in-homogeneous fields are combined together to deliver the full treatment dose. This is achieved by simultaneously optimizing all the Bragg peaks delivered from

all the applied fields. The IMPT approach best exploits the full potential of scanned proton therapy, as it provides even more flexibility in tailoring the dose distribution to the target - and in selectively avoiding critical structures -than the SFUD.

In order to ensure an optimal treatment quality it is necessary to measure and verify the accuracy of the dose distribution. Given that for the passive scattering approach and also for the SFUD approach, the plan dose is achieved by the simple combination of individually homogeneous fields, it is enough to experimentally verify the individual field dose distributions independently and extend the results to evaluate the total plan dose distributions. At the Paul Scherrer Institute (PSI) this is performed using a two-arm array of ionisation chambers positioned below a variable depth water phantom (Lomax et al 2004b). With this array, it is possible to simultaneously measure two orthogonal profiles of the dose distribution at one or more depths. These measurements can then be directly compared to the doses predicted from the treatment planning for each field when it is recalculated in a homogenous water phantom. This is the current approach at PSI for verifying the validity of the individual fields of each IMPT plan delivered at PSI.

However, this approach is not ideal for accurately measuring the total plan dose for IMPT treatments; in which highly in-homogeneous individual fields are combined together to achieve the final desired dose distribution. For instance, when combining the individual fields together, some slight misalignments of the high dose gradients could be present, leading potentially to quite severe under- or over- dosage in comparison to the calculated plan (see e.g. chapter 5 above). This effect could be even more important when complex density heterogeneities are present.

In 2004 Safai et al (2004) developed a dosimetric phantom to perform 3D measurements of IMPT plans using a scintillating mixture optimized to faithfully reproduce the proton Bragg peak. However, although this phantom can be potentially used to verify the accuracy of the dose distribution for multiple fields, the dose calculated from the treatment plan has to be recalculated in a homogeneous medium to be compared with the measured results. Therefore, to assess the accuracy in the delivery of full IMPT plans, whilst also simulating the effect of realistic density heterogeneities in the patient, in this work, we have developed a dedicated anthropomorphic phantom in which such measurements can be directly performed.

In addition, from a medical physics point-of-view, the ‘quality’ of a plan to be delivered is not only defined by the accuracy with which the fields can be delivered to a homogenous medium, but also on how robust the plan is to likely treatment uncertainties that are inherent to the process of fractionated radiotherapy. Range uncertainties represent the most critical source of

errors in the case of proton therapy (see discussions above) - as these are generally systematic and therefore will tend to propagate through the whole course of the treatment (Lomax 2008a). However, spatial errors can also not be ignored, especially when highly in-homogeneous IMPT fields are combined together.

As discussed in Chapter 3, given the huge degrees of freedom available to the optimization process in IMPT, there is a correspondingly huge degeneracy of solutions for the optimization process, all of which result in very similar dose distributions. However, the two extreme solutions are likely to be represented at one end of the spectrum by the 3D-IMPT (Lomax 1999) and at the other by distal edge tracking (DET) (Deasy *et al* 1997, Oelfke and Bortfeld 2000 and Nill *et al* 2004). For the first approach, Bragg peaks are distributed in three dimensions over the whole target volume from each incident field direction, whereas, for the DET approach, only Bragg peaks that lie on the distal edge of the target volume for each field are considered in the optimization process. In this case, and as has already pointed out by Lomax (Lomax 2008a), even if these two IMPT plans exhibit dosimetrically equivalent solutions, when analyzed in detail, they can exhibit different behaviors from the point of view of robustness analysis. Firstly, when assessing the accuracy of the analytical against the Monte Carlo dose calculation, there is a degradation of agreement for the DET plan (Lomax 2008a) and more importantly for this work, and as shown in Chapter 4, when plans are compared in terms of their robustness (or lack of) to delivery uncertainties, the 3D-IMPT plan resulted generally in more robust (less sensitive) plans than DET.

In this Chapter we extend the comparison between 3D-IMPT and DET by directly measuring the accuracy of the dose calculation in both cases and also the sensitivity of both approaches to range and angular uncertainties when delivered to an anthropomorphic phantom. Furthermore, through the direct comparison of the measured dose distributions in the presence of these uncertainties, we intend to verify the validity of the error-bar model presented in chapter 5 as a predictor of areas that may be particularly sensitive to such errors.

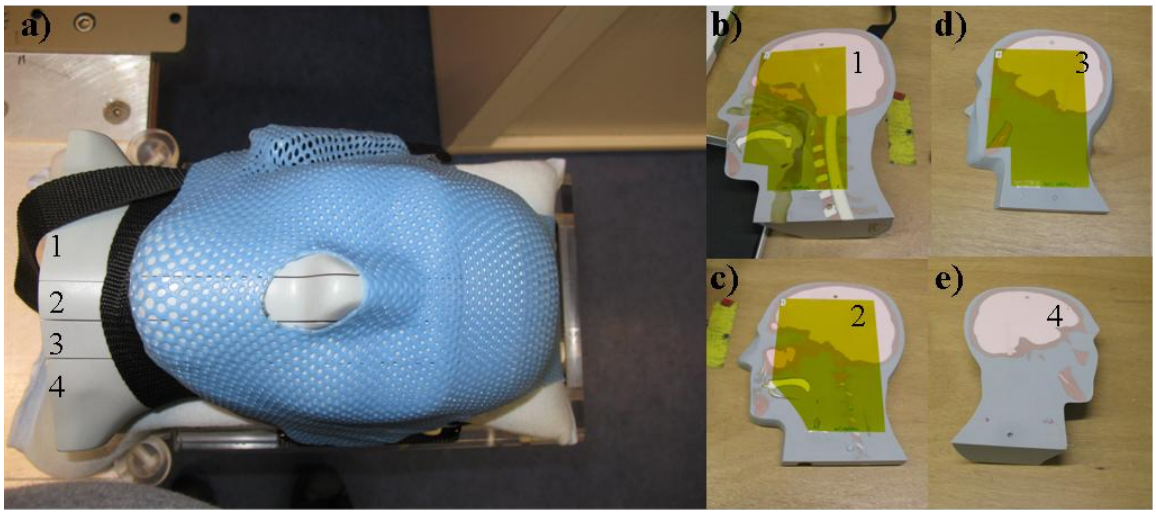


**Figure 6.1** Special anthropomorphic phantom customized based on the diagnostic head phantom model 711HN (CIRS company). The individualized fixation device used is shown, as well as the four (1-4) components of the phantom, both as combined together (a) and as individually sliced (b)-e)). Anatomical details and the Gafchromic EBT2 films used for dosimetrical purposes are also visible.

# 6.2 Material and methods

### *6.2.1 The anthropomorphic phantom.*

As mentioned above, density heterogeneities can have a large impact on the delivered dose distribution, as they affect both the range of the protons and the shape of the Bragg peak (Goitein 2008, Lomax 2009) and will also therefore play a role in the robustness of a plan to spatial and range uncertainties. In order to measure delivered dose in as realistic conditions as possible therefore, it is important that a realistic anthropomorphic phantom is used.

The most common such phantom used in radiotherapy is the Alderson® phantom (Radiology Support Devices, Long Beach, CA USA). However, for the measurements we want to perform here, this standard phantom has a number of drawbacks. First, the Alderson® phantom is cut in the wrong direction when imaging using x-ray CT. As it is divided into sections cut orthogonal to the superior-inferior axis of the phantom, then the inevitable air gaps between the different slices can significantly affect the quality of the CT images, particularly in the slices close to these gaps. Although not a major issue for photon verifications, these artifacts can cause significant problems for proton calculations. In addition, for verifying co-planar or near co-planar field directions, these air gaps will also significantly affect (in a negative way) the quality of the delivered proton beams, as the beams will be travelling parallel, or close to parallel, to these interfaces. As such, these gaps themselves will act as major density heterogeneities to the beam, whereas such heterogeneities will obviously not be present in real patients. Second, the

Alderson® phantom is rather limited in its composition, being composed of just two materials, bone (albeit real bone) and one type of soft tissue substitute material representing all non-bony tissues. For these reasons, in this work we have developed, together with the firm CIRS (Norfolk, Virginia USA), a new anthropomorphic phantom specifically for proton therapy verifications and measurements.

This new anthropomorphic phantom, shown in figure 6.1, has been customized based on the diagnostic head phantom model 711HN manufactured by CIRS. This phantom approximates an average male human head in both size and internal structure, and includes detailed 3D anthropomorphic anatomy such as brain, bone (with both cortical and trabecular components), larynx, trachea, sinus, nasal cavities and teeth. The phantom is constructed of tissue equivalent materials (ATOM®max resins) mimicking the CT attenuation coefficient of different human tissues. For comparison, lateral X-rays of an Alderson® phantom and our new phantom are shown in figure 6.2.

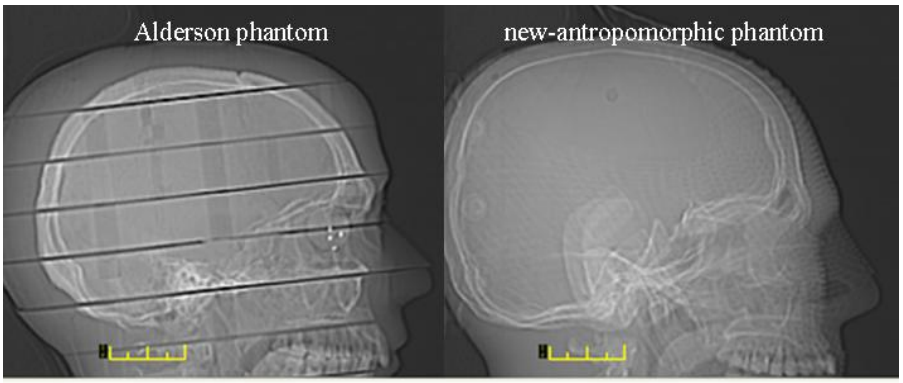


**Figure 6.2** Comparison of lateral x-rays of an Alderson phantom and of the new anthropomorphic phantom

This standard diagnostic phantom has been customized for our purposes in the following way. The standard phantom was sliced in two along the central lateral axis, and one half further sliced into 3 segments along the superior-inferior direction to allow for positioning of dosimetric films at different anatomical positions (see figure 6,1). The second and third slices were respectively cut 20 and 40 mm away from the center of the phantom. This allowed us to place up to three films in the phantom in different positions for each measurement, with two centimeters of phantom between each film.

### *6.2.2 Treatment planning*

In order to reproducibly place the phantom in relation to the beam for each measurement, as with our real patients, an individualized thermoplastic mask and vacuum pillow were used (see

figure 6.1). Using these devices, a planning CT was acquired using a 2mm slice separation and our standard acquisition protocol for planning CTs in this anatomical region. CT values have been used to determine the proton stopping power relative to water for each tissue substitute samples through the application of the clinically used CT stoichiometric calibration curve, introduced by Schneider et al (1996). These relative stopping power values have been further compared with those measured with a 138 MeV proton pencil beam. The measurement was done, as described in detail elsewhere (Schaffner and Pedroni 1998), by scanning the energy of the beam in discrete steps through the introduction of range shifters plates, and by acquiring for each step the integral Bragg peak through the use of a large parallel-plate ionization chamber (of 8cm diameter). Measurements are then repeated with the pencil beam passing through the different samples. By measuring the shift measured between these two curves, the relative stopping powers for each sample can be finally calculated. A good agreement (within 1.5% for the soft tissues and of about 2% for the bony structure) was found between the measurement and the calculation for all the samples analyzed (see Table 6.1).

| Material | Substitute for | Stopping power (calculated from CT) | Stopping power (measured) |
|---|---|---|---|
| LDT | sinus cavities | 0.187 | 0.191 |
| STG-955 | soft tissue | 1.021 | 1.036 |
| VSC 31 | spinal cord | 1.04 | 1.046 |
| BT 358-1 | brain | 1.046 | 1.059 |
| DTB 109 | trabecular bone | 1.194 | 1.176 |

**Table 6.1** Comparison between the stopping powers relative to water measured for various tissue substitutes and the ones calculated from the CT stoichiometric calibration curve

Using this phantom and 'treatment' set-up, two main measurements have been carried out. First, we have measured the accuracy of the delivered dose for the different IMPT plans in the nominal conditions (i.e. without deliberately introduced uncertainties). Second, we have additionally measured the effect of deliberately introduced range and spatial uncertainties on the delivered dose distributions.

For this analysis, two 'flavors' of IMPT treatments have been measured, namely 3D-IMPT (Lomax 1999) and distal edge tracking (DET) (Deasy et al 1997, Oelfke and Bortfeld 2000 and Nill et al 2004) (Figure 6.3). In 3D-IMPT, the Bragg peaks are distributed in three dimensions over the whole target volume for each beam direction. The initial beamlet fluences have been pre-weighted, for each beam direction, such as to deliver a flat 'spread-out Bragg-peak (SOBP)' throughout the target volume, as described in Chapter 4. In contrast, for the DET approach, only Bragg peaks lying on the distal edge of the target volume for each field are considered in the

optimization process. Apart from these different starting conditions, the optimization and dose calculation for the two cases was identical and based on the algorithms described in chapter 2.

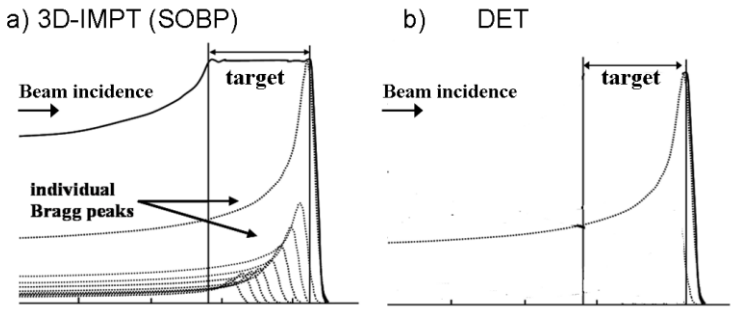


**Figure 6.3**. Schematic representation of the two flavours of IMPT plan analysed: (a) 3D-IMPT and (b) DET plan respectively. The initial beamlet weights for a single field approaching the target from left to right for both plans are here represented, with the 3D-IMPT plan, showing individual Bragg peaks with a reduced weighting from distally to proximally, thus resulting in an 'flat SOBP' dose distribution; and with the DET approach built so that only the most distal Bragg peak is selected, for a given lateral position.

To simulate our clinical situation as closely as possible, we followed the normal patient treatment procedure as follows: (i) definition of the phantom positioning through the customization of an individually tailored vacuum pillow and thermoplastic mask, (ii) acquisition of planning CT data using normal clinical acquisition protocols, (iii) delineation of a target volume and neighbouring critical structures (assuming a skull base chordoma) and (iv) calculation of the IMPT plans using the PSI in-house developed planning system.

The two plans (3D-IMPT and DET) to this target are shown in figure 6.4. Figures 6.4(a) and (e) show transaxial slices through the centre of the tumour for four field 3D-IMPT and DET plans, respectively. For both plans, the beam angles are all quasi-coplanar and incident from the lateral aspect of the patient (as indicated in the figure), as is our practice for such cases. Plans have been optimized with stringent constraints on the optic nerves, chiasm, brainstem and both cochlea structures. By setting challenging constraints on these organs—which partially overlap with the PTV—we have constructed complex IMPT plans, such that the individual fields have high in-field dose gradients. Note that, as expected due to the degeneracy of the optimization problem for IMPT, there is a great similarity in the resultant dose distributions for the two plans despite the very different starting conditions.

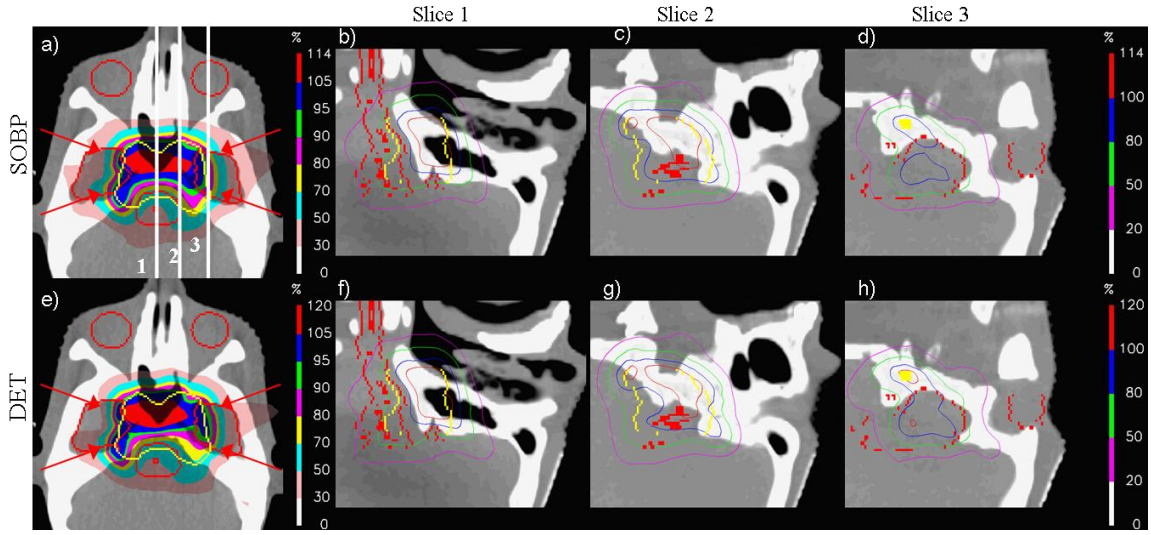


**Figure 6.4**. The two IMPT plans (resp. with the SOBP and the DET approach) are shown both in the axial and along the three slices of the sagittal projection. Beam arrangements are indicated by the red arrows. Yellow and red contours indicate the target volume and organs at risk respectively. In the sagittal views the iso-doses contours are reported as well as the complexity of the anatomical structures present at the three measurements positioning.

#### 6.2.2.1 Range and spatial errors

To evaluate the effect of range uncertainties, two IMPT plans have also been calculated (and optimized) on two modified versions of the planning CT data, one with the Hounsfield Unit (HU) systematically increased by 3%, and a second decreased systematically by 3%. By measuring the actually delivered doses in the phantom for these plans, we are then simulating the dose errors resulting from range uncertainties of -3% for the plan calculated on the +3% CT data set, and vice-versa for the -3% CT study. By subtracting, the two measurements, we then have a measured estimate of the error-bar distributions described in chapter 5. Although this is not exactly the same situation (we have optimized two in principle different plans on modified versions of the CT and have compared these to the nominal rather than optimizing on the nominal and then delivering to a modified CT data set), as the dose distribution for the two plans are very similar (see, for example, in Figure 6.5 the comparison of the two 3D-IMPT plans optimized on the incorrect CT data), the differences due to the re-optimization are almost negligible, indicating that the measured differences are only related to the range errors introduced.

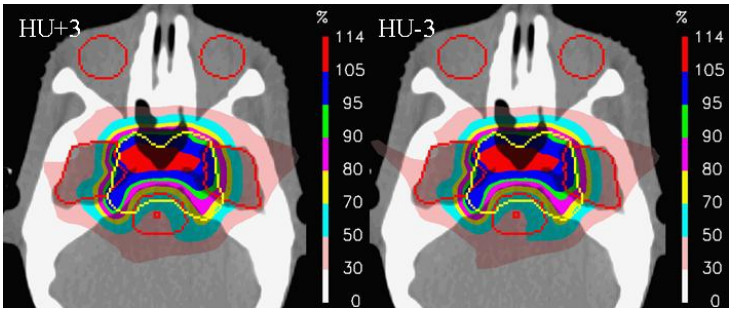


**Figure 6.5**. Axial views of the 3D-IMPT plans optimized on incorrect version of CT data (resp. with the HU systematically increased (decreased) by 3%).Yellow and red contours indicate the target volume and organs at risk respectively.

Finally, spatial uncertainties have been introduced by measuring the IMPT plans, with the phantom rotated by 3° along the cranio-caudal axes. Thus simulates the effect of a residual patient set-up error.

### *6.2.3 Radiochromic film measurment*

For all measurements, GafChromic® EBT2 films have been used to measure two-dimensional dose distributions of the delivered IMPT plans in the measurement planes of the anthropomorphic phantom. The suitability of radiochromic films in proton beams has been reported by different authors (Vatnitsky *et al* 1997a, 1997b, Daftari *et al* 1999, Piermattei *et al* 2000, Martišiková and Jäkel 2010, Mumot *et al* 2009, Zhao *et al* 2010).

The films were positioned within the phantom in a plane quasi-perpendicular to the delivered beam arrangement (see Figure 6.1(b)-(d)), The first is positioned in the middle of the target, the second 20mm away and the third one 40mm away from the central plane at the edge of the target volume. All films used in the measurements were from the same film batch.

After irradiation, each film was scanned and digitized using a commercial A3 size flatbed scanner (Expression 10000XL, Epson) in transmission mode using a transparency adapter. Each film is scanned separately in landscape orientation (shorter side of the whole film parallel to the scanner short side), at the same position in the middle of the scanner bed, approximately 48 h after irradiation. The software used for image acquisition was supplied with the hardware (Epson scan v3.04s). The images were acquired in 48-bit colour (RGB) mode, 16 bit per colour channel, with a resolution of 75 dpi and saved in uncompressed tagged image file format (TIFF). All the post processing parameters were switched off and a linear look up table (LUT) was used. Multiple scans of each film were taken to obtain a mean value over the regions of interest.

In-house developed software (using the Matlab 7.11 application) was used to import the stored images, perform the average of the three acquisitions and extract the red channel. The last operation is done to enhance the film read-out as the absorption spectrum for the EBT2 films has

a maximum at a wavelength of 636nm. To calibrate the non-linear response of EBT2 film to dose, a series of uniform square fields of dimension 3x3 $cm^2$ were applied to a single film placed orthogonally to the beam direction in a Plexiglas phantom (30x30$cm^2$ slabs) at 9.5 cm depth. Eight fields with different doses (from 0.2 to 3.8 Gy) were delivered using a modulated 138 MeV proton beam. The dose for each field was measured using a charge coupled device (CCD) detector, coupled with an absolute measurement carried out with an ionization chamber. Finally, a calibration curve, which converts grey levels into dose, was obtained by a third order polynomial fit of the measured data. This film response as a function of dose is shown in Figure 6.6.

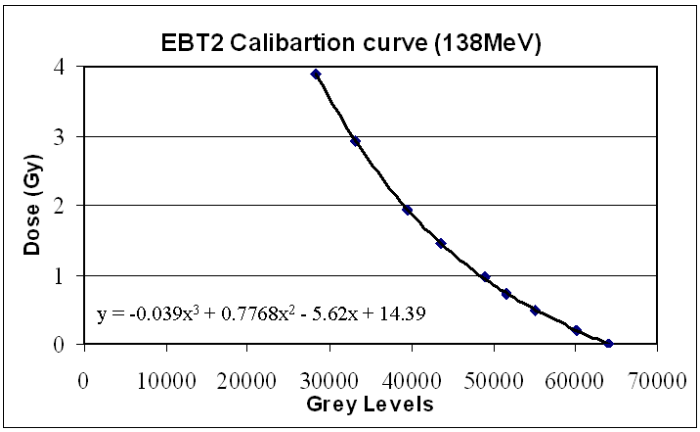


**Figure 6.6** Calibration curve converting Grey Levels into dose. Data are fitted with a third order polynomial function

### *6.2.4 Measurements and data analysis*

The EBT2 films in the above set-up have been used to measure the accuracy of the delivered dose against the calculated dose from the treatment planning system and to ascertain the robustness of the different IMPT plans to planned and known errors in the delivery (range and set-up errors).

To assess the accuracy of the different IMPT plans, the following steps have been performed. Two dimensional dose distributions have been measured with the films and then compared to the corresponding two dimensional distributions extracted from the treatment planning system. The measured dose distributions are manually aligned to the calculated ones and the two dose distributions have been normalized to a common point selected to be within a relatively uniform part of the dose distribution. The film images and calculated dose distributions have then been analyzed by visual assessment of the agreement between the isodose lines calculated by the treatment planning system and the measured iso-dose lines recorded by the films, and through a gamma evaluation (see below). Data analysis has been performed using the IDAS software ('IMRT Dose Analysis System'), developed at the Department of Radiation

Oncology, University of Ulm for the dosimetric verification of IMRT plans (Salk et al 2003). This utilizes the data storage and processing package HBOOK in conjunction with the Physics Analysis Workstation (PAW) system, developed at the European Organization for Nuclear Research (CERN) in Geneva (Switzerland). These software packages are part of the CERNLIB program library, which is offered under the terms of the GNU General Public License (GPL).

The gamma evaluation has been introduced by Low et al (1998) and is a widely used approach in radiotherapy for comparing two dimensional measurements with calculated dose distributions. The gamma index corresponds to the minimum multi-dimensional distance between the measurement and calculation points in a space composed of dose and distance, scaled by preselected tolerance limits for dose differences and distances to agreement (DTA). For our analysis, tolerance limits of 3% in dose and 3mm in DTA have been chosen, indicating that any point that is within 3% of the expected dose *or* within 3mm distance of a point with the same dose would have a gamma value of less than 1. Gamma values can then also be displayed as two-dimensional distributions (as the gamma value can be calculated for every measurement point) or as cumulative histograms.

To assess the effect of delivery uncertainties for the different IMPT plans, we have compared the two dimensional profiles measured with the errors included, as described in section 6.2.2, with the profiles measured in the nominal condition. Furthermore, the measured distributions have been analyzed to validate the error-bar distributions introduced in chapter 5. For this, the effect of range errors have been modeled by recalculating for both types of IMPT plans a general error of ±3% in the nominal HU values of the phantom CT data. The data have been then combined into one unique error distribution, namely the **Range Error- Bar dose distribution** calculated as

$$\Delta D^{i}{}_{range\_error} = Max\left(D^{i}_{+HU}, D^{i}_{-HU}\right) - Min\left(D^{i}_{+HU}, D^{i}_{-HU}\right)$$

where $\Delta D^{i}{}_{range-error}$ is the value at the dose grid point *i* in the resulting error distribution and $D^{i}_{+HU}$ ( $D^{i}_{-HU}$ ) is the dose at point *i* for the recalculated plan assuming a general error of +3% (respectively, -3%) in the HU values.

The calculated Range Error-bar distributions have been compared to the dose differences, measured with ±3% in the HU values, to verify the validity of this model to provide a good estimate of the areas mostly affected by uncertainties.

## 6.3 Results

### *6.3.1 3D-IMPT vs DET: Plan accuracy*

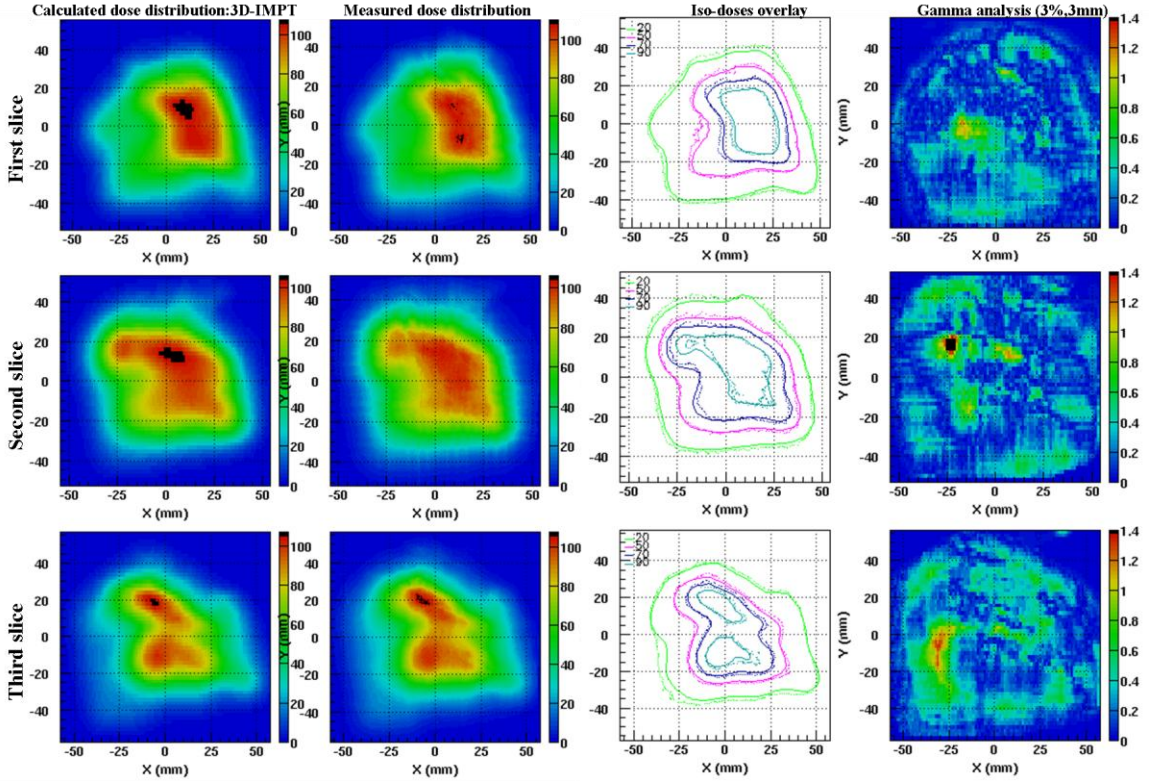


**Figure 6.7** Calculated and measured dose distributions, for the 3D-IMPT plan, are presented individually as two dimensional color maps, for the three positions of the phantom, in saggital way. Iso-doses contours are overlaid with solid and dashed lines representing the calculated and the measured dose respectively. Last column shows the gamma-index map. Tolerance limits of 3% in dose difference and of 3mm in distance-to-agreement were chosen for gamma evaluation.

Figure 6.7 shows the comparison between the nominal 3D-IMPT calculated and measured doses for the three slices of the anthropomorphic phantom shown in Figure 6.1. For both, doses are represented as color maps varying from blue to red as shown in the colour bar on the right of each distribution. As mentioned, the calculated and measured dose distributions are normalized to a common dose point in a uniform part of the distribution. In the third column the 20%, 50%, 70% and 90% levels of iso-dose contours for the calculated and measured are overlaid and finally, in the last column, the point-to-point gamma index is also presented in the form of a color map.

The results show an excellent agreement between the calculated and the measured dose distribution, particularly as these have been measured in a particularly heterogeneous part of the phantom. The iso-dose curves match very well for at all displayed iso-dose levels. Moreover, the

gamma index maps (using a 3%/3mm tolerance) indicate that, over all measured films, more than 99% of measured points agree to within 3%/3mm.

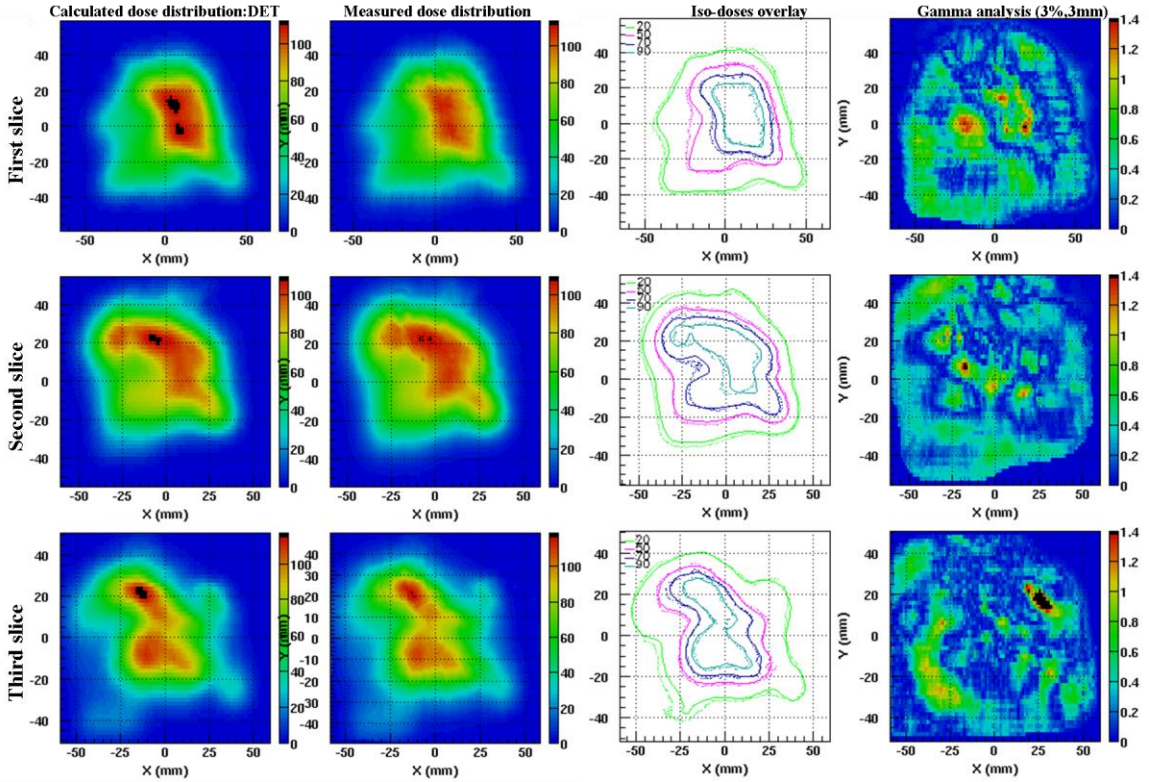


**Figure 6.8** Calculated and measured dose distributions, for the DET plan, are presented individually as two dimensional color maps, for the three positions of the phantom, in saggital way. Iso-doses contours are overlaid with solid and dashed lines representing the calculated and the measured dose respectively. Last column shows the gamma-index map. Tolerance limits of 3% in dose difference and of 3mm in distance-to-agreement were chosen for gamma evaluation.

Figure 6.8 shows a similar comparison for the DET. As for the 3D-IMPT plan, a good agreement can be seen between the calculated and the measured dose distributions, with almost 98% of points satisfying the gamma criteria. Nevertheless, these results are somewhat worse than those achieved for the 3D-IMPT plan (Figure 6.7). This is confirmed by a direct comparison of the cumulative gamma-volume histograms, as shown in Figure 6.9. For all the slices, the number of points with a γ-index <1 is always lower for the 3D-IMPT plan (red solid line) than for the DET plan (black dashed line). In particular, for the last slice (which is located at the target edge, Figure 6.4) only about 97% of points are satisfying the gamma tolerance limits (with $\gamma_{max}$ of 2.6) for the DET plan, which can be compared to the more than 98% agreement found on this slice for the 3D-IMPT plan ($\gamma_{max}$ of 1.3). This indicates a small, but perhaps significant advantage to the 3D-IMPT technique from the point of accuracy of delivery in heterogeneous regions. For the rest

of the chapter these measurements will be referred as the 'nominal' measurements (i.e. achieved in the absence of delivery uncertainties).

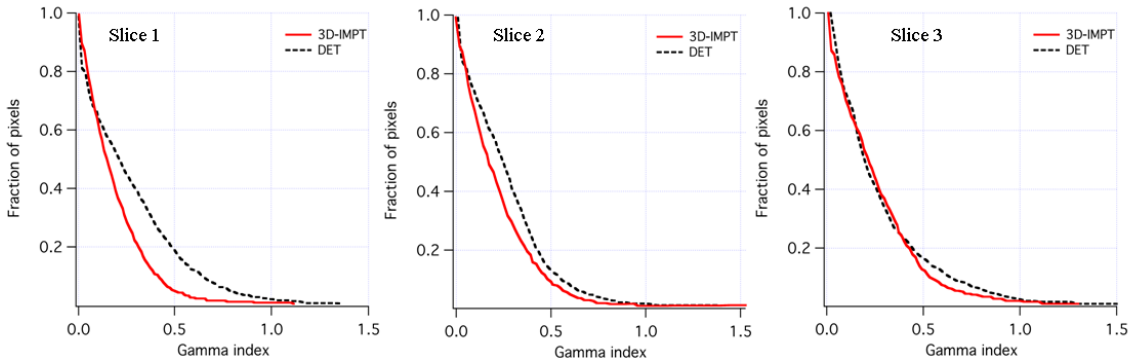


**Figure 6.9** Cumulative gamma-volume histograms for both the 3D-IMPT (red solid line) and the DET (black dashed line) plan. Tolerance limits of 3% in dose difference and 3 mm in distance-to-agreement were chosen for gamma evaluation.

## *6.3.2 3D-IMPT vs DET: robustness to uncertainties*

### *6.3.2.1 Range Uncertainties*

In order to experimentally assess the robustness of each plan to range uncertainties, the two films calculated using deliberately introduced range errors (+/-3%) for each plan have been subtracted from each other, and the differences displayed as difference distributions in Figure 6.10. These should be the direct measured equivalent of the error-bar distributions described in section 6.2.4. Moreover, Figure 6.11 shows the point-by-point difference histograms for both the 3D-IMPT plan (red solid line) and the DET plan (dashed line) as calculated from the distributions of Figure 6.10. A delta function centered at '0-difference' value would be the ideal result, indicating that even in the case of range errors, no differences are measured in the dose distribution. As such, the narrower the curve is, and the closer it is to the zero point, the more robust the plan is to range uncertainties. From these figures it is clear that for both IMPT plans here studied, the first slice (which is located in the middle of target volume) is more robust to these uncertainties than the last slice (located at the target edge). The reason for this result is connected to the position of the highly weighted Bragg peaks, which are more sensitive to range uncertainties, in respect to the target geometry. For both plans, only a few highly weighted Bragg peaks are contributing to the calculated dose for the slice located in the middle of the target volume (slice 1). In contrast, as the last slice (slice 3) is located at the target edge, for two fields it will coincide with the distal-end of the target volume, while for the opposite two fields it will correspond to the proximal portion of the target. Thus, given that, for both plans, the dose of each of the individual fields is obtained through a combination of highly weighted Bragg peaks

delivered at the distal edge of the target volume (see figure 6.3), the combined dose is then achieved by the superposition of the distal-fall off of many highly weighted Bragg peaks combined with the proximal edge of the Bragg peaks delivered from the two opposite fields.

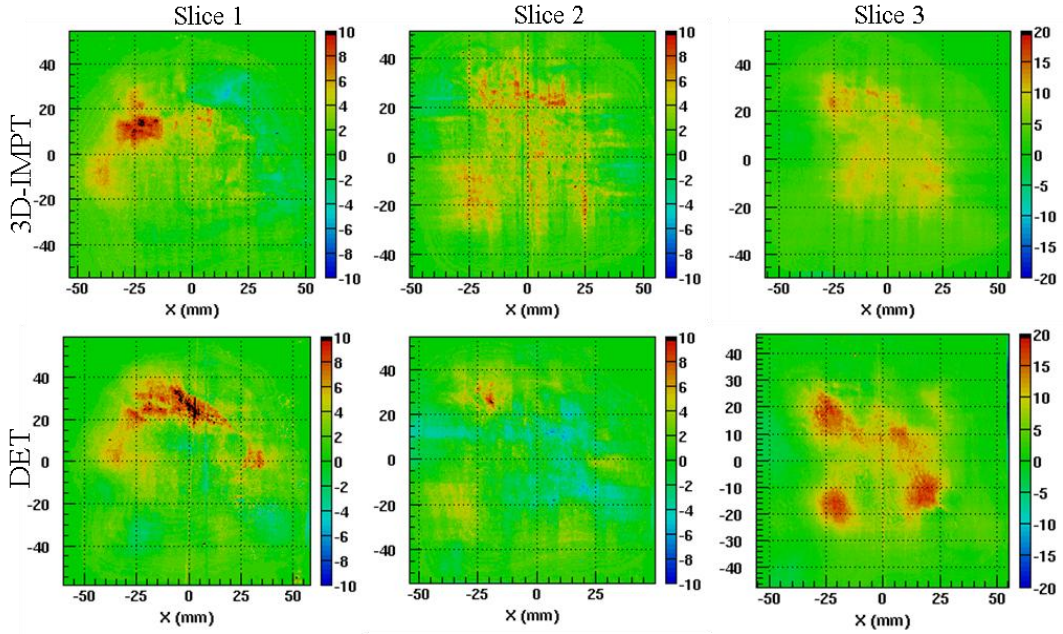


**Figure 6.10**Computed distribution differences (in %) between the doses measured with range errors included (±3% of HU) for both 3D-IMPT and DET plans, for the three depths. Note the different dose-scale used for the last depth

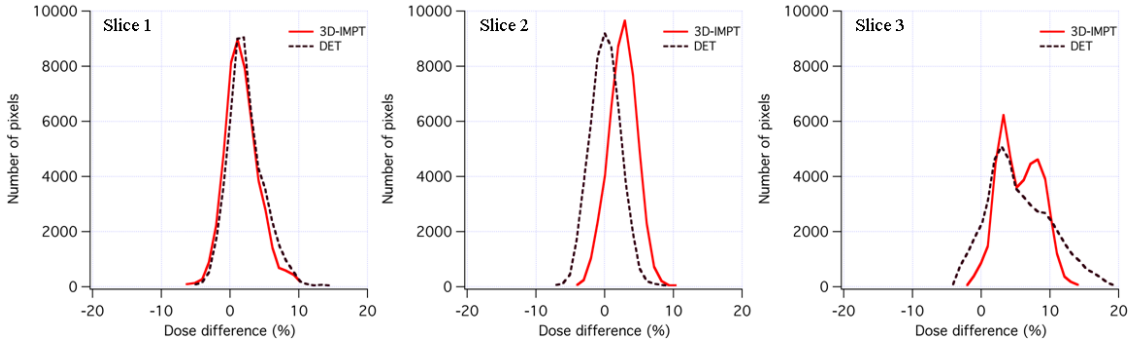


**Figure 6.11** Differential differences histograms between the doses measured with range errors included. Note how the curve for the 3D-IMPT (red solid line) plan is always narrow and closer to 0 than the one for the DET approach.

Moreover, by comparing the two IMPT plans, it has been found that the 3D-IMPT plan (red solid line in figure 6.11) is somewhat more robust than the DET approach (black dashed line), particularly in slice 3 (at the target edge). For two (out of three) slices, the red curve is indeed slightly narrower and closer to the "0 value" than the black one. The effect is however somehow reversed for slice 2, where although the curve for the 3D-IMPT plan is somewhat narrower than for the DET curve, it is systematically shifted about 3%. This could be related to a

systematic error that occurred in the measurement due, for example, to a difference in the film response.

Finally, Figure 6.12 shows the estimated Error-Bar distributions resulting from range errors, as described in section 6.2.5.2 for the slice positions in the phantom where the measurements have been made. Pleasingly, the shapes of the distributions are generally comparable with those presented in Figure 6.10 (see, for example, the slice 3 for the DET plan), thus indicating that our model for assessing plan robustness model provides, at least for this case, a reasonable estimate of the areas mostly affected by range uncertainties. The same held true when comparing Figure 6.11 with Figure 6.13, which displays the point-by-point difference histograms for both plans calculated from the distributions of Figure 6.12. As for the measurements, the calculated histogram resulting from the 3D-IMPT plan (red solid line) is indeed always narrower than that of the DET plan (black dashed line).

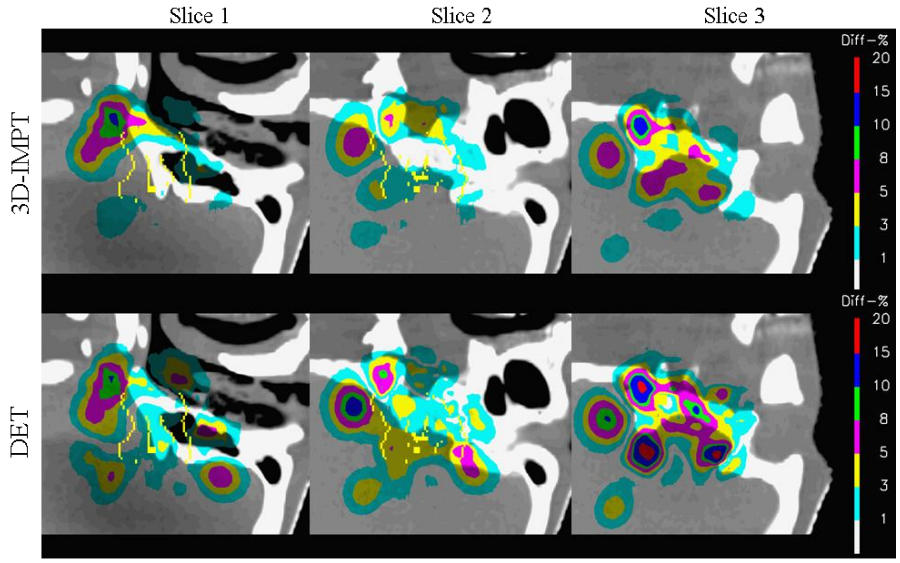


**Figure 6.12** Calculated error-bar dose distributions, for both IMPT plans for the three depths, for range uncertainties.Yellow contours represent the target volume.

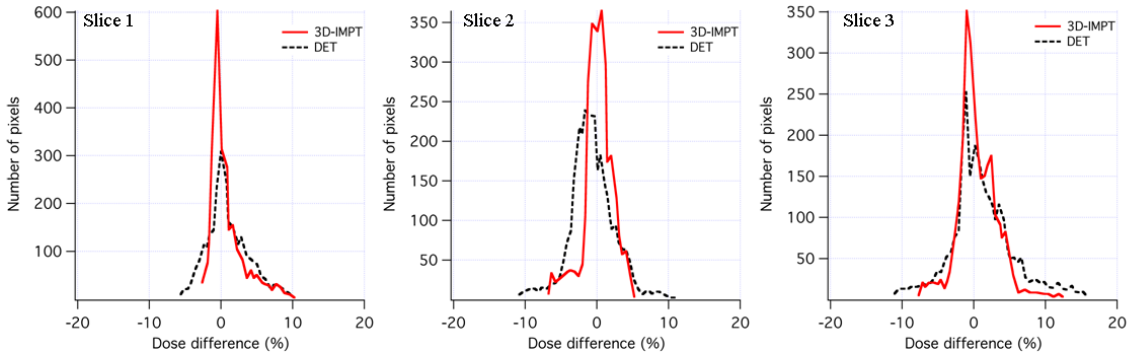


**Figure 6.13** Differential differences histograms for both the 3D-IMPT plan (red solid line) and the DET plan (dashed line) calculated from the distributions of figure 6.11.

### *6.3.2.2 Spatial Uncertainties*

Figure 6.14 shows similar comparisons for the three slices and both IMPT plans when the spatial uncertainties are introduced, i.e. the differences between the doses measured with a 3° of table rotations and the nominal doses. Results indicate, as before, that the 3D-IMPT plan is somewhat more robust than the DET plan, again particularly in slice 3, and that the doses in slice 1 are the most robust for both plans.

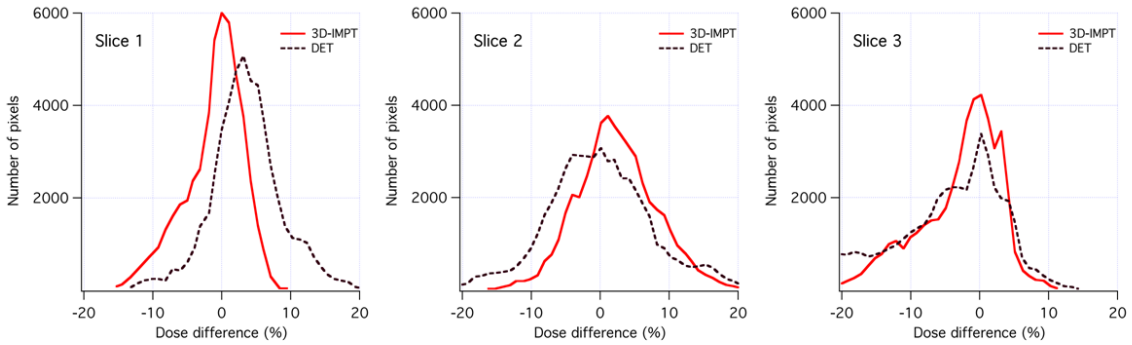


**Figure 6.14** Differential differences histograms between the dose measured with a spatial error (i.e. a 3° of table rotation) and the one measured in the absence of uncertainties for both IMPT plans. Note how the curve for the 3D-IMPT (red solid line) plan is always narrow and closer to 0 than the one for the DET approach.

## 6.4 Discussion

In this work we have verified IMPT plans delivered in an anthropomorphic phantom using GafCromic® film measurements and under different error conditions. In particular, we have determined the accuracy of the dose distributions delivered from highly in-homogeneous, 4-field 3D-IMPT and DET plans calculated with the standard field arrangements used clinically at PSI to treat tumors in the skull base region; and also experimentally compared the robustness of these plans to planned errors in the delivery of each plan.

In a specially developed anthropomorphic phantom, we have verified that, even in the presence of highly heterogeneous tissues (due to the presence of bony structures and air, see e.g. figures 6.1 and 6.4), the calculated dose is, at least relatively, a quite accurate predictor of the actually delivered dose with more than 99% of the points satisfying the gamma tolerance limits of 3%/3mm (Figure 6.7). Nevertheless, there are still small mismatches of the iso-doses contours visible in figures 6.7 and 6.8. As both IMPT plans were measured twice on different days, with excellent reproducibility of the results, these small differences may well originate from other effects, namely the film response and the nuclear interaction effects on the delivered dose distribution.

First, many authors have reported an under-response of GafChromic films in the Bragg peak region of up to 40% depending on the type of films used (Daftari et al 1999, Kirby 2010, Mumot et al 2009, Piermattei et al 2000, Vatnitsky et al 1997a,b). This is due to a quenching effect that occurs in areas with higher linear energy transfer along the incident particle track, as is the case in the Bragg peak region where the proton energy is very low. To properly deal with this effect it would be necessary to recalculate the dose after the introduction of a quenching correction factor in the depth dose curve, as suggested by Boon et al (2000) for a CCD/camera based dosimetry system which exhibited a similar effect. However, for all dose calculations performed here, we have used the pristine, unquenched depth-dose curve used in our treatment planning system (i.e that measured using a plane parallel ionization chamber). On the other hand, the effect of quenching is expected to mainly affect the measurements only in the distal region and not in the proximal and central regions, where the dose is primarily delivered by the plateau portion of the depth-dose curve and where generally no quenching effect is present. Indeed, Mumot et al (2009) recently measured a quenching effect of only about 12% in the Bragg peak for EBT films and below 5% in the plateau. In the same study, an under-dosage of only about 4% was reported in a SOBP field, due to the mix of proton energies contributing dose in the SOBP. Similarly for this work, as we have analyzed a full dose distribution achieved by a combination of many thousands of individually weighted Bragg peaks, incident from different fields, we assume that the quenching effect is to a large part reduced and smoothed out.

The second cause of differences could be due to dosimetric effects resulting from proton interactions with atomic nuclei (Pedroni et al. 2005). For the clinical dose calculation implemented at PSI, the effect of nuclear interactions of the incident protons is only dealt with in a one-dimensional manner and mainly through the reduction of the primary proton fluence as a function of penetration. In 2005, Pedroni described a new dose model which also more accurately takes into account the lateral beam halo of the secondary particles resulting from nuclear interactions (Pedroni et al 2005). Up to now however, this model has only been implemented as a way of calculating a global scaling factor to adjust the overall delivered monitor units on a field-by-field basis. Although such an approach is currently sufficient for most clinical applications, this dose ‘halo’ can nevertheless have local effects on the dose distribution, for instance by increasing the dose in the low dose tails of the distribution, as well as partially in dose ‘valleys’ in the middle of the distribution (see e.g. Bolsi et al 2005). Despite these limitations, the agreement between the dose calculation and film measurements for the nominal cases presented here are, however, extremely satisfactory.

An interesting consequence of these results is the small difference in the accuracy of the delivery between the 3D-IMPT and DET plans, with the agreement being somewhat better for the 3D-IMPT plan (see Figure 6.7 and Figure 6.8). The reason for this could be related to the number of Bragg peaks (spots) delivered with the two approaches. For instance, a single field of the DET plan consists of only about 180 beamlets, each of which is necessarily highly weighted (i.e. is delivered with a large fluence), which can be compared to the approximately 1500 beamlets delivered for each field of the 3D-IMPT plan (see scheme in Figure 6.3). Any uncertainty in the delivery process (due for example to small misalignments in the relative beamlet positions) could consequently have a larger impact on the DET than for the 3D-IMPT plan. Nevertheless, the dose measured for the DET plan also provides a more than satisfactory result (with 98% of points satisfying the gamma criteria when compared to the calculated dose) with the added benefit that the delivery time per field is reduced by about 30% in comparison to the 3D-IMPT plan, (35s vs 50s per field respectively).

In a second experiment, we have also compared the robustness of the 3D-IMPT and DET plans in the case of deliberately introduced spatial and range errors. For both types of errors, the tendency was for the 3D-IMPT plan to be more robust than the DET plan, particularly for the measurements performed at the periphery of the target volume (slice 3). This is connected to the position, in relation to the target volume, of the highly weighted Bragg peaks, which are more likely to be affected by uncertainties. These are mainly deposited, for each individual field of both plans, at the target edges (see the scheme in Figure 6.3), and thus, we would expect larger changes in the measured dose at the target periphery (slice 3) as a result of the delivery errors, particularly the range errors.

Additionally, and as hoped, both plans were found to be more robust to range than to the spatial (angular) uncertainties. As spatial uncertainties are likely to be random errors (i.e. will change day-to-day in a more or less random fashion throughout the treatment course), it is likely that the dose variations measured here would smooth-out through the treatment, assuming that there is no systematic component to the error. Range uncertainties on the other hand are very likely to be systematic in nature and could thus well be the same every day of treatment (Lomax 2008a). The fact that in particular the 3D-IMPT plan appears to be quite robust to these errors is an important result from the point-of-view of IMPT and its clinical applicability.

Moreover, the error based measurements performed here have provided a method by which the ‘error-bar’ robustness analysis tools described in chapter 5 can be validated experimentally. The rather good agreement between the measured differences of figure 6.10 and

6.12 shows the potential of this model for predicting the areas of a plan where dose uncertainties could be expected.

Finally, as described in section 6.2.2.1, range uncertainties have been here introduced through optimizing plans on two different CT data sets, with HU values amodified by +/-3%. This, if course, is not really the case for uncertainties during treatments. Despite this approach however, as the two 'error' plans resulted are extremely similar to each other (see Figure 6.5), we believe that, as a method to validate our error-bar concept described in chapter 5, our results are still valid. The measured dose distribution differences are indeed mainly due to the introduction of range errors. However, an alternative approach would have been to systematically change the energy of each delivered proton pencil beam in relation to the energy assumed for the nominal plan. This could possibly be done at the PSI Gantry1 system (Pedroni et al 2005) by simply adding or removing one of the polyethylene plates normally used to shift the proton range. However, although this solution possibly provides a better way to evaluate range uncertainties (as both plans will be optimized on the same CT data sets), with this approach only fixed range shifts of minimum 2.5mm of water can be simulated. Nevertheless, this is an approach which may be further investigated in the future.

Spatial uncertainties have here been measured assuming a rather simple, general rotation of 3° of the patient table. This is a valid first approach to simulate residual rotational positioning uncertainties which may occur assuming that the main translational errors have been corrected through the application of a daily patient positioning correction as described by Bolsi et al (2008). Nevertheless we intend to extend the validity of these results by measuring the dose in the anthropomorphic phantom also under different set-up uncertainties, in particular by varying the degree of rotations along the three major axis (i.e. antero-posterior, lateral-lateral and cranial-caudal), and by assuming residual translations. For this reason a mechanical motorized jig able to rotate the anthropomorphic phantom along three axes under computer control is currently under construction (see section 7.3 below).

## 6.5 Summary

In this chapter we have experimentally verified different 'flavors' of IMPT plans (namely 3D-IMPT and DET) using an anthropomorphic phantom customized to aid detailed measurements of complex proton treatments. We have verified the extremely high accuracy of the dose distribution delivered from a 4-field 3D-IMPT plan. Moreover we have verified that the dose calculated with the 3D-IMPT plan is somewhat more accurate than the corresponding one

achieved with the DET approach. Additionally we have experimentally measured the robustness to delivery uncertainties of both IMPT plans. The 3D-IMPT plan resulted to be more robust to both range and spatial uncertainties (at least in the form here implemented) and therefore confirmed to be, in general, a safer plan to deliver than the DET plan. Finally, experimental measurements confirmed (at least for range uncertainties) the validation of the theoretical method, introduced in Chapter 5, which can be used to calculate and evaluate the robustness of treatment plans for IMPT.

Note: This chapter has been published as part of the following paper: Albertini F, Casiraghi M, Lorentini S and Lomax AJ *Experimental verification of IMPT treatment plans in an anthropomorphic phantom in the presence of delivery uncertainties* in Phys. Med. Biol (56) 4415-31 (Albertini et al 2011b)

# 7 Future Work/ Further development

In this Chapter we would like to take the opportunity to mention some of the open questions raised during this PhD work and consequently of some areas of future development that will and could be pursued.

## 7.1 Controlling plan robustness

In this PhD work we have investigated the impact of varying the starting conditions on the quality and on the robustness of plans to both set-up and range errors. We have focused our attention either in varying the initial weight of pencil beams (see examples in Chapter 3.1 and Chapter 4) or in varying the beam configuration (see Chapter 3.2). However starting conditions used in the optimization process could be modified in many other ways, to affect (and possibly improve) the quality and the robustness of a plan. I will hereby mention only those that we are planning to investigate in the near future.

1. The 'Regions of Uncertainty' (ROU) concept. For many patients, changes in the anatomy can occur during the treatment course (e.g. changes in density of the nasal filling) which can have a significant impact on the range and therefore delivered dose to the patient. We plan to improve the robustness of plans to such uncertainties by the *a-priori* identification of such 'Regions of Uncertainty' (ROU's) by assigning them a level of 'proton transparency'. This transparency level will be used to reduce the initial weight (fluence) of pencil beams crossing these regions before starting the optimization process. For example, a transparency level of 0 would imply that all pencil beams passing through such a structure would be set a zero weight at the start, effectively rmoving them completely from the optimization. In contrast, a transparency level of 1 would have no effect on the pencil beam weight. Any value between 0 and 1, will reduce the weight of pencil beam accordingly. We

will then investigate if (and how) the plan quality (and its robustness to uncertainties) is affected by the introduction of this ROU and transparency concept.

2. Removal of spots from critical OAR's. Organs at risk (OARs) that are positioned at the distal end of the target volume are always quite critical when planning with protons as most highly weighted Bragg peaks are positioned at the distal end of the target volume to properly construct the spread out Bragg peak (SOBP) curve and it is here that often the steepest dose gradient can be acheived. Therefore any range uncertainty could potentially increase the dose drastically to these OAR's. Furthermore, the uncertainty in the definition of the Radio Biological Effect (RBE) in the distal fall-off of the Bragg-peak curve could also increase the effective biological dose in this part of the curve (see e.g. Paganetti et al 2002). Extending the concept of Regions of Uncertainty (ROU's) discussed above, we are planning to assign a weight of 0 (and therefore remove) all the highly weighted spots positioned distally against some defined OAR (e.g. the brainstem for children treated in the posterior fossa) before the optimisation. As for the point above, we will then investigate if the quality (and robustness to uncertainties) of the plan is affected by this change.
3. IMPT plan optimization using pre-optimised SFUD fields. As discussed in Chapter 4, starting the IMPT optimization algorithm with individual fields delivering a SOBP dose distribution can improve the quality of the plan and moreover its robustness to range errors. We want to further investigate how the quality (and robustness) of the IMPT plan can be improved if the fields are first pre-optimized to deliver an homogeneous dose distribution throughout the target volume. I.e. calculate SFUD type fields first and use the pencil beam weights of these fields as the starting conditions of the subsequent IMPT optimisation).
4. Optimisation of spot positions. Presently, in our treatment planning software, we calculate Bragg peak positions on a rectilinear grid, such that it is rare for spots to be delivered exactly at the edge of the target volume. Typically, we select Bragg peaks for the optimisation based on their position in relation to the target, by accepting all Bragg peaks completely within the target, and also those up to 5mm outside. However, as nearly all 'exterior' Bragg peaks are then delivered outside the target volume, this inevitably leads to more dose being delivered to the normal tissues outside the target volume, and leads to slower dose fall-offs. The aim here will be to see how we can improve dose fall-off outside the target by optimising Bragg peak positions on the exterior of the target such that they lie, as much as possible, exactly on the surface of the target volume and see how this conversely affects plan robustness.

## 7.2 Robustness analysis tool (Error-Bar Distributions)

The error-bar distribution tool described in Chapter 5 is a convenient and easy way to display (and directly compare) the robustness of different plans to both spatial and range uncertainties. We are planning to incorporate it as soon as possible in our clinical treatment planning system as we believe that the robustness analysis should be a standard toolset for any treatment planning system. Nevertheless, we are aware that bringing robustness deeply into the clinic (meaning with that, to be confident to affect a plan decision based on results of the robustness analysis) will be a learning curve for both physics and physicians, as there is no clinical experience of interpreting such displays.

As discussed in Chapter 5, the error-bar distributions for spatial uncertainties are applicable only to the error of delivery in a *single* fraction (assuming that spatial errors are randomly distributed over the treatment course). Although this information could be enough to directly compare the robustness to such uncertainties for two different plans, it will be perhaps even more useful to properly estimate the impact of these uncertainties on the whole treatment course for each plan. We are therefore currently working to extend this robustness analysis tool to be able to evaluate the error-dose distribution for the whole treatment course.

## 7.3 Anthropomorphic phantom developments.

### *7.3.1 Computer controlled rotational positioning jig.*

In the last year I have been involved with the development of a motorised jig for the accurate rotation about three axes of the anthropomorphic phantom presented in Chapter 6 (Figure 7.1). Currently, the mechanical components of the jig are in the construction phase and next year we will implement the on-line control of the system to steer the stepping motors remotely.

The idea behind the development of this jig is to be able to easily measure the impact of different residual set-up errors (due to a slight rotation of the patient) on the delivered dose distribution. This will add more information on the dosimetrical robustness of different plans to spatial misalignments (intra-fraction motion), as discussed in Chapter 6 for a single angle rotation. In addition, the capability to steer on-line (i.e. during the treatment) could be used for example to dosimetrically assess the effect of inter-fraction patient misalignments, which are achieved by assigning a different angle rotation to each field of the same treatment; Additionally, it could be used to emulate patient breathing motions. This can be achieved either just by rotating the phantom with a constant speed during the irradiation period, or by steering the phantom motion through a sinusoidal function, which is the most common way to simulate, in a rigid way,

the patient breathing pattern (see e.g. Knopf et al 2010). As for performing this measurement, it could be worth to substitute the anthropomorphic with a different phantom (for example of a patient's abdomen), the rotating jig has been designed to be as flexible as possible, by supporting different kind of phantoms

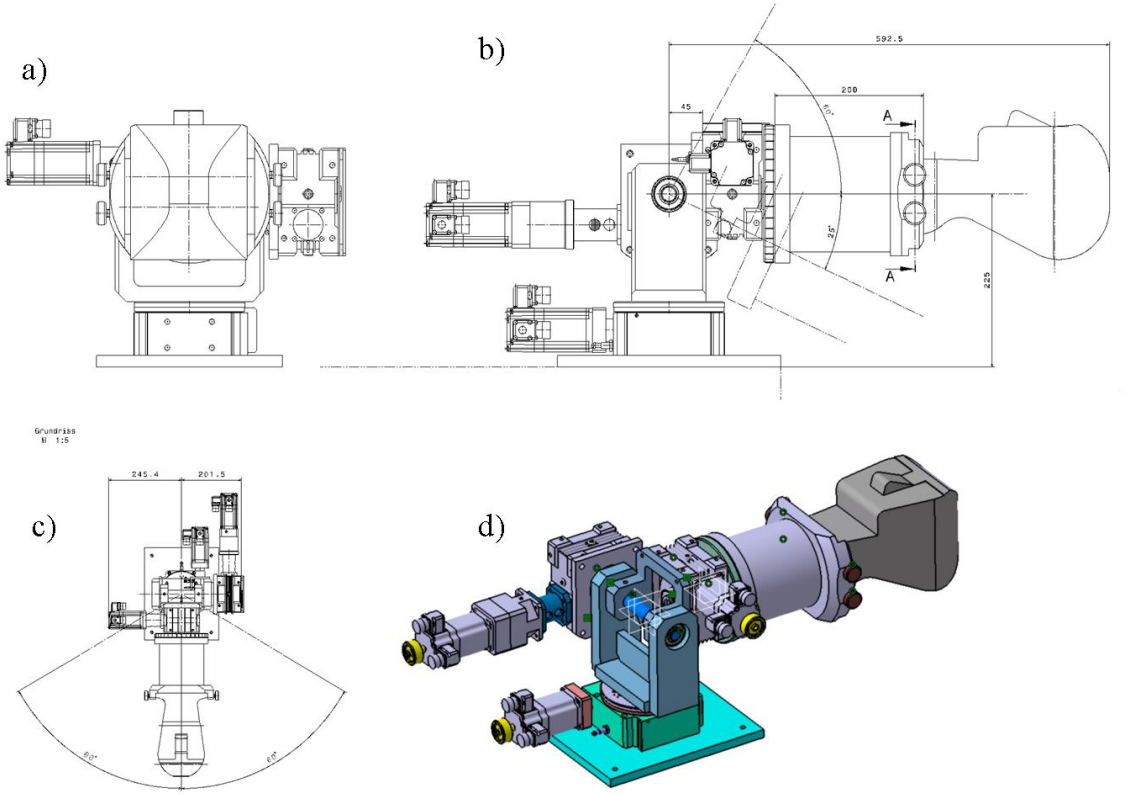


**Figure 7.1** a)-d) Back, lateral, frontal and a prospective view of the stepping motor developed to steer the anthropomorphic phantom to different angle off-sets. The angle freedom is visible on the corresponding views. Note: the support and the steering of the three stepping motors are currently under construction

### *7.3.2 Additional applications of the anthropomorphic phantom*

The anthropomorphic phantom developed in this PhD work will have a number of applications, with (and without) the use of the turning motor gear, for both clinical and research purposes in the near future. I hereby mention just those measurements that are either currently undergoing or that have already been scheduled.

1. Range probe measurements. The phantom is currently being used to experimentally assess the feasibility of the 'range probe' concept, as introduced by Mumot et al (2010) as an in-vivo measurement of range. The idea behind the 'range probe' is to directly assess the accuracy of range calculations by delivering a low-dose, high-energy proton pencil beam

that passes through the patient, and whose integral Bragg peak can be measured on the exit side using a multi-layer detector. The comparison between the measured integral Bragg peak with that calculated based on the patient's planning CT, will provide useful information about the accuracy of in-vivo range calculations. In addition, as the Bragg peak shape degrades depending on the density heterogeneities that the protons have passed through, it is hoped that the application of a few different range probes through well selected regions of the patient may allow for the quantitative assessment of the residual rotations of a patient's positioning. This hypothesis can be directly validated using the anthropomorphic phantom and the rotating jig currently being developed.

2. Clinical commissioning of Gantry2. The phantom will be used next year to run end-to-end tests for different IMPT plans during the commissioning phase of Gantry2, which is the second generation gantry of the PSI (Pedroni et al 2004).
3. VisionRT®. The phantom has already been used (and it will be re-used) to test the accuracy of the VisionRT® system, an optical surface imaging device used to measure patient position, motion and differences in the patient surface.
4. Dosimetry measurements in the presence of metal. The phantom will be used to dosimetrically assess the impact of the metal implants on proton dose distributions. There is clinical evidence indicating that patients treated in regions with metal implants (as a result of stabilizing surgery after tumour removal) have a less favorable prognosis in respect to patients treated without any implants (Rutz et al 2006, Staab et al 2011). Although there are many possible reasons for this, the presence of metal both degrades the CT data (causing significant streak artifacts) and presents an extremely sharp interface that could degrade the delivered Bragg peak shape (Tourovsky et al 2005).

Some work has been previously done at the Paul Scherrer Institute to estimate the effect of the metal artefacts on the quality of the delivered proton plan, by recalculating the plan on a 'artifact free' CT image (Albertini et al 2006). However, an in-vivo measurement in an anthropomorphic phantom will provide additional information that will help to understand the real impact of the metal itself and of the artifacts on the delivered dose distribution. For this purpose, the phantom has been customized with the insertion of clinically realistic metal implants. In particular we have inserted a 5mm tungsten sphere (red point in Figure 7.2b) in one of the lower molar tooth of the phantom to simulate the effect of a gold dental filling, being one of the most common sources of artifacts in CT images for patients treated in the head and neck region. In addition, we have inserted between C3-C7 (cervical vertebrae) a titanium implant (see Figure 7.2) composed of two

screws (each of 5mm thickness and 3cm length) and connecting rod (5mm diameter, 5cm length), to emulate a typical spine stabilization device implanted in patients undergoing extensive surgical tumor de-bulking. To properly evaluate the effect of the metal, the phantom has been sliced in the proximity of these metal implants. We are planning to measure the dose distribution in the proximity of metal, by using Gafcromic® films (as discussed in Chapter 6) and compare it with the calculated doses as one of the next projects with the phantom.

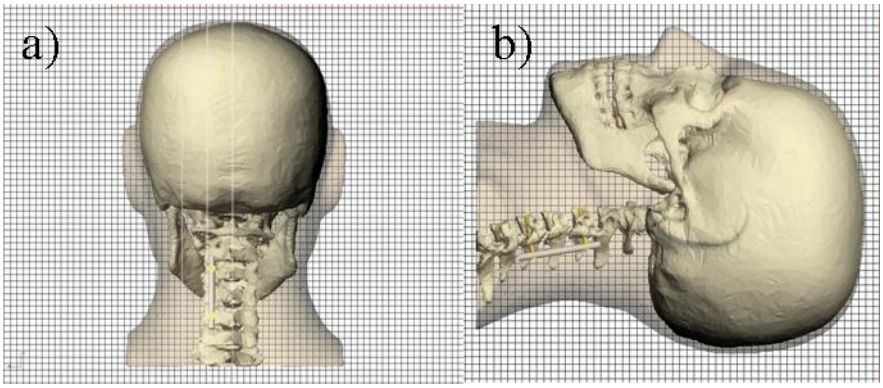


**Figure 7.2:** Back and lateral view of the anthropomorphic phantom customized at PSI with the support of the company CIRS. The white lines indicate the positions of the cranio-caudal cut of the slices. The titanium implant is visible in grey (rod) and yellow (the screws) and the tungsten filling of the tooth is in red.

# 8 Overall summary

In this PhD work we have studied in detail the planning and delivery of Intensity Modulated Proton Therapy (IMPT) plans from the point-of-view of plan degeneracy and the robustness of such plans to delivery uncertainties.

First, we have introduced the 'problem' of degeneracy for IMPT plans. Due to the high degree of freedom available and the basic goals assigned in the optimization process, the problem is extremely under-defined. Thus, there are very many different solutions providing very similar results, at least from the point-of-view of target coverage. However, when other, non-dosimetric characteristics are introduced, for example the concept of robustness, then the problem becomes more strictly defined and the degeneracy problem begins to be resolved. Moreover, as the use of a rather simple gradient based optimization algorithm provides solutions which are the closest to the starting conditions used, we have investigated how these starting conditions can be manipulated to 'steer' the result of the optimisation to solutions that are more robust (less sensitive) to delivery uncertainties. The aim here is to be able to provide 'class solutions' to the treatment planner so that depending on the position and shape of the tumour, particular starting conditions could be defined *a-priori* which are known to lead to more robust solutions. Nevertheless, we believe that such techniques can only be truly useful if complementary tools are provided for the analysis and evaluation of plan robustness. Thus, to efficiently assess the plan robustness of IMPT plans, we have introduced a new model based on multiple dose calculations under different error-conditions, and which collapses each error dose distribution into a single distribution, representing the two-sided error-bar of every point of the original, nominal dose distribution. We have further demonstrated that this method is a powerful tool to calculate, display and compare the plan robustness to uncertainties of different plans.

Second, we have experimentally measured two different types of IMPT plan, each optimized with different starting conditions, to validate the accuracy of the delivered dose and to directly measure their robustness to controlled delivery errors. Moreover, by comparing the measurements with the outcomes of the uncertainty model (i.e. the error-bar distributions), we have performed an initial validation of our model at least for the case of range uncertainties.

Finally, as part of the experimental validation of our IMPT calculations and robustness models, we have developed an anthropomorphic phantom which accurately mimics a real patient in the head and neck region and includes potential metal implants which can have a large impact on proton treatments. Together with a motorized jig currently under construction, this phantom will have a number of other applications within the proton project of PSI and we hope will attract interest from other proton therapy centers planning on treating patients using IMPT in the future.

# Acronyms

| | |
|---|---|
| 3D | 3 dimensions |
| 3D-IMPT | intensity modulated proton therapy plan, with spots positioned in 3D |
| 3F | 3 fields |
| 4F-Pi | 4 fields Π-arrangement |
| AP | antero-posterio |
| BS | brainstem |
| CC | cranio-caudal |
| CCD | charge coupled device |
| CT | computed tomography |
| CTV | clinical target volume |
| DET | distal edge tracking |
| DOF | degree of freedom |
| DRR | digitally reconstructed radiograph |
| DTA | distance to agreement |
| DVH | dose-volume histogram |
| $D_{98\%}$ | dose received by 98% of the volume |
| EUD | equivalent uniform dose |
| gEUD | generalized equivalent uniform dose |
| EVH | Error-Bar Volume Histograms |
| GPL | General Public Licence |
| GTV | gross tumor volume |
| $Gy_{RBE}$ | Grey, assuming a RBE of 1.1 for protons |
| H&N | head and neck |

| | |
|---|---|
| HU | Hounsfield unit |
| IDAS | IMRT Dose Analysis System |
| IMPT | intensity-modulated proton therapy |
| IMRT | intensity-modulated radiation therapy |
| LET | linear energy transfer (“stopping power”) |
| LR | left-right |
| LUT | look up table |
| MRI | magnetic resonance imagining |
| OAR | organ at risk |
| PAW | Physics Analysis Workstation |
| PET | positron emission tomography |
| PSI | Paul Scherrer Institute, Switzerland |
| PTV | planning target volume |
| QA | quality assurance |
| ROU | region of uncertainty |
| RBE | relative biological effectiveness |
| SC | spinal cord |
| SD | standard deviation |
| SFUD | single field, uniform dose |
| SP | stopping power |
| SOBP | spread-out-bragg-peak |
| TIFF | tagged imaged file format |
| TPS | treatment planning system |
| VOI | volume of interest |
| $V_{95\%}$ | % of the volume receiving 95% of the prescribed dose |
| $V40_{GyRBE}$ | % of the volume receiving 40 $Gy_{RBE}$ |
| WER | water equivalent range |

# References


Alber M, Meedt G, Nüsslin F 2002 "On the degeneracy of the IMRT optimisation problem" *Med. Phys.* **29** 2584-2589

Albertini F, Bolsi A, Ares C, Broggi S, Cattaneo GM, Lomax AJ 2006 "Advantage of using a MVCT for proton planning", proceeding of the 44th PTCOG conference

Albertini F, Lomax AJ, and Hug EB 2007 "In regard to Trofimov et al.: Radiotherapy treatment of early-stage prostate cancer with IMRT and protons: a treatment planning comparison" *Int J Radiat Oncol Biol Phys* **69**:444-453

Albertini F, Gaignat S, Bosshard M, and Lomax AJ 2009 "*Planning and Optimizing Treatment Plans for Actively Scanned Proton Therapy"* Biomedical Mathematics Promising Directions In Imaging, Therapy Planning, and Inverse Problems Y.Censor,M. Jiang,G.Wang, Editors

Albertini F, Hug EB, Lomax AJ 2010 "The influence of the optimization starting conditions on the robustness of intensity modulated proton therapy plans" *Phys. Med. Biol.* **55** 2863-2878

Albertini F, Hug EB, Lomax AJ 2011a "Is it necessary to plan with safety margins for active scanned proton therapy?" *Phys. Med. Biol.* **56**(14) 4399-413

Albertini F, Casiraghi M, Lorentini S and Lomax AJ 2011b "Experimental verification of IMPT treatment plans in an anthropomorphic phantom in the presence of delivery uncertainties" *Phys. Med. Biol* **56**(14) 4415-31

Bischel H 1972 "Passage of charged particles through matter *American Institute of Physics Handbook"*, 3d Ed Gray E.D, Ed (McGraw-Hill, New York) page 8

Bolsi A Lomax A J, Albertini F, Boehringer T, Coray A, Goitein G, Lin S, Pedroni E, Rutz HB, Timmerman B, Weber D 2005 "Nuclear interaction effects for IMPT plans" proceeding to the VII SASRO meeting (Basel)

Bolsi A Lomax A J, Pedroni E, Goitein G. Hug EB, 2008, "Experiences with a remote patient positioning procedure for high-throughput proton therapy". *Int J Radiat Oncol Biol Phys*. **71**(5) 1581-1590

Boon SN, van Luijk P, Böhringer T, Coray A, Lomax AJ, Pedroni E, Schaffner B and Schippers JM 2000 "Performance of a fluorescent screen and a CCD camera as a two dimensional dosimetry system for dynamic treatment techniques" *Med Phys*. 27:2198-2208

Bortfeld T, Buerkelbach J, Boesecke R and Schlegel W 1990 "Methods of image reconstruction from projections applied to conformation radiotherapy" *Phys. Med. Biol.* **35** 1423-1434

Bortfeld T, Kahler D L, Waldron T J, Boyer A L 1994 "X-ray field compensation with multileaf collimators". *Int. J. Radiat. Oncol. Biol. Phys.* **28** 723-730

Bortfeld T, Craft D, Dempsey JF, Halabi T and Romejin E 2008 "Evaluating target cold spots by the use of tail EUDs" *Int. J. Radiat. Oncol. Biol. Phys.* **71 (3)** 880-889

Bosshardt M, Lomax AJ. 2005 "*Optimising spot numbers for IMPT"*. ESTRO Physics Congress, Lisbon, Portugal, September 2005.

Chen G, Singh R, Castro J, Lyman J and Quivey J 1979 Treatment planning for heavy ion radiotherapy *Int. J.Radiat. Biol. Phys.* **5** 1809–19

Coray A.,Pedroni E., Böhringer T., Lin S., Lomax A., Goitein G 2002 "Dosimetry with the scanned proton beam on the PSI gantry, "Int. Symposium on Standards and Codes of Practice in Medical Radiation Dosimetry, IAEA, Vienna

Daftari I, Castenadas C, Petti PL, Singh RP. and Verhey LJ. 1999 "An application of Gafchromic Md-55 film for 67.5 MeV clinical proton beam dosimetry". *Phys. Med. Biol.* **44**, 2735–*2745*

Deasy, J.O., Shephard, D.M, Mackie, T.R. 1997 "Distal edge tracking: A proposed delivery method for conformal proton therapy using intensity modulation" In: Leavitt, D.D., Starkschall, G.S (Ed) *Proc XIIth ICCR (Salt Lake City)*. Medical Physics Publishing: Madison, Wisconsin, 406-409.

Drzymala R, Mohan R, Brewster L, Chu J, Goitein M, Harms W and Urie M 1991 Dose-volume histograms. *Int J Radiat Oncol Biol Phys* **21**:71–78

Ellerbrock M, Jaekel O, Kramer M, Nikoghosyan A, Schultz-Ertner D "Clinical implementation of intensity modulated heavy ion therapy", *Proc. 9th Biennial ESTRO meeting, Barcelona* 2007

Flynn RT, Kissick MW, Mehta MP, Olivera GH, Jeraj R, Mackie TR 2008 The impact of linac output variations on dose distributions in helical tomotherapy *Phys Med Biol*, **53** 417-430

Gensheimer MF, Yock TI, Liebsch NJ, Sharp GC, Paganetti H, Madan N, Grant PE, Bortfeld T.2010 "In vivo proton beam range verification using spine MRI changes". *Int J Radiat Oncol Biol Phys*. **78**(1):268-75

Goitein M 1983 "Nonstandard deviations" *Med Phys* **10**(5) 709

Goitein M. 1985 "Calculating of uncertainty in the dose delivered during radiation therapy". *Med Phys.* **12** 608-612

Goitein M, Lomax A.J. Lomax, Pedroni E 2002, "Treating cancer with protons" Phys. Today 45-50

Goitein M 2008 "*Radiation Oncology: a physicist's- Eye view"* Editor: Springer pages 113, 125, 217

Gonzales RC and Woods RE 1992 (p 199) "*Digital Imaging Processing"* Addison-Wesley Publishing Company

Hou Q, Wang J, Chen Y, Galvin J M 2003 "Beam orientation optimisation for IMRT by a hybrid method of the genetic algorithm and the simulated dynamics" *Med. Phys.* **30** 2360-2367

Hosang J, Chung H, Suh T-S, Palta J, Kim S, "*A new paradigm of IMRT plan evaluation with uncertainty volume histogram*," IFMBE Proceedings, 2007

ICRU report 50,"Prescribing, recording and reporting photon beam therapy" (1993)

Kanai T, Kanai K, Kumamoto Y, Ogawa H, Yamada T., Matsuzawa H 1980, "Spot scanning system for radiotherapy". *Med. Phys* (7):365--369

Kirby D, Green S, Palmans H, Hugtenburg R, Wojnecki C and Parker D 2010 "LET dependence of GafChromic films and an ion chamber in low-energy proton dosimetry" *Phys. Med. Biol.* **55** 417–33

Koehler A M, Schneider R J and Sisterson J M 1975 "Range modulators for protons and heavy ions *Nucl. Instrum. Methods.* **131** 437-440

Koehler A M, Schneider R J and Sisterson J M 1977 "Flattening of proton dose distributions for large field radiotherapy" *Med. Phys.* **4** 297-301

Knopf A, Bert C, Heath E, Nill S,Kraus K, Richter D, Hug E, Pedroni E, Safai A, Albertini F, Zenkluesen S, Boye D, Söhn M, Soukup M, Sobotta B, Lomax A 2010 "Special report workshop on 4D-treatment planning in actively scanned particle therapy: recommendations, technical challenges, and future research directions" *Med Phys* **37** (9) 4608-14

LLacer J, Deasy J, Bortfeld T, Solberg T D, Promberger C 2003 "Absence of multiple local minima effects in intensity modulated optimisation with dose-volume constraints" *Phys. Med. Biol.* **48** 183-210

Llacer J, Agazaryan N, Solberg T, Promberger C 2004 "Degeneracy, frequency response and filtering in IMRT optimization". *Phys. Med. Biol.* **49** 2853-2880

Lomax AJ, Pedroni E, Schaffner B. Scheib S, Schneider U, Tourovsky A 1996 "3D treatment planning for conformal proton therapy by spot scanning." *Proc. 19th L H Gray Conference* (BIR publishing London) 67-71

Lomax AJ 1999 "Intensity modulated methods for proton therapy" *Phys. Med. Biol.* **44** 185-205.

Lomax AJ 2001a "Methods for assessing the effects of inter- and intra-fraction delivery errors in radiotherapy", *ESTRO physics congress*, Seville.

Lomax AJ, Boehringer T, Coray A et al 2001b "Intensity modulated proton therapy: A clinical example" *Med. Phys.* **28** 317-324.

Lomax AJ, Pedroni E, Rutz HP, Goitein G 2004a "The Clinical Potential of Intensity Modulated Proton Therapy". *Z. Med. Phys.* **14** 147-152.

Lomax AJ 2004b '*Intensity Modulated Proton Therapy: The potential and the challenge'*, Habilitation thesis, Department of Physics, ETH, Zurich.

Lomax, A.J.,T. Böhringer, A. Bolsi, A. Coray, F. Emert, M. Jerman, S. Lin, E. Pedroni, H.P. Rutz, O. Stadelmann, B.Timmermann, J.Verwey, D.C.Weber, and G. Goitein. 2004c."Treatment planning and verification of proton therapy using spot scanning: Initial experiences."*Med Phys* **31**:3150–3157.

Lomax,A.J. 2008a. "Intensity modulated proton therapy and its sensitivity to treatment uncertainties 1:The potential effects of calculational uncertainties." *Phys Med Biol* **53**:1027–1042.

Lomax,A.J. 2008b. "Intensity modulated proton therapy and its sensitivity to treatment uncertainties 2:The potential effects of inter-fraction and inter-field motions." *Phys Med Biol* **53**: 1043–1056.

Lomax,A.J. 2008c "Intensity Modulated ProtonTherapy" in *Proton and Charged Particle Radiotherapy.* T. Delaney and H. Kooy (eds). Boston: Lippincott Williams & Wilkins.

Lomax A.J. 2009 "Charged particle therapy: the physics of interaction" *Cancer J.* **15**:285-291

Low DA, Harms WB, Mutic S and Purdy JA 1998 "A technique for the quantitative evaluation of dose distribution", *Med. Phys.* **25**(5):656-661

Maleike D, Unkelbach J, Oelfke U 2006 Simulation and visualization of dose uncertainties due to interfractional organ motion *Phys Med Biol* **51** 2237-2252

Martisikova M, Jäkel O 2010 "Dosimetric properties of Gafchromic EBT films in monoenergetic medical ion beams" *Phys Med Biol* 55:3741–51.

Matsinos E, Schaffner B and Kaissl W 2007 "Optimization Algorithms for Intensity Modulated Proton Therapy (IMPT)", AAPM Montreal

Moyers MF, Sardesai M, Sun S and Miller DW 2010 "Ion Stopping Powers and CT Numbers" *Medical Dosimetry*, **35** (3) 179-194

Mumot M, Mytsin GV, Molokanov AG, Malicki J 2009 "The comparison of dose measured by radiochromic films and semiconductor detector in a 75 MeV proton beam" Physica Medica **35** (3) 105-110

Mumot M, Algranati C, Hartmann M, Schippers JM, Hug E, Lomax AJ 2010 "Proton range verification using a range probe: definition of concept and initial analysis" *Phys Med Biol.* **55**(16):4771-82.

Mustafa A and Jackson D 1983 The relation between x-ray CT numbers and charged particle stopping powers and its significance for radiotherapy treatment planning *Phys. Med. Biol.* **28** 169–76

Negreanu C, Stadelmann O, Lomax A, "Robustness analysis study in fractionated proton radiotherapy", *Proc. ESTRO26*, Gothenberg, Sweden 2008

Niemierko A. 1997 "Reporting and analyzing dose distribution: a concept of equivalent uniform dose" *Med Phys* **24** 103-110

Niemierko A. 1999 "A generalized concept of equivalent uniform dose (EUD)" *Med Phys* [abstract] **26** 1100

Nill S, Bortfeld T and Oelfke U 2004 "Inverse planning of intensity modulated proton therapy" *Z. Med. Phys.* **14** 35-40

Oelfke U, Bortfeld T 2000 "Intensity modulated radiotherapy with charged particle beams: Studies of inverse treatment planning for rotation therapy" *Med Phys* **27** 1246-1257

Oelfke U and Bortfeld T 2001 "Inverse planning for photon and proton beams2 *Med Dos* **26 (2)** 113-124

Øverås H 1960 "Small angle multiple scattering in confined bodies" *CERN Yellow Report* 60-18

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

Parodi K, Enghardt W. 2000 "Potential application of PET in quality assurance of proton therapy". *Phys Med Biol*. **11**:N151-6.

Pedroni E, Bacher R, Blattmann H, Boehringer T, Coray A, Lomax A, Lin S, Munkel G, Scheib S, Schneider U and Tourovsky A 1995 "The 200 MeV proton therapy project at PSI: Conceptual design and practical realisation" *Med. Phys.* **22** 37-53.

Pedroni E., Bearpark R., Bohringer T., Coray A., Duppich J., Forss S., George D., Grossmann M., Goitein G., Hilbes C., Jermann M., Lin S., Lomax AJ, Negrazus M, Schippers M., Kotle G 2004 "The PSI Gantry 2: a second generation proton scanning gantry". *Z Med Phys* **14** 25-34

Pedroni E, Scheib S, Boehringer T, Coray A, Lin S, Lomax AJ 2005 "Experimental characterization and theoretical modeling of the dose distribution of scanned proton beams: the need to include a nuclear interaction beam halo model to control absolute dose directly from treatment planning". *Phys. Med. Biol.* **50** 541-561.

Pflugfelder D, Wilkens JJ, Nill S, Oelfke U 2008a "A comparison of three optimization algorithm for intensity modulated radiation therapy" *Z Med Phys.* **18** 111-119

Pflugfelder D "Risk adapted optimization in intensity modulated proton therapy" (IMPT) PhD dissertation 2008b, Heidelberg

Pflugfelder,D., J.J.Wilkens, and U.Oelfke. 2008c."Worst case optimization:A method to account for uncertainties in the optimization of intensity modulated proton therapy." *Phys Med Biol* 53:1689–1700

Piermantei A *et al* 2000 "Radiochromic film dosimetry of a low energy proton beam" *Med. Phys.* **27** 1655–60

Polf JC, Peterson S, McCleskey M, Roeder BT, Spiridon A, Beddar S, Trache L. 2009 "Measurement and calculation of characteristic prompt gamma ray spectra emitted during proton irradiation". Phys Med Biol. **54**(22):N519-27

Rutz HP, *et al* 2008 "Postoperative spot-scanning proton radiation therapy for chordoma and chondrosarcoma in children and adolescents: initial experience at paul scherrer institute" *Int J Radiat Oncol Biol Phys*. **71**(1):220-5

Rutz HP, Weber DC, Sugahara S, Timmermann B, Lomax AJ, Bolsi A, Pedroni E, Coray A, Jermann M, Goitein G 2007 "Extracranial chordoma: outcome in patients treated with function preserving surgery followed by spot-scanning proton beam irradiation" *Int J Radiat Oncol Biol Phys*. **67**(2):512-520

Safai S, Shixiong L and Pedroni E, 2004 "Development of an inorganic scintillating mixture for proton beam verification dosimetry", *Phys. Med. Biol.* **49** 4637-4655

Salk J, Kosta M. Blank P, Röttinger E.M 2003 "IDAS An Extensible Framework for IMRT Verification" proceeding of the 7th Biennal ESTRO Meeting on Physics and Radiation Technology (Geneve)

Sheib, S. Pedroni E, 1992 "Dose calculation and optimization for 3D conformal voxel scanning" Radiot. Environ. Biophy., **31**, 251-256

Schneider U, Pedroni E. 1995 "Proton radiography as a tool for quality control in proton therapy", Med. Phys. **22** (4) 353-363

Schneider U, Pedroni E, Lomax A. 1996 "The calibration of CT Hounsfield units for radiotherapy treatment planning" Phys. Med. Biol. **41** 111-124

Schaffner B and Pedroni E, 1998 "The precision of proton range calculations in proton radiotherapy treatment planning: experimental verification of the relation between CT-HU and proton stopping power" Phys. Med. Biol. **43** 1579-1592

Schaffner B, Pedroni E., Lomax A, 1999 2Dose calculation model for proton treatment planning using a dynamic beam delivery system: an attempt to include density heterogeneity in the analytical dose" Phys. Med. Biol. **44** 27-41

Staab A, Rutz HP, Ares C, Timmermann B, Bolsi A., Albertini F, Lomax AJ, Goitein G, Hug EB 2011 "Spot scanning based proton therapy for extra-cranial chordoma" *(submitted)*

Soukup M, Söhn M, Yan D, Liang J, and Alber M 2009 "Study of the robustness of IMPT and IMRT for prostate cancer against organ movement" *IJR BP* **75(3)** 941–949

Thieke C, Bortfeld T, Niemierko A, Nill S 2003 "From physical dose constraints to equivalent uniform dose constraints in inverse radiotherapy planning" *Med. Phys.* **30 (9)** 2332-39

Tourovsky A, Lomax AJ, Schneider U, Pedroni E 2005 "Monte Carlo dose calculations for spot scanned proton therapy" *Phys.Med.Biol* **50** 971-81

Trofimov A, Nguyen PL, Coen JJ, Doppke KP, Schneider RJ, Adams JA, Bortfeld TR, Zietman AL, DeLaney TF, Shipley WU 2007 "Radiotherapy treatment of early-stage prostate cancer with IMRT and protons: a treatment planning comparison" *Int. J. Radiat. Oncol. Biol. Phys* 69(2) 444-453

Unkelbach J.,T.C.Y. Chan, and T. Bortfeld. 2007. "Accounting for range uncertainties in the optimization of intensity modulated proton therapy." *Phys Med Biol* 52:2755–2773

Unkelbach J, Bortfeld T, Martin BC, and Soukup M 2009 "Reducing the sensitivity of IMPT treatment plans to setup errors and range uncertainties via probabilistic treatment planning" *Med Phys* **36** 149-163

van Herk M, Reneijer P, Rasch C, Lebesque JV 2000 "The probability of correct target dosage:dose-population histograms for deriving treatment margins in radiotherapy" *Int J Radiat Oncol Biol Phys*. **47**(4) 1121-1135

Vatnitsky S 1997a "Radiochromic film dosimetry for clinical proton beams" *Appl. Radiat. Isot.* **48** 643–51

Vatnitsky SM, Schulte RWM, Galindo R, Meinass HJ. and Miller D.W. 1997b "Radiochromic film dosimetry for verification of dose distributions delivered with proton-beam radiosurgery" *Phys. Med. Biol.* **42** 1887–98

Webb S 2003 "The physical basis of IMRT and inverse planning". *Br. J. radiol.* **76** 678-689

Wilson,R.R. 1946."Radiological use of fast protons." *Radiology* 47:487–491

Wu Q and Mohan R 2000 "Algorithms and functionality of an intensity modulated radiotherapy optimization system" *Med. Phys.* **27** 701-11

Wu Q Mohan R, Niemierko A, Schmidt-and Ullrich R. 2002 "Optimization of intensity-modulated radiotherapy plans based on the equivalent uniform dose" *Int. J.Radiat. Oncol. Biol. Phys.* **52** 224–35

Zhao L and Das IJ 2010 "Gafchromic EBT film dosimetry in proton beams" *Phys. Med. Biol.* **55** N291–N301