EPJ manuscript No.
(will be inserted by the editor)

Treatment planning in proton therapy

Marco Schwarz[1]

Agenzia Provinciale per la Protonterapia, Via Perini 181, 38122 Trento.

Received: date / Revised version: date

Abstract. Radiotherapy treatment planning is a procedure that, using radiation beam and patient's anatomy models as input data, produces as output the machine instructions to deliver the treatment and the expected dose distribution in the patient. Now that most proton therapy centers are moving from double scattered proton beams to active delivery systems such as pencil beam scanning (PBS), there is a need for treatment planning tools that could generate safe and effective dose distribution by taking full benefit of the potential of PBS degrees of freedom, and by avoiding the risks associated to this modality. The paper provides an overview of the current status of proton treatment planning techniques, from the creation of a patient model via imaging, to dose calculation, to the optimization of plans using intensity modulated proton therapy (IMPT). The issue of plan sensitivity to input data ('plan robustness') is emphasized and current approaches to robust optimization are presented. Finally, current developments in 'adaptive planning' and in the plan design for moving organs are shortly discussed.

PACS. PACS-key describing text of that key – PACS-key describing text of that key

1 Introduction

Radiotherapy treatment planning can be described as the procedure that, taking as input a model of both the radiation beam and the patient anatomy, produces as output two kinds of information: machine instructions to deliver the treatment (such as beam energy, beam shape and number of protons to be delivered in each beam) and the expected dose distribution in the patient, which allows to quantify the probability of tumor control and of complications to the normal tissues. Although treatment planning procedures have some common characteristics for different radiation types (e.g. photons vs. heavy charged particles), the use of protons has a number of specific implications when it

Send offprint requests to: schwarz@atrep.it

comes to designing and optimizing treatment parameters. In essence, the same reasons why protons are so interesting for radiotherapy applications (i.e. the finite range and the steep fall-off in dose deposition after the dose maximum) are also the reasons why it is so critical (and non trivial) to achieve a good correspondence between planned and delivered dose. With photon radiotherapy (XRT) one can safely assume that the dose distribution is invariant for small misalignments between beam and patient anatomy due to either positioning errors or small anatomy changes. With protons this assumption does not hold anymore: significant differences between planned and delivered dose can occur in case the relative position of beam and patient anatomy is different to what has been set during planning. This paper presents how these issues are dealt with in proton therapy and describes the current developments in the field of treatment planning. Although most proton treatments are now performed with scattered beams, the emphasis here will be on planning techniques for active beam delivery, like pencil beam scanning (PBS), which is likely to become the dominant proton therapy modality in the near future, and which represents the field where most developments in treatment planning are happening nowadays.

2 Creating a patient model

In general, the use of images in treatment planning does not depend on the kind of radiation that will be used during treatment and all procedures are based on a Computed Tomography (CT) dataset of the anatomical region to be treated. CT data, in addition to allow the visualization of the internal anatomy, provide a 3D map of patient's electron density, which is needed in order to calculate the beam attenuation, and thus the dose distribution within the patient. With protons there are at least two issues to be considered more carefully than with photons:

CT artefacts CT hardware and software are designed to perform at best in the most common situation, i.e. in presence of biological tissues. The presence of high Z material in the patient, e.g. due to gold fillings or hip prostheses, causes 'streaking artifacts' due to the strong absorption of X-rays in metal and a cut-off for large Hounsfield units (HU) in the software of most CT scanners, which may seriously deteriorate the image quality (see an example in fig. 1). This issue is particularly critical when metal implants are close to the treatment volume. The artifacts not only affect the possibility of identifying the anatomical structures of interest, but they also introduce errors in the electron density map of the patient, which will lead to errors in the estimated proton range in the patient. The dosimetric impact of streaking artefacts can be large (see e.g. [1]) and the presence of high Z material should, therefore, be handled with care. Software correction methods at the level of CT image generation have been proposed (e.g. [2,3]) and are now on the verge of becoming commonly available in clinical practice.

CT to stopping power calibration The usual CT calibration procedure produces a quantitative relation between Hounsfield units and electron density (or proton stopping power). In the calibration procedure, tissue substitutes are used to mimic biological tissues (e.g. soft tissue, lung, bone etc.). The choice of the specific tissue substitute (i.e. of its chemical composition) used for the calibration has an effect on the calibration curve. A new general stoichiometric calibration method has been proposed, which has been shown to be more accurate for both photon, and, especially, protons([4]). The stoichiometric calibration is based on the principle that the composition of each material, in particular in terms of the Z of its components, will influence the relative importance of the physical process involved in the absorption of a X-Ray in the 100 keV region, i.e. photoelectric effect, coherent and incoherent scattering. A proper chemical and physical characterization of the irradiated tissues is, of course, of particular interest for a Monte Carlo proton dose calculation (see e.g. [5]), which makes it possible to calculate the dose to medium instead of the dose to water. All uncertainties due to an imperfect estimation by CT imaging of beam attenuation through the patient contribute to what is referred to as 'range uncertainty'. In passive scattering techniques (see sec 4), beam attenuation devices are usually designed to produce safe plans against a $\pm$ 3% range uncertainty due to CT imaging. Given the quality of modern CT scanners and image reconstruction algorithms, this value may overestimate the actual errors occurring in most clinical situations, but it's still quite commonly applied. A whole different issue is the extent to which the planning CT scan is representative of the patient's anatomy throughout the treatment cycle, which typically lasts several weeks. While there are treatment sites, like the skull, where it is safe to assume that the anatomy will not be affected by changes, in other regions both physiological and pathological processes can lead to significant changes in patient anatomy and, as a consequence, in how the beam is going to be attenuated in the patient. In principle, anatomy changes (due e.g. to weight loss) can be considered as just another source of range uncertainty; in practice, current methods of compensating for range uncertainties are very simplistic, as they assume that the risk of uncertainty is solely a function of the nominal proton range at a given point. This is clearly insufficient to take into account the complexity of anatomy changes, which, therefore, should be handled in a different way (see sections 5.1 and 6.1 for more details).

3 Dose calculation

The development of dose calculation models has been an area of great interest for medical physicists, and proton therapy is not an exception. Several algorithms have been, and are still being, proposed, and a thorough review of their characteristics is beyond the scope of this paper. The current standard for both scattered and scanning beams is represented by pencil beam algorithms of different accuracy/complexity in the physical model (see e.g. ([6–9]). A

pencil-beam algorithm calculates the delivered dose via a generalized expression such as

$$Dose(x, y, z) = DD(z) \times T(x, z, \sigma_x(z)) \times T(y, z, \sigma_y(z)) \quad (1)$$

where z is the depth in medium along the beam direction, while x and y are the transversal directions. DD, i.e. the depth dose, can be obtained via measurements in a water phantom (this is the most common approach, in particular in commercial treatment planning systems), via a combination of first principles and data fitting ([7]), via analytical approximations[10] or via Monte Carlo calculations (see an example in fig2). While for scattered proton beams DD is measured with a small chamber on the beam central axis, in case of scanning beams an integral dose is obtained by measuring the whole dose envelope of a proton pencil beam with a sufficiently large detector (e.g. a 8-10 cm diameter ionization chamber). The transversal dose distribution is calculated starting from measured data about the pencil beam dimension, shape, divergence and direction. Some of these properties (e.g. pencil beam size) may vary according to the proton range, so in this case the energy dependence should also be characterized via experimental data. The beam broadening in patient is then calculated by modelling the result of the relevant physical processes (mainly multiple coulomb scattering and nuclear interactions), in most cases via two weighted components each described by a Gaussian distribution. If the beam has to traverse highly heterogeneous regions (e.g. the head and neck region), accurate beam tracing through the anatomy is needed: in fact, different section of the same pencil beam may interact with different material (e.g. air as opposed to bone), causing the dose distribution for the single pencil beam to be significantly different from the case of an homogenous media. To address this issue, approaches such as 'spot decomposition' have been proposed ([8]), i.e. a method allowing multiple ray-tracing for a single pencil beam for the sake of accurate pathlength and scattered dose calculation within highly spatially varying anatomies. Monte Carlo-based algorithms are in principle the gold standard for dose calculation, because of a superior capability of simulating the actual physical interaction processes (see e.g. ([11–13]); however, Monte Carlo codes in the daily routine are still an undelivered promise, due to calculation times that are often incompatible with the clinical needs. Things may change in the near future for at least two reasons; 1)The continuous increase in computer power and the interest for new computing approaches such as GPU-based calculation ([14, 15]); 2)Proton therapy with active beam delivery is likely to require a simpler Monte Carlo model of the incoming beam than treatments with scattered beam: a scattered beam interacts with much more material than a pencil beam before entering the patients (e.g. scatterers, modulating wheel, aperture, compensator), thus being more difficult to model.

4 Planning for beam scattering techniques

This paper is focused on planning techniques for PBS, but a few aspects of planning scattering techniques are nonetheless worth mentioning. With scattering, the crossplane dose distribution is controlled by shaping a customized aperture, while dose conformity at the distal end of the target is achieved via a customized compensator. A compensator is a device (typically made of plastic material) of variable thickness, which allows to compensate for different radiological depths traversed by the beam at different regions within the field. By compensating them, one can assure that the highest beam energy applied to the patient will conform around the distal end of the target along the beam direction. In scattering, the dosimetric effect of geometrical uncertainties is managed by enlarging apertures and 'smearing' compensators in such a way that the target will still receive an appropriate dose even in presence of small misalignments between beam and patient anatomy [16,17]. In analogy with the way photon therapy planning copes with geometrical uncertainties (see e.g. [18,19]), also in protons the concept of planning target volume (PTV) is often adopted in order to make sure that the target will receive a sufficient dose even in presence of geometrical uncertainties (i.e. positioning errors and organ motion). However, with protons one needs a field-specific PTV to take into account that the dosimetric effects of range uncertainties occur only along the beam direction, i.e. they are different from beam to beam. Proton therapy dose distributions obtained with scattering techniques have three main characteristics that make them different to what can be achieved with PBS:

1. Since scattering dose distributions are delivered via broad beams as opposed to pencil beams, the length of the Spread Out Bragg Peak (SOBP), i.e. the size of the high dose region along the beam direction, is fixed within a field. This means that dose conformity at the proximal end of the target can not be explicitely controlled.
2. Irradiation fields are mostly designed to deliver a homogeneous dose to the target volume. The exception is represented by the so-called 'patched' fields, where a quasi-homogeneous dose distribution is achieved via two or more fields delivering full dose to a part of the volume and no dose elsewhere. Thinking in terms of dose modulation, one can say that with patching only 'binary modulation' (i.e. either full dose or no dose) can be achieved within a field.
3. Since the proton range is modulated within a field via a fast-rotating wheel, for practical purposes one can assume that all points within an irradiated volume are irradiated at all times while the beam is on. In other words, patient motion on the timescale of a few seconds (e.g. breathing) does not introduce 'interplay effects' [20,21] with beam delivery.

5 Plan optimization and evaluation

Pencil beam scanning (PBS) is a beam delivery modality with a large number of degrees of freedom, since one can control position, energy and number of protons for every single pencil beam, whose size can get as small as 3mm (1 sigma). PBS treatment planning therefore implies pencil beam weight optimization, where the number of protons associated with each pencil beam is adjusted in order to achieve the best compromise between target coverage and organs at risk sparing. Since the pencil beams are typically in the order of one thousand, this process heavily relies on what is referred to as 'inverse planning' or, perhaps more appropriately, treatment plan optimization. For those who are familiar with photon radiotherapy, it may be important at this moment to clarify the terminology. While in photons 'plan optimization' uniquely refers to 'intensity modulated radiotherapy'(IMRT) (a technique where the photon fluence within a beam is not constant), the situation in protons is different. According to the commonly accepted proposal by Lomax[22], one should refer to 'intensity modulated proton therapy' (IMPT) as a planning and delivery approach where the final dose distribution (not the fluence) is the result of the contribution from different fields, each delivering an inhomogeneous dose distribution in the target. According to this definition, an active beam delivery system does not necessarily imply IMPT, nor does the use of computerized optimization to determine the pencil beams weight. For instance, plan optimization would still be needed if one decides to generate with PBS a uniform dose distribution from each field. This is what is usually referred to as 'single field uniform dose' (SFUD). PBS-based proton therapy is based on the minimization of a function that, at least in principle, quantifies all relevant aspects of the desired dose distribution and specifies the tradeoffs (e.g. between target volume coverage and organs at risk sparing) one is willing to accept, given that the ideal dose distribution (prescription dose in the target and no dose elsewhere) is physically unachievable. The most commonly applied optimization scheme can be described as

$$minimize f(d(x)), \, where \, x \geq 0 \tag{2}$$

where x is the vector of pencil beam weights, d is the dose vector, obtained as the sum of dose vectors associated with each pencil beam, and f is the total cost function. A clinically realistic optimization problem consists of several conflicting objectives, and they are typically represented in f as a weighted sum of cost functions, i.e.

$$f(x) = \sum_i w_i c_i \tag{3}$$

where w_i is the weight of the c_i cost function, which represents a specific treatment intent. According to the terminology proposed by Kessler[23], each c_i, whose numerical value is named 'costlet', has two components: a *parameter*, i.e. a metric quantifying a property of the dose distribution (e.g. the minimum dose in a given volume), and a *modifier*, i.e

a function associating a cost to the difference between the desired and the actual value of the parameter. In the vast majority of optimization schemes currently applied, the modifier assumes a quadratic form and the parameter is a purely dosimetric property, such as

$$c_{overdose} = \frac{1}{N}\sum_{i=1}^{N}\left(D_i - D_0\right)^2 \Theta\left(D_i - D_0\right) \tag{4}$$

$$c_{underdose} = \frac{1}{N}\sum_{i=1}^{N}\left(D_i - D_0\right)^2 \Theta\left(D_0 - D_i\right) \tag{5}$$

where N is the number of volume elements ('voxels') in the region of interest, D_0 is the prescribed dose and D_i is the dose in the i-th voxel. As one might expect, $c_{overdose}$and $c_{underdose}$ are cost functions applied to limit overdosage (in the target or in an organ at risk) and underdosage in the target, respectively. Recently, there has been increased interest for so called 'biological optimization', where the modifier describes a quantity more closely linked to the clinical response of an organ than the physical dose distribution within a single voxel of a region of interest. As a consequence, metrics such as the generalized Equivalent Uniform Dose (gEUD)[24] are included in the cost function, and they better describe the probability of complication of a healthy organ than a single physical dose index. gEUD is typically defined as

$$gEUD = \left(\frac{1}{N}\sum_{i=1}^{N} D_i^{1/n}\right)^n \tag{6}$$

where N is the number of voxels in the volume of interest, D_i is the dose in the i-th voxel and n is the so-called 'volume effect parameter', i.e. a measure of the capability for an organ to compensate the loss of functionality caused by the partial irradiation at high doses as long as the rest of the volume is kept at low enough doses.

5.1 'Robust' optimization

IMPT planning is often based on the use of PTV, which is an appropriate tool to generate and describe dose distributions that take into account geometrical uncertainties as long as

- The PTV margins are appropriately set given the level of uncertainties one wants to be protected against
- The dose distribution in the PTV is homogeneous
- The dose distribution in the PTV is invariant for uncertainties whithin the level one wants to be protected against

Unfortunately, the last assumption is in general not true with protons, raising the question wheter PTV-based planning is appropriate in IMPT. It is true that the effect of range uncertainties and setup errors is an issue regardless whether scattering or PBS is applied; pencil beam scanning, however, is potentially more affected by this problem for two main reasons:

1. For any given IMPT field, dose heterogeneities are present not only at the edges of the target volume but also within the target itself; while with scattering and SFUD range uncertainties and setup errors are in general going to result in dose fluctuations at the target edges, IMPT dose distributions are based on the assumption that anywhere within the target a homogeneous dose may be achieved via the sum of (highly) modulated fields.
2. Since spot weights are defined via computer optimization, and since cost function are composed only by dosimetric parameters (see e.g. eqs. 4 and 6), anything that will go in the direction of minimizing the residual cost will be 'accepted' by the minimization algorithm. As long as a cost component related to range uncertainty is not present in the cost functions, there is no way to explicit control the effect of range uncertainty in an IMPT dose distribution.

At the moment, the potential dosimetric errors due to the application of IMPT are mostly being handled in an implicit (or defensive) way, e.g. preferring SFUD when IMPT is not strictly needed and/or applying IMPT to anatomical sites where range uncertainties are less likely to happen and/or trying to use optimization starting conditions to limit the degree of dose modulation within a field [25].

In addition to that, an appropriate selection of beam number and directions, which is performed manually prior to pencil beam weight optimization, may greatly help in achieving robust plans. For instance, looking at a Bragg peak shape one may think that the obvious choice is to set a beam in such a way to spare the most critical organ with the distal dose fall-off, even if it is just 2-3 mm from the target volume. Actually, this is exactly what does *not* happen in nearly all clinical situations, due to the potential serious overdosage in the organ at risk due to a small underestimation in proton range. Another choice that typically helps robustness consists in avoiding directions where the beam traverses highly heterogeneous regions (e.g. the nasal cavities). Numerical methods have been developed to support the choice of beam angle, e.g. via the calculation of an 'heterogeneity number' which quantifies the degree of heterogeneity lateral to the beam direction [26].

Still the problem remains of how to 'steer' the optimization of IMPT plans taking into account additional parameters other than dosimetric indices on a nominal plan, and better optimization approaches are needed for IMPT to become a treatment option for most treatment sites. The issue of assessing and ensuring plan 'robustness' has then become a 'hot topic'. Qualitatively speaking, plan robustness can be defined as a metric quantifying to what extent a delivered dose distribution will change with respect to the planned dose distribution when either positioning errors and/or range errors and/or anatomy changes occur during the treatment course. From the optimization point of view, one can see robustness as an issue related to uncertainty in the input data. This uncertainty should be taken into account in the cost function, and different approaches have been developed in the recent years to accomplish that. These robust optimization methods can be divided into two broad categories:

– Approaches assuming previous knowledge of the probability distribution of errors. Most proposal for robust optimization in XRT belong to this category. A robust plan may then be obtained either via coverage probability approaches (see e.g. [27,28]), where each voxel has an assigned probability of presence for the target or an OAR, or via the optimization of the expectation value of tumor control and normal tissue complication probabilities [29]. This approach to robust optimization is in principle the most accurate and it is computationally affordable in photon therapy, when one can assume dose invariance for small translations. For proton therapy, robust optimization based on probabilistic planning has been proposed by Unkelbach et al[30]. In this setting, where only systematic errors due to set-up errors and range uncertainties are taken into account, the range of each pencil beam is considered to be a random variable, i.e.

$$\rho_j = \rho_j^n + \delta_j \sigma_J \tag{7}$$

Where ρ_j is the actual range, ρ_j^n is the nominal range value, σ_i is the magnitude of uncertainty and δ_j is the random variable. Additional hypothesis are then made on the correlation of uncertainties between ρ_js either in the same or in different beams, and (rigid) setup errors are expressed as a gaussian variable independent on range errors. Under this assumption, the quantity to be minimized is not a nominal plan, but rather the expected value of the residual cost over the possible values of range uncertainties and setup errors weighted according to their probability. The nominal plan can be thus interpreted as the robust plan where the probability of the nominal configuration is equal to one. By 'switching' on and off range uncertainties and setup errors, Unkelbach et al could clearly demonstrate that including information on the uncertainties drives the optimization towards different configurations of spot weights: for instance, the introduction of range uncertainties will cause the optimization to achieve organs at risk sparing more with the lateral penumbra than with the distal dose fall-off.

– Approaches that do not make assumptions of the probability distribution of errors. Two recent proposals of robust optimization in proton therapy [31,32] will be shortly described here to illustrate the features of this optimization approach. Plugfelder and colleagues implemented a so-called 'worst case optimization', initially proposed by Lomax et al[33] and applied to the evaluation of plan robustness (e.g [34]). The basic concept of worst case optimization is the following: via the simulation of range uncertainties and positioning errors one can obtain for each voxel the worst case scenario (e.g. minimum dose for a target voxel and maximum dose for an OAR voxel). The dose distribution composed by all worst case voxels is then included in the cost function, that becomes

$$\widetilde{F}(x) = F(D_{nom}(x)) + p_w F(D_w(x)) \tag{8}$$

where, x is the vector of spot weights, D_{nom} is the nominal dose distribution, D_w the worst case dose distribution and p_w a weighting factor to control the importance of the worst case dose distribution in the final result. The worst case optimization has evident advantages, being manageable from the computational point of view and being capable of providing safe dose distributions. On the other hand there are also disadvantages. First, the worst case dose distribution may be (and typically is) unphysical, because the weight configuration that creates the worst dose in the i-th voxel may or may not be the same that creates the worst dose in the j-th voxel. This means that the worst case optimization has the risk of being overly conservative, thus deteriorating the nominal dose distribution beyond what is actually needed. Second, p_w does not have an immediate physical meaning (how should one weight in the cost function the importance of an unphysical dose distribution?) and may become a parameter which is manipulated somewhat arbitrarily to steer the optimization results on a case-by-case basis. Fredriksson et al recently proposed to achieve plan robustness through minimax optimization [32]. The main difference with respect to the work of Pflugferlder et al is that only physically realizable dose distributions are considered, i.e .the correlation between uncertainties in different voxels is taken into account. As a result, dose distributions tend to achieve a better compromise between target coverage and organs at risk(OAR) sparing than alternative methods such as SFUD plans or plans obtained by overriding the CT number of low density tissue (e.g. lung).

It goes without saying that, regardless of how a robust plan is obtained, when uncertainties are included in the cost function the final result will be a tradeoff between different scenarios and therefore robustness will be in general associated to some deterioration in quality of the nominal plan. Thanks to the degeneracy of the solutions in IMPT, the deterioration is not necessarily significant. Clinical experience on how to weight nominal quality of a plan vs. robustness is largely missing, so the introduction of robust planning in the clinic will probably require an initial phase of *in silico* trial and error in order to appreciate feature and shortcomings of the new methods.

5.2 Dose reporting

One purpose of treatment planning is to accurately quantify the dose delivered to the patient, thus providing input data to the studies on dose–effect relations. In XRT, the common practice still consists of using the nominal target and OAR doses input for such studies. To what extent is this likely to change in the near future, both in photon and proton therapy? The emphasis placed throughout this paper on the effect of geometrical uncertainties on proton dose distributions might have convinced the reader that reporting the dose in terms of *static* dose values measured in an expansion of target volumes or organs at risk is not a good representation of the actual dose distribution delivered to the patient. If proton therapy aims at becoming a treatment option for all the lesions currently treatable with XRT, the

planning procedure should handle set-up errors and organ motion in a more appropriate way also in the phase of plan evaluation and dose reporting. In XRT there are examples showing the potential of probabilistic dose reporting (see e.g.[35]), i.e. an approach where the results of treatment planning are presented in terms of probability distributions of relevant dosimetric indices in a population. As a consequence, the dose in the target is not reported as, say, the minimum dose in the PTV, but as a probability that the minimum dose in the CTV will be greater or equal than a threshold. The computational burden needed to calculate the probability distributions is usually manageable with current desktop computers, so there are no technical obstacles for the introduction of probabilistic dose reporting in clinical practice. Another field where probabilistic dose reporting should become common practice is the comparison between proton and photon dose distributions. Treatment planning studies comparing protons and XRT are carried out quite often these days, and so far it is common practice to compare target dose distributions using the PTV, which might overestimate the benefits of protons in case the PTV is defined in the same way as for photon therapy, or it might underestimated them in case it is defined too conservatively. The use of probabilistic dose reporting would lead to more fair comparisons between the two radiotherapy techniques.

6 Where is (proton) treatment planning going?

6.1 Treatment planning in the radiotherapy workflow

Up to a few years ago, the radiotherapy workflow was quite straightforward: via a computed tomography (CT) scan, a 3D anatomical model of the patient was created, becoming the input for treatment planning. The planning procedure would produce as output the irradiation parameters, the 3D dose distribution and the reference images for patient positioning. The patient would then be treated for the prescribed number of fractions, his correct position with respect to the beam being periodically verified by comparing the reference images with X-ray projections taken in treatment position. More recently, the treatment workflow has become more complex, mainly due to the increasing amount of imaging data available both before and during the treatment cycle. It is now quite normal that, beside conventional CT imaging, either additional anatomical data (e.g. via magnetic resonance imaging (MRI)) or functional data (mostly via PET[36]) are used to create an accurate patient model. In addition, the availability of in-room volumetric imaging (e.g. via Cone-beam CT (CBCT)[37]) made it practical to monitor patient anatomy during the radiotherapy cycle and to decide whether the plan should be changed. Thanks to this develoment, the concept of Adaptive radiation Therapy (ART)[38] is now more and more applicable, in particular when aggressive hypofractionation is pursued and the 'plan of the day' becomes more a necessity than an interesting concept. Plan adaptation requires the availability of

both new planning tools (e.g. for non–rigid image, contour and dose registration) and, if applied online, a significant speed up of plan optimization and dose calculation. Most of the developments in the field of ART took place in XRT, and proton therapy has to catch up with this respect. On the one hand, in-room imaging prior to treatment should soon become a standard in protons too; on the other hand, again, in proton therapy one can not work under the general assumption hold by most ART protocols in photon therapy, i.e. that dose distributions are invariant under small geometrical and/or anatomical changes, so fast dose recalculation becomes a must.

6.2 Planning in the presence of intrafraction motion

Even the forefront of research in proton therapy treatment planning so far assumes that in most case a single representation of patient anatomy via CT is sufficient, and even the most advanced robust optimization schemes are often unable to include metrics concerning organ motion or anatomy changes in the cost function. This is one of the reasons why, at the moment, PBS-based proton therapy is typically not applied in presence of (second-to-second) organ motion, like in breathing. Being able to treat moving targets with PBS/IMPT is not entirely a planning issue, since beam delivery characteristics such as fast energy changes may be needed, but treatment planning needs to make steps ahead nonetheless. The methods to handle interplay effects (see example in figure 3) are essentially two: target tracking [39], where beam parameters are adjusted online to reflect target motion, and (volumetric) rescanning [40], where the statistical independence of beam delivery from respiratory motion is used in order to mitigate the effects of interplay by rescanning the treatment volume several times for every fraction and every field. Regardless which method is chosen, tracking or rescanning strategies should be evaluated first during planning. This means that treatment planning systems should provide the capability of time-resolved dose calculation based on a description of respiratory motion and beam delivery time structure. In addition, breathing and other kind of organ motion have a timescale that may vary from seconds to minutes/hours[41] to days/weeks[42]. Each of this motions requires appropriate imaging to be quantified and appropriate planning tools to combine the information from all patient models and produce a final cumulative dose distribution.

6.3 Radiobiological effectiveness

When proton plans are designed in clinical practice, it is still recommended that the radiation biological effectiveness (RBE) is set at a constant value of 1.1 for all treatments independent of dose, fraction size, position in the spread-out Bragg peak (SOBP), initial beam energy or tissue type [43]. There are indications that the value of RBE might

increase towards the end of range up to 1.3-1.4, but the relatively limited extent of variation and the uncertainty in the estimates *in vivo* has not yet lead to a change in the clinical practice. A recent study analyzed the possibility of taking advantage of the increased linear energy transfer (LET) of protons at the end of range to increase the effectiveness of treatments [44]. Although the increase of LET for protons while slowing down is nowhere near what is seen in heavier ions, and although the LET can not be translated to RBE in a straightforward way, the authors could demonstrate that different IMPT delivery schemes (i.e. 3D-IMPT and Distal Edge Tracking [22]) producing basically the same physical dose distribution are associated to different LET distributions, both in the target and in the OARs. The potential clinical implications of these differences are still to be explored, but, at least in principle, it is feasible to perform what the authors name 'biologically motivated optimization'.

7 Conclusions

Several developments occurred in the recent past in the field of proton treatment planning, but some problems have yet to be solved. Intensity modulated radiotherapy with protons is still rarely applied in the clinical practice, in part due to the difficulty of planning safe and robust dose distributions for disease sites where geometrical uncertainty is an important issue. The increasing interest in proton therapy is likely to lead in the next few years to further significant developments in treatment planning, both in research and in clinical practice.

8 Acknowledgments

Thanks to Valentina Piffer for her language editing of the manuscript.

References

1. Jikun Wei, George a Sandison, Wen-Chien Hsi, Michael Ringor, and Xiaoyi Lu. Dosimetric impact of a CT metal artefact suppression algorithm for proton, electron and photon therapies. *Physics in medicine and biology*, 51(20):5183–97, 2006.
2. CS Olive, MR Kaus, V Pekar, K Eck, and Spies L. Segmentation- aided adaptive filtering for metal artifacts reduction in radio-therapeutic ct images. In *Proceeding of SPIE 5370*, pages 1991–2002, 2004.
3. M. Bal and L. Spies. Metal artifact reduction in ct using tissue-class modeling and adaptive prefiltering. *Medical physics*, 33:2852, 2006.
4. U Schneider, E Pedroni, and A Lomax. The calibration of CT Hounsfield units for radiotherapy treatment planning. *Phys. Med. Biol*, 41:111–124, 1996.

5. H. Jiang, J. Seco, and H. Paganetti. Effects of hounsfield number conversion on ct based proton monte carlo dose calculations. *Medical physics*, 34:1439, 2007.

6. L Hong, M Goitein, M Bucciolini, R Comiskey, B Gottschalk, S Rosenthal, C Serago, and M Urie. A pencil beam algorithm for proton dose calculations. *Phys. Med. Biol*, 41:1305–1330, 1996.

7. E Pedroni, S Scheib, T Böhringer, a Coray, M Grossmann, S Lin, and a Lomax. Experimental characterization and physical modelling of the dose distribution of scanned proton pencil beams. *Physics in Medicine and Biology*, 50(3):541–561, 2005.

8. M Soukup, M Fippel, and M Alber. A pencil beam algorithm for intensity modulated proton therapy derived from Monte Carlo simulations. *Phys. Med. Biol*, 50:5089–5104, 2005.

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

10. T Bortfeld. An analytical approximation of the Bragg curve for therapeutic proton beams. *Med. Phys.*, 24:2024–2033, 1997.

11. M Fippel and M Soukup. A Monte Carlo dose calculation algorithm for proton therapy. *Medical Physics*, 31:2263–2273, 2004.

12. H. Paganetti, H. Jiang, S.-Y. Lee, and H. M. Kooy. Accurate Monte Carlo simulations for nozzle design, commissioning and quality assurance for a proton radiation therapy facility. *Medical Physics*, 31(7):2107, 2004.

13. A Tourovsky, A J Lomax, U Schneider, and E Pedroni. Monte Carlo dose calculations for spot scanned proton therapy. *Physics in Medicine and Biology*, 50:971–981, 2005.

14. Xun Jia, Xuejun Gu, Josep Sempau, Dongju Choi, Amitava Majumdar, and Steve B Jiang. Development of a GPU-based Monte Carlo dose calculation code for coupled electron–photon transport. *Physics in Medicine and Biology*, 55(11):3077–3086, June 2010.

15. A. Badal and A. Badano. Accelerating monte carlo simulations of photon transport in a voxelized geometry using a massively parallel graphics processing unit. *Medical Physics*, 36:4878, 2009.

16. M. Urie, M. Goitein, and M. Wagner. Compensating for heterogeneities in proton radiation therapy. *Physics in Medicine and Biology*, 29:553, 1984.

17. M.F. Moyers, D.W. Miller, D.A. Bush, and J.D. Slater. Methodologies and tools for proton beam design for lung tumors. *International Journal of Radiation Oncology* Biology* Physics*, 49(5):1429–1438, 2001.

18. ICRU. Prescribing, recording and reporting photon beam therapy (ICRU report 50). *ICRU report*, pages 357–360, 1993.

19. Marcel van Herk. Errors and margins in radiotherapy. *Semin Radiat Oncol*, 14:52–64, 2004.

20. M H Phillips, E Pedroni, H Blattman, T Boehringer, A Coray, and S Scheib. Effects of respiratory motion on dose uniformity with a charged particle scanning method. *Phys. Med. Biol.*, 37:223–234, 1992.

21. T Bortfeld, K Jokivarsi, M Goitein, J Kung, and S B Jiang. Effects of intra-fraction motion on IMRT dose delivery: statistical analysis and simulation. *Phys. Med. Biol.*, 47:2203–2220, 2002.

22. A Lomax. Intensity modulation methods for proton radiotherapy. *Phys. Med. Biol*, 44:185–205, 1999.

23. M L Kessler, D L McShan, M A Epelman, K A Vineberg, A Eisbruch, T S Lawrence, and B A Fraass. Costlets: A generalized approach to cost functions for automated optimisation of IMRT plans. *Opt. Eng.*, 2005.

24. A Niemierko. A generalized concept of equivalent uniform dose (EUD). *Med. Phys.*, 26:1100, 1999.

25. F Albertini, E B Hug, and A J Lomax. The influence of the optimization starting conditions on the robustness of intensity-modulated proton therapy plans. *Physics in medicine and biology*, 55(10):2863–2878, April 2010.

26. D. Pflugfelder, JJ Wilkens, H. Szymanowski, and U. Oelfke. Quantifying lateral tissue heterogeneities in hadron therapy. *Medical physics*, 34:1506, 2007.

27. J C Stroom, H C de Boer, H Huizenga, and A G Visser. Inclusion of geometrical uncertainties in radiotherapy treatment planning by means of coverage probability. *Int. J. Radiat. Oncol. Biol. Phys.*, 43:905–919, 1999.

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

29. Marnix G. Witte, Joris van der Geer, Christoph Schneider, Joos V. Lebesque, Markus Alber, and Marcel van Herk. IMRT optimization including random and systematic geometric errors based on the expectation of TCP and NTCP. *Medical Physics*, 34(9):3544, 2007.

30. Jan Unkelbach, Thomas Bortfeld, Benjamin C Martin, and Martin Soukup. Reducing the sensitivity of IMPT treatment plans to setup errors and range uncertainties via probabilistic treatment planning. *Medical physics*, 36(1):149–63, January 2009.

31. D Pflugfelder, J J Wilkens, and U Oelfke. Worst case optimization: a method to account for uncertainties in the optimization of intensity modulated proton therapy. *Physics in medicine and biology*, 53(6):1689–700, March 2008.

32. A. Fredriksson, A. Forsgren, and B. Hårdemark. Minimax optimization for handling range and setup uncertainties in proton therapy. *Medical Physics*, 38:1672, 2011.

33. Antony J Lomax, Eros Pedroni, Hanspeter Rutz, and Gudrun Goitein. The clinical potential of intensity modulated proton therapy. *Z Med. Phys.*, 14:147–152, 2004.

34. a J Lomax. Intensity modulated proton therapy and its sensitivity to treatment uncertainties 1: the potential effects of calculational uncertainties. *Physics in medicine and biology*, 53(4):1027–42, 2008.

35. M Schwarz, J der Geer, M Van Herk, J V Lebesque, B J Mijnheer, and E M F Damen. Impact of geometrical uncertainties on 3D CRT and IMRT dose distributions for lung cancer treatment. *International Journal of Radiation Oncology, Biology, Physics*, 65:1260–1269, 2006.

36. Ursula Nestle, Wolfgang Weber, Michael Hentschel, and Anca-Ligia Grosu. Biological imaging in radiation therapy: role of positron emission tomography. *Physics in Medicine and Biology*, 54(1):R1–25, January 2009. PMID: 19060363.

37. D. Létourneau, J.W. Wong, M. Oldham, M. Gulam, L. Watt, D.A. Jaffray, J.H. Siewerdsen, and A.A. Martinez. Cone-beam-ct guided radiation therapy: technical implementation. *Radiotherapy and Oncology*, 75(3):279–286, 2005.

38. D Yan, F Vicini, J Wong, and A Martinez. Adaptive radiation therapy. *Phys. Med. Biol.*, 42:123–132, 1997.

39. C. Bert, N. Saito, A. Schmidt, N. Chaudhri, D. Schardt, and E. Rietzel. Target motion tracking with a scanned particle beam. *Medical physics*, 34:4768, 2007.

40. SM Zenklusen, E. Pedroni, and D. Meer. A study on repainting strategies for treating moderately moving targets with proton pencil beam scanning at the new gantry 2 at psi. *Physics in medicine and biology*, 55:5103, 2010.

41. M von Siebenthal, G Székely, U Gamper, P Boesiger, A Lomax, and P Cattin. 4D MR imaging of respiratory organ motion and its variability. *Physics in Medicine and Biology*, 52:1547–1564, 2007.

42. Jan-Jakob Sonke, Joos Lebesque, and Marcel van Herk. Variability of four-dimensional computed tomography patient models. *International journal of radiation oncology, biology, physics*, 70(2):590–8, 2008.

43. ICRU. *Prescribing, recording, and reporting proton beam therapy (ICRU report 78)*. ICRU, 2007.

44. C. Grassberger, A. Trofimov, A. Lomax, and H. Paganetti. Variations in linear energy transfer within clinical proton therapy fields and the potential for biological treatment planning. *International Journal of Radiation Oncology* Biology* Physics*, 80(5):1559–1566, 2011.

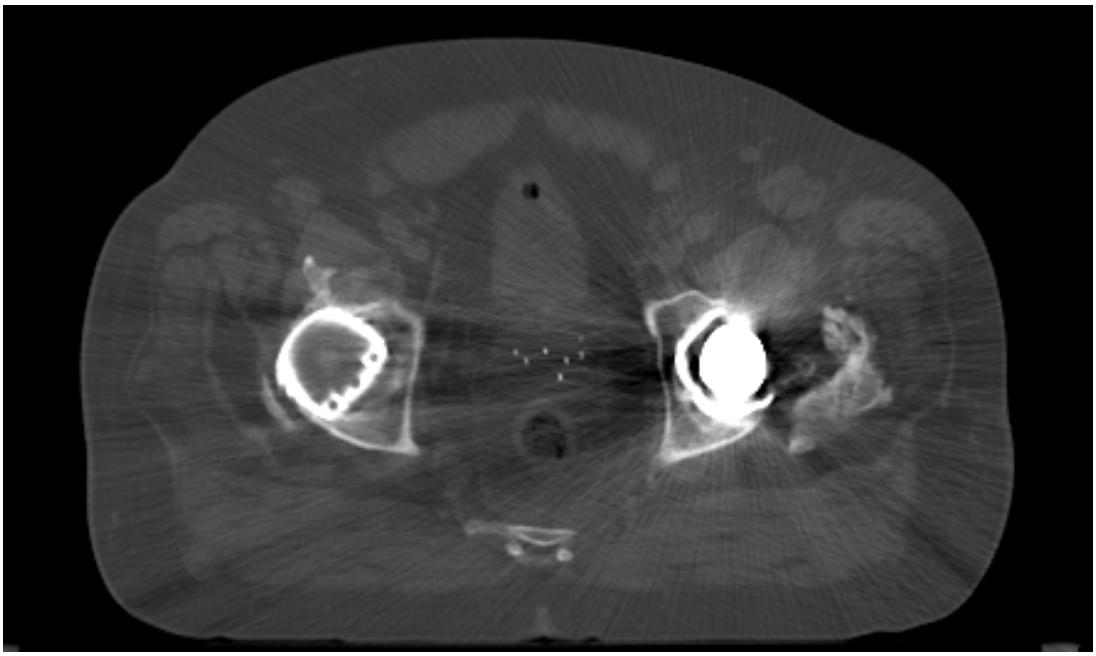

Fig. 1. CT image with artifacts due to a hip prosthesis.

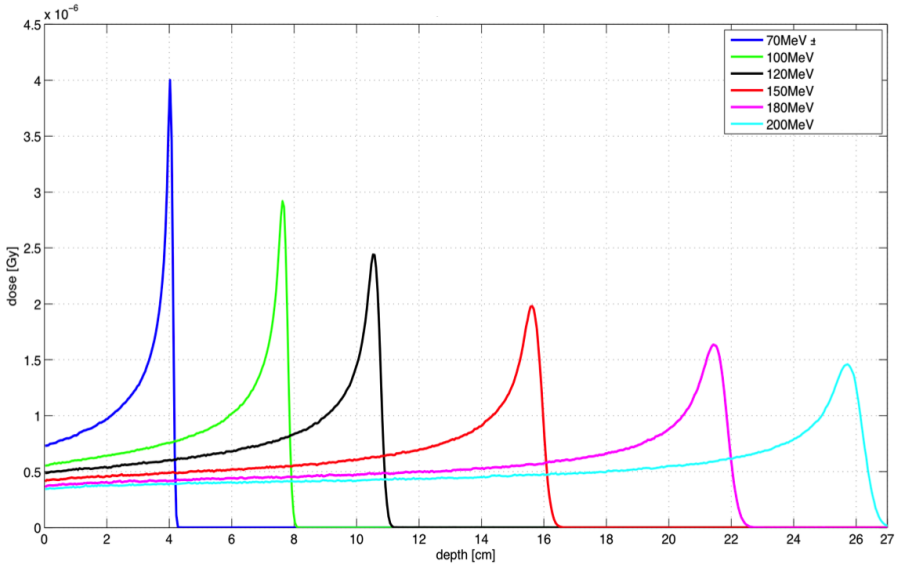

Fig. 2. Example of 'pristine' Bragg peaks based on Monte Carlo calculations to be used as input data for beam modelling and dose calculation

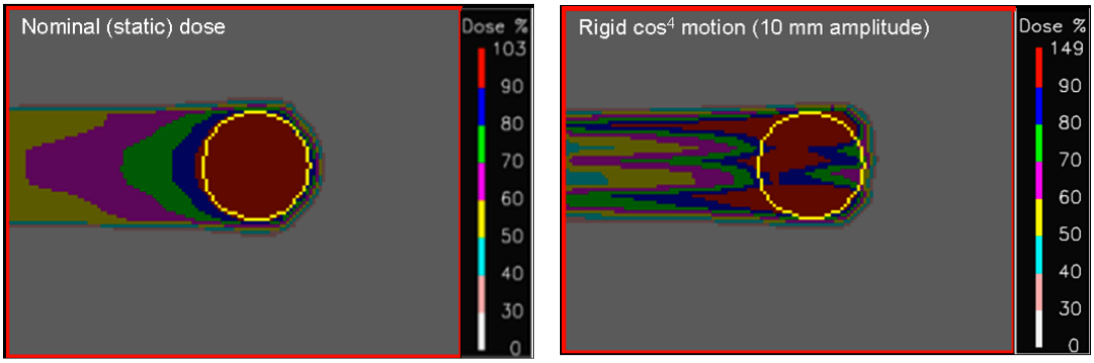

Fig. 3. Example of interplay effect. On the left hand side the nominal dose distribution for a static object is shown. One the right hand side the resulting dose distribution where the time structure of beam delivery is taken into account and a 10 mm peak-to-peak motion of the target is assumed.