IOPscience iopscience.iop.org

This content has been downloaded from IOPscience. Please scroll down to see the full text.

Download details:

IP Address: 18.117.165.66
This content was downloaded on 25/04/2024 at 07:39

Please note that terms and conditions apply.

You may also like:

Artificial Intelligence in Radiation Therapy

Monte Carlo Calculations in Nuclear Medicine (Second Edition)

Modern Applications of 3D/4D Ultrasound Imaging in Radiotherapy

Special section: Selected papers from the Fourth International Workshop on Recent Advances in Monte Carlo Techniques for Radiation Therapy
Jan Seuntjens, Luc Beaulieu, Issam El Naqa et al.

The impact of motion on onboard MRI-guided pencil beam scanned proton therapy treatments
Alisha Duetschler, Sairos Safai, Damien C Weber et al.

Characteristics of fiber-optic radiation sensor for passive scattering proton beams
J. Son, M. Kim, J. Jeong et al.

# Proton Beam Therapy

**Harald Paganetti** describes the use of proton beams in the treatment and management of cancer.

physicsworld Discovery

# Proton Beam Therapy

# Proton Beam Therapy

**Harald Paganetti**
*Director of Physics Research, Massachusetts General Hospital, Boston, MA, USA*
*Professor of Radiation Oncology, Harvard Medical School, Boston, MA, USA*

**IOP** Publishing, Bristol, UK

© IOP Publishing Ltd 2017

All rights reserved. No part of this publication may be reproduced, stored in a retrieval system or transmitted in any form or by any means, electronic, mechanical, photocopying, recording or otherwise, without the prior permission of the publisher, or as expressly permitted by law or under terms agreed with the appropriate rights organization. Multiple copying is permitted in accordance with the terms of licences issued by the Copyright Licensing Agency, the Copyright Clearance Centre and other reproduction rights organisations.

Permission to make use of IOP Publishing content other than as set out above may be sought at permissions@iop.org.

Harald Paganetti has asserted his right to be identified as the author of this work in accordance with sections 77 and 78 of the Copyright, Designs and Patents Act 1988.

ISBN 978-0-7503-1370-4 (ebook)

DOI 10.1088/978-0-7503-1370-4

Version: 20170101

Physics World Discovery
ISSN 2399-2891 (online)

British Library Cataloguing-in-Publication Data: A catalogue record for this book is available from the British Library.

Published by IOP Publishing, wholly owned by The Institute of Physics, London

IOP Publishing, Temple Circus, Temple Way, Bristol, BS1 6HG, UK

US Office: IOP Publishing, Inc., 190 North Independence Mall West, Suite 601, Philadelphia, PA 19106, USA

*To Anne and Enzo:*

*Thanks for brightening up my days!*

# Contents

# Abstract

Cancer therapy is a multi-modality approach including surgery, systemic or targeted chemotherapy, radiation (external beam or radionuclide), and immunotherapy. Radiation is typically administered using external beam photon therapy. Proton therapy has been around for more than 60 years but was restricted to research laboratories until the 1990s. Since then clinical proton therapy has been growing rapidly with currently more than 50 facilities worldwide. The interest in proton therapy stems from the physical properties of protons allowing for advanced dose sculpting around the target and sparing of healthy tissue. This review first evaluates the basics of proton therapy physics and technology and then outlines some of the current physical, biological, and clinical challenges. Solving these will ultimately determine whether proton therapy will continue on its path to becoming mainstream.

# Acknowledgements

Views expressed in this article are those of the author and do not necessarily reflect the opinion of Massachusetts General Hospital or Harvard Medical School.

# About the author

## Harald Paganetti

Prof Harald Paganetti PhD is the Director of Physics Research at the Department of Radiation Oncology at Massachusetts General Hospital and a Professor of Radiation Oncology at Harvard Medical School. He received his PhD in experimental nuclear physics in 1992 from the Rheinische-Friedrich-Wilhelms University in Bonn, Germany, and has been working in radiation therapy research on experimental as well as theoretical aspects since 1994. He has authored and co-authored more than 200 peer-reviewed publications and has edited two books on Proton Therapy. He has made significant contributions to the field of radiation oncology physics, many of which have found their way into clinical practice. Particularly, he is a pioneer in advanced Monte Carlo dose calculations for proton therapy treatment planning and is considered the world expert on the relative biological effectiveness of proton beams.

# Proton Beam Therapy

**Harald Paganetti**

## Introduction

Cancer therapies aim at destroying or removing cancerous tissues while limiting the damage to healthy organs. Treatment is typically based on a multimodality approach, including surgery, systemic or targeted chemotherapy, radiation (external beam or radionuclide) and immunotherapy. For many brain tumours, for example, surgery is used first. However, because surgery might be unable to conservatively remove a tumour without seriously compromising the patient's health, it is followed by radiation therapy to treat residual disease. On the other hand, some soft-tissue tumours may be treated with radiation first to cause shrinkage of the tumour so that it can be removed safely. Other cancers may be treated with just one modality or with different modalities concurrently or in succession.

### Radiation therapy treatment planning

Conformal approaches such as surgery or radiation are more successful if the target is clearly defined. However, the distinction between cancerous and healthy tissue is often difficult, particularly for tumours infiltrating healthy tissue—e.g. glioblastoma (a type of brain tumour). Even if the tumour can be clearly distinguished from healthy tissue, some surrounding tissue will be irradiated in radiation therapy because external beams typically have to penetrate healthy tissue to reach the tumour. Furthermore, healthy tissue is deliberately included as part of the target definition because of uncertainties in diagnostic imaging, treatment planning and delivery, as well as uncertainties with respect to disease spreading into healthy tissues potentially not visible on diagnostic images. Organ motion, as present in many tumours in the torso such as lung tumours, can require substantial target extension as well.

Target delineation and delineation of healthy tissue to be spared is done on CT images with additional MRI images, depending on the site. The target is defined as the Gross Tumour Volume (GTV), which is expanded into the Clinical Target Volume (CTV) to include potential residual disease. The CTV is then further expanded into the Planning Target Volume (PTV) to account for planning and delivery uncertainties. Furthermore, in the case of tumour motion, one might use the

doi:10.1088/978-0-7503-1370-4ch1  © IOP Publishing Ltd 2017

Internal Target Volume (ITV), encompassing the motion trajectory to ensure that the target always receives the full dose.

These target expansions emphasize that tumour coverage is typically given priority over normal tissue sparing. Tumour dose prescriptions are based on empirical evidence or limited by constraints to the maximum dose tolerable by surrounding tissues. The dose may also depend on whether the patient receives other treatments, such as chemotherapy, or whether there are other confounding factors. The main priority is to administer the necessary dose to ensure tumour control. If the required dose would have to be higher than the maximum allowed tolerance to critical organs without dire consequences, radiation can be given without curative intent.

As for prescription doses to the tumour, the dose constraints on critical organs (organs that are most susceptible to radiation and critical to the patient's well-being) are largely based on empirical evidence and not on a mechanistic understanding of the underlying biology. Dose constraints are either based on a mean dose or specific dose constraints to specific organ volumes being irradiated, depending on our empirical knowledge of dose response relationships. They may, as with prescription doses, be chosen patient specific based on the physician's knowledge of the patient's condition or prognosis.

Treatment planning for radiation therapy is performed by applying optimization algorithms that focus on administering the prescribed dose (energy deposited in the tissue per unit mass) to the target volume while limiting the dose to critical organs to an acceptable level. These algorithms consider various degrees of freedom, such as the number of beam angles or intensity modulation of the radiation (see below), to find a solution that meets all clinical constraints while delivering the prescribed dose. However, optimization algorithms may not find the truly optimized treatment plan because significant input from a treatment planner is required by setting beam angles and judging the trade-off between conflicting constraints. Depending on the complexity of the treatment, the quality of the plan may thus depend on the experience of the planner.

In the current clinical routine, the goal is to deliver a homogeneous, curative dose to the target. However, tumours are not homogeneous and their response to radiation varies across the tumour volume. Multiple studies are investigating the possibility of administering inhomogeneous dose distributions to the tumour and considering, for example, the difference in radiosensitivity due to oxygenation. These studies are motivated and supported by new advances in imaging technology that can visualize the tumour heterogeneities. Additionally, numerous research efforts aim at defining patient-specific constraints as well as patient-specific tumour doses based on imaging, blood or genetic biomarkers. So far, neither of these approaches is used in the clinical routine.

The ratio of damage to cancer cells and the damage to healthy cells is often referred to as the therapeutic ratio. The dose is typically administered in multiple fractions to allow healthy tissue to recover in between fractions. Most cancers are treated with 30 fractions and a dose of 2 Gy per fraction. These values were decided empirically decades ago. Fractionation takes advantage of the difference in non-linear dose response relationships between tumour and healthy tissues. In recent years, there has

been more research on the optimum fractionation and dose per fraction. With the advent of more conformal radiation therapy techniques, the dose to critical structures has been reduced so that doses per fraction can be increased for many treatment sites. Consequently, there is a trend towards hypofractionation, which provides not only a potential biological advantage but also a reduced burden on patients who, for example, welcome a one-week treatment course over a five-week course.

## Why proton beams?

### Conventional radiation therapy

Radiation therapy has been preformed for many decades. The most commonly used external-beam radiation therapy uses megavoltage photons from a gantry-mounted linear accelerator (LINAC). A gantry is a mechanical structure that allows the beam delivery system to rotate around the patient in order to treat from different angles. These LINACs attached to a treatment couch are standard equipment in most hospitals. Photons are uncharged and thus deposit the dose mainly indirectly via secondary electrons that are produced in ionization events. Owing to their limited attenuation in the patient, each beam impinging on the patient fully penetrates the patient (which can be utilized in transmission imaging to guide the treatment). Photon dose distributions as a function of depth show a maximum dose close to the entrance after a short build-up and then an exponentially decreasing energy deposition with increasing depth in tissue (figure 1). Except for small superficial tumours, only the use of multiple beam angles ensures a higher dose to the tumour compared to healthy tissues.

### Physics of proton-beam radiation therapy

The potential of proton beams for treating cancer was recognized by physicists in the late 1940s and the first patient was treated in 1954. Subsequently, proton therapy

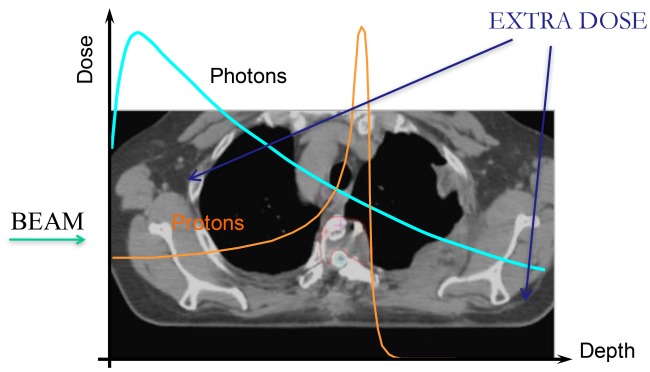


**Figure 1.** Depth–dose distributions from a mono-energetic photon beam (blue) and a mono-energetic proton beam (orange) with a beam entering from the left. The photon dose distribution shows a maximum close to the entrance and a decreasing dose with depth. The proton dose increases with depth, resulting in a Bragg curve with a peak at the end. For a single mono-energetic beam this results in significant extra dose when using photons. For modulated fields photons may however show a slight advantage in skin dose due to the build-up.

treatments were considered to be highly experimental and took place at physics research laboratories, with patients usually referred by local physicians. The first hospital-based proton therapy facility was not built until 1990. Since then, the number of hospital-based installations has steadily increased and to date more than 100 000 patients have been treated with proton therapy.

Electrons, protons (and heavier ions) are charged particles that undergo frequent interactions in material and have a finite range in tissue. Electrons show significant scattering and thus result in a very diffuse dose distribution. Protons, on the other hand, mostly undergo small angle scattering events and deposit the maximum energy per path length close to the end of range, causing a 'Bragg' peak (figure 1).

The distinct Bragg peak is a result of the energy transferred in electromagnetic interactions being inversely proportional to the velocity of the proton—i.e. as the protons slow down they lose more energy per path length. This causes an increase in linear energy transfer (LET)—i.e. the energy absorbed in tissue per unit length. The longitudinal profile is dominated by the inelastic electromagnetic interaction with atomic electrons (figure 2). This, together with the decreasing number of protons as a function of depth, causes the Bragg peak. The number of primary protons (beam intensity) is decreasing with depth because they stop once they run out of energy or

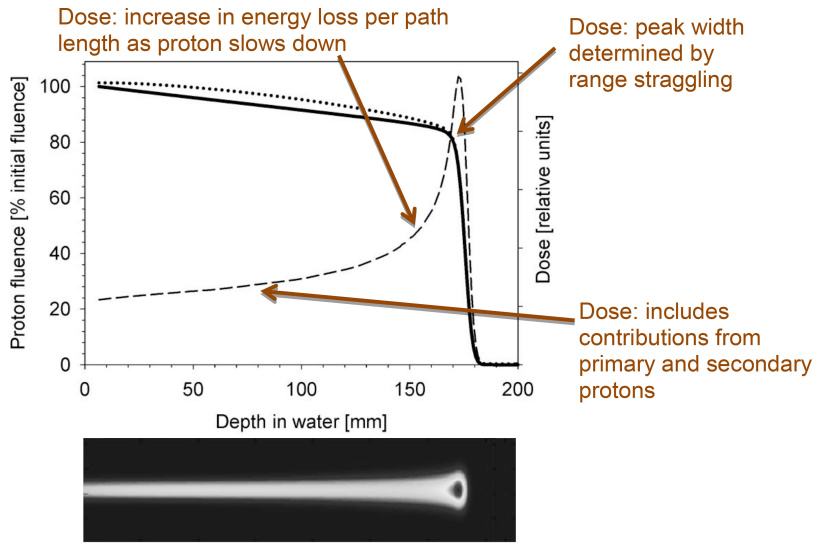


**Figure 2.** Dashed line and right axis: Bragg curve (dose deposited a function of depth for a 160 MeV proton beam). The Bragg peak width is determined by range straggling. Contributions from nuclear interactions affect mainly the dose in the entrance plateau. Left axis: total (dotted line) and primary (solid line) proton fluence as a function of depth, showing the contributions from secondary protons generated in nuclear interactions. The decrease in the entrance plateau is due to primary protons undergoing nuclear interactions whereas the sharp decrease at the Bragg peak is mainly due to the stopping of primary protons. The Bragg peak is a combination of the increasing energy loss per path length due to mainly electromagnetic interactions and the decreasing fluence as a function of depth. The lower graph shows the dose profile, illustrating the broadening of the beam due to multiple Coulomb scattering.

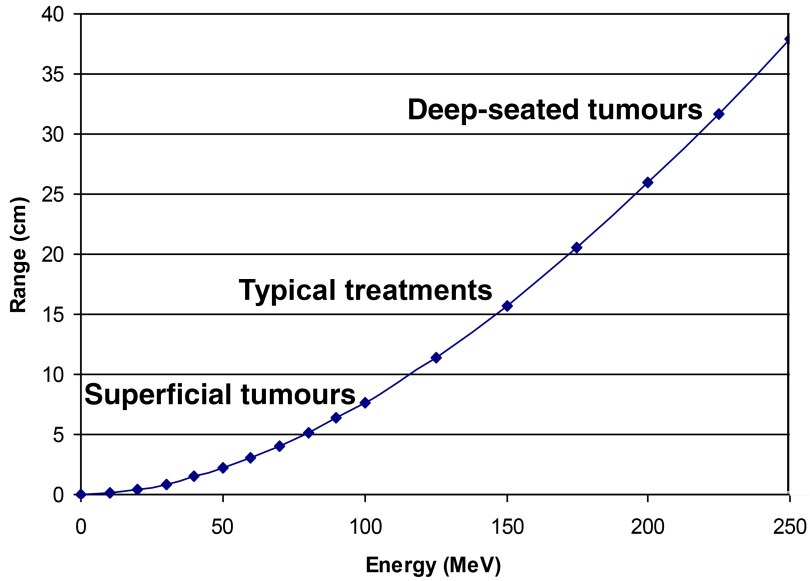


**Figure 3.** Proton range (in water) in cm as a function of proton energy in MeV.

because they undergo a nuclear interaction, which happens mostly in the entrance region before reaching the Bragg peak area. On average, about 1% of the protons undergo a nuclear interaction event per cm range in water. For a single proton, the peak would be extremely sharp but the combination of many protons, all having statistically slightly different ranges due to their interactions along the path (range straggling), causes a peak with a width of a few mm. This phenomenon is called range straggling.

Frequent elastic scattering on nuclei (multiple Coulomb scattering) leads to a lateral broadening of the beam as a function of depth (figure 2). For large ranges in tissue (more than ~15 cm), this can cause the lateral penumbra of proton beams to be less steep than with photon beams. This is also the reason why it is desirable to have the treatment-head exit (i.e. beam-shaping devices that constrain the field to the target area) as close to the patient as possible (so as to reduce beam broadening due to scatter in air). Figure 3 shows the range of a proton beam as a function of the beam energy, which is typically between 70 MeV and 250 MeV for most treatment facilities.

## Utilizing the physics of proton beams in radiation therapy

Owing to the absence of exit dose and the variable beam energy determining the depth of the Bragg peak position, protons offer more options for sculpting the dose distribution compared to photons. Further, the total energy deposited in the patient (often termed integral dose) is typically at least a factor of three lower than with any form of photon therapy (figure 1). This is mainly for two reasons: first, the fact that the maximum dose in a Bragg curve can be located in the tumour while the depth–dose curve with photons shows a maximum at the entrance; and second, there

is negligible dose downstream of the Bragg peak while photons fully penetrate the patient.

The advantage in integral dose allows substantial reduction in dose outside of the target, or, by maintaining the dose to healthy structures, allows escalation of target dose. The latter can be important particularly for those tumours where conventional techniques do not allow the delivery of doses sufficient for curative intent. Typically, proton therapy treatments exploit the reduction of dose outside of the target. This dose reduction can reduce toxicities and improve the patient's quality of life. The proton depth dose distribution also allows the use of fewer beam directions compared with photon techniques. By using intensity-modulated techniques, with movable collimator leaves and multiple beam angles, treatment plans from photon treatments can be comparable to proton therapy plans in terms of tumour coverage (i.e. the high dose region), but the low-dose envelope is substantially bigger with photons.

While proton treatments overall deliver less energy outside of the target, it is unclear whether this physical advantage matters in terms of clinical outcome. We often do not know the importance of low dose radiation with respect to serious toxicities. This is the main reason why proton therapy is controversial, considering that its cost is higher than conventional therapy. However, protons offer distinct advantages in cases where we do know that extra dose matters—e.g. for paediatric brain-cancer patients, where additional dose outside of the target can adversely impact cognitive development or growth of the child's bones.

### Dose shaping with passively scattered proton beams

As with other radiation modalities, proton beams have to be 'modulated' in order to conform the dose distribution as close as possible to the target. For instance, a tumour is typically larger than the tip of the Bragg peak, so that the Bragg peak has to be broadened by stacking several peaks with different energies. In recent decades, proton therapy planning and delivery has been based on single-field uniform dose—i.e. each field (beam direction) delivers a uniform dose to the target—employing a 'passive scattering' technique, which requires patient- and beam-specific hardware.

A proton beam entering the treatment room typically has a Gaussian shape with a width of a few mm (figure 2). In passive scattering, beam modulation is done in range to cover the thickness of the tumour in the beam's eye view and in width to cover the extent of the tumour in the beam's eye view (figure 4). This allows a homogeneous dose in the tumour. Note that this is possible with photon beams only if multiple beam angles are being administered.

Range modulation—i.e. the stacking of Bragg curves of different energies—is achieved by placing absorbers in the beam path. By using a fast rotating absorber with individual steps of absorber thickness to reduce the energy, the whole spread-out Bragg peak can be delivered within the time of one rotation (which at a rotation speed of typically around 10 Hz is more or less instantaneous). The widths of the absorber step controls the beam fluence and thus the height of each Bragg peak. Such rotating absorbers ('range modulator wheels') act like a propeller with blades that the beam goes through, the thinnest blade resulting in the largest range of the

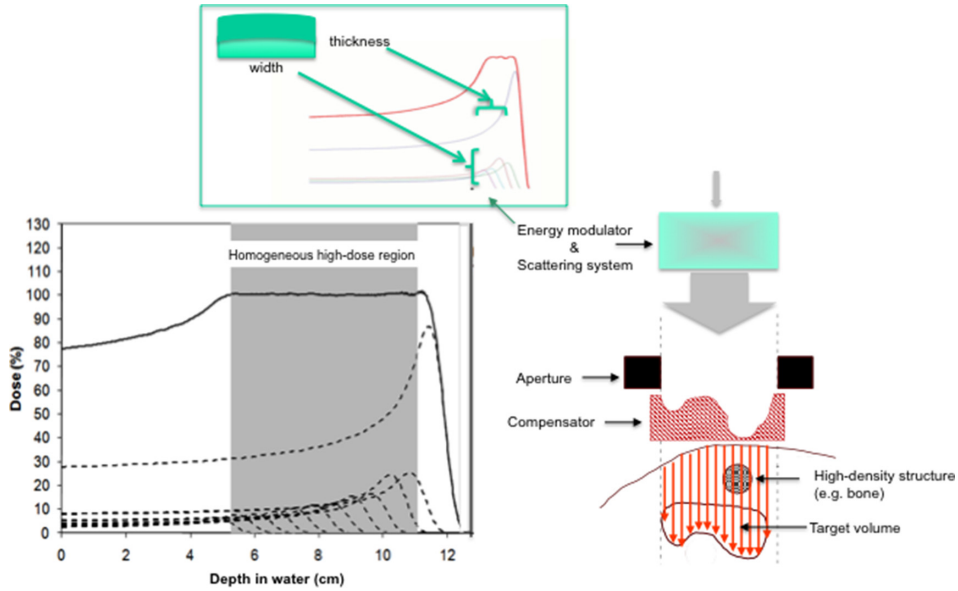


**Figure 4.** Lower left: Spread-out Bragg peak as a combination of multiple Bragg curves to homogeneously cover an extended region with the same dose as a function of depth. This is achieved with a rotating absorber (energy modulator) with different elements where the thickness of each element determines the range of each individual Bragg curve and the width determines the Bragg peak height by modulating the proton fluence (illustrated on top). Right: Beam shaping with the beam entering from the top as a small beamlet, which is subsequently broadened in a scattering system that also includes the energy modulator. The beam then passes an aperture (to reduce the lateral beam size) and a compensator to modulate the energy so that the distal peaks follow the shape of the target volume. The spread-out Bragg peak results in the target volume being covered with a homogeneous dose distribution as a function of depth.

beam. The resulting spread-out Bragg peak delivers the same depth laterally. Considering that the distal surface of the tumour has an irregular shape, the spread-out Bragg peak fall-off in lateral direction is further modulated using a range compensator—i.e. a patient- and field-specific plastic block with varying thickness (typically with a resolution of a couple of mm) depending on the position in the field (figure 4).

In order to cover the tumour in the lateral direction, the beam needs to be broadened as well. This is done with scattering material in the beam path. Such scattering systems can be quite complex as one tries to reach a uniform field both in terms of scattering as well as angular spread of the beam. This is typically achieved by using two scatterers ('double scattering system') and by using combinations of materials with high and low atomic number. To ensure that tumours up to a certain size can be treated with a single field (e.g. 30 cm × 30 cm), a large circular-shaped field is created that is then clipped with a field-specific aperture. The aperture is made of a material with high atomic number to stop the beam outside of an opening that resembles the projection of the tumour extension to the position of the aperture (figure 4).

The entire beam-shaping hardware is incorporated in the 'treatment head' in the treatment room and typically requires about 2 m in length. A treatment head in proton therapy not only contains beam-shaping devices, but also beam-monitoring devices such as ionization chambers that mainly serve two purposes—i.e. to monitor the accurate delivery of the treatment and to define the treatment prescription in terms of ionization chamber readings. Ionization chambers are calibrated so that their reading is proportional to the delivered dose at a certain position in water, which can be related to the dose in patients as predicted by the dose calculation algorithm in the treatment planning program.

The passive scattering technique has its limitations in terms of dose sculpting—e.g. the distribution cannot be made conformal to the proximal surface of the target. Furthermore, the scattering and collimation system causes beam efficiencies to be of the order of typically just 3%–15%. Therefore, most proton planning and delivery strategies are now transitioning towards beam scanning.

## Dose shaping with scanned proton beams

In beam scanning, the shaping of the dose distribution is done by magnetically scanning small Gaussian-shaped 'pencil' beams and thus individual Bragg curves ('beamlets'). Individual proton beams with a width (sigma) of a few mm (typically around 2–10 mm depending on the beam energy and delivery system) scan the tumour slice by slice with a varying magnetic field in the *x*- and *y*-directions (figure 5). Because the tumour can be 'painted' with small pencil beams, there is no need for a scattering system to expand the field laterally (as in the passive scattering

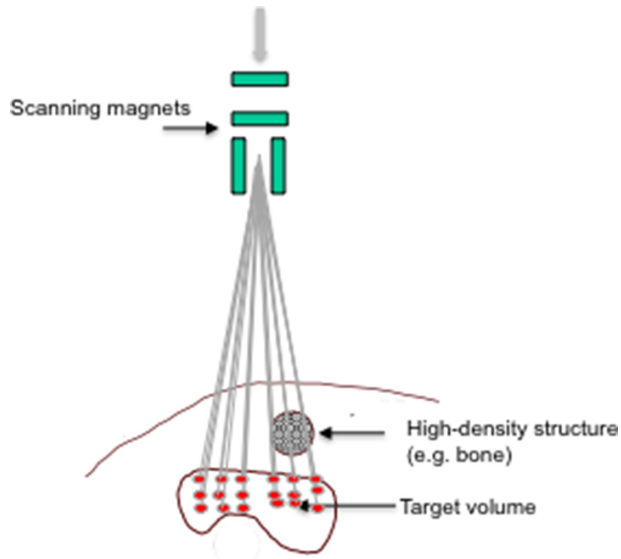


**Figure 5.** Beam scanning technique with two scanning magnets (scanning in the *x*- and *y*-directions) to deliver individual beamlets shaping the dose distribution. The depth is controlled by the beam energy.

technique). This causes less scattering material in the beam and thus a sharper Bragg peak in beam scanning than in passive scattered beams. Furthermore, the efficiency of the delivery system is close to 100%.

The depth dimension is manipulated by a change in beam energy as in passive scattering. However, instead of a modulator wheel, the beam energy for each set of pencil beams is controlled separately. The beam energy is typically selected outside of the treatment room. The time required to change the energy depends on the accelerator and delivery system and is an important clinical parameter as it determines the duration of treatments as well as the dose homogeneity due to potential patient motion. For very low energies or for energy fine-tuning, absorbers may also be incorporated in the treatment head. Beam scanning irradiates the tumour in iso-energy layers with facility- or treatment-planner-dependent energy steps as well as beamlet-layer spacing. Note that owing to inhomogeneous patient geometry, iso-energy layers do not necessarily translate in iso-depth layers. While apertures are not needed in beam scanning, they are sometimes used to improve the lateral penumbra of the field.

The biggest advantage of beam scanning is that it allows the delivery of highly inhomogeneous dose distributions by controlling each individual beamlet in terms of position and intensity. Intensity-modulated treatments—commonly known as intensity-modulated proton therapy (IMPT)—allow superior dose-shaping capabilities and fewer fields compared with any photon or electron technique (figure 6). This is even true for rotational therapies, which can produce highly targeted conformal dose distributions but at the cost of a considerable low-dose bath. As in intensity-modulated radiation therapy (IMRT), when treating with multiple photon beam directions, each individual field does not necessarily have to produce a homogeneous dose distribution as long as the combination does.

The quality of the dose distribution (i.e. its conformality to the target and sparing of organs at risk) depends on the beam characteristics of the delivery system. The latter is defined by the beamlet size as a function of energy, the spacing of the

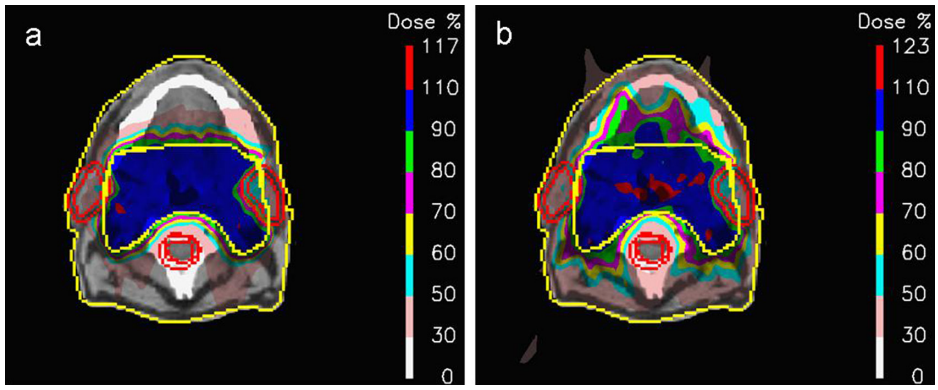


**Figure 6.** Reduction in integral dose when using proton compared to photon treatments. Left: three-field plan using intensity-modulated proton therapy (IMPT); right: nine-field plan using intensity-modulated photon therapy (IMRT). Reproduced from Paganetti (2012) with permission. Copyright 2012, Taylor & Francis.

individual beamlets, the energy spread determining the width of each Bragg peak, the scanning speed, and the time required by the system to change the beam energy in order to move from one iso-energy layer to the next.

Figure 7 shows an example of a proton dose distribution achievable with IMPT. In order to optimize the array of beamlets and their energies, mathematical algorithms performing inverse treatment planning based on dose prescriptions are being used. Often, a variety of beamlet maps fulfils the prescribed target dose and dose constraints. Note that for IMPT, the beam angles are still manually fixed by the treatment planner.

As with passive scattering, the treatment head contains beam monitoring devices. As far as quality assurance is concerned the requirements for these devices may be different for beam scanning owing to the time structure of the beamlet delivery.

## Technical challenges

### Size of facilities

Proton beams for radiation therapy are generated by using a particle accelerator, either a cyclotron or a synchrotron. Both have their pros and cons. Cyclotrons are smaller in size, but generally allow the extraction of only one fixed beam energy. This means that energy changes have to be achieved by placing absorber material in the beamline, typically right at the exit of the beam from the cyclotron. A synchrotron produces a significantly smaller beam (in cross-section) and can switch energy on sub-second timescales. On the other hand, its pulsed beam delivery makes it less flexible for some modulation techniques.

Currently most proton therapy accelerators are large, such that they are sited outside of the treatment room. In fact, a proton accelerator will typically feed several treatment rooms (figure 8), with a significant amount of space needed to transport the beam from the accelerator using magnets for bending, steering and focusing. In contrast, photon LINACS, including all beam shaping and monitoring devices, easily fit into a single room.

Proton facilities are therefore typically larger than photon treatment facilities, not least because a complete radiation delivery system also contains a gantry to rotate the entire treatment head around the patient. Most proton gantries have diameters of several meters, owing to the magnetic field strength required to bend the beam path of protons with energies of up to 250 MeV. The cost of a proton delivery system is also higher than that of a photon machine, in part because of room requirements when using a 360° gantry. A number of studies have assessed the need for a full rotational gantry and whether specific beam angles might be sufficient. A full 360° gantry might not be required for most treatments, for example, with the use of a robotic couch with six degrees of freedom and with patients being treated in a lying or seated position.

The size and cost of proton installations have implications for the widespread clinical adoption of proton therapy. A single-room proton machine, comparable in size and cost to a photon LINAC, would likely transform the field of radiation oncology and position proton therapy as the standard treatment option because of

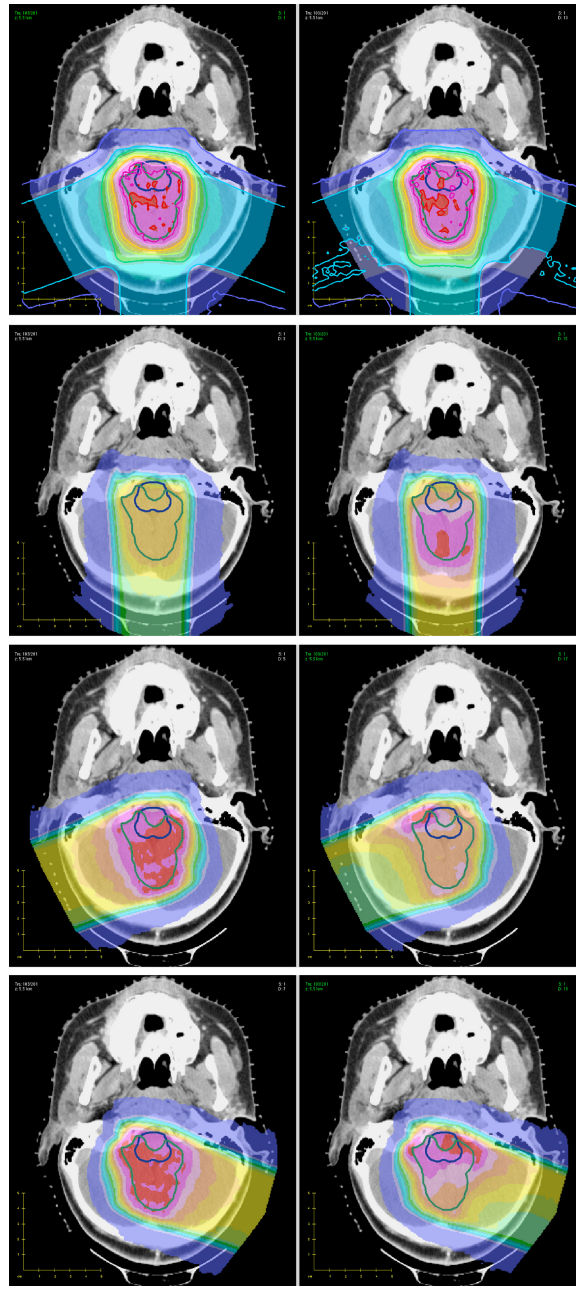

**Figure 7.** Target volume (CTV) coverage and tissue (brainstem) sparing using a 3-field plan with single field uniform dose (left) and a 3-field intensity-modulated plan (right). From top to bottom: composite dose distribution, beam 1, beam 2, beam 3.

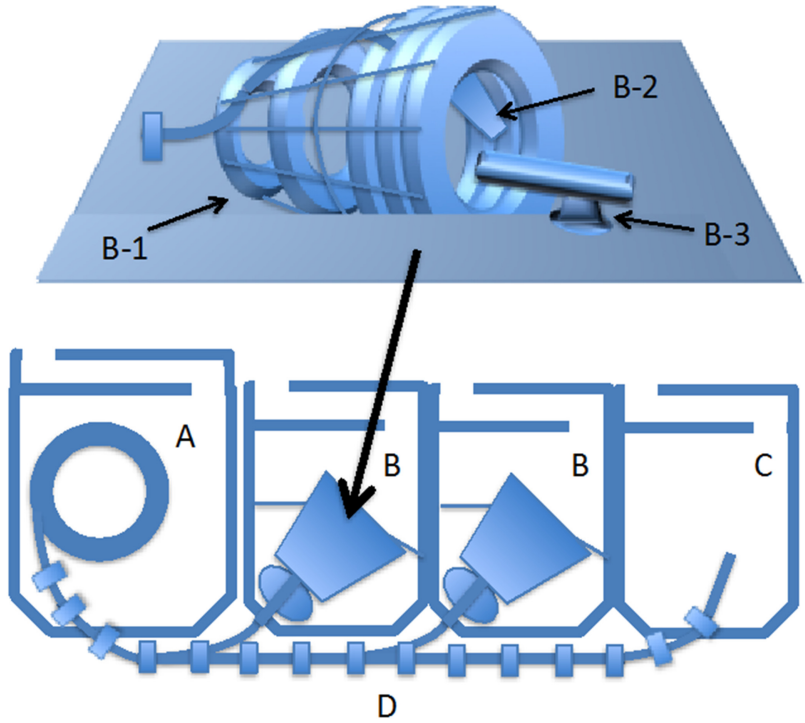


**Figure 8.** Bottom: typical proton therapy center with a single cyclotron feeding several gantry rooms (A: cyclotron; B: gantry room; C: fixed beam room; D: beam line. Top: gantry (B-1: mechanical structure with beam line (typically not visible to the patient); B-2: treatment head; B-3: treatment couch).

the dosimetric benefit. As proton therapy facilities have on average a higher capital cost, and most treatment centres are already equipped with LINACs, it is an open question whether a steady move towards proton therapy can be expected or whether this technique will be used for a subset of patients only. The size and cost of a single room treatment system is important for a widespread use of proton therapy. It is less of an issue for large academic centers where multiple rooms can rely on a single accelerator. Note that a single accelerator for multiple rooms does not cause limitations in workflow because the beam-on time is typically only a fraction of the time it takes to set up a patient for treatment.

Given this context, significant R&D efforts are being directed towards the development of smaller, single-room proton machines. Superconductivity will reduce the size of accelerators as well as gantries. For example, there are already cyclotrons with diameters of less than 2 m that can be mounted on a gantry (which nevertheless exceeds the size of a conventional treatment room). There are also small synchrotrons with a ring diameter of less than 6 m. In addition, cyclotrons as well as synchrotron systems with limited gantries (less than 360°) can now be fitted into rooms about twice the size of a photon treatment room. In summary: while further cost/size reductions can be expected for proton therapy installations, a reduction to the scale of the photon LINAC is likely to remain elusive over the near term.

### Beam quality

In addition to advances in accelerator and gantry technology, significant efforts are also under way to improve proton beam characteristics and delivery systems. There are important differences between current proton therapy facilities in terms of beam quality. For instance, superior beam-shaping capability can be achieved with a very small geometrical beamlet together with a small initial energy spread in beam scanning (owing to the resulting sharp Bragg peak in three dimensions).

Beamlets are never entirely monoenergetic and a change in energy using an absorber may increase the energy spread even further, which widens the Bragg peak. The beam energy is selected either by a multi-energy extraction accelerator (synchrotron) or a single-energy machine (cyclotron) with a degrader to decrease the energy after the beam extraction from the cyclotron. Independent of the accelerator type, it takes a finite amount of time to change the beam energy when moving from one iso-energy layer to the next. The lateral beamlet size also depends on the specifics of the delivery system as it is determined by the beam optics. The typical sigma of such a beamlet in most delivery systems is of the order of 12 mm for small beam energies and of the order of 3 mm for large beam energies.

While small beamlets (in all directions) offer superior shaping capabilities, they also have some disadvantages. For instance, the number of individual beamlets that have to be delivered in beam scanning increases with decreasing beamlet size, which in turn extends the delivery time. Furthermore, smaller beamlets are dosimetrically less robust towards delivery uncertainties and organ motion.

Patient motion is of particular concern if the delivery patterns are on a similar timescale as periodic patient motion (such as breathing). The optimal solution may therefore depend on the treatment site. Today's treatment delivery systems are capable of delivering high doses with high fluence. Ultrafast delivery might be desirable to avoid the impact of patient/tumour motion, but may be less robust towards hardware or software errors. Delivery parameters thus also impact the quality assurance requirements. Note, however, that delivery times are already much shorter than the times required to set up the patient for treatment in the treatment room (and in turn have little impact on patient throughput in a facility).

## Physics challenges

### Dose uncertainties

When prescribing and delivering dose we rely on the dose distributions shown by the treatment planning program after optimization. However, there is always uncertainty when predicting the dose and the dose distribution may not be delivered precisely as planned. In radiation therapy, the aim is to deliver the dose to within 2.5% of the prescribed dose. This value was suggested by international regulatory bodies.

Geometrical uncertainties are considered when defining the target volume (see above), though this does not cover potential uncertainties in absolute delivered dose. Standard radiation treatments deliver the target dose over 20–40 fractions in 2 Gy

fractions. This fractionation scheme causes set-up and motion uncertainties to be averaged over the number of fractions. However, with the trend towards hypo-fractionation, small errors for both uncertainties may no longer average out over many fractions. Some of the uncertainties in radiation therapy can be addressed by physics, such as uncertainties due to dose calculation.

Dose calculations are routinely performed using analytical algorithms that are fast enough to allow treatment optimization in minutes. While these techniques are typically sufficiently accurate in the photon world, they have significant shortcomings in proton therapy. The steeper dose gradients present in particle fields expose the approximations in analytical algorithms, in particular with respect to scattering of protons at interfaces such as bone and soft tissue. While too slow to be used in the clinic in the past, dose calculations based on particle-track simulations ('Monte Carlo') have recently achieved efficiencies that make them suitable for use in treatment planning. Monte Carlo codes offer superior dose calculations to analytical algorithms and are considered the gold standard.

Uncertainties can have more severe consequences in proton therapy versus photon therapy because of the finite range and the impact of proton scattering in an inhomogeneous patient geometry. Understanding uncertainties in proton therapy is therefore vital when making treatment planning decisions.

### Range uncertainties

In proton therapy, one of the main concerns is the impact of range uncertainties on tumour coverage (e.g. having a smaller range than prescribed). Miscalculating the range could potentially cause the treatment beam to miss the distal part of a tumour or irradiate more healthy tissue downstream of the tumour.

The finite range of protons in tissue offers a great advantage to photon therapies owing to the lower integral dose and the potential to point a beam towards a critical structure. However, this advantage can only be fully utilized if the range is indeed well known to within at least a few mm. Range uncertainties are unique to heavy-charged-particle therapy (such as proton therapy) and require margins along the beam path (whereas uncertainties in conventional therapy can be considered as isotropic expansion of the volume to be treated). As a consequence, the PTV concept is not sufficient in proton therapy and it is necessary to apply an additional margin in the beam direction.

Clinically, a substantial range is added to the prescribed range in order to ensure tumour coverage—i.e. of the order of 3.5% of the prescribed range plus an additional millimetre. This adds a substantial dose to healthy tissue. For instance, if 10 cm water-equivalent depth in tissue were required to cover the distal part of the target, one would deliver 10.45 cm—i.e. almost 5 mm of healthy tissue across the width of the target. Placing the Bragg curve, with its sharp distal fall-off, downstream of the Bragg peak right at the edge of a critical organ is often avoided due to range uncertainties, often resulting in non-optimal treatment plans. A reduction in uncertainties, particularly in terms of the proton range, can be expected in the near future due to Monte Carlo dose calculation.

For intensity-modulated therapy (using beam scanning and not delivering homogeneous doses per field), range uncertainties are often not explicitly considered. Instead the uncertainties are incorporated into a so-called robust optimization scheme that tries to prescribe treatment fields that are more forgiving in terms of each contributing beamlet. Such robust optimization techniques are currently incorporated in commercial planning systems for proton therapy.

To some extent, absolute uncertainties in range are reduced because the predicted range in water is adjusted so that it matches the measured range in water (this also takes care of uncertainties in considering the correct energy distribution from the accelerator). As a result, uncertainties are reduced to those relative to water.

However, this does not cover the uncertainty due to the parameterization of scattering in human tissues. Figure 9 demonstrates the effect of an underestimation of density in parts of the irradiated volume. The effect on the target volume is much smaller in photon therapy versus proton therapy. Furthermore, diagnostic imaging for treatment planning may not be sufficient to characterize tissue properties for dose calculation as CT scans provide a three-dimensional image of the patient resembling the photon attenuation in the material relative to water ('Hounsfield unit'). This then has to be translated into tissue properties for energy degradation and scattering.

Note that while in photon therapy dose calculation algorithms use a conversion from Hounsfield units to electron density, in proton therapy this conversion is done to stopping power relative to water (because of the difference in physics interactions between photons and protons). This does not necessarily lead to bigger uncertainties, but the uncertainties might have a bigger impact for proton beams owing to the finite range.

Other uncertainties originate from patient set-up for treatment delivery and the beam delivery reproducibility, while some uncertainties are related to the actual

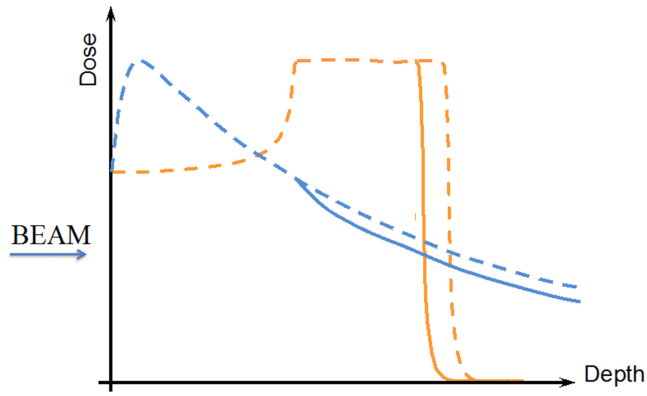


**Figure 9.** Dosimetric effect of density uncertainties on photon (blue) and proton (orange) depth dose curves. An overestimation of the density in the target would cause the dose to the target region to shift from the dashed line (planned) to the solid line (delivered). This would cause a slight underdosage in photon treatments but, because of the effect on the proton range, might result in missing parts of the target entirely.

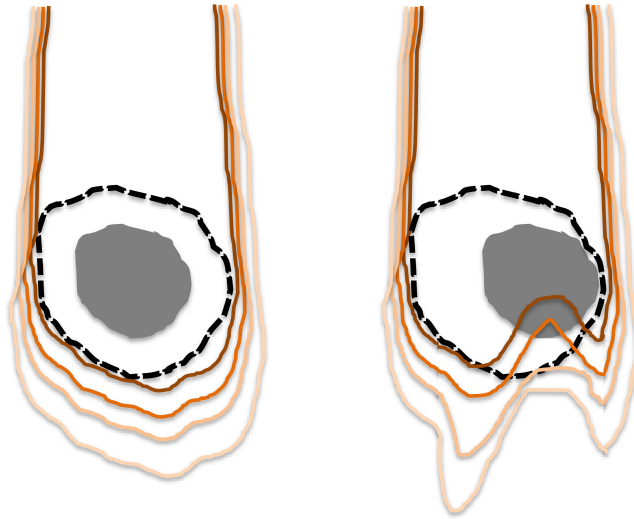

**Figure 10.** Range shift due to organ motion. Left: target volume (grey), margin to account for motion (dashed line), and iso-dose lines (dark brown (prescribed target dose) to light brown (penumbra doses)). Right: tumour motion to the right and dosimetric consequences because of low density lung tissue replaced by higher density tumour tissue (and *vice versa*).

patient. In a fractionated treatment, for example, the patient geometry may change from day to day (e.g. weight loss) or even during the beam delivery (e.g. breathing motion). A prime example is the filling of nasal cavities and the impact this has on the proton range. More frequent imaging and plan adaptation might therefore be necessary for proton treatments.

The situation becomes more complicated still in the presence of organ motion—for instance, when treating lung-cancer patients. Motion causes not only an overall blurring of the dose distribution but also local changes in range owing to high-density tumour tissue moving in and out of the field within low-density lung tissue. Ranges and their uncertainties are expressed in water-equivalent path length. This corresponds to substantial overshoot in the lung due to the low lung density (roughly a factor of three lower than the density of water). These fluctuations can result in an excessive dose to distal critical structures (figure 10). In addition to lung tissue being replaced by higher-density tumour tissue (or *vice versa*), rib-cage motion can result in significant changes in water-equivalent path lengths. These effects need to be considered carefully in treatment planning.

### Range verification *in vivo*

Range in patients can potentially be assessed using imaging technology. The use of imaging devices to monitor treatments is common practice in photon therapy, where each beam penetrates the patient so that the exit fluence can be utilized. Protons, on the other hand, stop in the patient and thus imaging can only be based on secondary radiation that is being created by the primary beam.

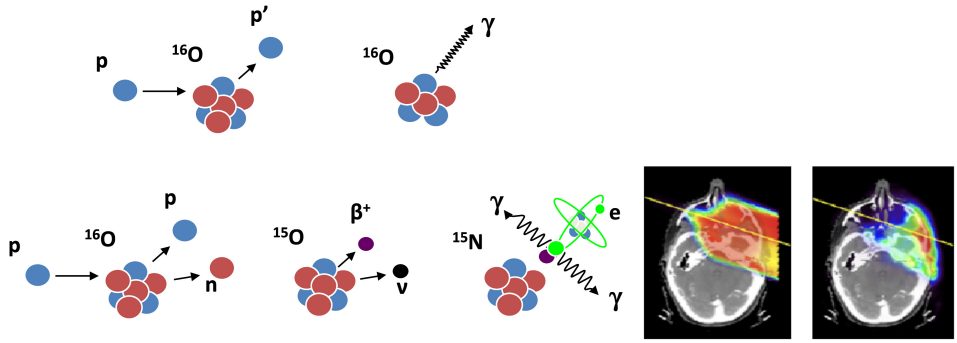


**Figure 11.** Concept of proton beam activation causing gamma rays outside of the patient that can be detected for treatment verification. The upper reaction shows an example leading to a high-energy (prompt) gamma. The lower reaction shows a reaction leading to beta decay and two annihilation photons that can be detected. The patient image illustrates the difference between dose distribution (left) and PET activity distribution (right).

Various methods have been proposed based on the detection of secondary radiation from nuclear interactions (figure 11). Two promising methods exploit the detection of gamma radiation created directly (after de-excitation of a nucleus, referred to as prompt gammas) or indirectly (by annihilation of positrons from carbon, oxygen or nitrogen isotopes created by nuclear interactions in the patient, used in positron emission tomography (PET)).

These methods currently form an active area of research, though they are not as yet being used in clinical practice because their accuracy in determining the range is still unclear. There are technical challenges in terms of detector developments to measure high-energy gammas leaving the patient; there are also physics challenges when interpreting the detected signal, due to uncertainties in the underlying physics cross-sections and the fact that the signal is not one-to-one related to dose (as dose is mainly deposited by electromagnetic and not by nuclear interactions). Figure 11 shows the difference between the delivered dose distribution and the detected PET image. The dose distribution can be verified indirectly by simulating (using Monte Carlo) the expected PET distribution determined from the treatment plan and subsequent comparison with the imaged PET data.

## Next steps

Other physics challenges in proton therapy include quality assurance and treatment monitoring. With higher precision comes the need for more accurate delivery verification and feedback systems to avoid treatment errors. Modern radiation therapy aims more and more at treatment adaptation during the course of treatment, with frequent monitoring of tumour physiology. With higher precision and concern for uncertainties, this aspect is more important for proton therapy compared to conventional therapy.

The described physics challenges address uncertainties in proton therapy, which for some sites do impact the potential dosimetric advantage of proton therapy versus

photon therapy. In general, the dosimetric advantage or disadvantage of one modality versus another is not constant over time. For instance, in-room CT imaging was first performed in proton therapy because the lack of exit dose and the beam stopping in the patient required more accurate patient set-up. Subsequently, however, in-room imaging was further developed in the photon world while proton therapy was lagging behind.

In recent decades, many developments in the photon world have reduced the dosimetric gap between photons and protons. But important advances have also found their way into proton therapy over the same period. While photon therapy continues to improve, we cannot expect it to catch up with protons dosimetrically because the integral dose difference cannot be overcome. And as we are just starting to utilize the dose-shaping capabilities of IMPT, the dosimetric advantage of proton therapy is likely to increase.

## Biological challenges

### Dose constraints to limit toxicities

When we treat patients with radiation we assume that a certain dose reflects a specific normal tissue complication on an organ level. Yet, clinically prescribed dose refers to energy deposited in a volume the size of a CT voxel, while biological effects are determined by energy depositions on a sub-cellular scale, including their biochemical consequences.

Starting from an energy deposition event, there are numerous steps—for instance, from creating a radical oxygen species to causing a double-strand break in the DNA to the failure of specific repair pathways to accumulation of these events eventually causing toxicity in an organ. We are just starting to understand the details of this multiscale problem. Most importantly, dose–response relationships are patient-specific and genetic profiling to determine patient-specific treatment options and doses is still in its infancy. Consequently, radiation therapy is prescribed based on physics, not biology, and uses empirical dose–response relationships. Current treatment planning methods generally rely on dose-based surrogates of biological response, such as the underlying dose prescription and constraints, instead of directly optimizing clinical endpoints such as tumour control probability (TCP) and normal tissue complication probability (NTCP).

The majority of our experience in radiation therapy has been obtained from photon treatments, mostly from three-dimensional photon treatments not using intensity modulation. As a consequence, our knowledge on dose constraints for organs-at-risk is typically based on very homogeneous doses to these organs. We are thus reasonably confident to define dose constraints based on mean dose. However, modern intensity-modulated photon therapy, and even more so proton therapy, causes highly inhomogeneous doses. This makes the definition of dose constraints more challenging because dose-volume effects need to be considered instead of simply relying on the mean dose to an organ. For example, ‘parallel’ organs, such as the liver, can still maintain functionality, even if a small region receives very high doses.

## The relative biological effectiveness

When using treatments with non-conventional radiation modalities such as protons, the same absorbed dose as with photons may produce a different clinical outcome (owing to differences in dose distribution but also due to differences in energy deposition pattern on the microscopic level). The latter effect is quantified by the relative biological effectiveness (RBE). The RBE is the ratio of the absorbed doses that produce the same biological effect between a reference radiation (e.g. photons) and a proton-beam irradiation. The RBE depends on the dose level, on the tissue endpoint and also on the LET. As a first order approximation, radiation is more effective if the LET is higher. Although the proton LET increases as protons slow down, proton therapy treatments are currently planned and delivered assuming a constant proton RBE relative to high-energy photons of 1.1. Thus, patients are prescribed 10% less physical dose when using protons as compared to photons. This value was deduced as an average of measured RBE values *in vivo* and *in vitro*. Importantly, the majority of experiments are based on the biological endpoint of cell survival *in vitro*, which may not be related to some normal tissue toxicities. Current models to calculate RBE suffer from the lack of mechanistic knowledge and a lack of robust input parameters. This is one of the reasons why treatment planning systems do not report RBE-weighted doses based on RBE models but instead use the constant RBE factor.

There are significant gaps in our knowledge of the biological effectiveness of protons, which is likely to vary between 1.0 and 1.5 in the relevant dose region close to the Bragg peak. There is substantial evidence that the RBE of protons may be a complex function of numerous factors such as treatment technique (proton energy, beam angles, scanning technique), dose, dose rate, cell type, oxygenation, intrinsic radiosensitivity, and the biological or clinical endpoint of interest (local tumour control or treatment complication). While these factors are currently not taken into account quantitatively, they are often considered indirectly when planners choose beam angles or margins. Importantly, these dependences are likely patient-specific and it is difficult to deduce them without improved understanding of the underlying processes, appropriate imaging and blood/DNA biomarkers.

The potential RBE uncertainties are important because RBE is increasing with LET and thus with depth in a Bragg peak. We would therefore see a potential dose increase at the edge of tumours if a spread-out Bragg peak is delivered, which due to safety margins is typically within healthy tissue. It is difficult to identify toxicities in patients as being correlated with uncertainties in RBE because side-effects are relatively rare. Statistical significance is further difficult to achieve considering intrinsic differences in patient radiosensitivity. Some treatment planning algorithms try to avoid high LET (or RBE) values in or close to critical structures. These efforts are related to robust optimization efforts for range uncertainties. Figure 12 shows how LET distributions can differ for almost identical dose distributions simply by rearranging the placement of beamlets. This is because of the ambiguity of beamlet placement in IMPT to create a specific dose distribution. Herein lies a great chance of indirectly taking advantage of LET and thus RBE variations. LET distributions

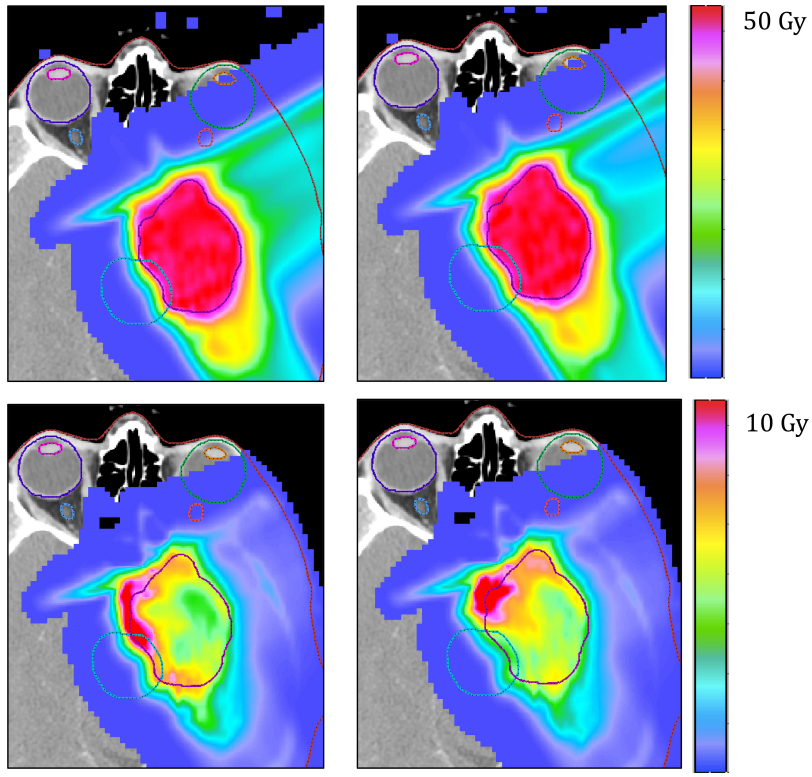


**Figure 12.** Target (purple outline) overlapping with the brainstem (blue outline). The upper row shows two different IMPT plans resulting in comparable dose distributions. The lower row shows the two corresponding distributions of LET times dose illustrating that the right plan might be beneficial in terms of RBE. LET times dose is in Gy as it includes a normalization factor to demonstrate the extra dose due to LET effects.

can be influenced in treatment planning and plan quality can thus be improved without knowing accurate RBE values by assuming a monotone relationship between LET and RBE for a given dose and tissue.

## Combined-modality treatments

Another challenge in radiation oncology in general is the multitude of treatment options being used together with radiation. Not only chemotherapy, but drugs targeting tumour-specific molecular pathways as well as immunotherapy are all being administered. Some of the targeted drugs and immunotherapy approaches are linked to specific radiation damage and repair phenomena on the cellular and molecular level and may impact the response differently for proton and photon therapy. The mechanistic understanding and prediction of these effects is challenging and might not only influence dose constraints, but also the timing of applying different modalities.

Patient specificity adds another level of complexity that can only be tackled with biomarkers (as deduced from blood samples, for example). The decoding of patient-related dose–response can influence the treatment prescription prior to treatment or

it may call for treatment adaptation during treatment, if biomarkers define themselves by dose–response relationships. Substantial research is currently being done towards personalized treatment planning, which is not limited to radiation therapy. The biological difference between photon and proton therapy may also provide an additional degree of freedom when combined with drugs.

## Clinical challenges

Clinical efficacy of a treatment method has both a scientific as well as an economic aspect. For instance, drug development is often very expensive and is mainly undertaken by pharmaceutical companies, resulting in considerable costs for drugs that have to be covered by healthcare providers or patients. In radiation therapy, the burden on proving efficacy is with the user, as treatment-machine vendors are not conducting or funding clinical trials. This is in part due to our reliance on dose (i.e. physics) when judging the advantages of radiation therapy treatment options. Proton therapy dose distributions seem advantageous in terms of sparing normal tissue. In contrast to the scenario for chemotherapy drugs, with proton therapy we are in a situation where clinical efficacy is being tested after the treatment agent is introduced in the clinic (not the other way around).

While most proton therapy treatment plans show a considerable dosimetric advantage, it is unclear for many sites whether this necessarily translates into improved outcome. Not only is there not a one-to-one relationship between dose and biological effect, the distribution of the dose within an organ also differs between photon and proton treatments. Depending on the magnitude of inhomogeneity, it is often difficult to define the appropriate normal tissue constraints in proton plans (given that those constraints were defined by the photon world, often based on organ mean dose). Concepts such as the equivalent uniform dose have been suggested to translate an inhomogeneous dose distribution into a biologically equivalent uniform dose. These are being used in outcome analysis but not in treatment planning.

Elsewhere, there are efforts to determine for which sites proton therapy might be most beneficial. It seems clear that this includes sites that benefit most from the reduced integral dose—e.g. in paediatric patients, where the low dose bath can cause severe toxicities due to developing organs/tissues, or head-and-neck tumour patients, where tumours are typically close to critical structures. Furthermore, as a result of the integral dose advantage, protons are expected to be beneficial for patients with larger tumour volumes, as well as tumours where photons do not allow dose escalation to levels needed for tumour control.

Proton beam therapy might not be the most cost-effective treatment for all patients or all cancers. Research towards personalized cancer therapies will help to identify those patients that benefit the most from proton therapy. Cost effectiveness is tied to this question. In order to determine the true cost of a treatment, not only the cost for the treatment itself but also long term costs due to reduced quality of life and toxicities need to be considered. It is here where proton therapy, due to reduced integral dose, might make up for the still higher up-front investment. The question whether proton therapy will be used in the future only for selected patient

populations or treatment sites or whether it has the potential to widely replace photon treatments impacts the design of clinical studies because the latter calls for randomized trials whereas the former might be based on trials designed to identify a subset of patients that would benefit from proton treatments. This also decides the technology that will ultimately prevail. It may be acceptable for selected patients to travel to large specialized treatment centres offering protons, whereas a widespread use of proton therapy will demand the use of single-room machines.

The bottom line here is that it is very important to learn from long-term follow-up, both for tumour control as well as toxicity. Efforts are underway to standardize the way we store and analyse data, not only for the purpose of studying the efficacy of treatment techniques but also to perform knowledge-based and automated treatment planning.

## Outlook

Right now, only about 1% of radiation therapy patients are treated with protons, though that number is growing rapidly due to the increased availability of proton facilities (figure 13). In recent decades, proton therapy has transitioned from research laboratories to the clinical setting—and the number of proton therapy facilities continues to scale significantly. This results in a lack of well-trained personnel, i.e. physicians, physicists, treatment planners etc. Unless the size and the cost of these facilities is of the same order as LINACs, protons will never entirely replace photon therapy (though it's fair to say that, depending on the outcome of many ongoing clinical trials, specific tumours or patient populations might only be treated with proton beams in the future).

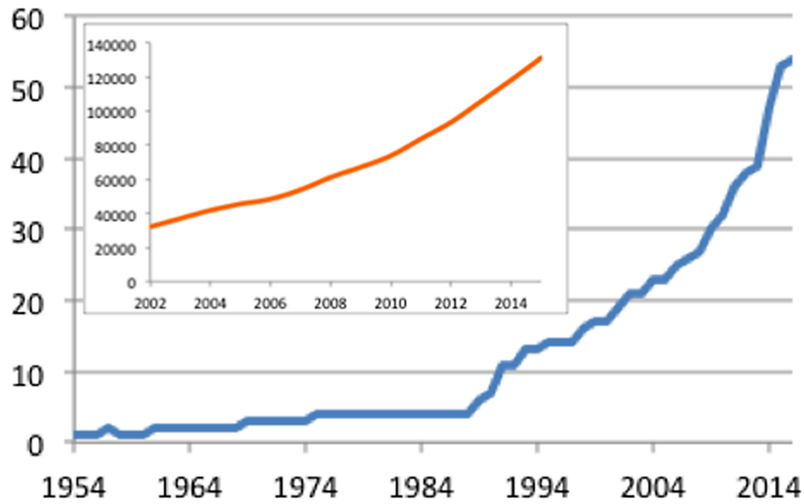


**Figure 13.** Increase in the worldwide number of proton facilities since 1954 (blue) and accumulated patient numbers since 2002 (red).

At the same time, there is also a growing interest in heavy-ion therapy (e.g. using carbon ions). Heavy ions have a number of advantages compared to protons, namely their reduced scatter (causing a much sharper lateral penumbra) and higher effectiveness in hypoxic tumours (which are very radio-resistant). Disadvantages include a fragmentation tail downstream of the Bragg peak (arising from nuclear interactions with tissue) and a larger uncertainty in RBE compared to proton beams. Also, RBE values are changing more significantly as a function of depth, requiring the delivery of inhomogeneous dose distributions while relying on models to predict RBE that have substantial uncertainties. Most importantly, the cost of a heavy-ion centre is much higher than a proton facility—an investment that can only be justified with either solid clinical evidence or by utilizing a centre with substantial research interests.

Overall, the options to treat cancer are constantly evolving, while delivery modalities register ongoing improvements. The trend towards extended life expectancy and increasing tumour control rates is therefore expected to continue.

## Further Reading

De Laney T F and Kooy H (ed) 2007 *Proton and Charged Particle Radiotherapy* (Philadelphia, PA: Lippincott) ISBN: 978-0781765527

Paganetti H (ed) 2012 *Proton Therapy Physics* (London: Taylor and Francis) ISBN-10: 1439836442

Linz U (ed) 2012 *Ion Beam Therapy: Fundamentals, Technology, Clinical Applications* (Berlin: Springer) ISBN: 978-3642214134

Ma C-M and Lomax T (ed) 2012 *Proton and Carbon Ion Therapy* (CRC Press) ISBN: 978-1439816073

Das I J and Paganetti H (ed) 2015 *Principles and Practice of Proton Beam Therapy* Medical Physics Monograph No. 37 (Madison, WI: Medical Physics Publishing) ISBN: 978-1-936366-43-9