# REDUCING RANGE UNCERTAINTY IN PROTON THERAPY

Joost M. Verburg

# Reducing Range Uncertainty in Proton Therapy

PROEFSCHRIFT

ter verkrijging van de graad van doctor aan de Technische Universiteit Eindhoven, op gezag van de rector magnificus prof.dr.ir. F.P.T. Baaijens, voor een commissie aangewezen door het College voor Promoties, in het openbaar te verdedigen op woensdag 20 mei 2015 om 16:00 uur

door

Joost Mathijs Verburg

geboren te Nijmegen

Dit proefschrift is goedgekeurd door de promotoren en de samenstelling van de promotiecommissie is als volgt:

| | |
|---|---|
| voorzitter: | prof.dr.ir. G.M.W. Kroesen |
| 1e promotor: | prof.dr.ir. P.F.F. Wijn |
| 2e promotor: | prof.dr. T. Bortfeld (Harvard Medical School) |
| copromotor: | dr. J. Seco (Harvard Medical School) |
| leden: | prof.dr.ir. F.W.J. Verhaegen (Maastricht University) |
| | dr.ir. M. Engelsman (Delft University of Technology) |
| | prof.dr.ir. O.J. Luiten |
| | prof.dr.ir. E.J.E. Cottaar |

# Reducing Range Uncertainty in Proton Therapy

Joost M. Verburg

This work was performed at the Department of Radiation Oncology, Harvard Medical School and Massachusetts General Hospital, Boston, MA, United States.

Reducing Range Uncertainty in Proton Therapy, by Joost M. Verburg.

A catalogue record is available from the Eindhoven University of Technology Library.
ISBN 978-90-386-3849-2

Cover art: Chris Verburg
Printing: Gildeprint – The Netherlands



20150912

# Contents

# 1

# Range uncertainties of proton therapy beams

This chapter introduces the topic of our research. We discuss the rationale of proton therapy, the delivery of proton beams, the issue of range uncertainty, and the sources of range uncertainties. Previous research on the reduction of range uncertainty is also reviewed. Finally, we provide an overview of the research in this thesis.

Radiotherapy is one of the main treatment modalities for cancer. It delivers ionizing radiation to kill malignant tumour cells, as primary therapy or in combination with surgery or chemotherapy. The goal of curative radiotherapy is to control the tumour, while minimizing severe side effects. When designing a radiotherapy treatment plan for a patient, the balance between the probability of tumour control and normal tissue complication, known as the therapeutic ratio, is the most important consideration. The principal factor that determines the therapeutic ratio is the distribution of the radiation dose. In many clinical settings, the dose that can be delivered to the tumour is bound by the dose that can be delivered to other tissues without an unacceptable chance of morbidity. Technological developments that enable the radiation dose to be better confined to the tumour can therefore improve the therapeutic ratio and thereby the treatment outcome.

## 1.1. Proton therapy

Currently, radiotherapy is most commonly delivered with external beams of x-rays generated with a linear accelerator. Proton therapy, in which proton beams are used instead of x-rays, is a radiotherapy technology that is receiving increasing interest. The physical interactions of protons with matter are very different as compared to x-rays, which allows for a superior dose distribution to be created.

### 1.1.1. Physical rationale

In figure 1.1, a comparison is shown between the depth-dose distributions of x-rays and protons in a typical case, which illustrates the key advantage of proton therapy. The figure shows beams traversing 25 cm of tissue with the target volume located at a depth between 10 cm and 15 cm. The proton beam delivers the same dose to target while delivering less than half of the integral dose of the x-ray beam.

The interaction of x-rays with matter is a stochastic process. The photons lose their full energy in a small number of interactions, in radiotherapy mostly due to Compton scattering. Their energy is transferred to secondary electrons, which deposit the radiation dose locally by creating ionization events along their paths. With the exception of the build-up region at the beam entrance, where electronic equilibrium has not yet been reached, the x-ray beam delivers less dose at larger depths because photons are removed from the beam. When aimed at a target, most of the radiation dose is typically delivered outside the intended target volume, as also displayed in figure 1.1.

To deliver a higher dose to the target as compared to surrounding tissue, multiple x-ray beams need to be used. Technological developments such as the multileaf collimator, intensity modulated radiotherapy [1] and arc therapy [2] have significantly improved the ability to customize the shape of the three-dimensional x-ray dose distribution. Although these techniques enable the volume receiving the highest

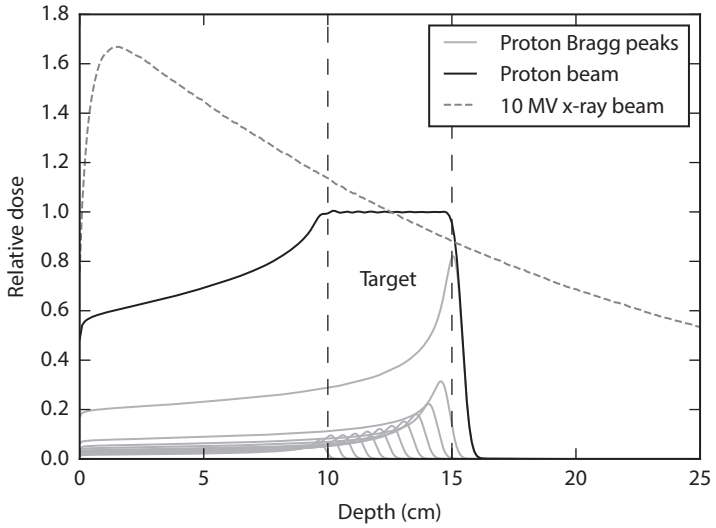



**Figure 1.1** | Comparison of the depth-dose distribution of an x-ray beam and a proton beam in water. The beams target a volume at a depth of 10 cm to 15 cm.

radiation dose to be precisely conformed to the target, the unwanted dose is mostly redistributed. The integral dose that is delivered to the patient remains determined by the physics of x-ray interactions.

The dose delivered by the proton beam, on the other hand, can be conformed to the target in depth. Because protons are massive and charged particles, they directly ionize atomic electrons. In each interaction, the proton loses only a small amount of its energy. Most of the protons travel in an almost straight line and traverse the body until all of their energy is lost. The depth at which the proton comes to rest depends on its initial energy. Most of the proton energy is depositioned just before the end of its range, which results in a peak in the depth-dose distribution that is known as the Bragg peak [3].

By delivering protons with a range of energies, Bragg peaks with different ranges as shown in figure 1.1 are combined to obtain a uniform high dose area within the target volume. Compared to the x-ray beam, the tissue distal to the tumour is spared from the harmful effects of radiation and the tissue proximal to the tumour receives significantly less dose. In the case shown in the figure, the x-ray beam delivers 2.1 times the integral dose of the proton beam. The dose fall-off at the end of range of the proton beam also provides a second potential advantage. If positioned accurately, it can be used to create a sharp dose gradient between a target and a nearby organ at risk that needs to be spared from the radiation. The dose gradient at the lateral edge of a typical x-ray or proton beam is not as sharp as this distal gradient.

### 1.1.2. Clinical history and use

The use of protons for radiotherapy was conceived at Harvard in 1946 by Wilson [4], when cyclotrons capable of delivering protons with sufficient energy to penetrate to tumours in the body were being developed. The first treatment, with high-energy protons traversing the entire body, was performed in 1954 at Lawrence Berkeley National Laboratory [5]. Treatments using the Bragg peak were performed starting in 1957 at the University of Uppsala, Sweden [6] and from 1961 at the Harvard Cyclotron [7–10]. During the last two decades, translation of proton therapy technology from research laboratories to hospitals has made it a more commonly available treatment option. As of 2013, approximately 10.000 patients receive proton therapy each year [11]. This is still only a very small fraction of the millions of patients worldwide receiving radiotherapy.

Proton therapy improves the therapeutic ratio. For the same dose delivered to the target, the integral dose to the patient is reduced by a factor of 2 to 3, which is beneficial to reduce side effects. It is particularly important in the case of curative treatment of paediatric and young patients, who are expected to have a long life expectancy and can therefore be affected by long term side effects such as secondary radiation-induced tumours. In the case of tumours that require high radiation doses to achieve local control, the radiation dose can be increased as compared to conventional radiotherapy. Proton therapy is very suitable for tumours that are located close to critical organs in the brain, head and neck, and spinal regions.

## 1.2. Proton interactions with matter

Protons interact with matter through different processes, which determine the dose distribution of a beam. Next, we briefly review these processes.

### 1.2.1. Stopping power

For radiotherapy with protons, the Bragg peak in the depth-dose distribution and the finite range of the beam are the most important characteristics. These are a result of the proton losing energy as it traverses matter, which is described with the stopping power. That is, the mean energy loss $S(E)$ per unit path length $z$ of the traversed material:

$$S(E) = -\frac{\mathrm{d}E}{\mathrm{d}z}. \tag{1.1}$$

The stopping power of water is plotted in figure 1.2 for proton energies $E$ up to 230 MeV, which is the energy range of relevance for clinical proton beams. The increase of the stopping power with a decrease in proton energy gives rise to the Bragg peak in the dose distribution. Stopping powers in the range of 0.5 MeV to 230 MeV can be calculated with an accuracy of a few percent using the Bethe formula [12] with Barkas [13] and Bloch [14] corrections. The energy loss is mostly

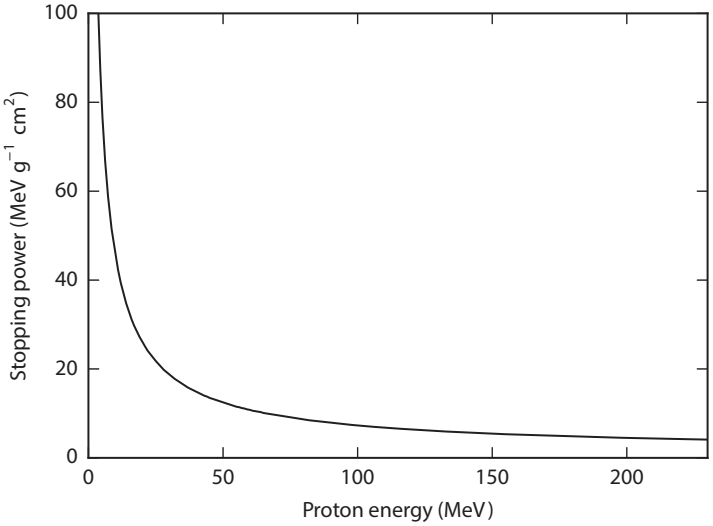



**Figure 1.2** | The stopping power of protons in water, as a function of the proton energy in the therapeutic energy range.

due to electromagnetic interactions of the protons with atomic electrons, which ionize or excite atoms. A small contribution is from elastic interactions with atomic nuclei.

### 1.2.2. Range

The range of a proton is defined as the distance traversed until it has lost all of its energy. After the proton has come to rest, no radiation dose is deposited further downstream in the beam path. The mean range $R$ of a proton with initial energy $E_p$ can be approximated by integrating the reciprocal stopping power:

$$R = \int_0^{E_p} \frac{1}{S(E)} \, \mathrm{d}E. \tag{1.2}$$

This mean range is the depth at which 50 % of the protons have stopped. In figure 1.3, the relation between the proton energy and the range in water is shown. A 230 MeV proton has a range of 33 cm in water, which is sufficient to reach tumours throughout the body. At this energy, a proton travels at 60 % of the speed of light.

Because of statistical fluctuations in the interaction processes, not every proton with the same energy has an identical range. This range straggling effect results in the finite dose gradient at the end-of-range of a proton beam. In clinical practise, the range of a beam is normally defined relative to a certain point in the dose distribution, for example at the distal 80 % of the maximum dose. The 80 % dose fall-off coincides approximately with the mean range.

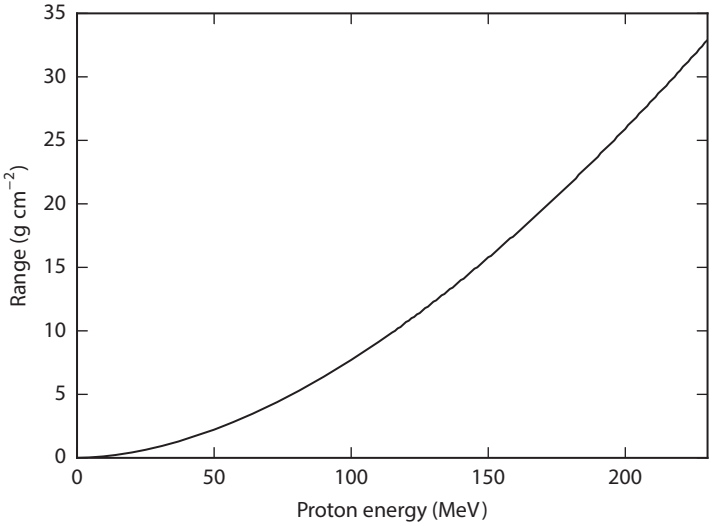



**Figure 1.3** | The projected mean range of a proton beam in water, as a function of the proton energy in the therapeutic energy range.

### 1.2.3. Nuclear reactions

Protons also interact with nuclei through elastic or inelastic reactions, which are not included in the stopping power. At energies above a few MeV, inelastic reactions can excite nuclei that occur in human tissue, such as $^{12}$C and $^{16}$O. At somewhat higher proton energies, nuclear spallation also occurs. As a rule of thumb, a nuclear reaction occurs for approximately 1 % of the protons per centimetre of water. Therefore, most protons in clinical beams never undergo a nuclear reaction. The loss of primary protons due to nuclear reactions does impact the dose distribution, because the proton fluence is reduced as a function of depth. Secondary protons from nuclear interactions also contribute to the dose. Neutrons and $\gamma$-rays from nuclear reactions result in a small background dose.

### 1.2.4. Scattering

Because of interactions with the Coulomb field, protons will experience a very small deflection as they pass close to a nucleus. Although the deflection from a single scatter is negligible, many of these interactions result in a random walk in angle. The proton angular distribution due to this multiple Coulomb scattering process is well described by Molière theory [15, 16]. The main part of the angular distribution is almost Gaussian. For clinical beams, proton scattering is relevant as it affects the lateral extent and dose gradient of proton beams. The radial spread of a proton beam due to scattering has a standard deviation of approximately 2 % of the range in water.

## 1.3. Clinical proton beams

Most radiotherapy treatments consist of multiple treatment fractions. The typical dose delivered to the target is about 2 Gy per fraction, although fewer fractions with higher doses are also used for certain treatments. For each fraction, multiple beams are delivered, which irradiate the target from different directions. This section gives an overview of the methods to plan and deliver clinical proton beams.

### 1.3.1. Treatment planning

Similar to conventional radiotherapy, a proton therapy treatment is based on a carefully designed treatment plan for the individual patient. For the treatment plan, a computed tomography (CT) scan is obtained with the patient in the same position as used during treatment. Based on the prescription of the physician, a treatment planning system is then used to design proton beams that together deliver a dose distribution that provides a good trade-off between target coverage and sparing of organs at risk. For the target, the goal is typically to create a high and uniform dose volume. In the case of organs at risk, the dose tolerance and the importance of the mean dose or the maximum dose varies depending on the type of organ.

Different options exist to design proton therapy treatments. Each beam can deliver a uniform dose to the target, similar to the proton beam shown in figure 1.1. Another option is to use additional sets of beams that each deliver a uniform dose to a different part of the target. This option gives more flexibility to spare organs at risk, particularly if the target is partially wrapped around such an organ [17]. Last, each beam can deliver an optimized inhomogeneous dose distribution to the target, which is a technique known as intensity modulated proton therapy [18]. The dose delivered by all beams combined then yields the desired uniform dose. Intensity modulation provides a large degree of freedom to optimize trade-offs. Due to the complexity of the optimization, this process is assisted by computer algorithms.

A unique requirement for the design of proton treatment plans is the need to accurately determine the finite ranges of the proton beam. For this purpose, proton stopping powers as defined in equation (1.1) need to be determined for the patient's anatomy. A conversion algorithm is used to determine the stopping powers from a CT scan, considering the typical composition of human tissues [19].

The range of the proton beam is normally different throughout the treatment field. In the case of a uniform dose, the beam is given a range at each position that is sufficient to reach the distal surface of the target volume. Intensity modulated beams may also stop within the target volume. Unlike beams that traverse homogeneous matter, these ranges cannot always be clearly defined as a single mean range. If the anatomy is inhomogeneous, protons that enter the patient with the same energy and at the same position may end up having different ranges, because differences in the scattering can result in different trajectories through the anatomy.

### 1.3.2. Acceleration and beam delivery

A cyclotron or synchrotron can accelerate protons with a sufficient beam current to achieve a dose rate of at least 1 Gy $min^{-1}$ for most targets. This dose rate is desired to minimize the time during which the patient has to remain immobilized and to maximize the number of treatment slots. A cyclotron accelerates a continuous fixed-energy beam and is used with an energy selection system to reduce the proton energy. A synchrotron can accelerate bunches of protons to different energies.

A particle accelerator can either be used with a single treatment room, or the beam can be shared between several rooms. Treatment rooms can be designed with fixed beams, which are used for specific treatments, or with a rotating gantry, which enables the patient to be easily treated with beams from multiple angles. From the particle accelerator, a small pencil-beam of protons is magnetically transported to the treatment room. Proton delivery methods differ in the way this pencil-beam is used to create the radiation field targeting the tumour.

#### Passive scattering

Passively scattering a pencil-beam is the conventional method of creating a proton field. A pencil-beam with a fixed energy enters the treatment head, in which a series of mechanical devices modulate the energy and broaden the beam [20]. Combined with a custom machined range compensator and aperture for each treatment field, this well established method provides a beam that is precisely shaped to the distal edge of the tumour. Because the energy modulation width of the proton beam is constant throughout the field, the extend of depth of the high-dose area is fixed for each beam. Therefore, passive scattering cannot fully conform the high-dose area at the proximal side of the tumour. The neutrons induced by proton interactions with the scattering and collimating devices are also a disadvantage [21].

#### Pencil-beam scanning

Pencil-beam scanning is a more recent method, that aims the proton pencil-beam directly at the target [22, 23]. A three-dimensional dose distribution is created by magnetically scanning the beam in the two-dimensional plane to the extends of the target volume, while changing the proton energy to irradiate at multiple depths. Compared to passively scattered delivery, pencil-beam scanning has more dimensions of freedom which allow for the dose distribution of each beam to conform to the proximal and distal surfaces of the target. Another advantage is the avoidance of the production of neutrons as a result of proton interactions with the components in the treatment head. For the delivery of intensity modulated proton therapy, pencil-beam scanning is a prerequisite. The ability to change the dose delivered by each spot enables the delivery of inhomogeneous dose distributions.

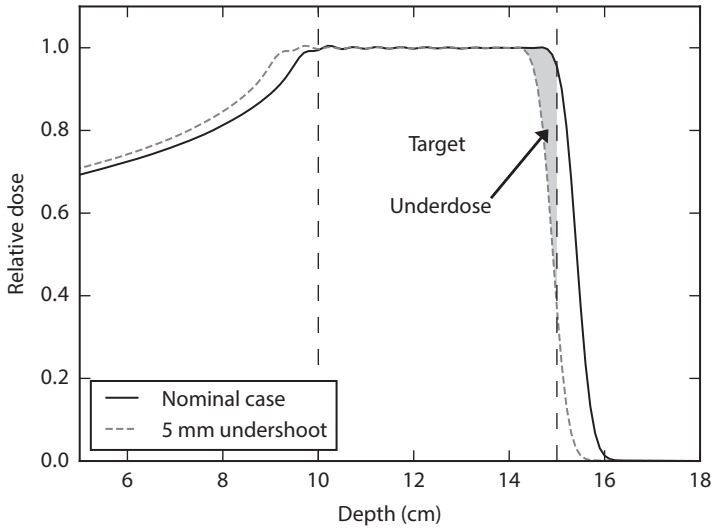



**Figure 1.4** | The effect of a range undershoot on the depth-dose distribution of a proton beam in water. This beam was designed with the maximum dose at a depth of 10 cm to 15 cm.

## 1.4. Range uncertainty

The depth-dose curve depicted in figure 1.1 shows the great potential of proton beams to deliver a high radiation dose to the tumour volume while sparing surrounding tissue. However, it is also a somewhat optimistic picture, as it does not display the uncertainties that are encountered when such a beam is delivered to a patient.

An important challenge in proton therapy is the uncertainty in the range of the proton beam. The end-of-range is where the beam features its sharpest dose gradient. Therefore, an undershoot of the proton range can lead to the distal edge of the tumour not receiving the intended dose, as shown in figure 1.4. For this beam with a range of 15 cm, a 5 mm undershoot of the beam reduces the dose delivered to the most distal part of the target by about 50 %.

### 1.4.1. Sources

There are multiple sources of uncertainties in the proton range, which were reviewed by Paganetti [24]. Although the exact magnitude of the range uncertainty is not completely known, reasonable estimates of the total magnitude can be made.

Many of the sources of range uncertainties relate to the determination of the proton range based on the patient CT scan. The method used to infer proton stopping powers from CT numbers is not exact. A conversion has to be made between the CT x-ray attenuation measurements and proton stopping powers, requiring certain

assumptions on the composition and ionization potential of the tissues [25]. There is also degeneracy, that is, the same x-ray attenuation may correspond to different stopping powers. Uncertainties related to image noise [26], calibration, and the spatial resolution [27] of the CT scan also impact the converted stopping powers to a lesser extent. These uncertainties generally are systematic for most of the tissue traversed by a typical beam. Therefore, the magnitude of the total uncertainty is proportional to the range of the beam.

Paganetti [24] estimated the range uncertainty due to these issues to be approximately ±2.5 % of the beam range in water for most tumour sites, evaluated at a $1.5\sigma$ confidence level. Uncertainties can be larger for specific cases, such as when metal implants cause artifacts, which degrade the quantitative precision of the CT numbers. These systematic uncertainties impact the entire course of treatment. In addition, a random uncertainty of about ±1 mm is attributed to the reproducibility of the beam and the patient setup.

Besides these uncertainties related to the determination of the properties of the patient anatomy, the algorithm used to calculate the range of the proton beams can be a source of uncertainty. In relatively homogeneous tissue, the additional uncertainty due to the dose calculation is almost negligible. However, in the case of interfaces between tissues with different densities, the calculation of the proton scattering is more complicated. Pencil-beam dose calculation algorithms, which are commonly used in current clinical systems, can introduce an uncertainty up to about ±2.5 % for specific treatment sites and beam arrangements [24].

Other uncertainties will play a role if the initial treatment plan does not accurately reflect the patient anatomy during treatment. Changes in the patient weight are one example [28]. For treatments of tumours in for example the lung and the abdomen, motion associated with the breathing cycle also needs to be considered. The aforementioned range uncertainties do not incorporate such factors.

### 1.4.2. Treatment margins

Range uncertainty is incorporated in the design of treatment plans, by making the plan robust against uncertainties. For beams delivering a uniform dose to the tumour, robustness can be achieved by adding a range margin to the distal surface of the target. If the delivered range remains within the margin, the target receives the planned dose. For intensity modulated proton beams, the situation is more complex, because parts of the proton fields may stop inside the target volume. Robust optimization, in which the uncertainty is incorporated in the treatment optimization process, is a good approach to plan intensity modulated proton beams [29].

In figure 1.5, a schematic drawing of a target, the range margin, and the position of the end-of-range is shown. While there is no universally accepted standard, indi-

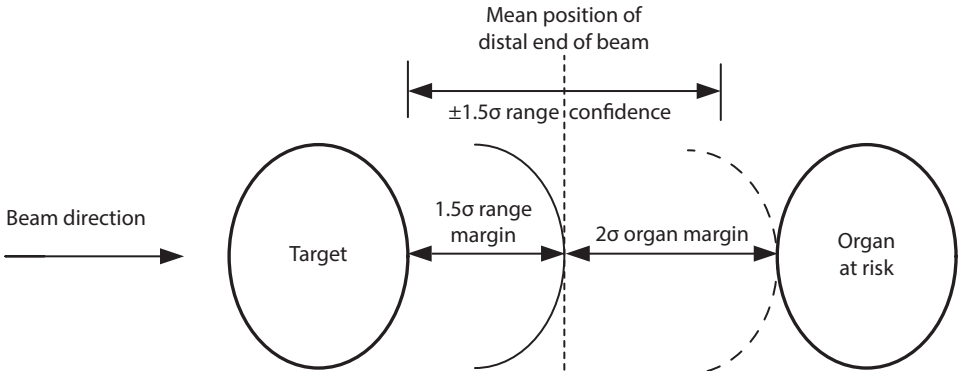



**Figure 1.5** | Typical range margins that ensure a proton beam delivering a uniform dose is robust against range uncertainty. The distal dose surface is positioned at a distance from the target of 1.5$\sigma$ of the expected range distribution. A separation of 2$\sigma$ between the mean end-of-range and a downstream organ at risk is also shown. If the organ at risk is located closer to the target, it may not be clinically possible to use a beam in this direction.

vidual treatment beams are generally made robust against a $-1.5\sigma$ variation of the end-of-range, which according to a normal distribution gives a confidence level of 94 % that the target is fully covered. This remains somewhat of an approximation, because it is not fully known whether the range in patients follows a normal distribution and whether the currently used methods to estimate the range yield the mean of the population.

The range uncertainties also need to be taken into account when selecting the directions from which to irradiate the target. If an organ at risk is located distal to the target in the beam path, a margin is required to be certain that the beam will not overshoot into the critical tissue. Unlike the target, an organ may be in the path of only a single beam. If a high confidence is required that the end-of-range position is proximal to the organ at risk, a separation of $+2\sigma$ between the mean position of the end-of-range and the edge of the organ could be a reasonable choice, which provides a 98 % confidence level. This margin is also drawn in figure 1.5.

This means that to use a beam in the direction of the organ at risk, an additional separation of 3.5$\sigma$ of the range distribution may be required between the target and the organ. Other compromises between target coverage and dose to the organ at risk may still exist for such a beam, due to the finite distal dose gradient or other nearby organs. However, as compared to the lateral edge of a typical proton or x-ray beam, the distal edge of the proton beam has a sharper dose gradient.

At our institution, we use a range margin of 3.5 % of the range in water for most treatment sites, which is an estimate of 1.5$\sigma$ of the range distribution based on studies such as discussed in the previous section. An additional 1 mm margin is added to

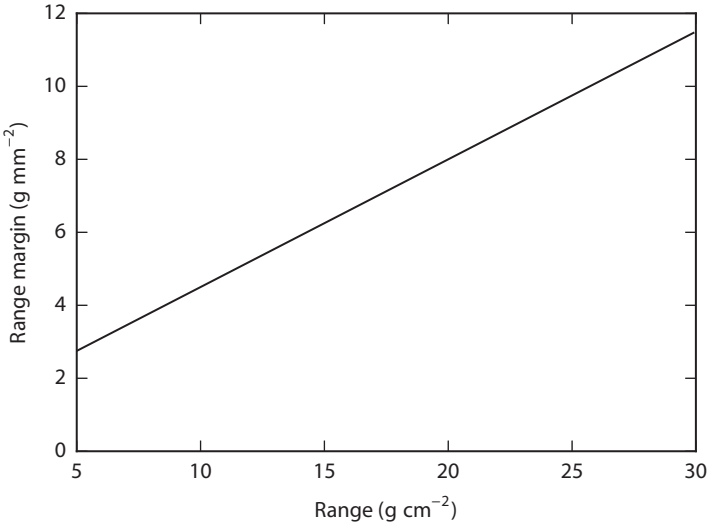



**Figure 1.6** | The magnitude of a 3.5 % + 1 mm proton range margin as a function of the range in water. This margin is used for clinical treatment plans at our institution.

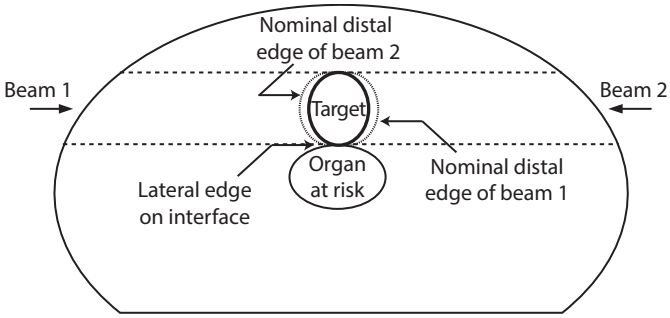
account for random errors in the beam delivery and patient setup. Figure 1.6 shows the magnitude of the 3.5 % + 1 mm range margin as a function of the proton range in water. A 3.5$\sigma$ separation between a target and an organ at risk amounts to 8 % of the range in water according to the same estimate.

This margin recipe is a general one that applies to most treatments. For specific treatment sites, additional margins may be added, because of range differences that are expected due to the difficulty of the patient immobilization and setup, or due to motion of the patient during treatment. In some cases, it may also not be clinically feasible to design a treatment with the standard margins for all beams.

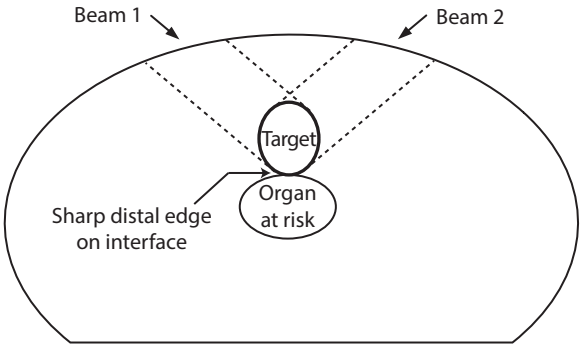
### 1.4.3. Clinical impact

Making a proton therapy treatment plan robust against range uncertainty does come at a cost. Additional dose is delivered because of the necessary range margins, and other compromises have to made when designing treatment plans. To illustrate the clinical impact of range uncertainties, we consider as an example a clinical case shown in figure 1.7, in which a target is irradiated that is located adjacent to an organ at risk. In this example, the proton beams deliver a uniform dose to the target and the treatment is designed to minimize the dose to the organ at risk.

The treatment plan in figure 1.7a, in which the target is irradiated from the left and right sides, is robust against uncertainties. Beam directions are selected that do not point towards the organ at risk, which avoids the risk of an overshoot of the beam into this organ. Also, the end-of-range of both beams is planned with an




(a) Treatment that is robust against range uncertainties.




(b) Improved treatment that is feasible if the range uncertainty is small.

**Figure 1.7** | Two potential treatment plan designs to irradiate a target adjacent to an organ at risk, with the aim of minimizing the to dose to the organ at risk. The dashed lines show the edges of the proton fields.

additional margin, such that it is positioned somewhat distal to the target in the nominal case. Therefore, the target still receives full dose as long as the range of the beam remains within the planned uncertainty margin.

There are several disadvantages of this treatment design. First, the margin on the end-of-range results in a volume of tissue around the target to receive almost the full target dose. Second, the distance between the target and the surface of the body is large for these beam directions, resulting in more normal tissue that is irradiated proximal to the target in the beam path. Last, the lateral edge of the beam is located on the interface between the target and the organ at risk. The dose gradient at the lateral edge of the beam increases with range from about 3 mm to 12 mm at the 20 % to 80 % dose level, which is less sharp as compared to the gradient at the distal edge of the beam of 3 mm to 5 mm. The shallower dose gradient results in a high-dose volume that is less conformal to the target volume.

A different treatment design is shown in figure 1.7b. The same target is irradiated from two oblique angles, with the end-of-range of the beams located on the interface between the target and the organ at risk. These beam angles result in less integral dose to be delivered to the patient. The distal dose fall-off is sharper because of the shorter beam range, and it is used to conform the high-dose area to the target. This alternative design clearly is more optimal, but it is not robust against range uncertainties. Therefore, it will only be feasible if the range uncertainty is significantly reduced as compared to the current clinical practise.

## 1.5. Reducing range uncertainty

Because of the limitations imposed on current proton treatments by the uncertainty in the beam range, a reduction of this uncertainty is highly desired. Next, we review the main areas of current research to reduce the range uncertainty.

### 1.5.1. *In vivo* beam monitoring

A direct way of reducing range uncertainty is to verify the range of the beam *in vivo* during patient treatment [30]. Based on range verification measurements, the treatment plan could be adjusted to ensure tumour coverage while maintaining reduced margins. *In vivo* range verification would also be an excellent quality assurance method to confirm the planned range is reproduced during treatment.

#### PET

Positron emission tomography (PET) is the first method that was investigated for *in vivo* range verification of proton and ion beams [31, 32]. The concept of measuring beam range with PET was proposed in the late 1980s [33]. A number of the isotopes that are produced in proton-induced nuclear reactions with the patient's tissue are positron emitters, in particular $^{11}C$ and $^{15}O$. A PET scanner detects the coincident

$\gamma$-rays that are emitted when the positron, which is emitted during the decay of these radioactive isotopes, annihilates with an electron. Tomographic reconstruction provides a three-dimensional distribution of the positron emission, which can be correlated to the delivered radiation dose. The main disadvantages of PET are related to the half lives of the main positron emitters, which are on the order of minutes. Real-time feedback is therefore not possible, and the signal represents an accumulation of the different treatment beams that have been delivered. Also, the positron emitting isotopes diffuse before they decay. Therefore, depending on the anatomical site, the correlation between the proton dose and the PET measurement is blurred, which limits the precision. Range verification using PET has been evaluated in a number of clinical trials [34, 35]. Whether the method has a clear clinical benefit is still debatable, because of the difficulty of obtaining quantitative information [30].

**Prompt γ-ray detection**

More recently, there has been interest in detecting prompt $\gamma$-rays for range verification, which are emitted almost instantaneously as a result of proton-nuclear interactions. These $\gamma$-rays are emitted while the excited residual nuclei decay to their ground state. Prompt $\gamma$-ray emission accompanies many different nuclear reactions. A major advantage of using prompt $\gamma$-rays is the potential for direct real-time range verification. The half lives of most of the relevant excited states is on the order of femtoseconds to picoseconds. Therefore, issues related to the delayed PET signal are entirely avoided. Experiments to show a correlation between the proton beam range and prompt $\gamma$-ray emissions were first published by Min et al. [36] in 2006. The detection of these $\gamma$-rays during the delivery of clinical proton beams is challenging, because of the neutron-induced background radiation. Range verification using prompt $\gamma$-rays has not yet been performed in a clinical setting.

**Direct dose measurement**

For very specific treatment sites, such as the prostate, a dosimeter could be placed in a body cavity [37]. Implantable dosimeters could also be an option for some treatment sites. While these measurements are limited to specific points within the patient's body, directly measuring the dose has advantages over the use of more indirect methods based on secondary radiation. This method could be complimentary to other *in vivo* range verification technology that is more generally applicable.

**MRI**

The use of MRI has been investigated to evaluate the range of the beam by analysing changes in the tissue that have been observed with certain MRI protocols [38, 39]. This method relies on effects which occur on longer time scales, which limits the potential for treatment adaptation as compared to other methods. There are also biological effects and differences between patients, which are not fully understood.

#### Acoustic

The acoustic signature of the proton beam could also be used for range verification. The measurement of this signal was already proposed in 1979 [40] and the application to proton therapy has recently received renewed attention [41]. If the proton beam is pulsed with a sufficient number of protons per pulse, sound waves due to the thermal effects of the radiation beam can be detected with ultrasound equipment. A potential advantage of this method is the direct relation between the dose, i.e. energy deposition, and the signal. This pulsed delivery however is not compatible with common proton acceleration methods. The propagation of the wave to the surface of the patient is also far from straightforward to model and will likely introduce errors in the quantification of the range.

### 1.5.2. Dose calculation

Clinical proton dose calculations currently use pencil-beam calculation algorithms, in which the proton field is split in small computational pencil-beams for which the dose distribution is determined [42]. Pencil-beam algorithms typically use measured dose distributions in water. Approximate methods are employed to determine to dose deposited to the patient based on these measurements. These methods are less reliable in the presence of heterogeneous tissue.

Monte Carlo methods, in which the interactions of many individual protons with the patient's anatomy are simulated, are a very accurate method to calculate proton dose. These methods are already widely used for research studies and developments towards clinical implementation are ongoing [24]. Because of the high computational cost of general purpose Monte Carlo simulations, which feature very detailed models of proton-matter interactions, simplified algorithms that focus on the more clinically relevant processes are also being investigated [43].

### 1.5.3. New imaging technologies

The ±2.5 % range uncertainty that results from the conversion of CT images, can be reduced by future improvements to imaging for treatment planning. Spectral and photon-counting CT imaging has the potential to provide more detailed information about the patient's anatomy and to reduce the degeneracy in the conversion of CT information to proton stopping powers [44]. This technology is also of interest for diagnostic purposes and can be adapted to proton therapy applications.

Another possibility is to perform radiography or tomography using protons instead of x-rays [45–49]. By using protons itself to image the patient, the indirect method of inferring proton interactions from x-ray based imaging is avoided. For full tomographic imaging, protons need to be accelerated to higher energies to traverse the entire body, which is only possible if the accelerator and beam lines are

specifically designed for this purpose. The scattering of protons is an issue, which limits the spatial resolution of proton radiography or proton CT. Combinations of x-ray based and proton-based imaging are also being considered. Similar to most technologies for *in vivo* range verification, the development of a system for proton imaging would be specific for therapy applications. Beyond range verification, there may be other applications, such as imaging during the delivery of the treatment.

## 1.6. This thesis

In this work, we investigate the reduction of range uncertainty of proton beams. Our aim is to develop physics-based methods to reduce the range uncertainty by improving the calculation of the proton range and by verifying the range during treatment. The goal of these developments is to work towards reducing the current uncertainties, which amount to about 3.5 % of the range in water at a $1.5\sigma$ confidence level, to an uncertainty of around one millimetre, independent of the range.

We believe *in vivo* range verification will be key to reduce and to better understand range uncertainties. Several sources of uncertainties in the treatment planning process will be difficult to fully eliminate. *In vivo* range verification can reduce these by enabling continuous monitoring and adaptation of the proton range. Moreover, before reducing treatment margins, there will be a desire to validate technologies to improve the accuracy of the calculated range. Safety will also be improved, because significant deviations from the intended treatment plan can be quickly determined.

Our research focuses on the use of prompt $\gamma$-rays for range verification. Compared to other methods, this approach has the unique potential to provide real-time verification of the range of proton beams. This is a major advantage, because it can allow for treatments to be directly adapted to variations of the range. Second, because these $\gamma$-rays are emitted nearly instantaneously as a result of proton-nuclear interactions, there are no complex biological effects.

For the *in vivo* range verification of clinical proton beams, a method will be required to reliably determine the absolute range of the proton beam from prompt $\gamma$-ray measurements performed during treatment. We hypothesized that a detailed spectroscopic measurement and analysis of the proton-nuclear reactions with tissue, can facilitate the development of a system and physics model for this purpose. Based on this concept, we investigate both the detection of prompt $\gamma$-rays and the modelling of nuclear reactions.

We study the feasibility of quantitative spectroscopy of prompt $\gamma$-rays emitted during the delivery of proton beams, by developing a pre-clinical prototype detection system. This system is also used to study the nuclear reactions of interest, based on which models are created to predict the $\gamma$-ray emission. In clinical use, the *in vivo* measurements will need to be compared with models that are based on the

treatment plan. Therefore, we investigate automated Monte Carlo simulation of the proton beam transport through the patient's tissue, using information from the clinical treatment planning system.

We also investigate the specific uncertainties that arise if patients have metallic implants. These implants are often used to stabilize the spine after surgical resection of a tumour. Implants cause artifacts in the CT scan that is used for treatment planning, which affects the accuracy of the calculated proton dose. We therefore study the reduction of these artifacts, by improving CT image reconstruction methods. Current clinical dose calculation methods also do not consider the specific effects of metal implants. For this reason, we study Monte Carlo dose calculations, which can fully model the proton interactions with the implants.

### 1.6.1. Outline

Chapter 2 contains a simulation and literature study of the prompt $\gamma$-ray emissions from nuclear interactions of protons in the therapeutic energy range. This study provides an overview of the relevant reactions, which is necessary to conceptualize and develop range verification methods based on prompt $\gamma$-ray detection. The data available in literature was analysed and simulations were performed with different nuclear reaction models.

Chapter 3 describes the development of a prototype prompt $\gamma$-ray detector and data acquisition system. The detection system was developed to resolve prompt $\gamma$-rays in dimensions of energy- and time, which allows specific nuclear reactions to be identified and quantified. We also performed initial tests of the prototype system to show the correlation between discrete $\gamma$-line excitations and the proton range in a water phantom.

Chapter 4 describes a novel method to verify the range of proton pencil-beams based on prompt $\gamma$-ray spectroscopy. The proton range and the elemental composition of irradiated tissue were determined from spectroscopic measurements through an optimization procedure. We measured differential cross section for 15 $\gamma$-ray excitations to support the model. Experiments with phantoms were performed to evaluate the performance of the method.

Chapter 5 reports on an experimental study to investigate the applicability of prompt $\gamma$-ray based range verification to passively scattered proton beams. This is of interest because passively scattered beam delivery is currently in widespread use in existing proton therapy centres. We developed a method to verify the consistent delivery of treatments using prompt $\gamma$-ray detection synchronized with the range modulation cycle.

Chapter 6 describes the development of a framework for automated Monte Carlo simulations of clinical treatment plans. The study focused on a systematic consideration of uncertainties in the physics models and the necessary calibrations

for clinical use. The Monte Carlo methods can reduce range uncertainty by improving dose calculation accuracy if proton beams pass through implants or inhomogeneous tissue. Second, they enable accurate modelling of prompt $\gamma$-ray emissions based on measured cross sections, to verify the consistency of the planned treatment with *in vivo* range verification measurements.

Chapter 7 describes CT image reconstruction methods that were developed to improve the visual and the quantitative accuracy of CT scans of patients with metal implants. The artifacts were reduced by better modelling the interactions between the CT x-rays and the implants.

Chapter 8 contains a study in which the developed CT reconstruction and Monte Carlo simulation methods were applied to investigate the uncertainties caused by metal implants. We also assessed the impact on current treatment plans for chordoma patients with titanium implants. Because post-operative proton therapy is used for these treatments, implants around the tumour are common.

Finally, in chapter 9, we summarize the main conclusions of our research and provide recommendations for future research and clinical implementation.

# 2

# Simulation of prompt γ-ray emission during proton therapy

The measurement of prompt $\gamma$-rays emitted from proton-induced nuclear reactions has been proposed as a method to verify the range of a clinical proton radiotherapy beam. A good understanding of the prompt $\gamma$-ray emission during proton therapy is key to develop a clinically feasible technique, as it can facilitate accurate simulations and uncertainty analysis of $\gamma$-ray detector designs. Also, the $\gamma$-ray production cross sections may be incorporated as prior knowledge in the reconstruction of the proton range from the measurements. In this chapter, we perform simulations of proton-induced nuclear reactions with the main elements of human tissue, $^{12}$C, $^{16}$O, and $^{14}$N, using the nuclear reaction models of the GEANT4 and MCNP6 Monte Carlo codes, and the dedicated nuclear reaction codes TALYS and EMPIRE. For each code, we made efforts to optimize the input parameters and model selection. The results of the models were compared to available experimental data of discrete $\gamma$-ray line cross sections. Overall, the dedicated nuclear reaction codes reproduced the experimental data more consistently, while the Monte Carlo codes showed larger discrepancies for a number of $\gamma$-ray lines. The model differences lead to a variation of the total $\gamma$-ray production near the end of the proton range by a factor of about two. These results indicate a need for additional theoretical and experimental study of proton-induced $\gamma$-ray emission in human tissue.

Based on published work: J.M. Verburg, H. Shih, and J. Seco. Simulation of prompt gamma-ray emission during proton radiotherapy. *Phys. Med. Biol.* **57**, 5459–5472 (2012).

The main advantage of proton radiotherapy is the finite range of the protons in the patient. However, uncertainties in the proton range currently limit the ability to make full clinical use of the sharp distal falloff of the proton beam. In an ideal scenario, the incident protons are given such energy to position the end-of-range of the beam exactly on the distal edge of the clinical target volume, sparing all tissue downstream in the beam path. Range uncertainties however, make necessary the use of additional margins to ensure tumour coverage. Also, if a tumour is located next to a critical organ, for safety reasons the lateral edge of the proton beam is placed on the tumour-organ interface instead of the sharper distal edge. Because of these limitations, it has been recognized that a means of verifying the proton range *in vivo,* can facilitate better treatment designs which could lead to reduced normal tissue complications or improved tumour control.

Although the primary protons stop inside the patient, a part of the secondary $\gamma$ radiation resulting from non-elastic nuclear interactions will escape the body, and could potentially be used to establish the range of the beam. These $\gamma$-rays consist of prompt photons, which are emitted during the nuclear reactions, and delayed emission from the decay of unstable nuclear reaction products.

Initial clinical trials have been performed on the use of positron emission tomography (PET), which detects the photon pairs produced due to the decay of positron emitters such as $^{11}$C and $^{15}$O and the subsequent annihilation of the positron [35]. A coincidence measurement of these 511 keV photons enables tomographic reconstruction of the distribution of the positron emitters, which is correlated to the proton range. However, this technique is indirect and does not facilitate an immediate verification of the range of the protons. Because of the delay between the production of the reaction products and their decay, the PET scan needs to be performed over an extended period of time. This delay limits accuracy, as metabolism will result in a washout of the reaction products [35]. Other issues include patient motion during the PET scan [35] and the uncertainty in the chemical composition of the tissue being irradiated, which affects the correlation between the positron emission and the proton range [50].

The detection of prompt $\gamma$-rays has been proposed as an alternative method, which can provide a direct and potentially more precise proton range verification [36, 51–53]. Since prompt $\gamma$-rays are emitted nearly instantaneous, treatment plan deviations could be determined prior to the actual treatment, by delivering only a small subset of the protons. Range errors could also be tracked continuously during treatment. Spectroscopy of the $\gamma$ emission may reduce uncertainties due to tissue composition, because the emitted $\gamma$ energies are unique to the nuclear structure of the reaction products.

The measurement and collimation of high-energy $\gamma$-rays is however challenging. Significant research and development is needed to determine the feasibility of this

method and the potential detector designs for range verification. Simulation studies play an important role in these developments.

Monte Carlo simulations in particular may be employed to design a complete simulation of both the radiation interactions as well as the geometry of the treatment hardware and patient anatomy. It is important to critically evaluate such simulations for each particular application under study. For dose calculations, the electromagnetic interactions of the protons are of primary importance, which are well known and can be reliably simulated. Simulation of prompt $\gamma$-ray emission depends on detailed modelling of the hadronic interactions. These processes are not as well understood, and nuclear reaction models therefore rely to a high degree on phenomenology. Such models cannot be expected a priori to have good predictive power for all reaction channels. A recent study on the production of positron emitters during proton therapy showed large differences between various Monte Carlo codes [54]. Compared to the positron emitters, the production of prompt-$\gamma$ rays depends on a far greater number of reaction channels.



In this chapter, we investigate the simulation of prompt $\gamma$-ray emission during proton therapy, using the Monte Carlo codes GEANT4 and MCNP6 and the dedicated nuclear reaction codes TALYS and EMPIRE. Proton-induced nuclear reactions in the 1 MeV to 200 MeV incident energy range on $^{12}$C , $^{16}$O and $^{14}$N were studied, which are the most abundant nuclides in human tissue. Of particular interest are the incident proton energies up to about 50 MeV, because of their impact on the $\gamma$-ray emission near the end-of-range of the proton beam. The simulation results are compared to experimental data reported in literature, and to evaluated nuclear data. The impact of the model differences on the simulation of proton range verification is also discussed.

## 2.1. Methods and materials

### 2.1.1. Prompt γ-ray emission

Prompt $\gamma$-ray emission due to proton-induced nuclear reactions is the result of a nucleus being brought into an excited state, which subsequently decays to the lower state accompanied by the emission of a photon. Above the particle separation energies, this process competes with the emission of other ejectiles such as neutrons, protons and alpha particles. If particle emission occurs, the residual nucleus may again be left in an excited state. Gamma emission therefore can originate from either the target nucleus or any of the reaction production created through fusion or fission.

The lower-lying nuclear levels of most nuclei have clearly distinct quantum states and their properties are well established [55]. The cross sections of certain transitions between these levels are sufficiently high to enable the discrete $\gamma$ emissions to be resolved from the background in a practical measurement. At high excitation

energies, many close nuclear levels exists whose properties and decay modes are not completely known.

At lower incident proton energies, only a small number of excited levels can be reached and the $\gamma$ spectrum therefore consists of a small number of resolvable discrete lines. At higher energies, many nuclear reaction channels and $\gamma$ emissions are possible, most of which cannot be resolved and are also referred to as a quasi-continuum [56].



### 2.1.2. Nuclear reaction modelling

Current nuclear reactions codes divide the nuclear reaction process into three stages, for which different models are used [57–59]:

- Direct reactions, in which the proton interacts directly with only one or two nucleons of the target. These reactions are associated with short reaction times and high incident proton energies.
- Pre-equilibrium interactions, which involve interactions with parts of the target nucleus before the target has reached statistical equilibrium.
- Compound reactions, which take place after the energy of the proton is shared statistically among the target nucleons.

In the energy range of therapeutic protons, all three stages are of relevance. Near the end-of-range of the protons, the compound reaction stage of main importance. At increasing proton energies, pre-equilibrium reactions and then direct reactions become dominant. Nuclear excitation and subsequent prompt $\gamma$-ray emission can accompany the nuclear reactions in all stages.

Nuclear reaction models require tabulated nuclear levels schemes and branching ratios as input data. Most of the known lower-lying levels are simulated as discrete levels. Above a certain cut-off energy, nuclear levels are simulated as a continuum based a level density model.

### 2.1.3. ENDF/B-VII evaluated data

Evaluated data for proton-induced reactions is available in the ENDF/B-VII library [60]. These are generated using nuclear reaction models benchmarked to experimental data. Some small manual corrections to better fit experimental data may sometimes also be performed. Data for proton interactions are provided for energies up to 150 MeV and originate from the LA-150 evaluation, which was performed in the 1990s at Los Alamos National Laboratory [61].

The data for proton reactions with $^{12}$C, $^{14}$N, and $^{16}$O were evaluated by Chadwick and Young [62], who specifically focused on their application for radiation transport simulation of particle therapy. Benchmarks to measured proton, neutron and alpha emissions are reported, but no mention is made of benchmarking to experimental data of discrete $\gamma$ lines. The evaluation was performed using the GNASH nuclear reaction code [63], and optical model parameters and level densities were analysed and fitted.

In the published ENDF/B-VII data, prompt-$\gamma$ ray emissions are only provided in the form of angle-integrated continuous spectra with a relatively coarse energy resolution. Specific cross sections for the discrete lines are not included. For proton energies up to 20 MeV however, we were able to extract cross sections for the main discrete lines from the continuous data. The total non-elastic cross section and the total $\gamma$-ray production cross section were also obtained.

### 2.1.4. Nuclear reaction simulations

The nuclear reaction simulations were performed using the nuclear reaction models of the Monte Carlo codes GEANT4 9.5 [64] and MCNP6 beta 2 [65]. Also, we used the dedicated nuclear reaction simulation codes TALYS 1.4 [58] and EMPIRE 3.1 [57]. We made efforts to use the best available models and parameters, as discussed in this section. The simulations were performed with a 1 MeV energy resolution for the incident protons.

#### Geant4

GEANT4 generates nuclear reaction events on-the-fly using built-in nuclear reaction models. We used the binary cascade model, which performs an intranuclear cascade followed by a precompound and de-excitation model [66]. This model has been recommended for proton therapy applications [67]. Moreover, the alternative Bertini cascade was not suitable as it only provides a simple model for prompt $\gamma$-ray emission, that does not consider the discrete nuclear energy levels.

For incident protons with an energy below 45 MeV, the cascade is not used and the simulation starts with the precompound model. In the de-excitation model, Fermi break-up is activated by default for light nuclei; an evaporation model based on Weisskopf and Ewing [68] theory is used for heavier nuclei. Gamma emission from the excited residual nuclei is handled by the photon evaporation model, which uses tabulated nuclear levels and branching ratios from the Evaluated Nuclear Structure Data File [69]. A Fermi gas model is implemented for levels densities above the discrete tabulated data.

The standard GEANT4 physics lists relevant to proton therapy use a parametrization of the total non-elastic reaction cross section by Wellisch and Axen [70].

Comparison with the ENDF/B-VII optical model analysis [62] for the nuclei considered, showed relatively large differences in the proton energy range up to about 20 MeV. The Tripathi et al. [71] cross sections for light systems were in better agreement and selected instead.



By default, Fermi break-up is activated for $A < 17$; therefore the production of excited residuals for the reactions on carbon and nitrogen is fully simulated using this model. The initial compound system of $p$+$^{16}$O is however handled by the evaporation model followed by the Fermi break-up for residuals. Initial simulations showed this combination to strongly underestimate the excitation of the $^{16}$O residual nucleus. This may be a result of fact that the evaporation model does not consider the discrete levels of the nuclei, which leads to a mismatch with the photon evaporation model that does include the discrete levels. We therefore decided to activate Fermi break-up also for the complete oxygen reaction, and augmented the list of break-up products with the relevant nuclear levels. In addition, we removed the $\gamma$ energy corrections that are normally performed by GEANT4 to enforce per-event conservation of energy. Although such corrections are needed for certain applications, they result in unphysical $\gamma$ emission energies which is undesirable for the use case under study.

The calculations of $\gamma$ emission in GEANT4 are performed in the centre-of-mass frame, and a Lorentz boost to the lab frame is performed to simulate Doppler broadening. This boost was disabled in order to obtain the $\gamma$ energies in the centre-of-mass frame, which allows for the $\gamma$ emissions to be directly assigned to specific transitions.

To obtain cross sections, a simple geometry was simulated in which mono-energetic protons pass through 50 μm of the target material. This range is sufficiently small to neglect proton energy loss and the possibility of multiple nuclear reactions. All secondary photons generated where scored directly when produced. This process was repeated for all incident proton energies with a total of $10^9$ histories simulated per energy. By scoring the results using narrow energy bins, the discrete $\gamma$ lines could be resolved.

### MCNP6

For the MCNP6 simulations we used the default Bertini intranuclear cascade model followed by the multi-step pre-equilibrium and the evaporation model. Fermi break-up is used by default for light excited nuclei with $A \leq 20$ and an excitation below 44 MeV. MCNP6 includes a nuclear structure library by Prael [72], and by default uses the Gilbert-Cameron-Cook-Ignatyuk level density model [73].

The simulated geometry consisted of a sphere with a 50 μm radius, in the centre of which the primary protons were generated. A tally was activated on the surface to score the energy distribution of all photons passing through. The difference between

the MCNP6 and GEANT4 geometry was only for implementation reasons; for the proton energy range under study, both can be considered to represent an infinitely thin target. Similarly to the GEANT4 simulations, narrow energy bins were used, Doppler broadening was turned off, and mass-energy balancing was disabled.

It should be noted that MCNP6 also allows for the ENDF/B-VII data to be used, in which case no nuclear model calculations are performed, and the tabulated cross sections are sampled and interpolated instead. The cross section results will then be identical to the values that were obtained directly from the ENDF/B-VII database. Also, no detailed $\gamma$ spectra will be available.



### TALYS

The TALYS code is a modern dedicated nuclear reaction simulation code, which supports proton interaction with target nuclei $A \geq 12$ and a projectile energy up to 200 MeV. It features several models to determine the entire chain of possible nuclear reactions and their associated cross sections.

The optical model calculations, which determine the reaction cross section and transmission coefficients, are performed using the integrated ECIS-06 code. TALYS by default uses the Koning and Delaroche [74] global optical model potential for proton-induced reactions, which was not designed for light nuclei. Therefore, the specific optical model potentials by Young [75] where used when applicable (up to about 50 MeV), and the Madland global potential was used for higher energies [76]. The only exception is the $p + {}^{12}C$ reaction above 10 MeV incident energy. In this energy range, Young uses a Gaussian form factor for the imaginary surface derivative potential, which is not supported by ECIS. We therefore used for these energies the Madland global potential, and enabled the build-in TALYS feature to scale the total reaction cross section and transmission coefficients to the Tripathi universal parametrization [77].

The default exciton pre-equilibrium model was used for the pre-equilibrium stage. The decay of the compound nucleus was simulated with a Hauser-Feshbach model [78]. Levels densities above the discrete level region were calculated using the Fermi gas model.

TALYS includes a database of nuclear structure which is based on RIPL-3 [55]. The default database does not include deformation parameters for elements with $Z < 10$, which led to an underestimation of the direct contribution to the discrete levels at higher proton energies. Therefore, recommended parameters for carbon and oxygen were obtained from RIPL-3 and added into the database. The nuclear structure database was also adjusted to reflect the fact that several lower-lying levels of the nuclei being studied decay only through alpha emission [79, 80].

#### EMPIRE

The EMPIRE code [57] is also dedicated to the simulation of nuclear reactions. The PCROSS exciton pre-equilibrium model was selected as this model is recommend by the authors for proton-induced reactions. Simulations of direct reactions were also enabled.

The code provides the option to use any optical model potential from the RIPL-3 database [55]. The Young and Madland optical model potentials were used similarly to TALYS. For the $p + {}^{12}C$ reactions above 10 MeV incident energy, the same ECIS limitation exists and we used the manual scaling option to scale the reaction cross section to the published values obtained with the Young potential [62].

The discrete level and deformation parameter data originates from RIPL-3 and is mostly similar to the TALYS database. EMPIRE also implements a Hauser and Feshbach [78] model to simulate the compound nucleus. The EMPIRE-specific level densities were used.

### 2.1.5. Experimental cross sections

A literature search yielded a number of experimental studies of $\gamma$-ray emission during proton-induced reaction on $^{12}C$ , $^{16}O$ and $^{14}N$ [81–86]. Details of these studies are listed in table 2.1. These studies were all performed in the context of $\gamma$-ray astronomy, in which $\gamma$-ray lines are analysed to study nuclear reactions in the astrophysical environment.

Most previous measurements used protons with energies up to 50 MeV, which is also the most important energy range for proton therapy range verification. These published experimental cross sections describe the resolvable $\gamma$-ray lines resulting from transitions between the main lower-lying discrete levels. In our study, the angle-integrated cross sections reported in these experimental studies are compared. The reported statistical and systematic errors were combined to obtain a total estimate of the measurement error.

### 2.1.6. Proton Bragg curve simulations

The impact of the nuclear reaction model differences on prompt $\gamma$-ray emission during proton therapy was assessed by simulating proton irradiation of soft tissue and lung tissue. Using GEANT4 9.5, the proton energy spectrum was determined as a function of depth in the tissue. These energy spectra were then convolved with the cross section data obtained from the models, which allows for the impact of the cross section differences to be analysed without introducing additional uncertainties due to other models. The elemental compositions of the tissues as defined by ICRP were used [88]. The density of soft tissue was assumed to be 1.06 g cm$^{-3}$, and the density of lung tissue was set to 0.30 g cm$^{-3}$.

**Table 2.1** | Experimental $\gamma$-ray emission cross sections available in literature. These cross sections were reported by Dyer et al. [81], Narayanaswamy et al. [82], Lang et al. [83], Lesko et al. [87], Kiener et al. [84], Belhout et al. [85], and Benhabiles-Mezhoud et al. [86]. g.s.: ground state.

| Target | Emitter | $\gamma$ energy (MeV) | Transition | Study | Proton energy (MeV) |
|---|---|---|---|---|---|
| $^{12}$C | $^{12}$C | 4.44 | $2^+$ 4.44 → $0^+$ g.s. | Dyer et al. [81] | 5–23 |
| | | | | Lang et al. [83] | 40, 65, 85 |
| | | | | Lesko et al. [87] | 9–50 |
| | | | | Belhout et al. [85] | 5–25 |
| | $^{11}$C | 2.00 | $\frac{1}{2}^-$ 2.00 → $\frac{3}{2}^-$ g.s. | Lang et al. [83] | 40, 65, 85 |
| $^{16}$O | $^{16}$O | 6.13 | $3^-$ 6.13 → $0^+$ g.s. | Narayanaswamy et al. [82] | 23.7, 44.6 |
| | | | | Dyer et al. [81] | 5–23 |
| | | | | Lang et al. [83] | 40, 65, 85 |
| | | | | Lesko et al. [87] | 9–50 |
| | | | | Belhout et al. [85] | 20.0, 22.5, 25.0 |
| | | 6.92 | $2^+$ 6.92 → $0^+$ g.s. | Kiener et al. [84] | 9–19 |
| | | 7.12 | $1^-$ 7.12 → $0^+$ g.s. | Kiener et al. [84] | 9–19 |
| | | 2.74 | $2^-$ 8.87 → $3^-$ 6.13 | Lang et al. [83] | 40, 65, 85 |
| | | | | Kiener et al. [84] | 9–19 |
| | $^{12}$C | 4.44 | $2^+$ 4.44 → $0^+$ g.s. | Dyer et al. [81] | 14–23 |
| | | | | Lang et al. [83] | 40 |
| | | | | Lesko et al. [87] | 20–50 |
| | | | | Belhout et al. [85] | 20.0, 22.5, 25.0 |
| | $^{15}$N | 5.27 | $\frac{5}{2}^+$ 5.27 → $\frac{1}{2}^-$ g.s. | Lang et al. [83] | 40, 65, 85 |
| | | | | Lesko et al. [87] | 30, 33, 40 |
| $^{14}$N | $^{14}$N | 1.64 | $1^+$ 3.95 → $0^+$ 2.31 | Dyer et al. [81] | 5–20 |
| | | | | Lesko et al. [87] | 9–40 |
| | | | | Benhabiles-Mezhoud et al. [86] | 6–26 |
| | | 2.31 | $0^+$ 2.31 → $1^+$ g.s. | Dyer et al. [81] | 4–23 |
| | | | | Lang et al. [83] | 40, 65, 85 |
| | | | | Lesko et al. [87] | 9–40 |
| | | | | Benhabiles-Mezhoud et al. [86] | 6–26 |
| | | 5.11 | $2^-$ 5.11 → $1^+$ g.s. | Benhabiles-Mezhoud et al. [86] | 7–26 |
| | | 0.73 | $3^-$ 5.83 → $2^-$ 5.11 | Benhabiles-Mezhoud et al. [86] | 7–26 |
| | | 3.38 | $1^-$ 5.69 → $0^+$ 2.31 | Benhabiles-Mezhoud et al. [86] | 7–14 |
| | | 2.79 | $2^-$ 5.10 → $0^+$ 2.31 | Benhabiles-Mezhoud et al. [86] | 7–14 |
| | | 3.89 | $1^+$ 6.20 → $0^+$ 2.31 | Benhabiles-Mezhoud et al. [86] | 8–14 |

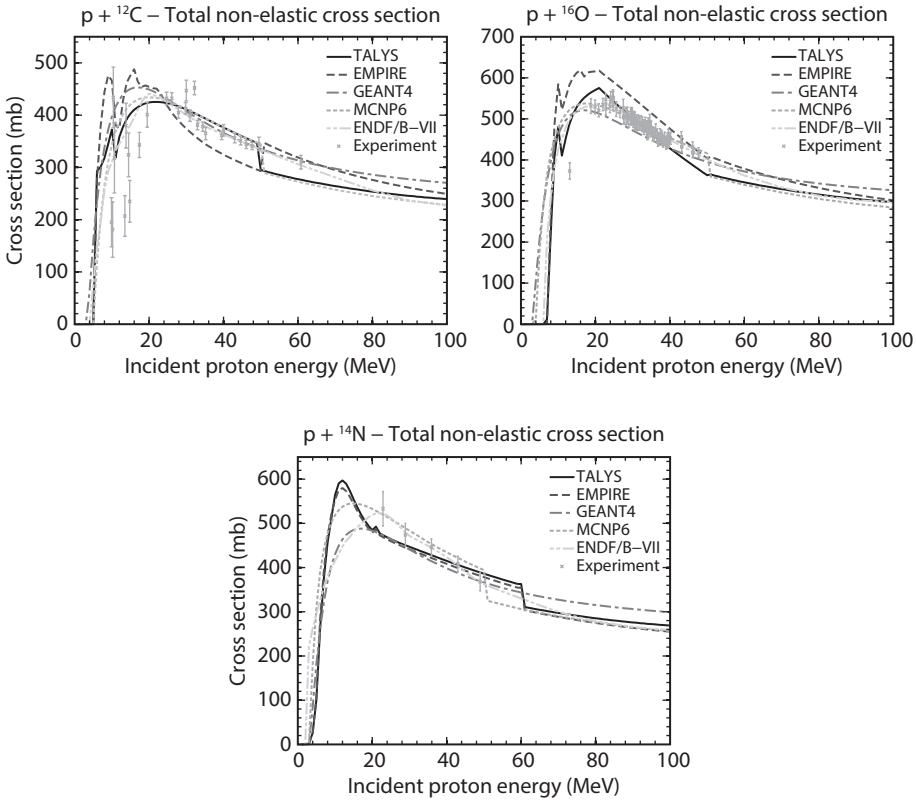



**Figure 2.1 |** Calculated total non-elastic cross sections of the proton-induced reactions. Experimental data were obtained from Bauhoff [89] and references therein.

## 2.2. Results

### 2.2.1. Total non-elastic cross sections

A comparison of the total non-elastic cross sections as calculated by the models is shown in figure 2.1 along with experimental data (Bauhoff [89] and references therein). Overall, with the previously discussed parameter adjustments, a reasonable agreement between the models and experiments was achieved. The experimental cross sections for the $p + {}^{12}\mathrm{C}$ reaction shows evidence of narrow resonances in the lower energy region. None of the simulations incorporate such details.

### 2.2.2. Discrete line of $^{12}$C

At lower proton energies, the $\gamma$-ray emission due to $p + {}^{12}\mathrm{C}$ reactions is dominated by the 4.44 MeV $\gamma$ emission. This emission is due to the ${}^{12}\mathrm{C}\,(p, p')\,{}^{12}\mathrm{C}^{*\,4.439}$ reaction, and also includes a small contribution of the ${}^{12}\mathrm{C}\,(p, 2p)\,{}^{11}\mathrm{B}^{*\,4.445}$ reaction. Because

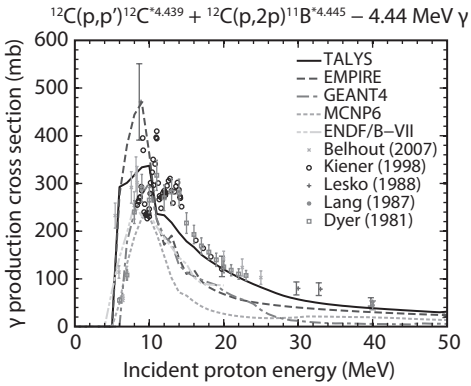



**Figure 2.2** | Cross section of the 4.44 MeV $\gamma$-ray emission due to proton-induced reactions on $^{12}C$.

of the kinematic Doppler broadening, these two lines cannot be resolved from each other. Most levels above the 4.44 MeV state decay through alpha emission [79].

The model results and experimental data are shown in figure 2.2. It is important to consider that, as shown by Kiener et al. [84], several narrow resonances exist at low incident energies. The models do not include such effects and aim to reproduce the general trend. The TALYS simulations provide the best fit to the experimental data at higher proton energies. The other models underestimate the cross sections in this energy region. For the lower incident energies, GEANT4 and MCNP6 seem to better fit the experimental data, although due to the resonances it is somewhat difficult to determine the trend in the measurements. The ENDF/B-VII results are similar to GEANT4.

### 2.2.3. Discrete lines of $^{16}O$

The first three excited levels of $^{16}O$ that can decay through $\gamma$-ray emission are at 6.13 MeV, 6.92 MeV, and 7.12 MeV [80]. Gamma emission from these states to the ground state results in most $\gamma$ emission at lower incident proton energies, for which the simulation results and measurements are depicted in figure 2.3. The cross sections for the important 6.13 MeV $\gamma$ line have been the subject of a large number of experimental studies. As the detailed measurements of Kiener et al. [84] show, various narrow resonances exist at low proton energies.

For the 6.13 MeV line, the TALYS and EMPIRE codes produce quite similar results, which mostly fit well to the experimental data, expect for an overestimation of the cross section at the lowest proton energies. GEANT4 gives similar results up to 20 MeV. Above 20 MeV, the cross section decreases rapidly to zero, which indicates

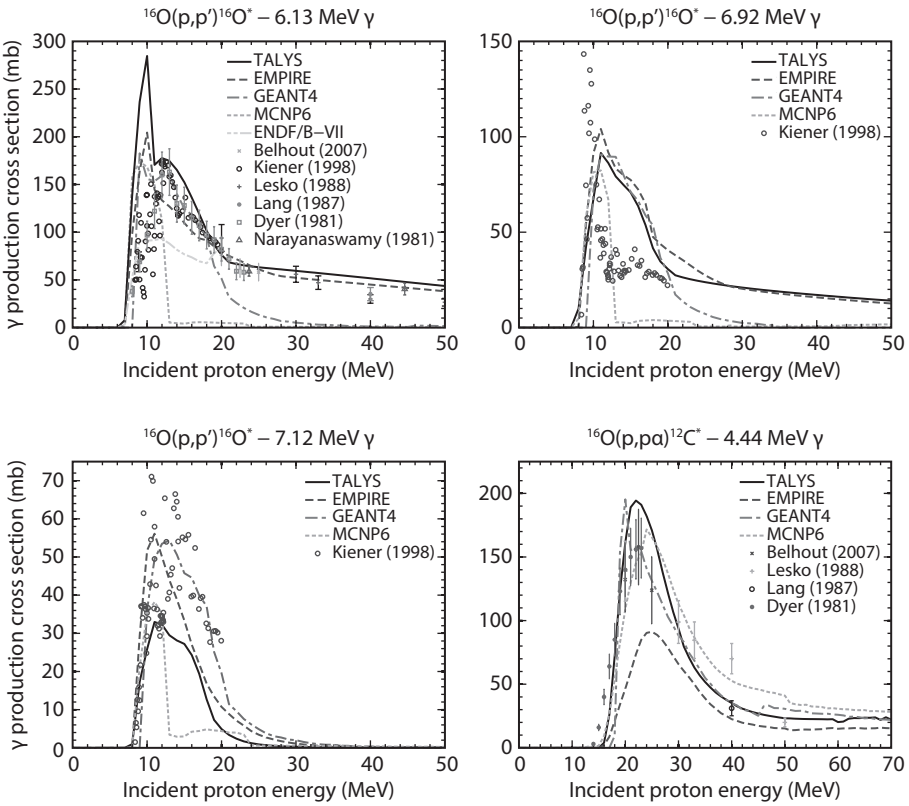



**Figure 2.3** | Cross sections of the 6.13 MeV, 6.92 MeV, 7.12 MeV, and 4.44 MeV $\gamma$-ray emissions due to proton-induced reactions on $^{16}$O.

the contribution of direct reactions to this discrete level is not simulated accurately. MCNP6 fits the data poorly, predicting almost no $\gamma$ production above 12 MeV. The proton separation energy of $^{16}$O is 12.13 MeV, which suggests a possible model deficiency in the competition between particle and $\gamma$ emission. The ENDF/B-VII data shows an overall underestimation of the cross section.

The experimental data show a narrow peak in the excitation function of the 6.92 MeV line, which is not reproduced by any of the models. The simulations of the 7.12 MeV $\gamma$ line follow the main trend in the experimental data, except for the MCNP6 results which again show a large discrepancy at higher incident energies.

For incident proton energies above 15 MeV, the $^{16}\mathrm{O}\,(p, p'\alpha)\,^{12}\mathrm{C}$ reactions result in significant emission of 4.44 MeV $\gamma$s from the residual $^{12}$C nucleus, which is also shown in figure 2.3. This cross section is quite well reproduced by TALYS, GEANT4 and MCNP6. EMPIRE underestimates the cross section.

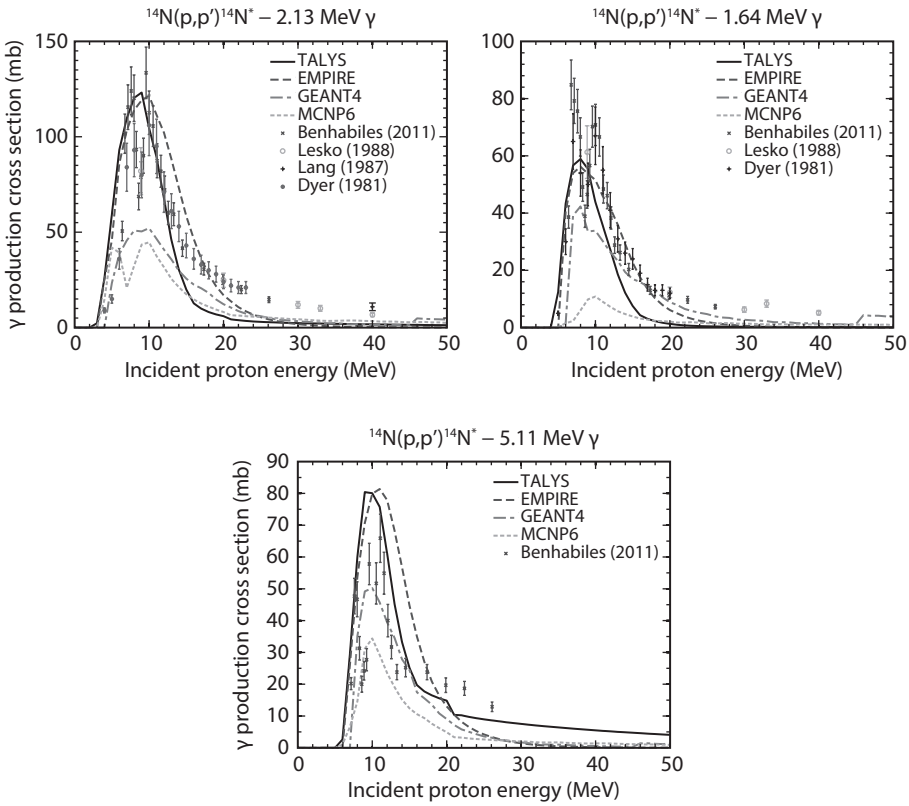



**Figure 2.4** | Cross sections of the 2.31 MeV, 1.64 MeV, and 5.11 MeV $\gamma$-ray emissions due to proton-induced reactions on $^{14}$N.

### 2.2.4. Discrete lines of $^{14}$N

The role of $^{14}$N in $\gamma$-ray emission from proton interactions with human tissue is more limited, because the number of nitrogen atoms in tissues is lower as compared to carbon and oxygen. Simulations on $^{14}$N can however also serve as additional validation of the nuclear reaction models for light elements. Because of its role in astrophysics, experimental data is available for many of the discrete $\gamma$ lines.

The three most important $\gamma$-ray lines have energies of 2.13 MeV, 1.64 MeV, and 5.11 MeV (transitions listed in table 2.1), for which we show the models and data in figure 2.4. For all three lines, TALYS and EMPIRE reasonably reproduce the main trends, while MCNP6 and GEANT4 underestimate the cross sections.

### 2.2.5. γ-ray emission during proton therapy

Shown in figure 2.5 is total $\gamma$-ray yield as a function of depth in tissue, for 150 MeV protons irradiating soft tissue and 70 MeV protons irradiating lung tissue. The

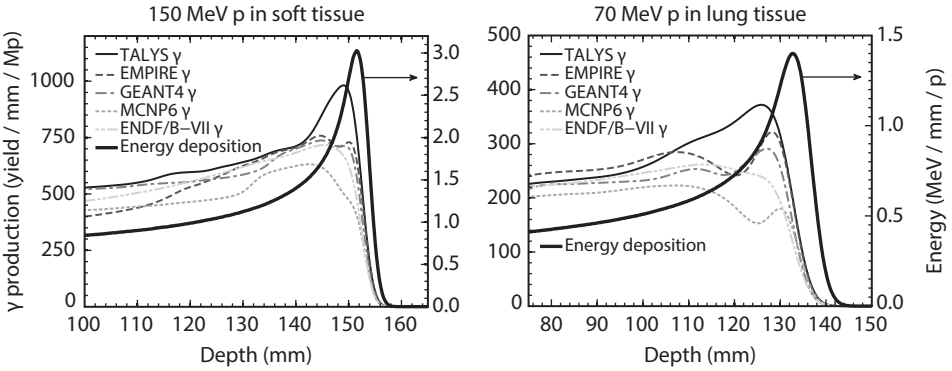



**Figure 2.5** | Simulated total $\gamma$-ray emission during irradiation of tissue. For reference, the energy deposited by the protons (Bragg curve) is also shown.

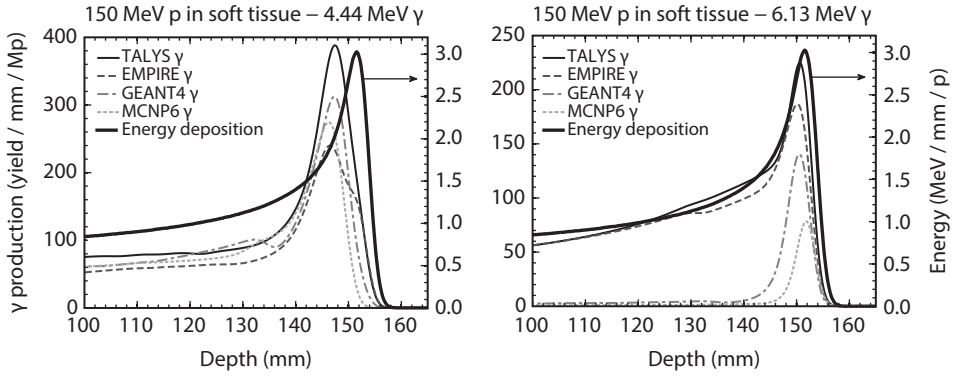



**Figure 2.6** | Simulated 4.44 MeV and 6.13 MeV $\gamma$-ray emissions during irradiation of ICRP soft tissue with 150 MeV protons. For reference, the energy deposited by the protons (Bragg curve) is also shown.

difference near the end of the proton range is on the order of a factor of two. Also, the slope of the fall-off of $\gamma$ emission differs. As compared to soft tissue, the lower density of lung tissue increases this difference relative to the depth.

The discrete $\gamma$ emission at 4.44 MeV and 6.13 MeV, which dominates at low incident proton energies, is plotted in figure 2.6. The 6.13 MeV $\gamma$ lines originates from the $p$+$^{16}$O reaction. The 4.44 MeV line correspond to the decay of the first excited level of $^{12}$C which is due to several reactions such as $^{12}\mathrm{C}\,(p, p')\,^{12}\mathrm{C}^*$, $^{14}\mathrm{N}\,(p, 2pn)\,^{12}\mathrm{C}^*$ and $^{16}\mathrm{O}\,(p, p\alpha)\,^{12}\mathrm{C}^*$. As can be seen in the figure, the correspondence to the dose delivered of the specific discrete $\gamma$ lines is more strongly impacted by the model differences.

## 2.3. Discussion and conclusions

Prompt $\gamma$-ray emission during proton therapy is dominated by proton-induced nuclear reactions on $^{12}$C, $^{16}$O, and $^{14}$N. Knowledge of these reactions is key to an accurate simulation of potential proton therapy range verification methods employing prompt $\gamma$-rays. In this study we found considerable differences in the prediction of the prompt $\gamma$-ray emission due to these elements, comparing the nuclear reaction models of GEANT4, MCNP6, TALYS, and EMPIRE, and the evaluated ENDF/B-VII cross sections.

The model estimates of the total $\gamma$ emission during proton irradiation of soft tissue and lung tissue, differed by a factor of about two near the end-of-range of the protons. At higher incident proton energies, the models agreed within about 25 %. If specific discrete $\gamma$ lines are considered, the models variations can be larger. The dedicated nuclear reaction codes reproduced the general trends of experimental data more consistently as compared to the Monte Carlo codes. The GEANT4 and MCNP6 models showed a number of larger discrepancies. In particular, the important 6.13 MeV $\gamma$ emission due the $^{16}\mathrm{O}\,(p, p')\,^{16}\mathrm{O}^*$ reaction is not simulated accurately at higher energies where the direct reaction processes are of importance. The evaluated ENDF/B-VII data includes only a continuous $\gamma$ spectrum with a limited energy resolution. The main $\gamma$ line cross section that we could extract from this data showed uncertainties similar to some of the other models.

These findings are important to consider in the design of simulations of prompt $\gamma$-ray detection systems. Because experimental data is limited, a complete simulation of prompt $\gamma$-ray detection needs to rely on phenomenological nuclear reaction models. Practical detectors used in clinically realistic scenarios will likely have to deal with low count rates, therefore the model uncertainties may lead to sub-optimal design choices. The cross section uncertainties can also affect the reconstruction of the proton range from the detected $\gamma$ radiation. The impact of the model uncertainly will depend on the extend in depth in which the $\gamma$ production is analysed as well as the $\gamma$ energy range considered. In addition, accurate cross sections are important to study the sensitivity of a particular technique to factors such as the elemental composition of the tissue being irradiated.

Our results may also be used to guide the choice of simulation models and parameters. It is important to note that for each code, several model adjustments were required to obtain reasonable simulations. We recommend to critically assess the nuclear reaction model used and to consider the uncertainties when interpreting the results. If specific resolvable $\gamma$-ray lines are considered, the use of a direct parametrisation of experimental data would be advisable.

There is a clear need for additional theoretical and experimental studies of the nuclear reactions of relevance to proton therapy. An updated cross section evaluation based on the most recent experiment data, and with additional focus on

the production of prompt $\gamma$-rays and positron emitters, would certainly be desirable. Improved input parameters for the nuclear reaction models most likely can improve the agreement between models and measurements, and result in a better predictive power of the models for $\gamma$ lines for which no experimental data is currently available. In general, the differences between the models exceeded the reported uncertainties in the available experimental data, which shows room for model improvement even based on current data.

A number of other issues should be noted. We analysed angle-integrated cross sections, but prompt $\gamma$-ray emission in general is not isotropic. Depending on the quantum properties of excited nuclear level, various angular dependencies are possible. Each $\gamma$ line has a different double differential cross section, which also depends on the proton energy. Also, the Doppler broadening due to nuclear reaction kinematics is a factor that is specific to each nuclear reaction.

To summarize, we have shown that nuclear reaction simulations of prompt $\gamma$-ray emission during proton therapy are subject to considerable uncertainties. Even with our attempts to identify the best models and parameters available, the difference in the total $\gamma$ production near the end-of-range of the protons was approximately a factor of two. These uncertainties should be considered when simulations are performed for the design of $\gamma$-ray detection and proton range reconstruction methods.

# 3

# Energy- and time-resolved detection of prompt γ-rays

In this chapter, we present experimental results of a novel prompt $\gamma$-ray detector for proton beam range verification. The detection system features an actively shielded cerium-doped lanthanum(III) bromide scintillator, coupled to a digital data acquisition system. The acquisition was synchronized to the cyclotron radiofrequency to separate the prompt $\gamma$-ray signals from the later-arriving neutron-induced background. The detector features a high energy resolution and an effective reduction of background events, enabling discrete proton-induced prompt $\gamma$-ray lines to be resolved. We present the results of experiments in which a water phantom was irradiated with proton pencil-beams in a clinical proton therapy gantry. A slit collimator was used to collimate the prompt $\gamma$-rays, and measurements were performed at 27 positions along the path of proton beams with ranges of 9.0 g cm$^{-2}$, 16.0 g cm$^{-2}$, and 23.0 g cm$^{-2}$ in water. The magnitudes of discrete $\gamma$ lines at 4.44 MeV, 5.2 MeV, and 6.13 MeV were quantified. The prompt $\gamma$ lines were found to be clearly resolved in dimensions of energy and time, and had a reproducible correlation with the proton depth-dose curve. We conclude that the measurement of discrete prompt $\gamma$-rays for in vivo range verification of clinical proton beams is feasible.

Based on published work: J.M. Verburg, K. Riley, T. Bortfeld, and J. Seco. Energy- and time resolved detection of prompt gamma-rays for proton range verification. *Phys. Med. Biol.* **58**, L37–L49 (2013).

The finite range of proton beams is the most important characteristic that is exploited in proton radiotherapy, because it allows for healthy tissue distal to the tumour to be spared from the harmful effects of radiation. Positioning the end-of-range of the proton beam with millimetre accuracy on the distal end of the tumour has however remained elusive for deeper-seated tumours, because of uncertainty in the proton stopping power of tissue in the beam path. Most proton therapy centres use a 3.5 % range margin to ensure tumour coverage [24], which translates to a distal margin up to 1 cm. Also, beams stopping close to organs-at-risk are generally avoided because of the range uncertainty, although these may be favourable for avoidance of other organs. For these reasons, a method to verify the proton range *in vivo* would be highly desirable.

While the primary protons normally stop inside the patient, protons produce secondary prompt and delayed radiation that may be used for *in vivo* range verification. The two main methods that have been proposed for this purpose are positron emission tomography (PET) and prompt $\gamma$-ray detection.

A PET scan can detect delayed $\gamma$-rays from the annihilation of positrons, which are emitted following the decay of certain radioisotopes produced as nuclear reaction products, such as $^{11}$C and $^{15}$O. Since these isotopes have half-lives on the order of 1 min to 20 min, PET cannot be used for an immediate range verification. Also, metabolic processes result in diffusion of the produced isotopes prior to their decay, which limits the spatial accuracy [90]. The uncertainty in the tissue composition is another issue, as it affects the production of positron emitters.

Prompt $\gamma$-rays, on the other hand, are emitted almost instantaneously during the decay of the excited nuclear reaction products to their ground state. Therefore, they could potentially be used for an immediate and more accurate proton range verification. A direct range verification could facilitate tuning of the proton range after a subset of the treatment has been delivered. Furthermore, the $\gamma$ emission cross sections are higher near the end-of-range of the proton beam compared to that for producing the positron emitters [91], so the spatial pattern of the induced activity correlates better with absorbed dose.

A number of prompt $\gamma$-ray measurements along the path of proton and carbon ion pencil-beams have been reported, using detectors with a single scintillator [36, 52], as well as a $\gamma$ camera with a knife-edge slit aperture to measure $\gamma$-rays at multiple depths in a single measurement [92]. In these studies, measured events were integrated over a wide $\gamma$ energy range with a lower threshold of 2 MeV to 4 MeV. Other detectors, such as a multi-slit detector [93] and Compton camera [94] have been proposed, and prototype detectors are under development by different groups.

Particularly at higher beam ranges, these previous measurements showed the neutron-induced background to be an important issue limiting the signal-to-noise

ratio. Testa et al. [52] found time-of-flight measurements to be effective to reduce the background in measurements with carbon ions. Biegun et al. [95] performed a simulation study and concluded this can also be beneficial in the case of proton beams with a microstructure that is typical of cyclotron accelerated protons.

We previously studied nuclear reaction simulations and various known cross sections of prompt $\gamma$-rays induced by protons [96]. This study showed that the emitted prompt $\gamma$-rays near the end-of-range of the proton beam are mostly due to a limited number of transitions between nuclear levels, resulting in specific discrete $\gamma$-ray energies.



Several authors have performed spectroscopy of prompt $\gamma$-rays induced by protons. Earlier studies aimed to measure cross sections for use in $\gamma$-ray astronomy [81, 83–87]. Recently, Polf et al. [51, 97] resolved discrete lines in the integral prompt $\gamma$-ray emission from a 48 MeV proton beam stopping in matter, using a high-purity germanium detector. The focus of their study was to determine the elemental composition of tissue being irradiated during proton therapy.

Identifying discrete $\gamma$-ray lines, attributed to specific nuclear level transitions, can provide several benefits to improve the accuracy and sensitivity of range verification. First, each of these discrete $\gamma$-ray emissions has a unique correlation to the proton energy, and nuclear reaction cross sections may be used as prior knowledge in the range verification. Second, using the prompt $\gamma$-ray production cross sections, the concentrations of the target nuclei for the proton-induced reactions in the irradiated tissue may be estimated based on the specific $\gamma$-ray emissions, which can make range verification more robust if the proton beams stops in tissue with an uncertain composition.

In this chapter, experimental results are presented for a novel prompt $\gamma$-ray detection system designed to resolve discrete proton-induced prompt $\gamma$-rays. First, we describe the design of the detector and data acquisition system, and the methods used for data analysis. We then analyse the energy spectrum and time of the prompt $\gamma$-ray events. Finally, we show the measured magnitudes of discrete $\gamma$ lines, as well as the energy-integrated prompt $\gamma$-ray emission, at various depths along the path of proton beams in a water phantom.

## 3.1. Methods and materials

The prompt $\gamma$-ray detection system was developed as a small-scale prototype of a potential clinically applicable design. We therefore aimed to design a system that is practical to use, operates at room temperature and functions with clinically realistic proton pencil-beams. Because we aim to resolve discrete prompt $\gamma$-ray lines, both high energy resolution and effective background reduction are essential.

### 3.1.1. Detector

The primary detector consisted of a cylindrical $LaBr_3$ (5% Ce) scintillator with a diameter and length of 76 mm, coupled to a photomultiplier tube (PMT). Four optically separated BGO ($Bi_4Ge_3O_{12}$) crystals formed a mantle surrounding the primary detector. These crystals had a length of 152 mm and a thickness of 18 mm, and a curved shape to directly surround the cylindrical primary detector. Through anti-coincidence measurements, the surrounding detector acts as an active shield to reduce both the Compton background and the neutron-induced $\gamma$-ray background from the treatment room. Each of the four crystals of the shield was readout by two PMTs.



### 3.1.2. Proton beam

The experiment was performed in a clinical proton therapy gantry at the Francis H. Burr Proton Therapy Center. At our institute, protons are accelerated with an IBA C230 isochronous cyclotron (Ion Beam Applications SA, Louvain-la-Neuve, Belgium) which operates at a radio frequency (RF) of 106.3 MHz, corresponding to an RF period of 9.4 ns. During each cyclotron RF period, a bunch of 230 MeV protons is extracted, which are slowed down to the desired energy by the energy selection system, and transported along the beam lines to the treatment rooms.

### 3.1.3. Data acquisition

A custom digital data acquisition system was developed for the experiments, which enables the system to function with high count rates. The pulses from the PMTs were amplified by wideband amplifiers close to the detector and connected to the data acquisition system that was located outside the treatment room.

The signals were digitized with high-speed flash analog-to-digital converters (ADC). The sampling clocks of the ADCs were driven by a low-jitter clock generator that was phase-locked to the frequency synthesizer providing the cyclotron RF signal. This provides a time reference for sorting events in the detector and enables separation of $\gamma$-ray signals from the later-arriving neutron-induced background. Four interleaved ADC channels were used to sample the signal from the primary detector at a sampling rate of 36 times the cyclotron RF, which corresponds to 3.8 giga samples per second. The signals from the four crystals comprising the active shield were each sampled 9 times per cyclotron RF period, using one ADC channel.

All signals were continuously acquired and stored in a ring buffer. The transfer of signals from the ring buffer to a memory buffer was triggered by the signal from the primary detector reaching a threshold, after which the signal from all PMTs acquired during a predefined time interval before and after the trigger time was stored. The time interval was defined to store the baseline prior to the pulse for a short period of time, and the entire duration of the pulse until it decays to a level where it becomes

indistinguishable from the baseline. A double memory buffer was used to allow simultaneous acquisition and readout. The buffer content was transferred through an optical interface to a standard desktop computer on which the analysis of the digital pulses was performed.

### 3.1.4. Digital pulse analysis

The quantities of interest were obtained from the digitized PMT pulses using digital pulse processing algorithms. The baseline of the acquired signals was continuously established using a moving average. After subtracting the baseline from the pulse signals, they were filtered to remove high-frequency noise.



The area below the pulses of the primary detector was integrated to obtain the deposited energy, which was calibrated using several radioactive sources. To determine event time, a digital constant fraction discriminator was implemented. Several samples near a constant fraction of the signal maximum were interpolated to obtain timing information with a sub-sample precision.

The timing of the events in the active shield was established using leading edge discrimination. Any events in the active shield that were recorded with a estimated time within 10 ns of an event in the primary detector were considered to be coincident. The time window was selected to include almost all coincident events. By analysing the rate of events in the active shield that were not in coincidence with events in the primary detector, the normalization of the primary detector counts was corrected for random coincidences. The normalization was also corrected for pile-up in the detector.

### 3.1.5. Quantification of γ lines

To obtain the magnitude of discrete prompt $\gamma$ lines, the data were first binned in two-dimensional histograms $\mathbf{H}(t, e)$ with dimensions of $\gamma$ energy $e$ and time $t$. Because both the prompt $\gamma$-ray emission and the background depend on the measurement position, we consider both simultaneously by performing least-squares fitting to the parts of the histogram containing the discrete lines. The least-squares fitting was performed using the Levenberg-Marquardt algorithm [98] as implemented in the levmar solver [99].

We approximated the prompt $\gamma$ lines as Gaussian in dimensions of energy and time. Both the standard deviation of the particular $\gamma$ line in time, $\sigma_t$ and standard deviation of the energy, $\sigma_E$, were pre-fitted based on the entire dataset for a proton energy. The background as a function of time was approximated as a second order polynomial with coefficients $n_1, n_2, n_3$, and the baseline of unresolved prompt $\gamma$-rays was also fitted using a polynomial with coefficients $u_1, u_2, u_3$, considering the time dependence of the prompt $\gamma$ signal. These were all modelled as free parameters for each measurement point. With $A$ the amplitude of the discrete line, $E$ the energy

of the $\gamma$ line and $t_\gamma$ the position of the prompt $\gamma$ signal relative to the cyclotron RF period, we obtain:

$$\begin{aligned}\mathbf{H}(t,e) &= A\exp\left(-\left(\frac{\left(t-t_\gamma\right)^2}{2\sigma_t^2}+\frac{(e-E)^2}{2\sigma_E^2}\right)\right)+n_1+n_2t+n_3t^2 \\ &+ \exp\left(-\frac{\left(t-t_\gamma\right)^2}{2\sigma_t^2}\right)\left(u_1+u_2e+u_3e^2\right).\end{aligned} \tag{3.1}$$



To determine the total $\gamma$-ray emission in a certain energy range, the same procedure was used without fitting the energy dependence. In one case, a discrete neutron-induced $\gamma$ line without a clear time dependence was located close to the proton-induced prompt $\gamma$ line, and therefore an additional term was added to account for this line:

$$\mathbf{H}'(t,e) = \mathbf{H}(t,e) + A_n\exp\left(-\frac{(e-E_n)^2}{2\sigma_{E_n}^2}\right), \tag{3.2}$$

in which $A_n$ is the amplitude of the line, $E_n$ is the energy, and $\sigma_{E_n}$ is the standard deviation of the energy measurement.

### 3.1.6. Experiments in water phantom

The experimental setup is shown in figure 3.1. A proton pencil-beam was stopped in a water phantom, with the beam centre 7.5 cm below the surface in the direction of the detector. The detector was positioned at a 90° angle relative to the proton beam direction and 30 cm from the surface of the phantom. A 15 cm thick lead collimator with a 4.8 mm slit opening was used to collimate the prompt $\gamma$-rays, which was located centrally between the proton beam and the primary detector. The collimator-detector response to high energy $\gamma$-rays produced along the beam path is approximately Gaussian with a full width at half maximum of 11 mm.

To facilitate measurements at different positions along the beam path, the phantom was positioned using a platform attached to a motorized linear stage. Measurements were performed at 27 positions separated by 3.0 mm, ranging from 50 mm before the end-of-range, which we define at the 80 % dose fall-off level ($R_{80}$), to 30 mm beyond $R_{80}$. The systematic positioning error is estimated as ±1 mm, and the accuracy as ±0.1 mm.

The absolute number of protons delivered to the phantom was monitored with the clinically commissioned ionization chamber close to the exit of the treatment head. All measurement points were acquired with $10^{10}$ protons incident on the phantom and repeated five times. The beam current was approximately 0.2 nA, i.e. each measurement with $10^{10}$ protons was performed in about 10 s. This rate was selected to facilitate the collection of all raw data from the digitizer for off-line

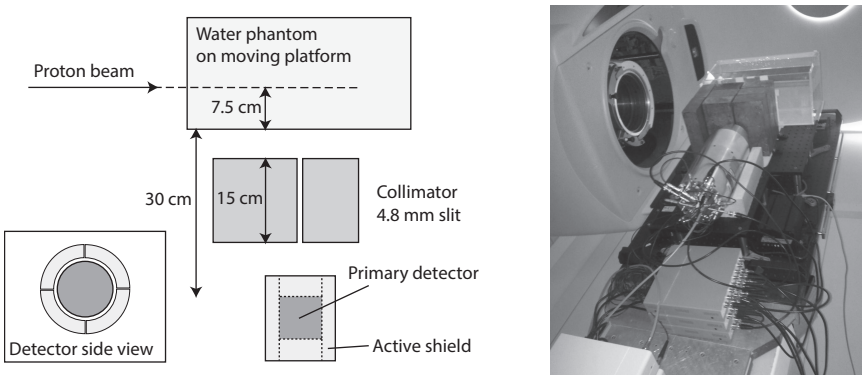



**Figure 3.1** | Schematic and photograph of the experimental setup.

development and tuning of the digital pulse processing algorithms. The detection system however is able to handle higher dose rates and this is limited principally by the 30 ns decay time of $LaBr_3(Ce)$, so count rates in excess of 1 MHz can be attained.

## 3.2. Results

### 3.2.1. Prompt γ-ray energy spectrum

Shown in figure 3.2 is the measured energy spectrum of prompt $\gamma$-rays in a 2 ns time window, measured at depths 9 mm proximal and distal to the $R_{80}$ for a proton beam with a range of 9.0 g cm$^{-2}$ in water. The spectra were obtained by integrating the five measurements, i.e. with a total of $5.0 \times 10^{10}$ incident protons.

A clear structure can be seen in the energy spectrum measured before the protons have reached the end of their range, which almost disappears in the second measurement beyond the end-of-range. As expected from available cross section data and theoretical models [96], the two strongest $\gamma$ lines are at 6.13 MeV, from the $^{16}O(p, p')^{16}O^*$ reaction, and at 4.44 MeV from the $^{16}O(p, p'\alpha)^{12}C^*$ reaction. The single escape peaks at 511 keV below these lines are also resolved, although small in magnitude since active shielding eliminates most of these events. A third strong $\gamma$ line is visible around 5.2 MeV, which is likely a combination of two unresolved lines: $^{16}O(p, pp)^{15}N^*$ (5.27 MeV) and $^{16}O(p, x)^{15}O^*$ (5.24 MeV).

### 3.2.2. Time structure of γ-ray emission

In figure 3.3, the number of detected events within the cyclotron RF period is shown, for both the distal and proximal locations, at proton ranges of 9.0 g cm$^{-2}$, 16.0 g cm$^{-2}$, and 23.0 g cm$^{-2}$. Based on the spectrum as shown in figure 3.2, events in the 3.0 MeV to 7.0 MeV energy range were selected.

A well defined peak of prompt $\gamma$-rays is visible in the measurements 9 mm before $R_{80}$. The width of the peak strongly depends on the proton range: the full width at

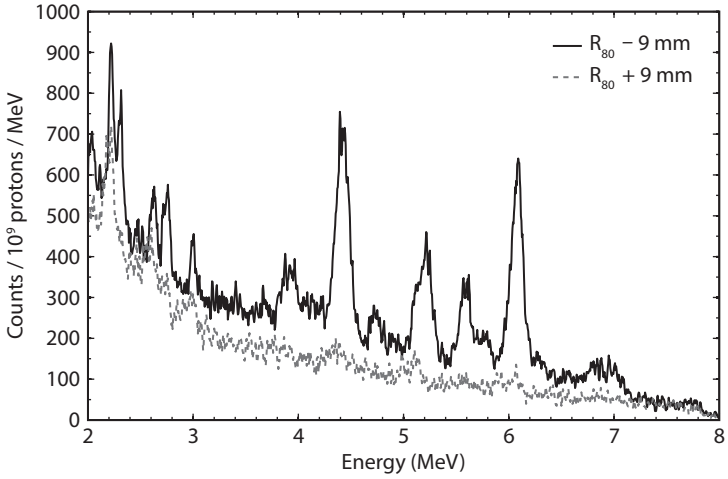



**Figure 3.2 |** Energy spectrum of prompt $\gamma$-rays integrated over a 2 ns time window, measured 9 mm before $R_{80}$ (——) and 9 mm beyond $R_{80}$ (- - - -) with a proton range of 9.0 g cm$^{-2}$ in water. The spectra were acquired with $5.0 \times 10^{10}$ incident protons.

half maximum in a Gaussian approximation is 2.8 ns, 1.9 ns, and 1.2 ns for respective ranges of 9.0 g cm$^{-2}$, 16.0 g cm$^{-2}$, and 23.0 g cm$^{-2}$. The magnitude of this energy dependence is consistent with the energy spread of the pencil-beam as measured by Clasie et al. [100]. During the transport of the protons from the cyclotron through the energy selection system to the treatment room (total beam line length 35 m), the small difference in proton velocities results in a widening of the proton bunch in time. This effect becomes smaller as the proton range increases, mainly because the proton velocity in the lab frame becomes less dependent on energy as protons reach relativistic velocities.

Since the width of the prompt $\gamma$-ray peak remains considerably shorter than the cyclotron RF period of 9.4 ns, a synchronization of the data acquisition to the cyclotron effectively improves the signal-to-noise ratio.

### 3.2.3. Background reduction

The effect of the various background reduction methods is illustrated in figure 3.4. Energy spectra are shown with and without active shielding and time-resolving the events within a 1.5 ns window. The spectra are shown for a proton beam with a range of 23.0 g cm$^{-2}$, because background reduction is more important at longer proton ranges.

Remarkably, discrete peaks can be observed even without any background correction due to the relatively good energy resolution of $LaBr_3(Ce)$. The results show that

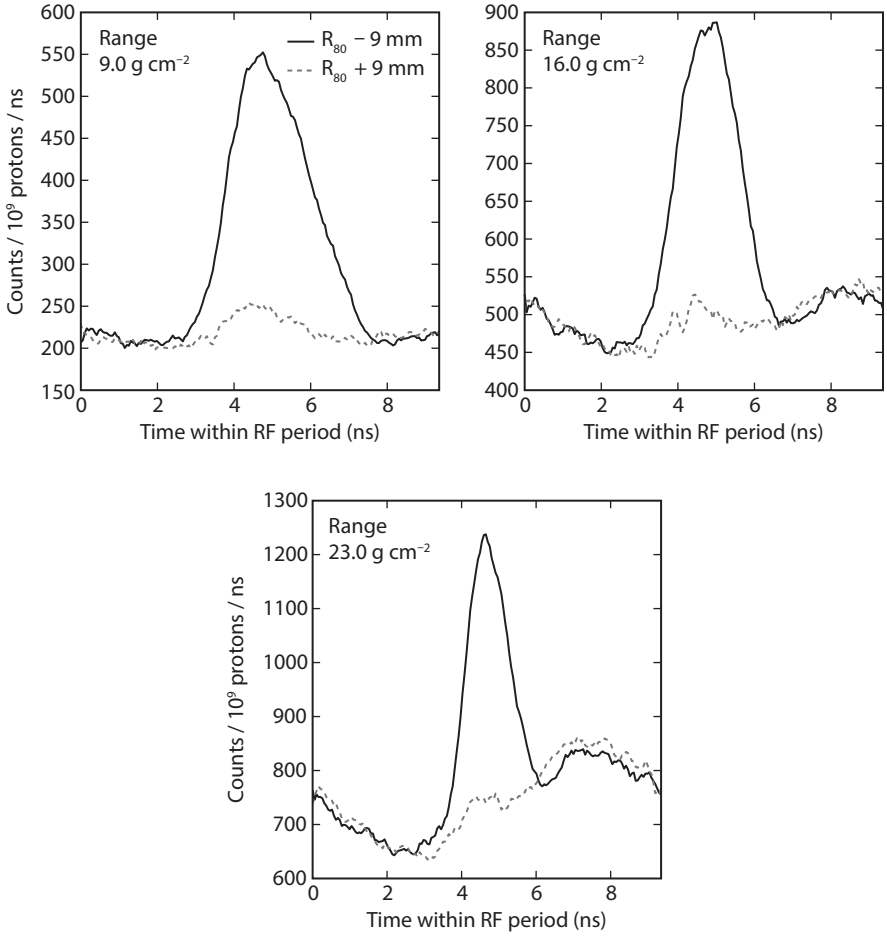



**Figure 3.3** | Time structure of the prompt $\gamma$-ray emission measured with proton ranges of 9.0 g cm$^{-2}$, 16.0 g cm$^{-2}$, and 23.0 g cm$^{-2}$ and $5.0 \times 10^{10}$ incident protons per measurement. Shown are the measurements 9 mm before $R_{80}$ (——) and 9 mm beyond $R_{80}$ (- - - -). The phase of the signals relative to the cyclotron RF is arbitrary; the prompt $\gamma$-ray peak is centred in each plot for presentational purposes. The counts are integrated over the 3.0 MeV to 7.0 MeV energy range.

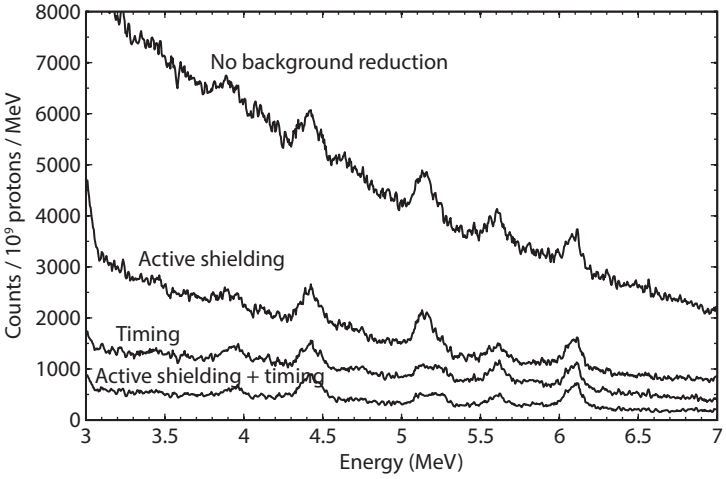



**Figure 3.4** | Energy spectra with and without the background reduction methods applied. The spectra were acquired with a proton range of 23.0 g cm$^{-2}$ and $5.0 \times 10^{10}$ incident protons. A time window of 1.5 ns was used for the spectra with timing.

active shielding and timing are both effective in reducing background. Combining the two methods reduces the background level in the spectra by about 90 %.

### 3.2.4. Energy–time histogram

Combing both the dimension of $\gamma$ energy and time with respect to the cyclotron RF, a density histogram as depicted in figure 3.5 is obtained. From such histograms, the discrete proton-induced $\gamma$ lines were quantified, as described in section §3.1.5. We quantified the strong $\gamma$ lines at 4.44 MeV, 5.2 MeV, and 6.13 MeV, which were previously described.

The $\gamma$ lines that are emitted throughout the entire RF period, appearing as as horizontal lines in the histograms, correspond to neutron-induced reactions with the phantom, collimator and the shielding of the treatment room. The strongest neutron-induced lines are at 2.22 MeV, which can be attributed to neutron capture on hydrogen, and at 2.61 MeV, which is presumably due to a neutron-induced prompt $\gamma$-ray from $^{208}$Pb in the collimator.

In figure 3.6, we show an example of the data fit to quantify the $\gamma$ lines. The model of the histogram surrounding the $\gamma$ line is fitted to the acquired data, and the magnitude of the $\gamma$ line is obtained from the fit parameters.

### 3.2.5. γ-ray emission along proton beam path

The measurements of prompt $\gamma$-ray emission along the path of the proton beams in water is shown in figure 3.7, again with $R_{80}$ at 9.0 g cm$^{-2}$, 16.0 g cm$^{-2}$, and 23.0 g cm$^{-2}$.

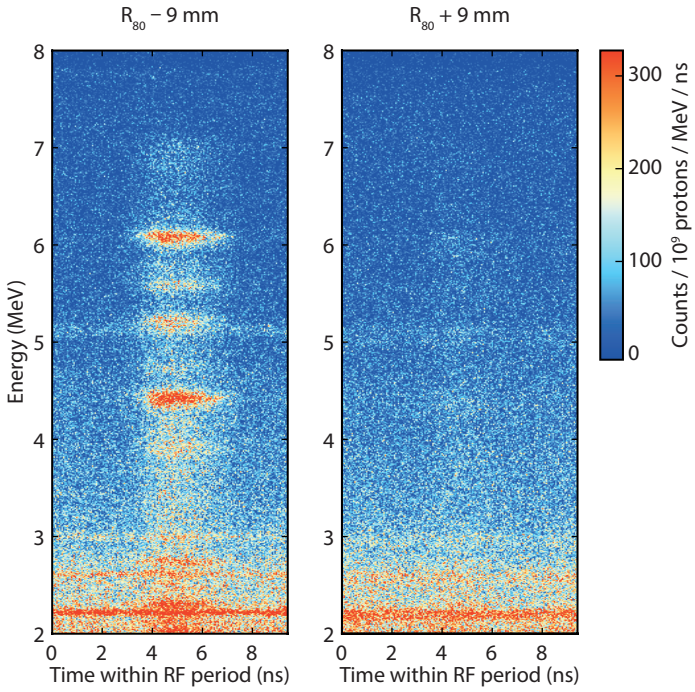



**Figure 3.5** | Density histograms of the energy and time, measured 9 mm before $R_{80}$ and 9 mm beyond $R_{80}$ for a proton beam with range of 9.0 g cm$^{-2}$ in water. The histograms were acquired with $5.0 \times 10^{10}$ incident protons.

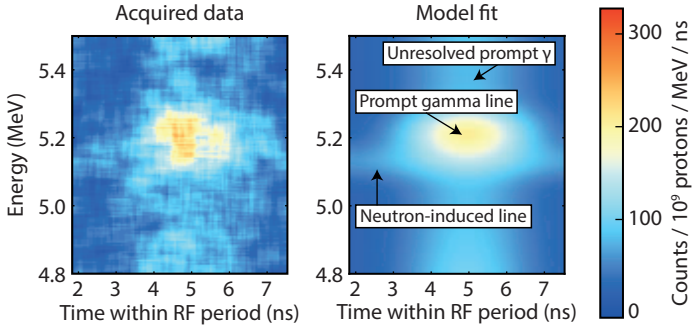



**Figure 3.6** | Example of the two-dimensional fit of an area of the density histogram to quantify the magnitude of a prompt $\gamma$ line. This example shows the 5.2 MeV $\gamma$ line in one of the measurements 9 mm before $R_{80}$ for a proton beam with a range of 9.0 g cm$^{-2}$ and $10^{10}$ incident protons.

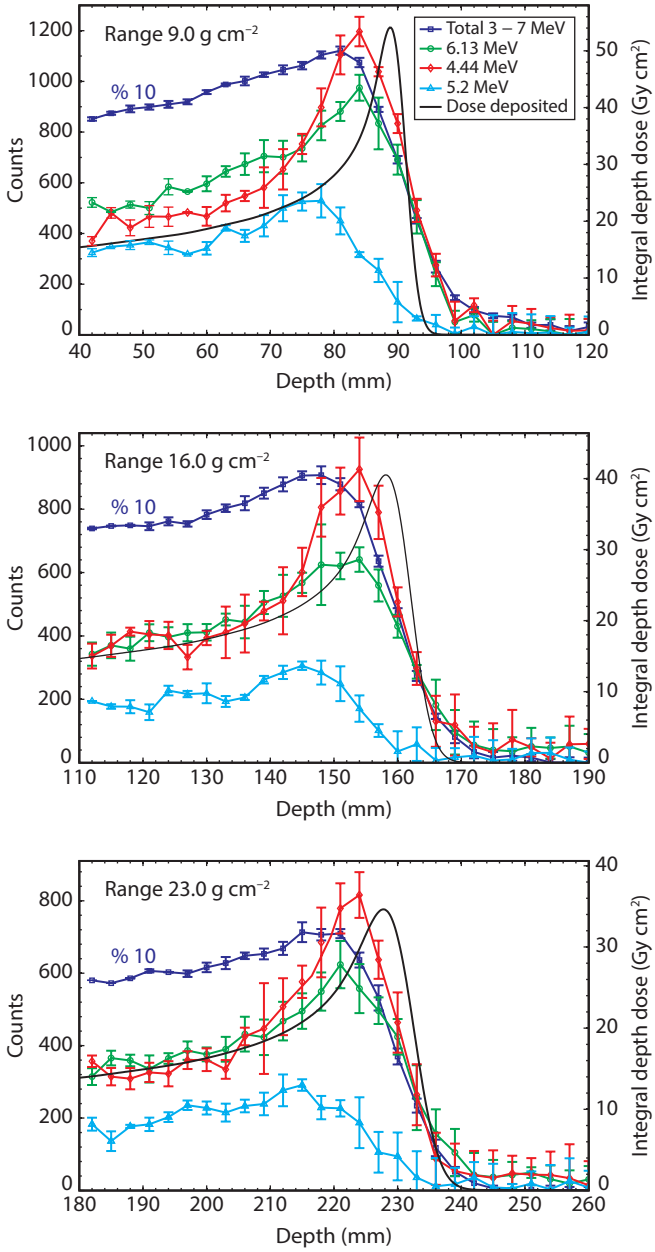



**Figure 3.7** | Energy-integrated (□) and discrete (○ ,◇,△) prompt $\gamma$-ray emissions along the path of proton pencil-beams with ranges of 9.0 g cm$^{-2}$, 16.0 g cm$^{-2}$, and 23.0 g cm$^{-2}$ in water. The mean and ±1σ statistical uncertainty of 5 measurements are shown. The lines connecting the points serve to guide the eye. For each measurement, $10^{10}$ protons were delivered. The dose deposited by $10^{10}$ protons is also depicted.

Shown in the figure is the mean of the five measurements with $10^{10}$ incident protons, and the $\pm 1\sigma$ uncertainty due to counting statistics. The dose deposited by the protons is also shown, which is described as integral depth dose in units of Gy cm$^2$. An integral depth dose of 25 Gy cm$^2$ for example corresponds to 1 Gy delivered to a 5 cm × 5 cm field.

For all three proton ranges, the different prompt $\gamma$-ray emissions show a distinctive correlation with the depth-dose curve. The total prompt $\gamma$-ray emission integrated over the 3.0 MeV to 7.0 MeV energy window is relatively flat in the proton dose plateau and decreases rapidly at the proton end-of-range. The discrete $\gamma$-ray profiles from $^{16}\mathrm{O}\,(p, p')\,^{16}\mathrm{O}^*$ (6.13 MeV) and $^{16}\mathrm{O}\,(p, p\alpha)\,^{12}\mathrm{C}^*$ (4.44 MeV) exhibit a peak near the end of the proton range, and this is consistent with the cross sections of these nuclear transitions, which have a maximum around a proton energy of respectively 13 MeV and 23 MeV [81, 85]. The 5.2 MeV $\gamma$ line has a peak at a depth further from the end-of-range.



## 3.3. Discussion and conclusions

In this chapter, we presented the first absolute measurement of discrete prompt $\gamma$-rays for proton beam range verification. Discrete peaks were resolved by using a scintillator with high energy resolution, active background suppression, and time-resolving events within the short cyclotron RF period of 9 ns. A clear correlation of the prompt $\gamma$-ray emission with the proton depth-dose curve was observed.

Several discrete $\gamma$ lines were resolved and quantified. The ability to identify discrete prompt $\gamma$ lines, originating from specific nuclear level transitions, can have several advantages for *in vivo* proton range verification. The magnitude of these discrete $\gamma$ lines can be related directly to nuclear reaction cross sections, enabling such prior knowledge to be incorporated in the determination of the proton range. This approach can also select the reactions that best correlate with proton absorbed dose and can eliminate the confounding effects of other reactions that may behave differently. It can also measure the concentration of target nuclei for proton-induced nuclear reactions. A determination of these concentrations can help establish the range of the beam and may also provide useful information for tumour characterization and optimization of the treatment plan. Polf et al. [97] showed the feasibility of determining the elemental concentration of oxygen based on the integral prompt $\gamma$-rays produced by a proton beam stopping in a sample. Our results indicate it is also possible to determine concentrations at different positions along the path of the proton beam.

We found that time-of-flight measurements, previously studied experimentally for range verification in carbon ion therapy [52] and simulated for proton beams [95], improve the signal-to-noise ratio relative to the neutron-induced background for proton beams. The time profile of the prompt $\gamma$ peak was resolved and decreased in

duration with higher proton energies because of the lower spread in proton velocity along the beam line. This enabled separating the desired prompt-$\gamma$ events from neutron or neutron-induced events within the short frequency of the cyclotron. In addition, active shielding was effective to reduce background due to both Compton scatter out of the primary detector, as well as neutron-induced $\gamma$-rays. These methods for background reduction are particularity important to verify higher proton ranges, since the neutron background increases in proportion with the proton range.



The additional information provided by the energy- and time resolved detection is anticipated to facilitate robust range verification with a smaller number of measurement points as compared to detection techniques that do not provide this information. The fluence of $10^{10}$ incident protons used in our measurements is on the order of one layer of a typical pencil-beam scanning fraction. Multiple measurements could be performed at the same time by combining multiple detector modules similar to the one used in this study, which could potentially be placed in a arc-like configuration around the patient. Both the design and shape of the detector modules can be optimized to obtain the best counting efficiency. Also, a more accurate determination of the slightly non-Gaussian time structure of the prompt $\gamma$-ray emission, can reduce the uncertainties that presently exist in the analysis. The time structure of the protons exiting the treatment head can be considered constant for a given proton energy and treatment room. It is therefore possible to characterize the structure more accurately in a separate measurement. Analysis of carbon beam microstructure was previously performed to eliminate random coincidences from prompt $\gamma$-rays during in-beam PET measurements [101].

Based on the dramatic reduction in background, improved signal to background ratio and reproducible correlation of the measured profiles with absorbed dose, the proposed technique appears suitable for future clinical use. Future improvements in detector design and larger-scale clinical prototypes that will determine figures of merit for the end of range using lower proton fluences are anticipated. The quantitative results of this study will also be used to benchmark nuclear reactions models in simulation codes [96, 102]. These improved models will enhance the overall accuracy of not only prompt $\gamma$-ray related studies but proton therapy simulations in general.

# 4

# Proton range verification through prompt γ-ray spectroscopy

We present an experimental study of a novel method to verify the range of proton therapy beams. Differential cross sections were measured for 15 prompt $\gamma$-ray line excitations in proton-nuclear interactions with $^{12}C$ and $^{16}O$ at proton energies up to 150 MeV. These cross sections were used to model discrete prompt $\gamma$-ray emissions along proton pencil-beams. By fitting detected prompt $\gamma$-ray counts to these models, we simultaneously determined the beam range and the oxygen and carbon concentration of the irradiated matter. The performance of the method was assessed in two phantoms with different elemental concentrations, using a small scale prototype detector. Based on five pencil-beams with different ranges delivering $5 \times 10^8$ protons and without prior knowledge of the elemental composition at the measurement point, the absolute range was determined with a standard deviation of 1.0 mm to 1.4 mm. Relative range shifts at the same dose level were detected with a standard deviation of 0.3 mm to 0.5 mm. The determined oxygen and carbon concentrations also agreed well with the actual values. These results show that quantitative prompt $\gamma$-ray measurements enable knowledge of nuclear reaction cross sections to be used for precise proton range verification in the presence of tissue with an unknown composition.

Based on published work: J.M. Verburg and J. Seco. Proton range verification through prompt gamma-ray spectroscopy. *Phys. Med. Biol.* **59**, 7089–7106 (2014).

Proton beams used for cancer radiotherapy have the physical advantage of a finite range, which results in a sharp distal dose gradient. Therefore, healthy tissue distal to the target is spared from radiation. However, the advantage of the sharp distal dose gradient cannot yet be fully exploited, because the end-of-range of the beam is subject to uncertainty. To ensure complete tumour coverage, a range margin is currently added. Also, because of the risk of beam overshoot, sparing of organs-at-risk in close proximity to the tumour is normally achieved with the lateral edges of the beam instead of the sharper distal edge. A method to verify the range of proton beams *in vivo* is therefore desired to further improve proton therapy treatments. If the range of the beam is known more precisely, range margins can be reduced and new treatment plan designs would become feasible, leading to a lower integral dose to the patient. Continuous monitoring of deviations in the end-of-range can also facilitate adaptations over the course of treatment.



Research on range verification of proton beams has focused mainly on the use of secondary prompt and delayed $\gamma$-rays [30]. Positron emission tomography (PET) has been proposed to detect coincident $\gamma$-rays from the decay of positron emitters produced as a result of proton-nuclear interactions [32, 103]. The delayed decay of these radioactive isotopes is however a disadvantage. Because of diffusion, the location of the $\gamma$-ray emission does not fully coincide with the location of the nuclear interaction. This biological washout effect limits the precision of range verification depending on the tumour location [35]. More recently, detection of prompt $\gamma$-rays during beam delivery has gained interest [36, 52, 92, 104]. Prompt $\gamma$-rays are emitted as a result of the decay of excited nuclear reaction products, which occurs on extremely short timescales. Therefore, they show the potential to provide range verification immediately during treatment.

Because positron emitters and prompt $\gamma$-rays are produced through specific nuclear interactions with the elements in human tissue, the correlation between the detected $\gamma$-rays and the distal dose fall-off is dependent on the composition of irradiated tissue [51, 96, 102, 105]. Range verification methods that cannot distinguish different nuclear reactions therefore require the estimation of tissue compositions. This leads to uncertainty because of the lack of a one-to-one relation between the computed tomography images used for treatment planning and elemental composition [34, 50]. Prompt $\gamma$-rays have different discrete energies which correspond to transitions between excited states of the residual nuclei, such that the possibility does exist to identify specific nuclear reactions. Positron emitting isotopes cannot be directly identified because the resulting $\gamma$-rays have the same energy, however differences in half lives can be exploited to identify contributions from different isotopes [106].

Cross section measurements of prompt $\gamma$-ray lines from proton interactions with carbon and oxygen, which are the main elements in human tissue, have been

performed by various authors, initially for nuclear structure studies [107, 108] and later for applications in astrophysics [81, 84–86]. We have previously compared literature data and simulations of cross sections for reactions with $^{12}C$, $^{14}N$, and $^{16}O$ [96]. Discrete $\gamma$-ray lines have also been identified in experiments with proton beams stopped in matter [92, 97]. Recently, we performed the first measurement of $\gamma$-ray lines along proton Bragg peaks with a collimated detector [104]. We found that resolving the detected $\gamma$-rays in both dimensions of energy and time is beneficial to separate neutron-induced background radiation.

In this chapter, we present a new method to verify the range of proton beams by directly relating measurements of discrete prompt $\gamma$-ray lines to nuclear reaction cross sections. Our aim is to develop a method that is entirely based on quantitative physics models and is robust in the presence of tissue with an unknown elemental composition. Differential cross section measurements were performed to characterize the prompt $\gamma$-ray emission from proton-nuclear reactions with carbon and oxygen up to 150 MeV proton energy. Combining these cross sections with detector models, an optimization method was developed to simultaneously determine the proton range and the composition of the irradiated matter from prompt $\gamma$-ray measurements. Using a small scale prototype detector, the method was tested experimentally with different phantoms. We analysed the accuracy of the detected range and the statistical precision at different dose levels.



## 4.1. Methods and materials

### 4.1.1. Experimental setup

#### Prompt γ-ray detector and collimator

The prompt $\gamma$-ray detector used was an improved version of the detector previously described by Verburg et al. [104]. Designed as a small scale prototype of a potentially clinically applicable system, it consisted of a cerium-doped lanthanum(III) bromide crystal as primary detector surrounded by a segmented annulus of four bismuth germanate (BGO) crystals which served as an active anti-coincidence shield. A photograph and a schematic drawing of the experimental setup with a water phantom are shown in figure 4.1. Different phantoms were also used. The primary detector was cylindrical with a diameter and length of 76 mm and was located with the front at a distance of 284 mm from the proton beam axis. The surrounding crystals had a thickness of 18 mm and a length of 152 mm. All crystals were readout with photomultiplier tubes. Current-to-voltage conversion and signal amplification were performed near the detector, from which the signals were transferred to the data acquisition system outside the treatment room. A tungsten collimator with a thickness of 127 mm and a slit opening of 9.5 mm was used to collimate the prompt $\gamma$-rays. Based on calliper measurements, the alignment error of the collimator relative to the phantom was

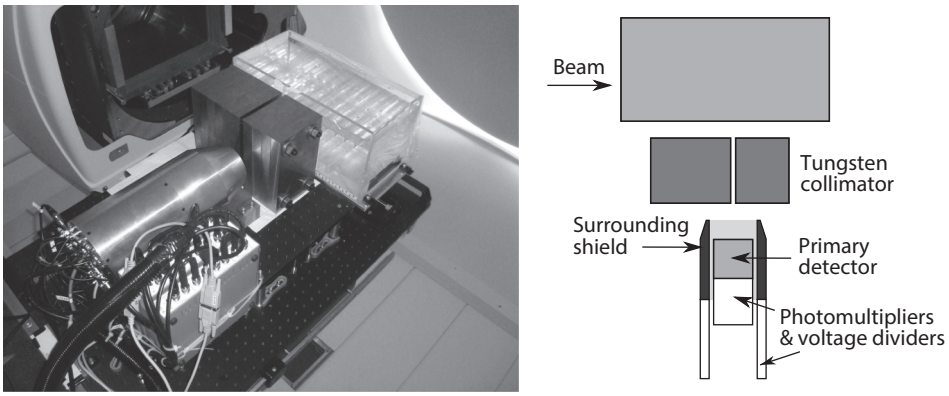



**Figure 4.1** | Experimental setup showing the prompt $\gamma$-ray detector, tungsten collimator and a water phantom in the proton therapy gantry.



determined as < 0.3 mm. The targets were mounted on a motorized linear stage for precise positioning with an accuracy < 0.1 mm.

### Beam delivery

The experiments were performed in a proton therapy gantry using proton pencil-beams. Protons were accelerated with a 230 MeV cyclotron (IBA C230, Ion Beam Applications SA, Louvain-la-Neuve, Belgium) that features a radiofrequency (RF) of 106.3 MHz, which corresponds to an RF period of 9.4 ns. Lower proton energies were obtained with an energy selection system [109]. The beam current and position were monitored with ionization chambers which have been calibrated against Faraday cup measurements to obtain the absolute number of protons delivered. We used a current of approximately 0.5 nA incident on the phantoms. All pencil-beams were delivered along the central axis of the treatment head.

### Data acquisition

The data acquisition system digitized all detector signals with 10-bit free running analog-to-digital converters of which the sampling clocks were phase locked to exact multiples of the cyclotron RF frequency, sampling the primary scintillator pulses 9 times per cyclotron RF period and the signals from the four shielding crystals 4.5 times per RF period. Triggering, beam monitoring and synchronization with the pencil-beam delivery system were handled in real time using field programmable gate arrays. The detector signals and beam monitoring data were continuously buffered and transferred to a computer for processing, storage and on-line visualization. The energy of a detected event was determined from the area under the pulse and the time was obtained by performing polynomial interpolation of the leading edge of the pulse around a constant fraction of the pulse height. Calibration of the $\gamma$-ray energy was

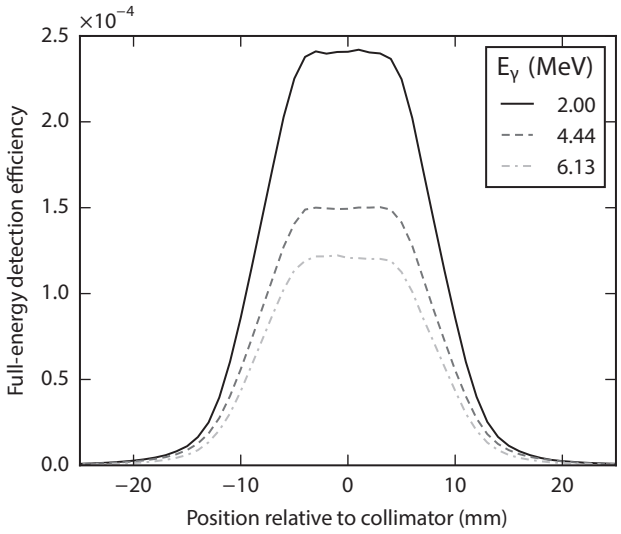



**Figure 4.2** | Simulated collimator-detector transfer function for three $\gamma$ energies. The simulation included 5 cm of water between the proton beam and detector.

automated though continuous monitoring of neutron-induced $\gamma$ lines with known energies. Corrections for pile-up and random coincidences between the primary detector and active shield were also applied, the occurrence of which amounted to a few percent of the detected events. Pile-up was detected through analysis of the pulse shapes. Random coincidences were quantified by fitting the baseline of a histogram of the time differences between events detected by the primary detector and the active shield.

### 4.1.2. Histogram data analysis

We quantified the number of detected proton-induced prompt $\gamma$-rays for a series of $\gamma$ lines with discrete energies. This analysis was performed on two-dimensional histograms of the detected events, with the dimensions being the $\gamma$ energy and the time within the cyclotron RF period. Analytic models were fitted to these histograms, parametrizing the proton-induced signal of interest along with the neutron-induced contribution, continuous background and other nearby proton and neutron-induced $\gamma$ lines. Because of the very short half lives of the excited states, we assumed a constant time structure of the proton-induced $\gamma$-rays independent of $\gamma$-ray energy. This time structure was fitted independently for each measurement. The energy resolutions of the resolved $\gamma$ lines were free parameters due to the proton energy dependence of Doppler broadening and the small dependence on experimental conditions. To fit the models to the measurement data, we used the Levenberg-Marquardt algorithm

as implemented in the Ceres Solver [110]. An example of such a fit is shown by Verburg et al. [104].

### 4.1.3. Detector-collimator transfer

Monte Carlo simulations of the experimental geometry as shown in figure 4.1 were performed using GEANT4 10.0p01 [64] to determine the probability of a $\gamma$-ray emitted along the proton beam path in a random direction to pass through the collimator and deposit its complete energy in the primary detector. We denote this transfer function $t(x)$, in which $x$ is the position of the $\gamma$-ray emission relative to the centre of the collimating slit. The function is specific for a $\gamma$ energy and depends on the location of the proton beam relative to the detector. It also included $\gamma$ attenuation between the source and detector. Examples of the transfer function are shown in figure 4.2.



### 4.1.4. Prompt γ-ray detection model

The proton energy distribution along the delivered proton pencil-beams was also modelled using GEANT4. For each proton pencil-beam, a function $p(w, E)$ was determined that describes the probability of a proton having energy $E$ at water equivalent depth $w$. With the water equivalent depth of the irradiated matter at the detector position denoted $d$, the combination of this simulation with a detector-collimator transfer function

$$g(E) = \int_0^\infty t(x - d)\, p(x, E)\, \mathrm{d}x \tag{4.1}$$

gives the probability density of an emitted $\gamma$-ray being detected with the proton having energy $E$.

The number of emitted proton-induced prompt $\gamma$-rays depends on the cross sections for the nuclear interactions. Because the prompt $\gamma$-rays are emitted non-isotropically [111], the differential prompt $\gamma$-ray production cross sections at a 90° angle with the proton beam direction are of relevance. The small variation of the angle over the primary detector area was neglected. We considered proton interactions with $^{12}$C and $^{16}$O that resulted in discrete prompt $\gamma$-ray emissions. For a given beam and detector model $g(E)$ and a differential cross section $\frac{\mathrm{d}\sigma}{\mathrm{d}\Omega}(E)$, the number of detected $\gamma$-rays $m$ per incident proton is described by

$$m = \frac{4\pi\rho N_A}{A_r} \int_0^\infty g(E)\, \frac{\mathrm{d}\sigma}{\mathrm{d}\Omega}(E)\, \mathrm{d}E, \tag{4.2}$$

in which $\rho$ is the mass concentration of the target element, $N_A$ is the Avogadro constant, and $A_r$ is the relative atomic mass of the target element.

**Table 4.1** | Prompt $\gamma$ lines included in the cross section optimization with nuclear reaction assignments.

| Target | $\gamma$ energy (MeV) | Assignments | Other data |
|---|---|---|---|
| $^{16}$O | 1.89 | $^{16}\mathrm{O}(p, pp\gamma_{1.89})^{15}\mathrm{N}$ | |
| | 2.0 | $^{16}\mathrm{O}(p, x\gamma_{2.04})^{15}\mathrm{O}$<br>$^{16}\mathrm{O}(p, x\gamma_{2.00})^{11}\mathrm{C}$ | |
| | 2.31 | $^{16}\mathrm{O}(p, x\gamma_{2.31})^{14}\mathrm{N}$ | Foley et al. [108]<br>Lang et al. [83] |
| | 2.8 | $^{16}\mathrm{O}(p, p'\gamma_{2.74})^{16}\mathrm{O}$<br>$^{16}\mathrm{O}(p, x\gamma_{2.79})^{14}\mathrm{N}$<br>$^{16}\mathrm{O}(p, x\gamma_{2.80})^{11}\mathrm{C}$<br>$^{16}\mathrm{O}(p, x\gamma_{2.87})^{10}\mathrm{B}$ | Lang et al. [83]<br>Kiener et al. [84] |
| | 3.68 | $^{16}\mathrm{O}(p, x\gamma_{3.68})^{13}\mathrm{C}$ | |
| | 4.44 | $^{16}\mathrm{O}\,(p, x\gamma_{4.44})\,^{12}\mathrm{C}$ | Foley et al. [108]<br>Lang et al. [83]<br>Belhout et al. [85] |
| | 5.2 | $^{16}\mathrm{O}(p, x\gamma_{5.24})^{15}\mathrm{O}$<br>$^{16}\mathrm{O}(p, pp\gamma_{5.27})^{15}\mathrm{N}$<br>$^{16}\mathrm{O}(p, x\gamma_{5.18})^{15}\mathrm{O}$<br>$^{16}\mathrm{O}(p, pp\gamma_{5.30})^{15}\mathrm{N}$ | Foley et al. [108]<br>Lang et al. [83]<br>Belhout et al. [85] |
| | 6.1 | $^{16}\mathrm{O}\,(p, p'\gamma_{6.13})\,^{16}\mathrm{O}$<br>$^{16}\mathrm{O}\,(p, x\gamma_{6.18})\,^{15}\mathrm{O}$ | Foley et al. [108]<br>Narayanaswamy et al. [82]<br>Lang et al. [83]<br>Kiener et al. [84]<br>Belhout et al. [85] |
| | 6.32 | $^{16}\mathrm{O}(p, x\gamma_{6.32})^{15}\mathrm{N}$ | |
| | 7.0 | $^{16}\mathrm{O}(p, p'\gamma_{6.92})^{16}\mathrm{O}$<br>$^{16}\mathrm{O}(p, p'\gamma_{7.12})^{16}\mathrm{O}$ | Foley et al. [108]<br>Kiener et al. [84] |
| $^{12}$C | 2.0 | $^{12}\mathrm{C}(p, x\gamma_{2.00})^{11}\mathrm{C}$ | Clegg et al. [107]<br>Lang et al. [83] |
| | 2.1 | $^{12}\mathrm{C}(p, pp\gamma_{2.12})^{11}\mathrm{B}$<br>$^{12}\mathrm{C}(p, x\gamma_{2.15})^{10}\mathrm{B}$ | Lang et al. [83] |
| | 2.8 | $^{12}\mathrm{C}(p, x\gamma_{2.80})^{11}\mathrm{C}$<br>$^{12}\mathrm{C}(p, x\gamma_{2.87})^{10}\mathrm{B}$ | |
| | 4.44 | $^{12}\mathrm{C}(p, p'\gamma_{4.44})^{12}\mathrm{C}$ | Clegg et al. [107]<br>Lang et al. [83]<br>Kiener et al. [84]<br>Belhout et al. [85] |
| | 4.80 | $^{12}\mathrm{C}(p, x\gamma_{4.80})^{11}\mathrm{C}$ | |

### 4.1.5. Cross section measurements

To enable quantitative modelling of prompt $\gamma$-ray emissions, we measured and optimized differential prompt $\gamma$-ray production cross sections at 90° for a total of 15 $\gamma$ line excitations in proton-induced reactions on $^{12}C$ and $^{16}O$, as listed in table 4.1. The cross section optimization was performed for the 0 MeV to 150 MeV proton energy range. Some of the $\gamma$-ray emissions comprised contributions from multiple nuclear level transitions that could not be reliably resolved individually, which was either due to the energy resolution of the detector or due to Doppler broadening in the case of excited nuclear states with a very short half-life. The nuclear reaction assignments in table 4.1 were identified as resulting in $\gamma$-ray production with an energy consistent with the observed $\gamma$-ray lines [55].

The main data used to determine the cross sections consisted of a series of measurements in which we stopped a 165 MeV proton beam in homogeneous targets of $H_2O$ and $CH_2$ with known proton stopping powers. Gamma yields during proton irradiation were measured with the collimated detector positioned at 90 different depths in the targets in the range of 30 mm to 210 mm. Each of these $\gamma$-ray yields therefore resulted from interactions of protons in a specific energy range. Because of the proton energy resolution of these measurements, which depends on the width of the collimator opening, it is difficult to deconvolve the exact energy dependence of the cross sections in the low energy range. Therefore, we additionally used previous cross section data from literature in the cross section optimization procedure.

#### Previous data

Several authors have performed thin target cross sections measurements at a number of proton energies, which are referenced in table 4.1. Differential cross sections at 90° were derived from these data, using the standard compound nucleus model in which the angular $\gamma$ emission is described with Legendre polynomials [111]. Some early thick target measurements at 90° and energies around 150 MeV are also available [107, 108], the accuracy of which may however be limited by the low energy resolution of the NaI detectors used at the time.

#### Optimization procedure

An optimization procedure was used to obtain differential cross sections that agreed with our experimentally determined $\gamma$ yields as well as trends obtained from the previous data. This procedure was followed for each individual $\gamma$ line and target element. For our measurements, we used target materials with a known composition and proton stopping power, therefore the function $g(E)$ as given in equation (4.1) was known a priori for all measurement points. Using equation (4.2), a discrete model was created of the expected $\gamma$ yields for a given cross section vector $\mathbf{c}$ that was discretized in 1 MeV energy bins. This model was pre-calculated at the 90 measurement positions in the form of a system matrix $\mathbf{A}$.

The NESTA solver [112] was employed to solve the following optimization problem at proton energies above the lowest nuclear reaction threshold of the contributing reaction channels:

$$\underset{\mathbf{c}}{\operatorname{argmin}} \left( \|\mathbf{Ac} - \mathbf{m}\|_2^2 + \lambda \|\nabla (\mathbf{c} - \mathbf{s})\|_1 \right). \tag{4.3}$$

The first term in this function is the least squares consistency of the cross sections with the vector **m** containing the measured $\gamma$ yields at the 90 positions. Because each measured $\gamma$ yield corresponded to a range of proton energies, optimizing the consistency of the solution with the measurement data alone resulted in non-physical solutions with large gradients. A second regularization term was therefore used that sets a small penalty on the gradient magnitude of the solution. For the $\gamma$ lines for which previous literature data was available, a shape estimate of the cross section denoted **s** was created by linear interpolation of the available data points. This estimate was subtracted from the cross section in the regularization term, thereby creating a penalty on the gradient magnitude of the difference between the cross section and the estimate from literature data. Therefore, solutions with gradients that are known to exist were favoured. Smooth systematic differences were however not penalized, i.e. we relied on our own data for a consistent normalization. A single value for the parameter $\lambda$ was used for all $\gamma$ emissions, which was empirically adjusted to the minimum value such that the inverse problem was effectively regularized but without significantly increasing the first data consistency term.

A final local smoothing procedure was applied to the optimized cross sections to remove small variations due to experimental uncertainties, which are unlikely to relate to actual resonances in the cross section. We applied a mean filter to the cross section above an energy threshold that was determined such that the increase of the data consistency term was limited to a small fraction. This process was repeated with a larger window size of the filter until it could not be applied to any part of the energy range.

### 4.1.6. Proton range verification

To verify the range of proton pencil-beams, we placed the prompt $\gamma$-ray detector at a fixed position. Instead of verifying the range of each individual pencil-beam, we propose to combine measurements during the delivery of pencil-beams to the same lateral location within the field to verify the water equivalent depth of the irradiated matter at the detector position. A comparison of this detected water equivalent depth with the expected depth then reveals a deviation of the proton range. If the detector is positioned close to the end-of-range of the beam, the uncertainties from the remaining downstream part of the beam path will be very small. Because the measured $\gamma$-rays result from proton-induced interactions with the same matter, the elemental oxygen and carbon concentrations at the measurement position can be

determined, which makes the range verification method generally applicable without requiring prior knowledge of the elemental composition.

**Prompt γ-ray model**

A model of the expected prompt $\gamma$-ray detection was created for each pencil-beam, $\gamma$ line, and target element. These model calculations were obtained by evaluating equation (4.2) as a function of $d$, which is the water equivalent depth of the phantom at the detector position that is now considered unknown. This calculation required a probability density function of the $\gamma$-ray detection $g(E)$ as defined in equation (4.1) to be determined for each water equivalent depth. Simulations in water were used to determine this function; the assumption was therefore that the energy spread of the beam was comparable to a proton beam traversing water. As discussed in section §4.3, different models derived from computed tomography scans can be applied in a clinical setting. The calculations of the expected $\gamma$-ray detection were were performed with 0.1 mm steps in depth and followed by spline interpolation to obtain continuous models of the $\gamma$-ray counts denoted $f^t_{b,l}(d)$ for a pencil-beam $b$, prompt $\gamma$ line $l$, and target element $t$. The models were normalized per unit concentration of the target element. Nuclear reactions with multiple target elements can contribute to the same $\gamma$ line.

**Model fitting**

Range verification was performed by fitting the calculated models to the measurement data $\mathbf{m}_{b,l}$ to determine the water equivalent depth at the detector position. As discussed earlier, the elemental concentrations of the target elements of the nuclear reactions were determined together with the water equivalent depth such that the range verification required no prior knowledge of the composition of the irradiated matter. All $\gamma$ lines were considered simultaneously in this optimization procedure by least squares minimization:

$$\underset{d,\rho_O,\rho_C}{\operatorname{argmin}} \left( \sum_{b\in B} \sum_{l\in L} h\left(\rho_O f^O_{b,l}(d) + \rho_C f^C_{b,l}(d) - \mathbf{m}_{b,l}\right)^2 \right), \tag{4.4}$$

in which the optimization parameters are the water equivalent depth $d$ of the phantom at the detector position, the $^{16}O$ concentration $\rho_O$ and the $^{12}C$ concentration $\rho_C$. These parameters were optimized for the set of pencil-beams $B$ and the sets of $\gamma$ lines $L$. The Huber loss function $h$ was applied to reduce the impact of outliers [113].

## 4.1.7. Detection of relative range shifts

The model described in the previous section was designed with the aim of determining the absolute range of the proton pencil-beam without requiring knowledge of

tissue composition. Once such a measurement has been performed, relative range shifts at the same location can be detected with a higher accuracy by considering all detected proton-induced prompt $\gamma$-rays. From a first measurement, the correlation between the beam range and the total detected number of proton-induced prompt $\gamma$-rays was established. The range shift of a second measurement relative to the first can then be determined, because the water equivalent range shift should equal the difference in beam range to match the previously measured proton-induced $\gamma$ counts. The total number of proton-induced $\gamma$-rays in a wide energy window was quantified by separating proton- and neutron-induced events. The average event count outside a predefined time window relative to the cyclotron RF, which contained the proton-induced events, was subtracted from the average count inside the window. This method provided a sufficient removal of neutron background while being well reproducible with limited statistics. Range shifts between the two measurements were detected by creating a spline interpolation of the relation between the beam range and $\gamma$ counts in the first measurement and finding the range shift that minimizes the least square difference between the two measurements.



### 4.1.8. Range verification experiments in phantoms

The proposed range verification methods were tested experimentally with two phantoms. The first phantom consisted of a water tank; for the second experiment a volume of 'solid water' material (Gammex Inc., Middleton, WI) was placed in the water tank, encompassing the entire proton beam in view of the detector. Although the solid water material has a stopping power similar to water, it has an elemental composition (67 % C, 20 % O, 8 % H by mass) which is very different from water (89 % O, 11 % H by mass). Such differences can also be found in human tissue; brain- and soft tissue consists mostly of oxygen, while adipose tissue has a high carbon content. To introduce range shifts, thin slabs of plastic were placed in the beam path. The measurements were repeated 90 times to enable statistical analysis.

## 4.2. Results

### 4.2.1. Prompt γ-ray spectrum

To show the characteristics of the measurements acquired with the detection system, an example of an energy- and time-resolved prompt $\gamma$-ray measurement is depicted in figure 4.3. This measurement was performed 20 mm upstream of the end-of-range of a 165 MeV proton beam stopped in water and polyethylene targets. A clear proton-induced $\gamma$ peak is seen in which discrete $\gamma$ lines are well resolved in both dimensions. Because of the velocity distribution of secondary neutrons, events from neutron-induced reactions appear as a horizontal continuum and horizontal lines in the histogram. Both the unique timing and energy characteristics of the proton-induced $\gamma$-rays enabled their separation from the background.

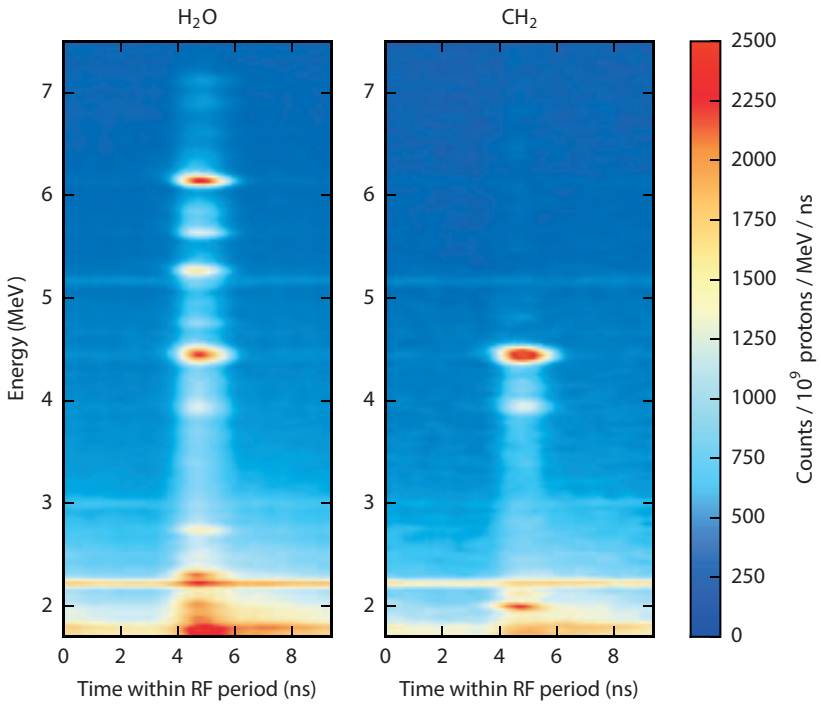



**Figure 4.3** | Histograms of energy- and time resolved prompt $\gamma$-ray measurements in $H_2O$ and $CH_2$ targets, performed 20 mm upstream of the end-of-range of a 165 MeV proton beam.

Figure 4.4 shows prompt $\gamma$-ray energy spectra in which the neutron-induced component was removed by fitting the background below the proton-induced peak. These spectra were averaged over a number of measurements performed at different positions along last 30 mm of the path of the same proton beam, in order to show all relevant $\gamma$ lines with different nuclear reaction thresholds. The $\gamma$ lines were assigned to specific residual nuclei as identified in the figure.

To confirm that the peak in figure 4.3 corresponds to $\gamma$-rays from the primary protons, a measurement was performed at different positions around the end-of-range of a 151 MeV proton beam in a water phantom, which had a range at 80 % distal dose fall-off in water of $R_{80}$ = 15.93 g cm$^{-2}$. The measurement was repeated with a 2.3 mm water equivalent slab in the beam path. Per measurement point, $10^{11}$ protons were delivered. The prompt $\gamma$-ray counts in the proton-induced peak from 1.8 MeV to 7.2 MeV are shown in figure 4.5, in which the range shift can be identified as a shift between the two curves. The slope of the fall-off is mainly dependent on the width of the collimator opening.

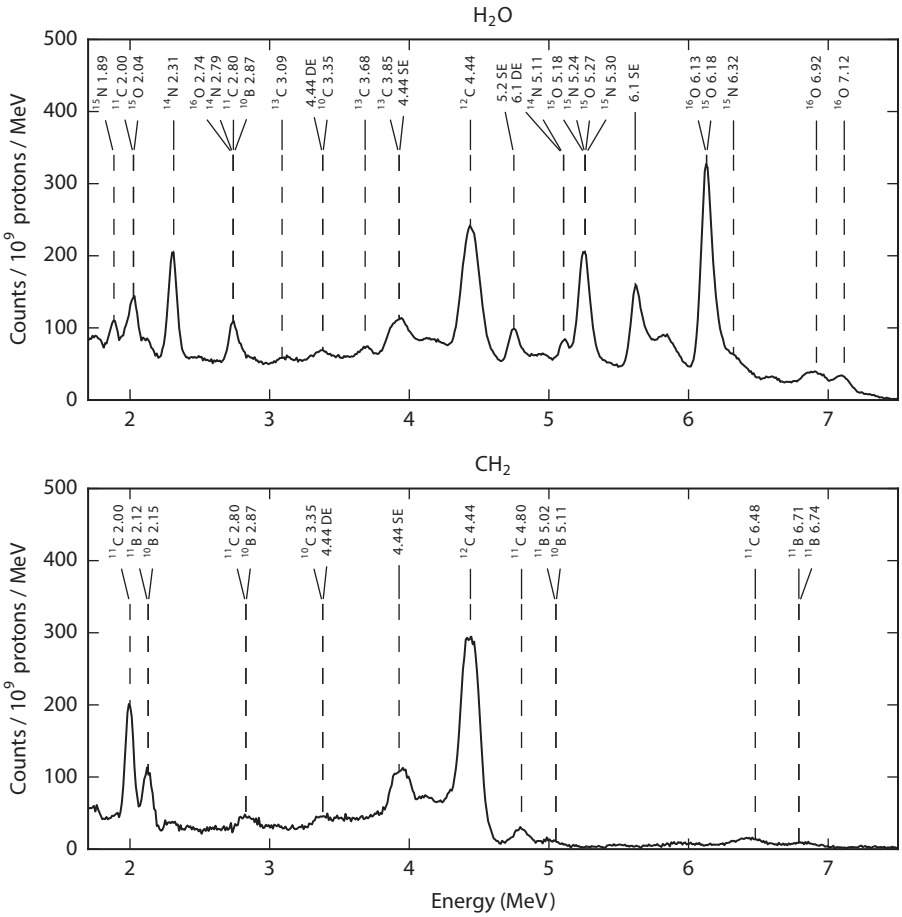



**Figure 4.4** | Energy spectra of proton-induced prompt $\gamma$-rays measured along the last 30 mm of the range of a 165 MeV proton beam stopped in water and polyethylene. Neutron-induced contributions were removed. The line assignments list the residual nucleus and the $\gamma$ energy in MeV (SE = single escape; DE = double escape).

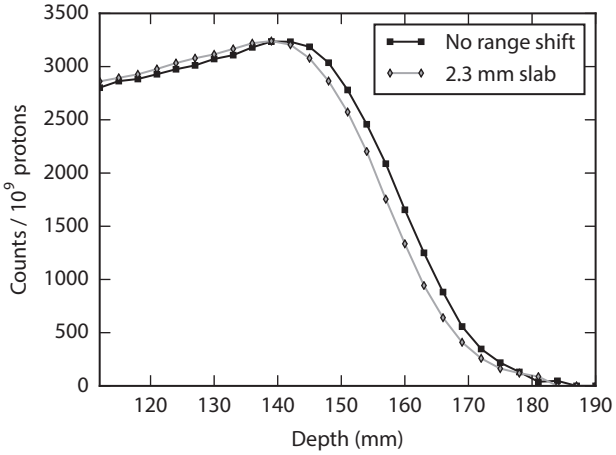





**Figure 4.5** | Total detected proton-induced prompt $\gamma$-rays in the 1.8 MeV to 7.2 MeV energy window, measured at different positions along a 151 MeV proton beam in water. The measurement was performed with and without a slab in the beam path, which had a water equivalent thickness of 2.3 mm.

### 4.2.2. Cross sections and prompt γ-ray model

The data acquired for cross section optimization, consisting of measurements along a 165 MeV proton pencil-beam stopped in water and polyethylene, are given in figure 4.6. These measurements were respectively performed with $2\times10^{11}$ and $5\times10^{10}$ protons per point. Prompt $\gamma$-ray production cross sections derived from the data, using the optimization procedure described in section §4.1.5, are shown in figure 4.7. For comparison, the data points obtained from the previous literature referenced in table 4.1 are also shown. We found a reasonable agreement between our cross sections and the previous measurements. The lines in figure 4.6 represent modelled $\gamma$ yields derived from the optimized cross sections as described in section §4.1.6. The agreement of the modelled $\gamma$ yields with the measurement data shows the consistency of the models.

### 4.2.3. Verification of absolute proton range

For the range verification experiments, we positioned the detector at a fixed physical depth of 155 mm. In the water phantom, this corresponded to a water equivalent depth of 156.5 mm because of a small correction for the plastic wall of the phantom. In the solid water phantom, which had a slightly higher stopping power as compared to water, the same physical depth corresponded to a water equivalent depth of 157.1 mm. Five pencil-beams were delivered with ranges $R_{80}$ of 15.54 g cm$^{-2}$, 16.07 g cm$^{-2}$, 16.58 g cm$^{-2}$, 17.09 g cm$^{-2}$, and 17.59 g cm$^{-2}$. We evaluated two dose

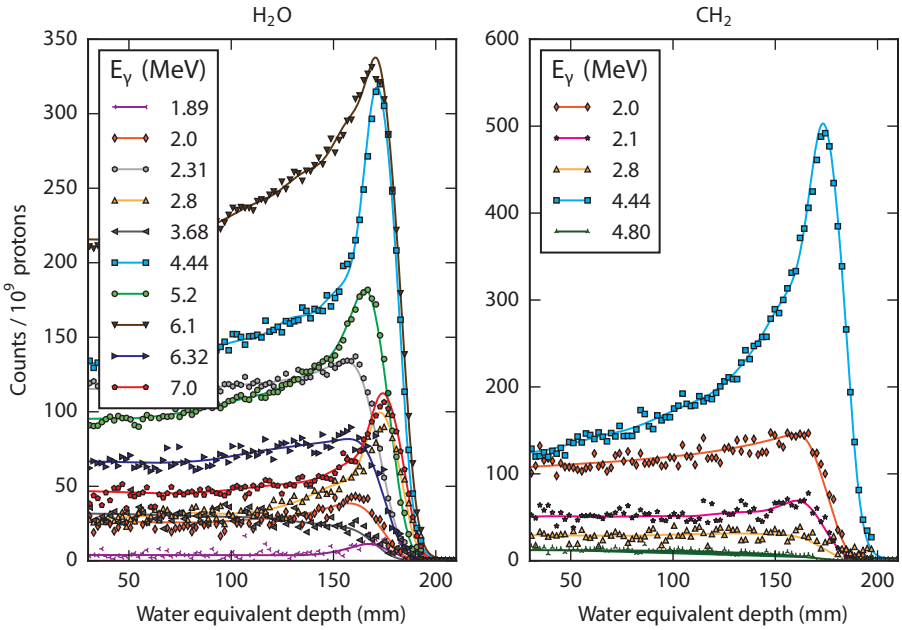



**Figure 4.6** | Measured discrete prompt $\gamma$-ray yields along a 165 MeV proton beam stopped in water and polyethylene. The lines are model calculations using the optimized cross sections and the detector-collimator model.

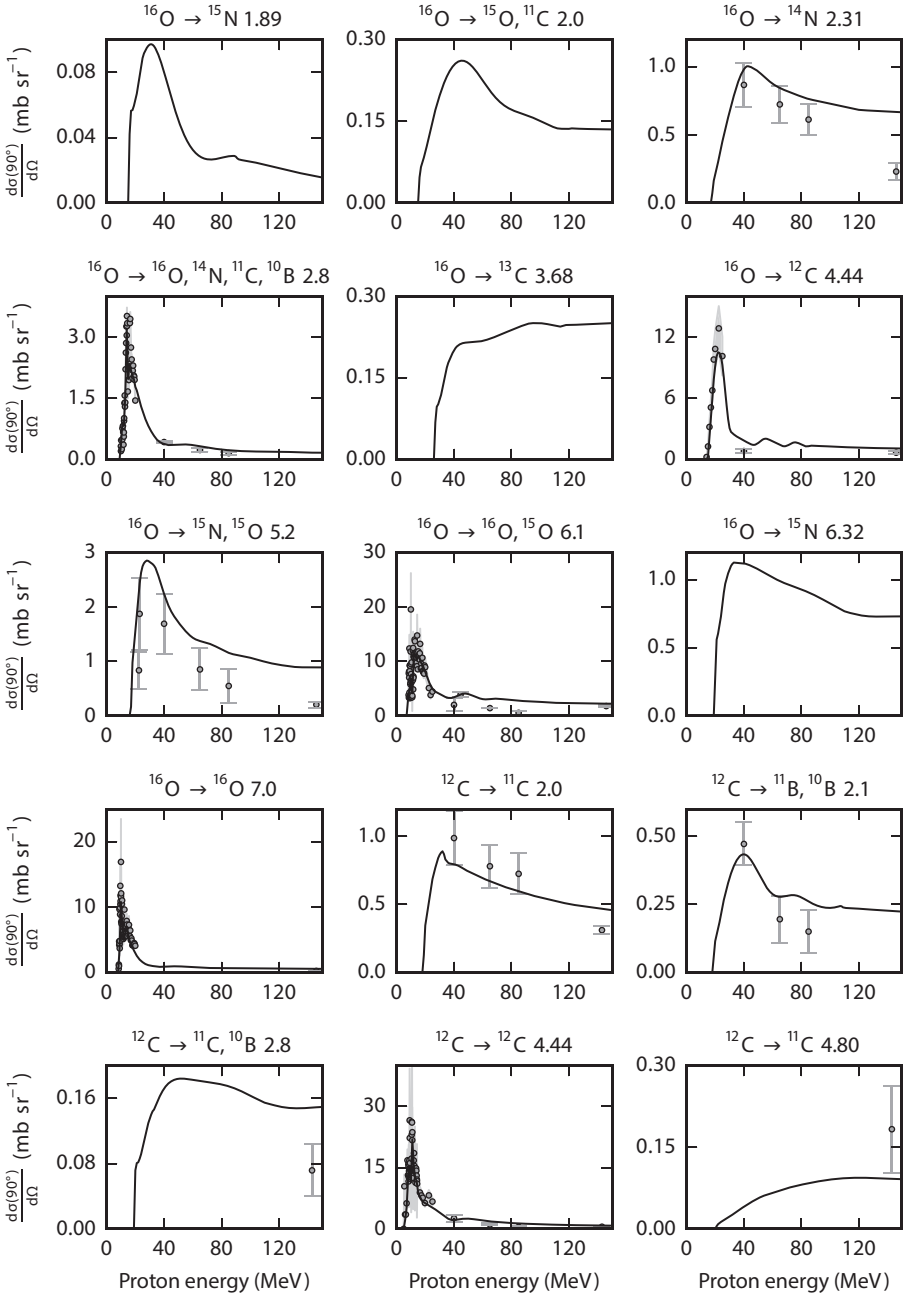



**Figure 4.7 |** Optimized differential cross sections (lines) of prompt $\gamma$-ray production from proton interactions with $^{16}O$ and $^{12}C$ at a 90° angle with the proton beam direction, as listed in table 4.1. The plot titles display the target and residual nuclei and the $\gamma$ energy in MeV. The circles are data points from literature with uncertainties shown as error bars or bands.

**Table 4.2** | Proton range verification using 5 pencil-beams with $R_{80}$ between 15.54 g cm$^{-2}$ to 17.59 g cm$^{-2}$. Listed are the detected water equivalent depth at the detector position, the difference between the detected depth and the calculated value, and the determined $^{12}$C and $^{16}$O concentrations of the irradiated matter. Uncertainties are reported as $\pm 1\sigma$ and were determined from 90 measurements.

| | Detected depth (mm) [a] | Range error (mm) [b] | $^{16}$O (g cm$^{-3}$) [c] | $^{12}$C (g cm$^{-3}$) [d] |
|---|---|---|---|---|
| Water, $5 \times 10^8$ protons | 156.3 ± 1.3 | 0.1 ± 1.3 | 0.94 ± 0.07 | 0.05 ± 0.07 |
| with 2.3 mm slab[b] | 158.6 ± 1.0 | −2.2 ± 1.0 | 0.90 ± 0.07 | 0.02 ± 0.06 |
| with 5.2 mm slab[b] | 161.6 ± 1.0 | −5.2 ± 1.0 | 0.88 ± 0.07 | 0.04 ± 0.06 |
| Solid water, $5 \times 10^8$ protons | 157.1 ± 1.4 | 0.0 ± 1.4 | 0.17 ± 0.04 | 0.70 ± 0.07 |
| with 2.3 mm slab[b] | 159.3 ± 1.4 | −2.3 ± 1.4 | 0.18 ± 0.03 | 0.70 ± 0.07 |
| with 5.2 mm slab[b] | 162.4 ± 1.3 | −5.4 ± 1.3 | 0.18 ± 0.04 | 0.69 ± 0.08 |
| Water, $1 \times 10^8$ protons | 156.4 ± 2.4 | 0.1 ± 2.4 | 0.91 ± 0.14 | 0.08 ± 0.12 |
| with 2.3 mm slab[b] | 159.0 ± 2.5 | −2.6 ± 2.5 | 0.87 ± 0.16 | 0.10 ± 0.16 |
| with 5.2 mm slab[b] | 161.8 ± 2.9 | −5.4 ± 2.9 | 0.86 ± 0.18 | 0.08 ± 0.16 |
| Solid water, $1 \times 10^8$ protons | 157.5 ± 3.6 | −0.4 ± 3.6 | 0.20 ± 0.09 | 0.72 ± 0.19 |
| with 2.3 mm slab[b] | 159.2 ± 3.1 | −2.2 ± 3.1 | 0.20 ± 0.08 | 0.68 ± 0.18 |
| with 5.2 mm slab[b] | 162.1 ± 3.0 | −5.0 ± 3.0 | 0.22 ± 0.10 | 0.67 ± 0.17 |

[a] Calculated water equivalent depth at detector position: water 156.5 mm, solid water 157.1 mm
[b] Water equivalent thickness
[c] Actual $^{16}$O concentration: water 0.89 g cm$^{-3}$, solid water 0.21 g cm$^{-3}$ typical.
[d] Actual $^{12}$C concentration: solid water 0.70 g cm$^{-3}$ typical.

levels: $5 \times 10^8$ protons per pencil-beam and $1 \times 10^8$ protons per pencil-beam. As described in section §4.1.6, no assumptions were made on the elemental composition of the irradiated phantom and the carbon and oxygen concentration were determined from the prompt $\gamma$-ray emissions.

The detected water equivalent depths and elemental concentrations are shown in table 4.2. With $5 \times 10^8$ protons per pencil-beam, a $1\sigma$ precision on the water equivalent depth of about 1 mm was achieved. The carbon and oxygen concentration determined by the model also agreed well with the actual values. Range shifts that were introduced by placing a slab of plastic in the beam path were accurately detected as a change in the water equivalent depth. With a reduced dose level of $1 \times 10^8$ protons per pencil-beam, the absolute depth was also detected reliably and the statistical precision was 2 mm to 4 mm. The difference in statistical precision between the experiments in the water and solid water phantoms is due to the difference in carbon content. Proton-nuclear reactions on $^{16}$O result in about twice as many prompt $\gamma$-rays as compared to $^{12}$C.

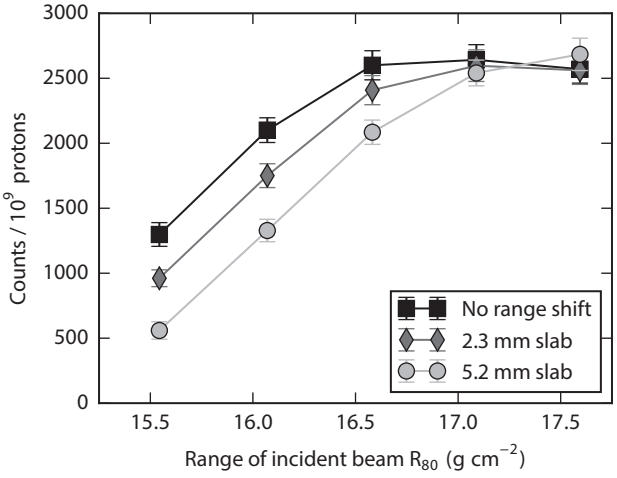





**Figure 4.8** | Total prompt $\gamma$-rays in the energy range of 1.8 MeV to 7.2 MeV, measured in water with the detector at a water equivalent depth of 156.5 mm. Only the events within a time window of 1.25 ns relative to the cyclotron RF period were considered. The baseline of neutron-induced events, determined by averaging the counts outside this time window, was subtracted. The error bars are $\pm 1\sigma$ uncertainty derived from 90 measurements with $5 \times 10^8$ protons per point. The lines serve to guide the eye.

**Table 4.3** | Relative range shifts detected using the difference between prompt $\gamma$-ray measurements. The 5 pencil-beams were delivered with $R_{80}$ between 15.54 g cm$^{-2}$ and 17.59 g cm$^{-2}$. The results for a single pencil-beam were based on the beam with a range of 16.07 g cm$^{-2}$. The detector was positioned at a water equivalent depth of 156.5 mm in water and 157.1 mm in solid water. Uncertainties are reported as $\pm 1\sigma$ and were determined from 90 measurements.

| | **Range error (mm)** [a] **$5 \times 10^8$ protons** | **Range error (mm)** [a] **$1 \times 10^8$ protons** |
|---|---|---|
| Water, 5 pencil-beams | 0.0 ± 0.4 | 0.0 ± 0.8 |
| with 2.3 mm slab[a] | −2.3 ± 0.3 | −2.5 ± 0.8 |
| with 5.2 mm slab[a] | −5.2 ± 0.3 | −5.5 ± 0.7 |
| Solid water, 5 pencil-beams | 0.0 ± 0.5 | −0.3 ± 1.3 |
| with 2.3 mm slab[a] | −2.2 ± 0.5 | −2.4 ± 1.1 |
| with 5.2 mm slab[a] | −5.1 ± 0.4 | −5.3 ± 1.0 |
| Water, 1 pencil-beam | 0.0 ± 0.6 | 0.0 ± 1.4 |
| with 2.3 mm slab[a] | −2.3 ± 0.5 | −2.4 ± 1.2 |
| with 5.2 mm slab[a] | −5.0 ± 0.5 | −5.1 ± 1.1 |
| Solid water, 1 pencil-beam | 0.0 ± 0.8 | 0.0 ± 1.9 |
| with 2.3 mm slab[a] | −2.1 ± 0.8 | −2.2 ± 1.7 |
| with 5.2 mm slab[a] | −5.0 ± 0.8 | −5.2 ± 1.6 |

[a] Water equivalent thickness

### 4.2.4. Detection of relative range shifts

The detection of range shifts relative to a previous measurement was tested using the same setup as in the previous section, following the procedure described in section §4.1.7. Again, the detector was at a fixed position and pencil-beams were delivered with same five ranges between 15.54 g cm$^{-2}$ to 17.59 g cm$^{-2}$. The total number of $\gamma$-rays was quantified within an energy window of 1.8 MeV to 7.2 MeV and a time window of 1.25 ns was used to separate proton- and neutron-induced reactions.

The first measurement was performed with the five pencil-beams delivering $5 \times 10^8$ protons, which is depicted in figure 4.8. For comparison, we also show two measurements in which a range shift was introduced, showing that the horizontal distance between these curves, i.e. the change in beam range to obtain identical $\gamma$ counts, corresponds to the range shift. We analysed the precision of range verification relative to this measurement using the same five pencil-beams and also using only a single pencil-beam with a range $R_{80}$ of 16.07 g cm$^{-2}$. Two different dose levels were again evaluated, for which the results are listed in table 4.3. The results show a good agreement between the detected relative range shifts and the thickness of the slab that was placed in the beam path. At a dose level of $5 \times 10^8$ protons per pencil-beam, the shifts were detected with a statistical precision of 0.3 mm to 0.5 mm based on five pencil-beams and 0.7 mm to 1.3 mm based on a single pencil-beam. These values approximately doubled with a lower dose of $1 \times 10^8$ protons. Compared to the absolute range verification, a higher statistical precision was observed because the available prior measurement enabled all detected events to be considered.

## 4.3. Discussion and conclusions

We have presented a new method for proton range verification through spectroscopy of proton-induced prompt $\gamma$-rays. For the first time, the range of proton beams was obtained by directly relating measured prompt $\gamma$-ray spectra with models of proton-nuclear interactions. The proposed range verification method was enabled by quantitative measurements of prompt $\gamma$-ray emissions from specific nuclear reactions. Discrete $\gamma$ lines were resolved and proton- and neutron-induced contributions were separated. Models derived from measured differential cross sections were used to simultaneously determine the proton range and the oxygen and carbon concentrations of the irradiated tissue. Therefore, the method is robust in the presence of tissue with an unknown composition.

Our results show that, instead of performing simultaneous measurements at many positions along the path of a single pencil-beam, it is possible to accurately verify the proton range using a single detector acquiring data at a fixed position during the delivery of a few distal pencil-beams to the same location. In a clinical scenario, the measurement plane would normally be positioned near the distal end

of the target. The method would provide a two-dimensional measurement at the detector position of the deviation of the proton ranges from expected values. While only a single detector is required, it is also possible to use multiple detectors to improve statistical precision and to obtain simultaneous measurements at different positions along the beam path. This may be advantageous to ensure coverage of the entire field when the distal surface of the tumour is highly curved.

Differential prompt $\gamma$-ray production cross sections at 90° were optimized for proton interactions with carbon and oxygen up to 150 MeV proton energy, based on new measurements combined with previous literature data. These cross sections provided the basis of our model that was directly based on fundamental nuclear interactions. An important advantage of this approach is that physical effects which will be of relevance for the verification of clinical treatments can be incorporated in the model. Because the proton range is verified based on $\gamma$-rays with known energies that did not scatter in the patient, precise attenuation corrections can be performed. The models can be created for individual pencil-beams to incorporate factors such as the energy spread due to surface curvature and inhomogeneities in the tissue traversed within the area of a pencil-beam, which can be derived from a computed tomography scan.



We focused on proton interactions with $^{12}C$ and $^{16}O$ because they are the main isotopes in human tissue. Because no other isotopes in human tissue result in nuclear reactions that significantly populate the excited states of these isotopes, the model is expected to be accurate within any tissue. Incorporating cross sections for other isotopes such as $^{40}Ca$ can improve precision when verifying the range of beams stopping near bones, and will be subject to future study. Also, there is certainly room for additional study of the differential cross sections. Thin target measurements throughout the entire therapeutic proton energy range would be desired.

We have emphasized that the determination of the *absolute* proton range without prior knowledge of tissue composition is different from detecting *relative* differences in proton range during the course of treatment. After a first measurement, relative range errors can be established with higher statistical precision by monitoring the change in the total number of detected proton-induced prompt $\gamma$-rays. A combination of methods could be applied in fractionated treatments. For example, absolute range verification could be based on measurements performed during the entire treatment fraction, while relative range verification could signal errors compared to the previous fraction after a subset of the dose has been delivered.

The achieved statistical precision with our small scale prototype detector was about 1 mm standard deviation for absolute range verification and 0.4 mm for relative verification based on five pencil-beams delivering $5 \times 10^8$ protons, which is encouraging in view of clinical application. In a proton field delivering 1 Gy using pencil-beams with a size of 10 mm sigma, a typical pencil-beam with a range near

the distal end of the tumour delivers about $10^8$ protons. We anticipate that improving detection efficiency by an order of magnitude is feasible for a full scale system, which brings the required dose for millimetre accuracy well within clinically useful levels.

# 5

# Range verification of passively scattered proton beams

This chapter describes an experimental study to verify the range of passively scattered proton beams, by detecting prompt $\gamma$-rays emitted from proton-nuclear interactions. A method is proposed using a single scintillation detector positioned near the distal end of the irradiated target. Lead shielding was used to attenuate $\gamma$-rays emitted along most of the entrance path of the beam. By synchronizing the prompt $\gamma$-ray detector to the rotation of the range modulation wheel, the relation between the $\gamma$ emission from the distal part of the target and the range of the incident proton beam was determined. In experiments with a water phantom and an anthropomorphic head phantom, this relation was found to be sensitive to range shifts that were introduced. The wide opening angle of the detector enabled a sufficient signal-to-background ratio to be achieved in the presence of neutron-induced background from the scattering and collimating devices. Uniform range shifts were detected with a standard deviation of 0.1 mm to 0.2 mm at a dose level of 30 cGy (RBE) to 50 cGy (RBE). The detectable magnitude of a range shift limited to a part of the treatment field area was approximately proportional to the ratio between the field area and the area affected by the range shift. We conclude that the consistency of the range of passively scattered proton beams can be verified using a relatively simple prompt $\gamma$-ray detection system.

Based on published work: J.M. Verburg, M. Testa, and J. Seco. Range verification of passively scattered proton beams using prompt gamma-ray detection. *Phys. Med. Biol.* **60**, 1019–1029 (2015).

Prompt γ-ray detection is a promising method for *in vivo* range verification of proton therapy beams [36, 52, 92, 114]. Detectors are currently being developed and studied experimentally by several groups, which so far have focused on the verification of proton pencil-beams [30]. Because passive scattering systems are in widespread use in current proton therapy centres, it is also worthwhile to investigate if the technique has potential to be applied to scattered proton beams. It could serve as a quality assurance tool to verify the consistency of treatments.

A Monte Carlo simulation study of prompt $\gamma$-ray detection during passive-scattering proton therapy was recently performed by Testa et al. [115]. They proposed a synchronization of the prompt $\gamma$-ray detector with the rotating range modulation wheel, which is a method that has been previously applied to *in vivo* dosimeters in body cavities [37]. Prompt $\gamma$-rays induced by the primary protons could be resolved in the presence of simulated radiation background from the treatment nozzle and hardware, which indicates the detection of prompt $\gamma$-rays during the delivery of scattered beams may be experimentally feasible.

Prompt $\gamma$-ray detection systems with slit collimators that are being developed for one-dimensional range verification of proton pencil-beams, such as multi-slit and knife-edge collimator systems [92, 116], could in principle be applied to passively scattered beams, either using a synchronization with the beam delivery or by using data integrated over the entire range modulation cycle. However, background radiation levels will be significantly higher because of neutron-induced radiation from proton interactions with the nozzle, in particular due to protons that are stopped in the snout base and aperture. Changes to the system designs will be needed to achieve a sufficient signal-to-background ratio for scattered proton beams. The optimal collimator and detector design is likely to be different.

The use of a Compton camera combined with iterative image reconstruction methods is another option. A major advantage of this detector is the potential to provide three-dimensional imaging of emitted prompt $\gamma$-rays [117]. Several groups are currently working on prototype detectors [94, 118–120]. However, handling high dose rates is challenging with a Compton camera, because of the high count rates and the need for coincidences between multiple detector layers to be analysed [120]. The increased background radiation from scattered beams would make it more difficult to perform range verification with clinically acceptable dose rates.

In chapter, we propose an alternative and relatively simple prompt $\gamma$-ray detection setup that was specifically designed for the passive scattering modality, with the aim of verifying whether the range of the beam is consistent throughout the course of treatment. For the first time, experimental measurements of prompt $\gamma$-rays during the delivery of passively scattered proton beams are presented. Similarly to the simulation study of Testa et al. [115], we used the relation between the range of the incident proton beam and the number of detected prompt $\gamma$-rays to detect variations

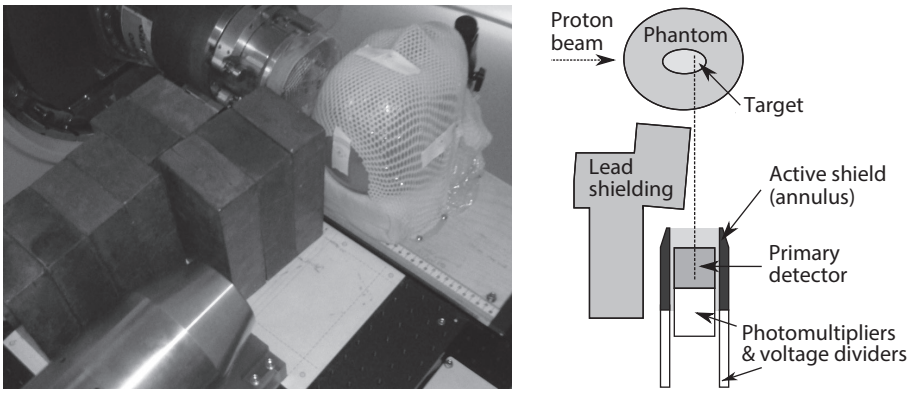



**Figure 5.1** | Photograph and schematic drawing of the experimental setup in the proton therapy gantry. Shown are the prompt $\gamma$-ray detector, lead shielding, and an anthropomorphic head phantom.

of the depth of the distal dose fall-off. Experiments to assess clinical feasibility were performed using different phantoms.



## 5.1. Methods and materials

### 5.1.1. Beam delivery

A passive scattering proton therapy nozzle (IBA, Ion Beam Applications SA, Louvain-la-Neuve, Belgium) in a proton therapy gantry at the Francis H. Burr Proton Therapy Center at Massachusetts General Hospital was used to deliver proton spread-out Bragg peaks for the experiments. The system uses a rotating range modulation wheel and a double scattering system [121]. During a rotation of the modulation wheel, proton Bragg peaks are delivered to sequential layers within the target volume, starting from the deepest layer. Each dose layer is delivered once per rotation. Because the wheel rotates with a frequency of 600 rpm, the entire dose distribution is delivered ten times per second. Protons were accelerated with a 230 MeV isochronous cyclotron (IBA C230) with a radiofrequency of 106.3 MHz, corresponding to a time of 9.4 ns between the accelerated proton bunches.

### 5.1.2. Prompt γ-ray detection and data acquisition

The prompt $\gamma$-ray detection system, as shown in figure 5.1, consisted of a cylindrical cerium-doped lanthanum(III) bromide scintillator with a diameter and length of 76 mm, which was actively shielded by a segmented bismuth gemanate scintillator in the form of an annulus with a thickness of 18 mm. Events in which a coincidence was detected between the primary detector and the active shield were rejected. The

crystals were readout with photomultiplier tubes and the signals were amplified using wideband amplifiers located next to the detector.

The detector signals were then acquired using a custom digital data acquisition system, based on free running analog-to-digitial converters and digital pulse analysis. The signal from the primary detector was sampled 18 times during each radiofrequency period of the cyclotron and the signals from the four segments of the active shield were sampled 4.5 times per radiofrequency period. We synchronized the sampling clock of the analog-to-digitial converters to the cyclotron radiofrequency to facilitate a separation of prompt $\gamma$-rays induced by the primary protons from background radiation through time-of-flight measurement [52, 95, 104].

Real-time synchronization between the range modulation wheel and the data acquisition system was implemented to establish the angle of the wheel for each detected event. Because of count rates differences as a function of the range of the incident proton beam, the $\gamma$ energy calibration and corrections for pile-up and random coincidences were also dependent on the modulation wheel angle. Further details about the data acquisition are described in our previous work [104].



### 5.1.3. Detector positioning and shielding

The prompt $\gamma$-ray detector was positioned to detect $\gamma$-rays emitted from a volume around the distal end of the irradiated target volume, as also depicted in figure 5.1. Lead shielding with a thickness of 15 cm was used to attenuate prompt $\gamma$-rays emitted along most of the entrance path of the proton beam. The shielding was placed at a 5° angle relative to the central axis of the detector to increase the detection sensitivity to $\gamma$-rays emitted downstream of the shielding position. The rationale for this shielding geometry is the fact that the prompt $\gamma$-ray emissions along the entrance path are relatively constant independent of small differences in the range of the beam. By shielding these $\gamma$-rays, the measurements are more sensitive to differences in the range. The shielding also serves to attenuate $\gamma$-rays from the aperture and range compensator.

Additional shielding was placed between the detector and the proton nozzle to shield $\gamma$-rays produced by protons stopping in the base of the snout. Because the scattering system creates uniform proton fields with a diameter of 25 cm [121], most protons are typically stopped in the snout base and result in the emission of secondary $\gamma$-rays. The snout base has a thickness sufficient to stop the primary protons, but the self-attenuation of high energy $\gamma$-rays is limited. We used 10 cm of lead shielding between the nozzle and the detector. The background events from the nozzle and the treatment room were further reduced by the active shielding.

Neutrons resulting from proton interactions with the nozzle and $\gamma$-rays induced by these neutrons remain a significant source of background. However, unlike $\gamma$-rays

induced by the primary protons, this background is relatively constant throughout the cyclotron radiofrequency period, because the spread in neutron velocities results in different arrival times at the detector.

### 5.1.4. Phantom experiments

We performed experiments with two different phantoms. For the first experiment, a water phantom was irradiated with a circular proton field having a diameter of 12 cm. The beam delivered a dose of 30 cGy (RBE) within the spread-out Bragg peak with a range $R_{80}$ at 80 % dose fall-off of 16.0 g cm$^{-2}$ and 6.0 g cm$^{-2}$ of range modulation. The detector was positioned at a depth of 135 mm, which was 25 mm upstream of the end-of-range. This experiment with a large field size and a uniform range was intended as an initial feasibility test.

To investigate a more realistic clinical scenario, a head phantom (SK150, The Phantom Laboratory, Salem, NY, USA) as shown in figure 5.1 was used for the second experiment. The phantom consisted of a human skull surrounded by a soft tissue density plastic. A computed tomography scan was made of the phantom and the XiO treatment planning system (Elekta AB, Stockholm, Sweden) was used to design an aperture and range compensator to irradiate a centrally located 175 cm$^3$ target with a dose of 50 cGy (RBE) . The target had a curved and somewhat irregular distal surface to mimic an actual clinical target volume. The range of the beam was 12.5 g cm$^{-2}$ and the range was modulated by 7.5 g cm$^{-2}$. The aperture had an opening area of 51 cm$^2$. The position of the detector was 20 mm proximal to the end-of-range at the centre of the field and the distance from the beam entrance surface to this position was 84 mm. A beam's eye view of the dose distribution at the detector position is shown in figure 5.2.

To investigate the detectability of differences in the distal dose fall-off, range shifts were introduced by placing slabs of plastic in the beam path between the nozzle and the phantoms. The water equivalent thickness of the range shifters was determined with Bragg peak measurements in a water tank. A dose rate of approximately 10 cGy min$^{-1}$ was used for the experiments. This dose rate was selected to enable real-time storage of the acquired detector pulses, enabling the digital pulse processing algorithms to be retrospectively fine-tuned to the specific conditions during the delivery of scattered proton beams. Each experiment was repeated five times for statistical analysis.

### 5.1.5. Time structure of the proton-induced γ-rays

To analyse and quantify range differences, only the prompt $\gamma$-rays induced by the primary proton beam are of interest. Therefore, a separation is needed between the proton- and neutron-induced events, by considering the arrival time of the

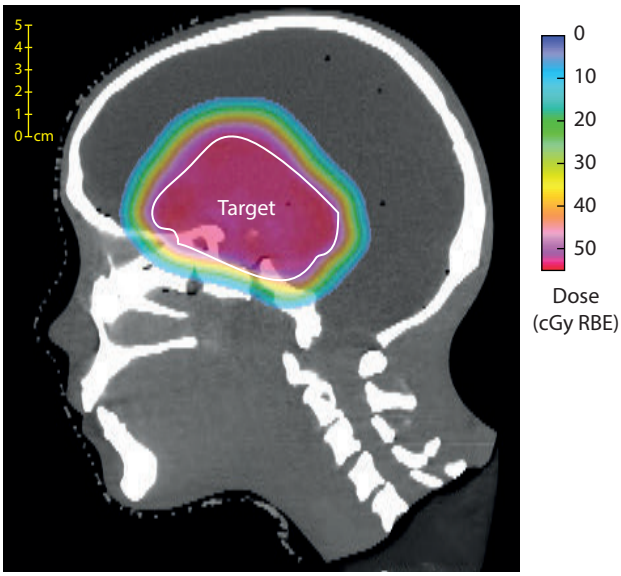



**Figure 5.2** | Sagittal beam's eye view of the computed tomography scan of the head phantom. The planned target volume is outlined and the dose distribution is shown as a colourwash. This slice was located at the centre of the detector.



detected events relative to the cyclotron radiofrequency period. Because the time duration and phase of the prompt $\gamma$-rays induced by the primary protons depend on many different factors, both were determined for each specific measurement. This determination was based on a time histogram of the detected events.

A high ratio of the proton-induced $\gamma$-rays relative to neutron-induced events is desired to accurately determine the time structure of the $\gamma$-rays from primary protons. This ratio is the highest when the most distal layer of dose is delivered. Therefore, we used a time histogram of the events detected during the delivery of this layer. We further increased the ratio by limiting the analysis to events with a detected energy within 100 keV of the 4.44 MeV level of $^{12}C$ [79], because of the high cross section for the excitation of this level in reactions with both carbon and oxygen nuclei [114].

The time histogram contains a peak that comprises the events induced by the bunches of primary protons accelerated by the cyclotron [104]. This proton-induced peak was modelled with a Gaussian function that was fitted to the time histogram, on top of the background that was parametrized with a second degree polynomial. For layers other than the deepest layer, the phase of the peak relative to the radiofrequency period was adjusted considering the flight time of the protons from the range modulation wheel to the target volume.

### 5.1.6. Determination of range shifts

To detect range differences within the phantoms, we used the relation between the range of the proton beam incident on the field-specific hardware and the number of detected proton-induced prompt $\gamma$-rays. Detected events within a $\gamma$ energy window of 2.0 MeV to 7.2 MeV were binned according to the angle of the range modulation wheel. Each bin corresponded to segment of the modulation wheel encompassing a 15° angle, for which the range of the incident proton beam was calculated by subtracting the water equivalent thickness of the modulation wheel and the scatterers from the initial range of the beam transported to the treatment room. In the case of multiple modulation wheel steps within a single bin, the ranges were averaged weighted by the spanned angles.

The events within each bin were further sorted to create histograms of the time of the events relative to the cyclotron radiofrequency period. These histograms were normalized per degree of modulation wheel rotation and a mean filter was applied to smooth small statistical variations. To quantify the number of proton-induced prompt $\gamma$-rays from a time histogram, we estimated the neutron-induced background with a linear interpolation between the time points at $\pm 2.5\sigma$ from the previously determined proton-induced peak described in section §5.1.5, and summed the events above this background between the two time points.



Combining the analysis of the series of time histograms, the relation between the range of the incident beam and the proton-induced $\gamma$-ray counts was obtained. The magnitude of a uniform range shift between two measurements was determined by comparing these relations. The difference in the range of the incident proton beam that is required to match the $\gamma$-ray counts between the measurements, corresponds to the water equivalent range shift. This range shift was determined by finding the value that minimizes the sum of the squared differences between the detected prompt $\gamma$-ray counts. If the range shift is limited to a subset of the treatment field, the relative size of this subset as compared to the field size will determine the magnitude of the difference between the measurements. In a first order approximation, it corresponds to the fraction of the field area affected by the range shift multiplied by the magnitude of the shift.

## 5.2. Results

To show the characteristics of the prompt $\gamma$-ray signal detected during the delivery of scattered beams, histograms are shown in figure 5.3 of the events detected while the deepest dose layer was delivered to both phantoms. The events were binned in dimensions of $\gamma$-ray energy and time within the cyclotron radiofrequency period. The peak in the histograms corresponds to the prompt $\gamma$-rays induced by the primary proton beam [104]. These prompt $\gamma$-rays of interest could be distinguished from

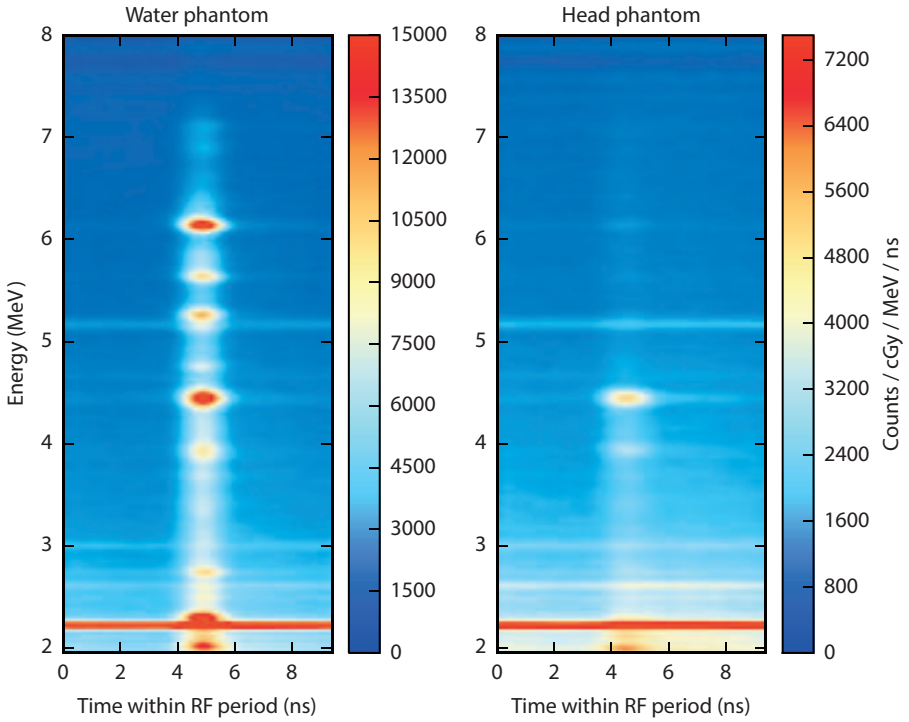



**Figure 5.3** | Histograms of events detected during the delivery of the deepest dose layer to the water phantom and the head phantom. These two-dimensional histograms show the detected events as a function of the $\gamma$-ray energy and the arrival time at the detector relative to the cyclotron radiofrequency (RF) period. The phase of the RF period is arbitrary.

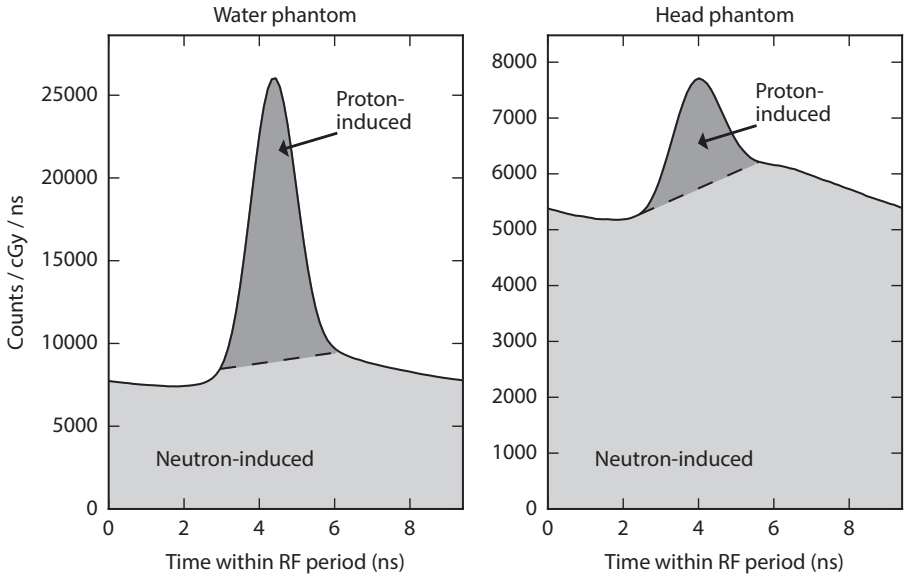



**Figure 5.4** | Time histograms of the events detected during the delivery of the deepest dose layer to the phantoms. The events are sorted according to the arrival time at the detector relative to the cyclotron radiofrequency (RF) period. Shaded areas show the determined separation between proton-induced and neutron-induced events. The phase of the RF period is arbitrary.



the radiation background in dimensions of time and energy. Discrete energy lines were observed from $\gamma$-ray transitions in $^{12}$C, $^{16}$O and other residual nuclei, which are discussed in more detail in our previous work [114]. The magnitude of the different lines is dependent on the elemental composition of the irradiated matter.

Examples of the time histograms that were used to quantify the proton-induced prompt $\gamma$-rays, which are of interest for range verification, are given in figure 5.4. These are histograms of the time relative to the cylotron radiofrequency period of all events within the 2.0 MeV to 7.2 MeV $\gamma$ energy window, which were measured during the delivery of the most distal dose layer. Similar histograms were created for the other dose layers. In the figure, we show the separation of the proton-induced and neutron-induced events following the procedure described in section §5.1.6. Because of the smaller field size, the fraction of proton-induced events relative to the total number of detected events was significantly smaller in the measurement performed with the head phantom. The plastic used in the head phantom also contained a large fraction of carbon. Proton reactions with carbon nuclei result in less prompt $\gamma$-ray emission as compared to reactions with oxygen nuclei [96].

For the verification of the proton range in the water and head phantom, we analysed the series of time histograms acquired during the range modulation cycle

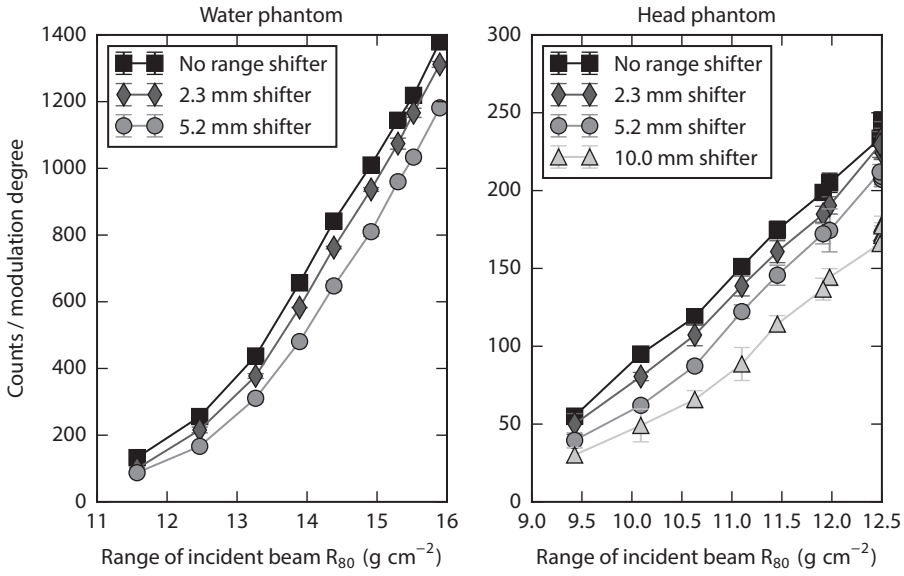



**Figure 5.5** | Detected proton-induced prompt $\gamma$-rays as a function of range of the proton beams incident on the field-specific hardware and phantoms. The range of the incident beam changes during the range modulation cycle. Measurements were performed with the water phantom and the head phantom. Different range shifters were introduced in the beam path, of which the thickness is expressed in water equivalent units. The markers and errors bars are the mean and standard deviation of five measurements.

**Table 5.1** | Detection of range shifts introduced by a uniform range shifter placed in front the irradiated water phantom and head phantom. The range shifts are expressed in water equivalent units. The mean and standard deviation of the detected shifts were determined from five repeat measurements.

| | **Range shifter (mm)** | **Detected range shift (mm)** |
|---|---|---|
| Water phantom | 0.0 | 0.0 ± 0.1 |
| | 2.3 | −2.0 ± 0.1 |
| | 5.2 | −5.4 ± 0.1 |
| Head phantom | 0.0 | 0.0 ± 0.2 |
| | 2.3 | −2.1 ± 0.2 |
| | 5.2 | −4.7 ± 0.2 |
| | 10.0 | −9.9 ± 0.2 |

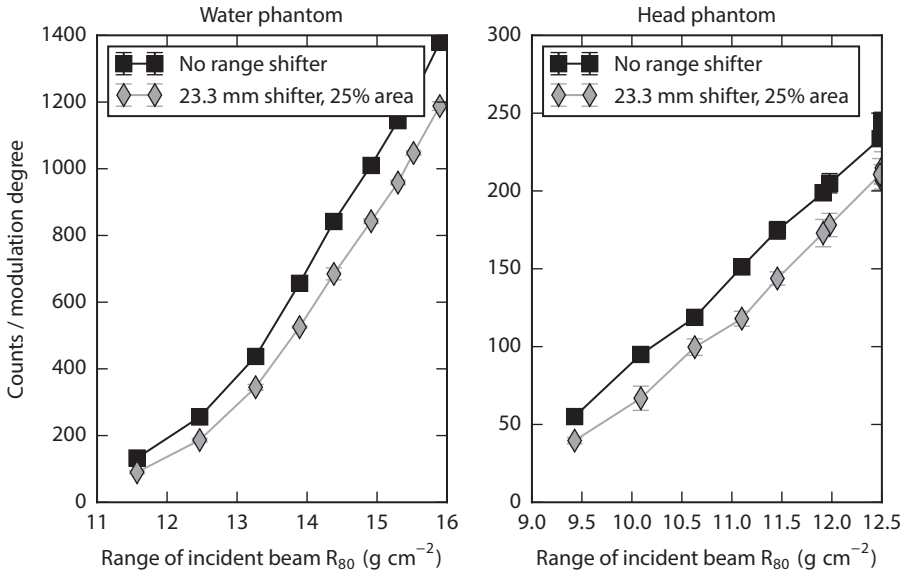



**Figure 5.6** | Detected proton-induced prompt $\gamma$-rays as a function of the modulated range of the incident proton beam. Measurements were performed with the water and head phantom, with and without a 23.3 mm water equivalent range shifter on a quarter of the field area. The markers and errors bars are the mean and standard deviation of five measurements.



of the incident proton beams, which yields the curves depicted in figure 5.5. This figure shows the relation between the range of the incident proton beam and the number of detected proton-induced prompt $\gamma$-rays. Range shifters were introduced in the beam path to investigate the detection of range differences. Clearly, if a range shifter is present, a higher range of the incident proton beam is required to yield the same prompt $\gamma$-ray counts as compared to a measurement without the range shifter. Listed in table 5.1 are the range shifts that were determined by analysing the shift between the curves, using the method given in section §5.1.6. The standard deviation of the detected range shift was on average 0.09 mm for the water phantom and 0.20 mm for the head phantom. The absolute magnitude of the detected shifts also agreed well with the thickness of the range shifters.

In a clinical setting, range differences will often occur within a certain part of the treatment field. In figure 5.6, measurements are shown in which a range shifter of 23.3 mm water equivalent thickness was placed in the path of one quarter of the proton field area. The resulting shifts between the curves were similar to the ones observed with the 5.2 mm uniform range shifter, which are depicted in figure 5.5, showing that the shift is as an approximate measurement of the volume affected by a range undershoot or overshoot.

## 5.3. Discussion and conclusions

We have showed for the first time the detection of prompt $\gamma$-rays for range verification of scattered proton beams. Differences of the end-of-range of the beam were successfully identified using a single prompt $\gamma$-ray detector synchronized to the range modulation wheel. Uniform range shifts were detected with a standard deviation of 0.1 mm to 0.2 mm at dose levels of 30 cGy (RBE) to 50 cGy (RBE). Range shifts that were limited to a part of the treatment field area could also be detected. The detectable magnitude of the range shift in this case was approximately proportional to the ratio between the field area and the area affected by the range shift.

The proposed method can be suitable as a quality assurance tool to ensure consistency of the delivered proton range during the course of a fractionated treatment. A limitation of the a method is the fact that the range differences that are detected cannot be directly attributed to a specific location within the treatment field. This is because complete energy layers are irradiated simultaneously by passive scattering delivery systems. A small range difference within a larger fraction of the treatment field area results in a very similar detected difference as compared to a large range difference within a small fraction of the field. If an inconsistency of the range is detected, further investigation will be needed to identify the cause. First, the setup and alignment of the patient could be verified. Further, the dose distribution could be recalculated based on a new computed tomography scan of the patient to determine if adaptations of the treatment plan and the patient specific hardware are needed to correct for anatomical differences.



In addition to the total proton-induced prompt $\gamma$-ray counts that were used to identify range differences, the spectroscopic information that can be obtained with our system could provide information about changes within the volume that is monitored. For example, a difference in the ratio between the magnitudes of specific discrete prompt $\gamma$-ray emissions, which are shown in figure 5.3, would likely indicate an anatomical change that led to different concentrations of elements such an oxygen, carbon and calcium around the target area. In such a case, follow-up imaging to analyse the anatomical differences could be performed.

The dose rate used for our first experiments was an order of magnitude below clinical dose rates. Detector count rates on the order of $10^5$ counts per second were observed. With the present experimental setup and beam delivery system, the count rate would be on the order of $10^6$ counts per second at a dose rate of $1\,\mathrm{Gy\,min^{-1}}$. At this count rate, one event occurs every 1000 ns on average, while the width of a detector pulse was around 100 ns. Therefore, we expect the system to be compatible with higher dose rates, which we plan to confirm with future experiments.

We conclude that prompt $\gamma$-ray detection for *in vivo* range verification during the delivery of passively scattered proton beams is feasible. The potential for treatment adaptation is naturally limited as compared to pencil-beam scanning, because of

the use of physical patient-specific hardware when scattered beams are delivered. On the other hand, the continuous range modulation makes it relatively straightforward to use a simple detector setup as a tool to verify the range consistency during fractionated treatments.

# 6

# Automated Monte Carlo simulation of proton therapy treatment plans

Simulations of clinical proton radiotherapy treatment plans using general purpose Monte Carlo codes have been proven to be a valuable tool for basic research and clinical studies. They have been used to benchmark dose calculation methods, study radiobiological effects, and to develop new technologies such as *in vivo* range verification methods. Advancements in the availability of computational power have made it feasible to perform such simulations on large sets of patient data, resulting in a need for automated and consistent simulations. A framework called MCAUTO was developed for this purpose. Both passive scattering and pencil-beam scanning delivery are supported. The code handles the data exchange between the treatment planning system and the Monte Carlo system, which requires transfer of plan and imaging information but also translation of institutional procedures, such as output factor definitions. Simulations are performed on a high performance computing infrastructure. The simulation methods were designed to use the full capabilities of Monte Carlo physics models, while also ensuring consistency in the approximations that are common to both pencil-beam and Monte Carlo dose calculations. While some methods need to be tailored to institutional planning systems and procedures, the described procedures show a general roadmap that can be easily translated to other systems.

---

J.M. Verburg, C. Grassberger, S. Dowdell, J. Schuemann, J. Seco, and H. Paganetti.
Submitted for publication.

Monte carlo simulations using general purpose toolkits such as geant4 [64] allow detailed tracking of proton interactions with matter in the energy range used for cancer radiotherapy. The increased availability of computational power has made it possible to perform simulations of complete proton therapy treatment plans with high statistical precision. These simulations have been proven to be a valuable tool for research and development. Applications include improved dose calculation, out-of-field dose estimation, modelling of radiobiology, and the study and detection of nuclear interactions in the patient for *in vivo* beam monitoring.

General purpose simulation codes are considered a 'gold standard', as they aim to provide a complete simulation of all interactions with matter. Therefore, these codes are employed for research studies when a complete model of the complex radiation fields is needed. While not usually considered a direct replacement for clinical algorithms, optimized methods for clinical application are benchmarked against general purpose codes [43].

The Monte Carlo simulation of clinical proton therapy treatment plans is a complex process [122–125]. An accurate description of the treatment plan, patient anatomy and beam delivery system is required. Simulations need to consider the transport of protons through the treatment head in the patient-specific configuration, through field specific hardware such as apertures and range compensators, and the beam delivery to the patient anatomy obtained from computed tomography (ct) images. A unique aspect of proton therapy simulations is the high accuracy that is required for clinically relevant results, mainly the dose delivered and the range of the proton beams. Monte Carlo simulations are based on the fundamental description of individual particle trajectories and interactions. Macroscopic quantities such as the deposited dose follow from the simulation of the interactions of many protons. Clinical calculation methods, on the other hand, normally use an actual ground truth based on experimental data as a reference model. It is needed that the Monte Carlo results closely match a clinical algorithm in such reference conditions, to avoid introducing a systematic bias. The uncertainties that exist in the physics models describing the transport and interaction of the protons can however exceed the maximum acceptable difference [126]. Therefore, taking into account this limitation, simulation methods need to be designed such that the uncertainties in the models and parameters are cancelled out or compensated by empirical adjustments.

In addition to dose calculation, the simulation of proton-nuclear interactions is receiving increasing interest because of the application to *in vivo* range verification. The main range verification methods under development use either coincident $\gamma$-rays from the annihilation of positrons, for which the production of positron emitters in the patient is of relevance, or prompt $\gamma$-rays emitted directly during the decay of excited nuclear reaction products. The required accuracy for these applications may

be well beyond the capabilities of current nuclear reaction models. General Monte Carlo methods can be used with specifically measured and optimized cross sections in such cases.

To routinely use Monte Carlo simulations of clinical proton therapy treatment plans, consistent methods, automation, and integration with the clinical treatment planning system are desired [127, 128]. The purpose of this chapter is to describe the development of a simulation framework to provide fully automatic Monte Carlo simulation of clinical proton therapy treatments using both passive scattering and pencil-beam scanning delivery. We systematically consider the relevant models and uncertainties to provide a consistent set of methods that can be applied to studies involving large sets of patient data.

In section §6.1, we describe the set of Monte Carlo simulation methods that were developed, as well as the various sources of uncertainty in absolute range and dosimetry and how these were dealt with. Then, in section §6.2, we briefly discuss the implementation of these methods in our clinic. We show two examples of clinical applications in section §6.3, in which we used the framework to perform simulations of the delivered dose and the occurrence of specific nuclear reactions in head and neck cancer patients. Finally, section §6.4 concludes the chapter.

## 6.1. Simulation methods

The Monte Carlo simulation methods were developed to fully automate the simulation of proton treatments designed using a clinical treatment planning systems. Simulations to calibrate different models were performed once for the full range of relevant energies, and stored for use in the clinical simulations.



### 6.1.1. Simulation processes

The Monte Carlo simulation of a clinical treatment plan was divided in several processes. Before describing the methods in detail, in this section we first provide an overview of all processes involved.

#### Input data

The simulation starts with a series of processes that generate input data. A schematic overview of these processes is given in figure 6.1. First, data was obtained from the clinical treatment planning system. Both the CT scan and the contours of the patient anatomy were transferred. The CT images were then preprocessed to optimize dose calculation accuracy, as is described later. For all beams used in the treatment plan, we transferred the beam properties such as the isocentre, gantry angle, couch angle and beam weight.

For passively scattered beams, the associated aperture and range compensator were obtained, and the configuration of the treatment head was determined based

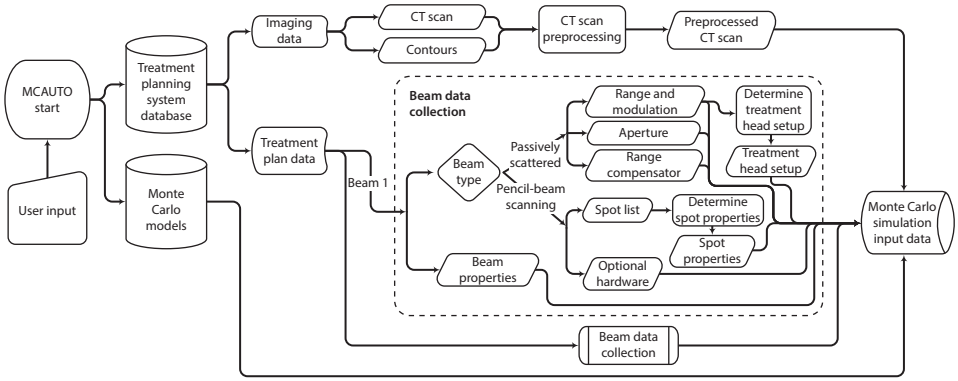



**Figure 6.1** | Processes to obtain the input data for a Monte Carlo simulation.

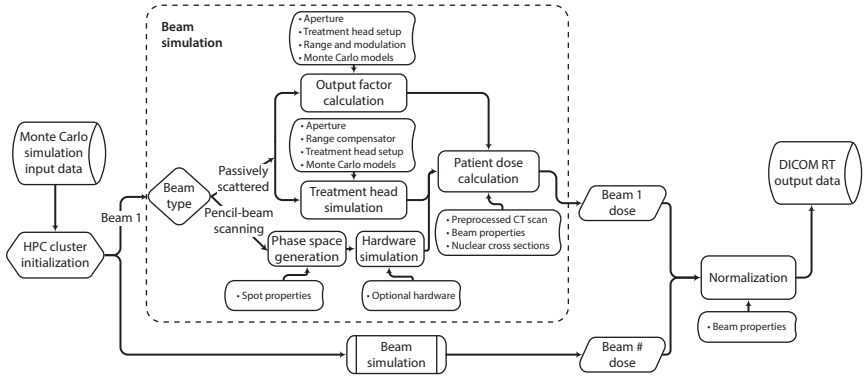



**Figure 6.2** | Processes involved in the Monte Carlo simulation on a high-performance computing infrastructure. Also shown are main parts of the input data that are used in each process.

on the prescribed range, modulation and field size. Also, the applicable geometric models of the treatment head were added to the input data.

In the case of a scanned proton beam, we transferred the list of pencil-beam spots and the specifications of optional hardware such as a range shifter or aperture. The list of pencil-beam spots was combined with the commissioned characteristics of the spot scanning system to determine the physical properties of each spot.

### Beam simulations

The simulation input data were transferred to a high performance computing infrastructure on which the Monte Carlo simulations were performed. Different simulation methods were developed for passively scattered beams and beams delivered using pencil-beam scanning, as shown in figure 6.2. For a passively scattered proton beam, separate Monte Carlo simulations were performed to determine the

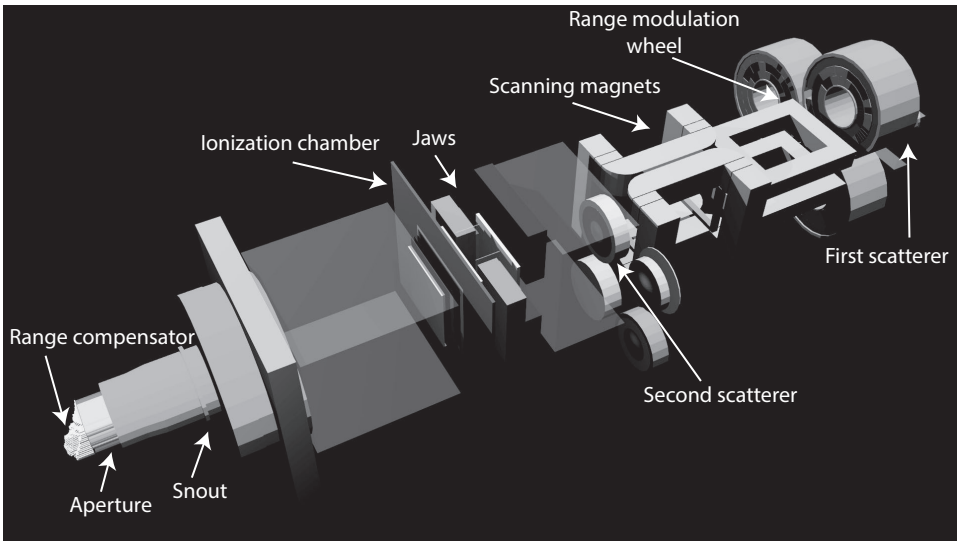



**Figure 6.3** | Model of the universal treatment head at the Francis H. Burr Proton Therapy Center, which can be configured for passively scattered delivery and pencil-beam scanning.

phase space at the end of the treatment head, and to determine the absolute output factor. In the case of pencil-beam scanning, both were determined using analytical models.

On the computer cluster, the dose delivered to the patient anatomy and other quantities of interest were simulated for all beams, and the results were combined and normalized. The phase space information at the end of the treatment head was also saved to facilitate simulations in which multiple patient dose calculations are performed, for example based on different CT images. Finally, we converted the results of the simulations to the standard DICOM RT format.

### 6.1.2. Passively scattered beams

Passively scattered proton beams are formed from a single pencil-beam entering the treatment head, which is broadened, modulated and shaped by a number of devices to create a spread-out Bragg peak (SOBP). The treatments heads at the Francis H. Burr Proton Therapy Center, as shown in figure 6.3, feature a double scattering system, with the most important components being the range modulator, first and second scatterer, snout and the patient specific aperture and range compensator. While the initial pencil-beam is well defined and its properties can easily be parametrized, the phase space of particles entering the patient depends in many complex ways on the configuration of the treatment head. For this reason, the entire treatment head was included in the Monte Carlo simulation.

### Treatment head simulation

The Monte Carlo models of the treatment heads were initially developed by Paganetti et al. [122], and have recently been redesigned to match with even higher accuracy the manufacturer blueprints. All components were modelled based on blueprints provided by Ion Beam Applications SA (IBA, Louvain-la-Neuve, Belgium). The time dependencies of the dynamic range modulation wheel were incorporated, as well as clinical parameters such as the beam current modulation [129]. For the materials in the treatment head, we used the recommended values for the mean excitation energies from ICRU [130] in the determination of the proton stopping powers.

In principle, the phase space upstream of the patient hardware could be pre-simulated for all possible combinations of range, modulation and field size. Given the tens of thousands of possible treatment head configurations, we currently opted not to create such a library, as it would require large computational resources and storage space, and render our simulations inflexible to future improvements. Instead, to reduce the computational cost of this part of the treatment head simulation, a variance reduction method was used. A reduced number of protons was simulated at treatment head entrance, and 8-fold particle splits were performed downstream of the second scatterer and upstream of the aperture. This method is described in further detail by Ramos-Mendez et al. [131].

### Output factor calculation



In our clinic, each beam is prescribed a beam weight, which is defined as the dose that would be delivered by the said beam to water in the high dose region, with the centre of the SOBP located exactly at isocentre. The output factor, defined as the relationship between the beam weight and the number of monitor units to be delivered, depends on the configuration of the scatterers in the treatment head. Also, in the case of small fields, such as those used in radiosurgery, the field size needs to be considered because of effects of the aperture edge. The output factor may be determined experimentally for each treatment field, which was the practice at our facility during the first years of operation. Nowadays, the characteristics of the beam line are very well understood and the output factor is usually derived using an analytic model [121].

For the Monte Carlo simulation, one could use these analytic models and simulate the delivered monitor units through simulation of the ionization chambers [132]. However, except for the specific application of validating clinical output factor calculations, it is generally not desired to introduce the approximations of the clinical output factor determination methods only in the Monte Carlo dose calculation method, because of the resulting bias when comparing dose calculation methods. Instead, we used a Monte Carlo based field-specific output factor calibration, similar to the procedure used to determine the output factor experimentally. As described in more detail in section §6.1.4, this method also has the advantage of providing more

accurate absolute dosimetry, as it cancels out uncertainties in the physics models. A separate Monte Carlo simulation was performed, in which the treatment head was configured as for the delivery of the field to the patient. The beam was impinged on a simulated water phantom, which was positioned according to the definition of the beam weight, i.e. the centre of the SOBP was at isocentre, and the dose averaged over the high dose region was scored. Combined with the simulated number of protons, this calibration establishes the relationship between the simulated dose in the patient and absolute dose. Based on the simulated dose $D_{sim}$, the beam weight $W$, the SOBP dose in the output factor simulation $D_{ref}$, the number of protons simulated $P_{sim}$, and the number of protons simulated for the output factor simulation $P_{ref}$, the absolute dose $D_{abs}$ is given by:

$$D_{abs} = W \frac{D_{sim}}{D_{ref}} \frac{P_{ref}}{P_{sim}}. \tag{6.1}$$

**Proton entrance energy**

A clinical proton beam is normally prescribed a range in water. Proton therapy systems can deliver beams to a water phantom with a sub-millimeter reproducibility of the range [121]. Instead of a secondary parameter such as range, Monte Carlo simulations require the energy and momentum direction of the initial protons to be specified. The simulated range of the beam in water is a result of interactions of the protons with the treatment head, and most importantly the simulated interactions in water. However, the relationship between proton energy and range in water is not known within the accuracy that we aim to achieve for our simulations. Tables such as Janni [133] and ICRU [130] can differ by > 1% for clinical proton energies. These tables have been calculated using the Bethe-Bloch equation; the differences are a result of the uncertainty in parameters used in this equation, such as the mean excitation energy of water. These uncertainties need to be considered to ensure the Monte Carlo simulations reproduce the experimentally determined range of the beam.

To obtain an exact match of the simulated proton range with the requested range, the proton entrance energy was determined such that the Monte Carlo simulation reproduces the requested range in water $R$. The mean excitation energy of water was fixed at 75 eV. We obtained the water-equivalent thickness $R_{th}$ of the combination of scatters, ionization chambers, and the range modulator step used for the longest range. Then, using the proton entrance energy as a function of range in water $e_w(r)$, derived through Monte Carlo simulations, the total water equivalent thickness was used to determine the proton entrance energy $E = e_w(R + R_{th})$.

After the entrance energy has been optimized to reproduce the range, it should be verified that the flatness of the SOBP and the modulation width agree with experimental data, as these can also be affected by uncertainties in the stopping power of the materials in the treatment head [134]. The sensitivity of these features to small

uncertainties in the models is however not as large as compared to the end-of-range, because of the more shallow dose gradients.

#### Energy loss in range compensator

For passively scattered proton therapy, field specific range compensators made of plastics or wax are used to conform the field to the distal edge of the tumour. Therefore, the simulated energy loss of protons in the range compensator also needs to match the experimentally determined values. The range compensators used at our institute are made of poly(methyl methacrylate) (PMMA) and have a proton stopping power of 1.15 relative to water. The material to simulate the range compensator was configured to use the mean excitation energy of PMMA from ICRU [130], and we tuned the mass density to reproduce the desired proton stopping power at 100 MeV.

### 6.1.3. Pencil-beam scanning

The simulation of magnetically scanned proton beams is significantly different as compared to passively scattered beams. Practical clinical systems feature a discrete set of deliverable pencil-beam energies, the properties of which are characterized experimentally at isocentre. Incorporating the geometry and components of the treatment head in the Monte Carlo simulations is not generally necessary [135, 136]. Because of the limited interactions of the pencil-beam with matter before reaching the patient, experimentally determined characteristics of the pencil-beams were well reproduced by analytically modeling the beam phase space. Some facilities do use small passive devices, such as ridge filters, to optimize the energy spectrum of the pencil-beam. In such cases, it may be advantageous to incorporate the treatment head in the simulation model to obtain the most accurate results [137].



#### Spot energy and output

For pencil-beam scanning, experimental Bragg peak measurements of the the deliverable pencil-beams were directly matched to Monte Carlo simulations by determining the parameters that result in the best agreement. The energy of the beams needs to be chosen such that the range in water agrees with high accuracy. In addition, the energy spread is of importance to reproduce the depth-dose curve. The treatment planning system used at our institute defines pencil-beams in absolute units of the number of incident protons, based on Faraday cup measurements. If treatment planning is based on monitor units, these need to be converted to an absolute number of incident protons. Small corrections to the absolute number of protons remain necessary to obtain the best possible agreement with simulations.

#### Spot geometry

Pencil-beam spots at the exit of the treatment head were geometrically modelled as Gaussian and with a Gaussian energy spread. More complex mathematical descriptions can be used if needed, depending of the characteristics of the pencil-beam

scanning system. From the macroscopic properties of the pencil-beam, individual protons were generated for the Monte Carlo simulation by randomly sampling these distributions

The positions of pencil-beam spots are defined relative to isocentre. Knowing the location of the scanning magnets and the position of the treatment head relative to isocentre, they can be transformed to position and momentum direction at treatment head exit. If *SAD* denotes the source-axis-distance of the scanning magnet and $Z_{phsp}$ the distance of the treatment head exit from isocentre, the transformations for position $(x_i, y_i)$ and momentum direction $(u_i, v_i)$ for spot $i$ are

$$\begin{pmatrix} x_i' \\ y_i' \end{pmatrix} = \begin{pmatrix} (SAD_x + Z_{phsp})/SAD_x \\ (SAD_y + Z_{phsp})/SAD_y \end{pmatrix} \cdot \begin{pmatrix} x_i \\ y_i \end{pmatrix} \tag{6.2}$$

$$\begin{pmatrix} u_i \\ v_i \end{pmatrix} = \left\| \begin{pmatrix} x/SAD_x \\ y/SAD_y \end{pmatrix} \right\|. \tag{6.3}$$

#### Hardware simulation

Additional hardware may be used for certain treatments with the pencil-beam scanning modality. For example, an aperture may reduce the lateral penumbra [136] or a range shifter can be used to obtain proton ranges below the minimum proton energy that can be delivered. For these cases, we implemented a Monte Carlo simulation process between the phase space generation and patient dose calculation, in which the proton beam was transported through the relevant hardware. A new phase space downstream of the hardware was then generated for patient dose calculation.

### 6.1.4. Dose calculation

After the phase space upstream of the patient had been determined, the dose calculation in the patient anatomy was performed. The patient was modelled based on the CT scan using a voxelized geometry [138]. For a clinical Monte Carlo simulation, it is important to develop a consistent model of the stopping power of the patient's tissue. The stopping power of tissue in general is subject to uncertainty [24], which leads to an uncertainty in the proton range. In order to facilitate comparisons between Monte Carlo simulated dose distributions and a clinical planning system, the modelled stopping powers need to fully agree.

#### Patient-specific CT calibration

Similar to our clinical protocol, a preprocessing method was applied to the CT images to improve range accuracy. First, any objects not present during the actual treatment, such as the CT scanner couch, were removed from the images. This was done using the outer contour of the patient and immobilization hardware as created by the treatment planner. The CT numbers in the areas outside this contour were replaced by the

CT number of air. Further, to account for systematic differences in the CT numbers due to factors such as patient size, a linear patient specific calibration of the CT numbers was applied. This calibration was also carried out by the treatment planner. A recalibrated CT number $H_{air}$ for air, nominally defined as –1000 Hounsfield units (HU), was determined by taking the average of several sample voxels outside the patient, which were manually selected. For water, which is nominally at 0 HU, $H_{water}$ was derived by subtracting a predefined CT number difference from surrogate tissues that have a consistent x-ray attenuation slightly higher as compared to water, such as the brain, eye, cerebrospinal fluid and muscles. The original CT numbers of the scan $CT_o$ were then renormalized to obtain the recalibrated CT scan $CT_c$:

$$CT_c = (CT_o + 1000)\,\frac{H_{water} - H_{air}}{1000} + H_{air}. \tag{6.4}$$

**Elemental composition and density of tissues**

Pencil-beam dose calculations for proton therapy are based on a relationship between the CT number and the proton stopping power relative to water, obtained using the stoichiometric calibration method [19]. Calibrations are performed specifically for a CT scanner, in which the relative proton stopping powers are normally assumed constant for therapeutic energies.

For the Monte Carlo simulation of particle transport through the patient anatomy, a model of a physical material was created for each value of the CT number. The materials were described in terms of their basic properties: the physical density and elemental composition. The Monte Carlo code then derives the proton stopping power and many other simulation parameters from these material properties. We used the tissue materials according to the model of Schneider et al. [139].

Although the proton stopping power of the tissue materials is not an input parameter of the Monte Carlo code, it remains the most important property in proton therapy. If the stopping power determined by the Monte Carlo code does not match exactly the stopping power for clinical dose calculation, an unbiased comparison of the methods is not possible. We therefore matched the stopping powers by slightly adjusting the mass densities of the materials in the Monte Carlo simulation, which also serves as a way of incorporating CT scanner specific calibrations. The density $\rho_i$ for the Monte Carlo simulation was set as:

$$\rho_i = P_i \frac{\rho_w S_w}{S_i}. \tag{6.5}$$

in which $S_i$ is the mass stopping power of the tissue from the Monte Carlo code, $S_w$ is the mass stopping power of water used by the code, $\rho_w$ is the density of water, and $P_i$ is the proton stopping power relative to water obtained using the stoichiometric calibration.

### Dose to water conversion

Doses in radiotherapy are conventionally reported as the water-equivalent dose, or dose to water, instead of the dose that is delivered to the medium. The reason for this convention is that pencil-beam dose calculation algorithms model human tissue as water with varying density, i.e. stopping power. Therefore it is necessary to convert the physical dose deposition as calculated using the Monte Carlo method to dose to water. Based on previous work [140], we implemented an on-the-fly conversion of dose to medium $D_m$ to dose to water $D_w$, using the following relationship:

$$D_w = D_m \frac{S_w}{S_m}, \tag{6.6}$$

in which $S_w$ is the mass stopping power of the particle (proton or electron) in water, and $S_m$ is the mass stopping power of the medium. The stopping powers were pre-calculated for simulation performance. For neutral particles which do not have a stopping power, we assumed the same conversion factors as for 100 MeV protons. This is a reasonable as their contribution to the dose is small (< 1 %) [141].

### Stopping power uncertainty

The stopping powers of all patient tissue materials assigned to the CT numbers were defined relative to the stopping power of water. Uncertainties in the stopping powers for these materials thus become relative to the stopping power of water, which was accounted for by tuning the initial parameters of the beam. For passively scattered beams, we used the same stopping power for water in the output factor calculation and the stopping powers assigned to the patient anatomy, which means that uncertainties that would lead to differences in absolute dose also cancel out. Similarly, for scanned beams, the number of protons used in the Monte Carlo simulation was adjusted to match the Bragg peaks used for treatment planning. The only quantity that is ultimately affected by small model uncertainties is the number of protons used to deliver unit dose. This error has about a 1 % effect on the rate of nuclear reactions, which is well below the uncertainty in the total nuclear reaction cross sections.

### Systematic uncertainties

While the stopping power is the most important parameter determining the depth-dose distribution of a proton beam, a number of other physical effects need to be accounted for as well. Primary protons are lost to nuclear reactions at a rate of approximately 1 % per centimetre of water equivalent tissue. Although systematic uncertainties in total reactions cross sections that are consistent throughout the proton energy range are cancelled out by the corrections on the absolute number of protons, a small energy-dependent uncertainty remains. A reasonable magnitude for this uncertainty is around 5 %, which leads to in the dose uncertainty up to 1 %.

The multiple Coulomb scattering of protons, which is well described by Molière theory [16], is also subject to a small uncertainty. The combined systematic effects of these uncertainties can be assessed by comparing Monte Carlo simulations of proton Bragg curves to experimental measurements, which are routinely acquired for quality assurance purposed in proton therapy centres.

### 6.1.5. Proton-nuclear interactions

The simulation of proton-nuclear interactions in patient treatment plans is of interest because of the use of secondary prompt and delayed $\gamma$-rays for *in vivo* range verification. While the general nuclear reaction models incorporated in Monte Carlo codes are useful to obtain a complete approximation of radiation from secondary particles, it is oftentimes also needed to simulate the production of prompt $\gamma$-rays or positron emitters from specific nuclear reactions. In this case, the accuracy of the phenomenological nuclear models may be insufficient [96, 102, 142]. Also, a very large number of histories needs to be simulated to obtain a reasonable statistical precision for relatively rare events.

Instead, if the production rate and location of such secondary particles are the quantities of interest, these can be determined from cross section data based on the fluence and energy spectrum of the primary particles, by accumulating for each primary particle track the small possibility of the occurrence of the nuclear reaction. Similar to dose calculation, these quantities can then be determined by simulating only a small subset of the actual protons delivered to the patient. Also, custom cross section data can be used and optimized based on experimental data. The production $p$ of secondaries from a reaction was determined during each simulation step in the volume of interest using:

$$p = \sigma(E)\, L \frac{N_A \rho}{A_r V} \tag{6.7}$$

in which $\sigma(E)$ is the cross section, which is dependent of the proton energy $E$, $L$ is the length of the simulation step, $N_A$ is the Avogadro constant, $\rho$ is the mass density of the target element in the volume, $A_r$ is the relative atomic mass of the target element, and $V$ is the size of the simulation volume.

### 6.1.6. Statistical uncertainties

For routine Monte Carlo simulations of complete treatment plans, it is impractical to simulate as many histories as protons used to deliver the physical dose. Therefore, statistical uncertainties have to be considered. For dose calculation, a number of histories 4-5 orders of magnitude below the number of protons will yield accurate results inside the treatment field. Quantities resulting from rare events, such as dose deposited far outside the treatment field, require a larger number of histories to be simulated. By default, we simulated about $10^8$ protons reaching the patient.

**Table 6.1** | GEANT4 physics builders used for Monte Carlo dose calculation.

| Builder name | Simulated interactions |
|---|---|
| G4EmStandardPhysics_option3 | Electromagnetic interactions, high-precision option. |
| HadronPhysicsQGSP_BIC_HP | Nuclear interactions using binary cascade model, high-precision neutron models below 20 MeV. |
| G4DecayPhysics | Decay of excited residual nuclei. |
| G4IonBinaryCascadePhysics | Nuclear interactions of light ions using binary cascade model. |
| G4HadronElasticPhysicsHP | Elastic scattering, high-precision neutron models below 20 MeV. |
| G4StoppingPhysics | Capture of charged particles at rest. |
| G4RadioactiveDecayPhysics | Decay of radioactive nuclear reaction products. |

## 6.2. Software implementation

Our software implementation MCAUTO automates all simulation processes. The user is only required to provide the identifier of the treatment plan. All parameters are set to appropriate values for a routine calculation, but these may also be customized to change various aspects of the simulation.

### 6.2.1. Treatment planning system interface

Passively scattered proton therapy treatment plans at our institute are created using the XiO (Elekta AB, Stockholm, Sweden) treatment planning system. The patient CT scan, contours, beam parameters and hardware were automatically obtained from the database of this system. For pencil-beam scanning plans, we currently use XiO in combination with the in-house Astroid system [143].

### 6.2.2. TOPAS Monte Carlo application

The Monte Carlo simulations were performed using TOPAS [144], which is a Monte Carlo application designed for proton therapy, based on the established GEANT4 toolkit [64]. The results in this chapter were obtained using TOPAS 1.0 beta 8 and GEANT4 9.6p02. TOPAS includes features tailored to radiotherapy and proton therapy simulations. Models of the proton treatment heads, patient geometry and simulations parameters are specified in a human-readable format.

The GEANT4 physics models were selected based on the applicability and accuracy in the therapeutic proton energy range, as well as validation studies [67]. We used a combination of GEANT4 physics builders listed in table 6.1. Range cuts for all particles were set to 0.05 mm, which means that any particles, such as secondary electrons, with an estimated range below this threshold are not transported, and their energy is deposited locally. The dimension of 0.05 mm was selected as being much smaller than a CT voxel. The dose calculation grid may be configured as desired and is independent of the CT geometry.

### 6.2.3. Cluster computation

Due to the high computational demand of Monte Carlo simulations, MCAUTO was designed to make use of a high-performance computing cluster. Since independent particle histories are simulated, the simulations can be split in many parallel tasks to be performed on separate computer nodes. The clusters at our institute use the Load Sharing Facility (LSF) software to distribute computational tasks over the available resources, for which MCAUTO was designed to provide the necessary inputs. Through dependency conditions on the various computational jobs, all processes are performed in the right sequence. Simulations of a typical treatment plan were completed in about an hour when divided in 1000 parallel calculations.

### 6.2.4. Validation

We compared the principal characteristics of simulated proton beams in water to experimental data. In the case of passively scattered beams, it is impractical to test every possible combination of range, modulation and field size. Similarly to our clinical quality assurance protocol [121], we selected a representative set of 14 fields that span the entire range of delivery options. For pencil-beam scanning, all deliverable energies were simulated. We found the beam ranges agreed with prescription parameters within 1 mm, and the error in absolute dose to the high-dose area was below 0.5 %. The error in the modulation width of passively scattered fields was below 3 mm. In addition to these validations specific to MCAUTO, our Monte Carlo models of the treatment head and the physics models of GEANT4 have been validated through quantitative comparisons with experimental data in both homogeneous and heterogeneous geometry [145].



## 6.3. Applications

In this section, we illustrate two different applications of MCAUTO and we show some typical results of the Monte Carlo simulations. MCAUTO has also recently been applied in studies of site-specific range margins [146], biologically optimized treatment plans [147] and four-dimensional dose calculations for lung cancer patients [148].

### 6.3.1. Dose calculation

An important application of general purpose Monte Carlo codes is to benchmark clinical dose calculation methods [149–151]. Here, we show a comparison between MCAUTO simulations and pencil-beam dose calculations for head and neck cancer patients. The pencil-beam algorithm by Hong et al. [42] is currently used in our clinic. Ten clinical proton treatment plans with a total of 120 proton fields were simulated. These patients were treated with passively scattered proton therapy. We selected this tumour site because of the presence of bony anatomy and interfaces

**Table 6.2** | Dosimetric analysis of head and neck treatment plans (PB: pencil-beam algorithm, MC: Monte Carlo). All doses are reported in Gy (RBE). The statistical uncertainty in these results is below 0.1 Gy (RBE).

| Site & prescription | Proton fields | Dose calculation | CTV $D_{mean}$ | CTV $D_5$ | CTV $D_{95}$ | Spinal cord $D_{max}$ | Brain stem $D_{max}$ | Chiasm $D_{max}$ | L. retina $D_{max}$ | R. retina $D_{max}$ | L. parotid $D_{mean}$ | R. parotid $D_{mean}$ |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Patient 1 | 17 | PB | 69.0 | 72.6 | 64.8 | 10.9 | 2.3 | 50.0 | 46.7 | 55.1 | 18.2 | 20.0 |
| Nasal | | MC | 68.1 | 72.6 | 63.9 | 11.2 | 3.1 | 50.5 | 51.6 | 54.2 | 18.6 | 20.9 |
| 70.0 Gy (RBE) | | MC – PB | −0.9 | 0.0 | −0.9 | 0.3 | 0.8 | 0.5 | 4.9 | −0.9 | 0.4 | 0.9 |
| Patient 2 | 8 | PB | 69.8 | 71.3 | 66.7 | – | 66.1 | 55.9 | 17.4 | 47.0 | – | – |
| Nasal | | MC | 68.6 | 70.7 | 64.7 | | 65.4 | 54.6 | 16.5 | 44.8 | | |
| 70.0 Gy (RBE) | | MC – PB | −1.2 | −0.6 | −2.0 | | −0.7 | −1.3 | −0.9 | −2.2 | | |
| Patient 3 | 14 | PB | 71.9 | 76.8 | 66.4 | 30.1 | 37.2 | 17.7 | 57.0 | 49.3 | 20.4 | 20.9 |
| Nasal | | MC | 70.0 | 75.8 | 64.6 | 30.3 | 40.1 | 12.1 | 56.0 | 49.6 | 20.8 | 21.5 |
| 70.2 Gy (RBE) | | MC – PB | −1.9 | −1.0 | −1.8 | 0.2 | 2.9 | −5.6 | −1.0 | 0.3 | 0.4 | 0.6 |
| Patient 4 | 13 | PB | 70.2 | 75.2 | 63.4 | 24.8 | 65.5 | 53.6 | – | – | 57.5 | 26.4 |
| Nasopharyx | | MC | 69.0 | 75.1 | 62.6 | 23.6 | 65.5 | 51.2 | | | 56.9 | 27.0 |
| 70.0 Gy (RBE) | | MC – PB | −1.2 | -0.1 | -0.8 | −1.2 | 0.0 | −2.4 | | | −0.6 | 0.6 |
| Patient 5 | 6 | PB | 61.4 | 63.2 | 59.2 | – | 18.9 | 13.8 | – | – | 12.5 | 1.3 |
| Nasal | | MC | 60.3 | 62.3 | 57.5 | | 28.9 | 11.7 | | | 13.0 | 1.8 |
| 60.0 Gy (RBE) | | MC – PB | −1.1 | 0.9 | −1.7 | | 10.0 | −2.1 | | | 0.5 | 0.5 |
| Patient 6 | 18 | PB | 66.9 | 71.3 | 56.3 | 32.5 | 67.0 | 39.7 | 4.3 | 4.4 | 23.3 | 22.7 |
| Nasopharyx | | MC | 66.5 | 70.8 | 56.2 | 34.2 | 67.5 | 40.0 | 3.7 | 3.9 | 23.9 | 23.0 |
| 70.0 Gy (RBE) | | MC – PB | −0.4 | −0.5 | −0.1 | 1.7 | 0.5 | 0.3 | −0.6 | −0.5 | 0.6 | 0.3 |
| Patient 7 | 15 | PB | 40.1 | 43.9 | 35.1 | 34.1 | 43.1 | 39.4 | – | – | – | – |
| Nasal | | MC | 39.7 | 43.5 | 34.7 | 34.8 | 43.2 | 39.0 | | | | |
| 40.0 Gy (RBE) | | MC – PB | −0.4 | −0.4 | −0.4 | 0.7 | 0.1 | −0.4 | | | | |
| Patient 8 | 10 | PB | 70.7 | 73.3 | 67.8 | – | 63.1 | 52.1 | – | – | – | – |
| Nasopharyx | | MC | 70.3 | 73.3 | 67.5 | | 62.4 | 52.1 | | | | |
| 70.0 Gy (RBE) | | MC – PB | −0.4 | 0.0 | −0.3 | | −0.7 | 0.0 | | | | |
| Patient 9 | 7 | PB | 70.3 | 72.4 | 65.8 | – | 58.5 | 43.1 | 43.4 | – | – | – |
| Oral cavity | | MC | 69.2 | 71.1 | 64.5 | | 58.0 | 44.6 | 44.3 | | | |
| 70.0 Gy (RBE) | | MC – PB | −1.1 | −1.3 | −1.3 | | −0.5 | 1.5 | 0.9 | | | |
| Patient 10 | 12 | PB | 66.6 | 72.8 | 54.4 | 32.7 | 9.1 | – | – | – | 45.2 | 35.8 |
| Base tongue | | MC | 65.2 | 71.5 | 52.9 | 32.9 | 9.7 | | | | 44.8 | 35.8 |
| 70.0 Gy (RBE) | | MC – PB | −1.4 | −1.3 | −1.5 | 0.2 | 0.6 | | | | −0.4 | 0.0 |

between lower- and higher-density tissue, which means the Monte Carlo method may have an advantage over pencil-beam dose calculations.

The analysis of dose-volume parameters for the treatment plans of all 10 patients is presented in table 6.2. For the primary CTV (clinical target volume), listed are the mean dose ($D_{mean}$) and the maximum dose received by 5 % and 95 % of the volume ($D_5$, $D_{95}$). The maximum dose to the brain stem, spinal cord, retina and chiasm was also determined, as well as the mean dose to the parotid glands.

Overall, we found a good agreement between the pencil-beam and Monte Carlo dose calculations. The differences in the mean dose to the CTV were between 0–2 Gy relative biological effectiveness (RBE). The Monte Carlo simulations showed a lower homogeneity of the dose delivered to the target. The largest differences were observed for beams with bone-tissue interfaces oriented parallel to the beam direction. As an example, depicted in figure 6.4 is an axial slice of the dose calculation for patient 1. In this case, several fields passed through bone-tissue and bone-air interfaces, resulting in hot and cold spots in the target because of multiple Coulomb scattering. The pencil-beam dose calculation algorithm does not reproduce these local dose differences, because it only considers the integrated radiological depth for scatter calculations [42]. In figure 6.5, we show the dose-volume histograms (DVH) for the CTV and critical organs of this patient. Similar differences were seen in the CTV coverage for the other 9 patients.



The differences in the doses to organs-at-risk were more dependent on the specific patient anatomy. In most cases, the maximum dose to the brain stem, spinal cord, retina and chiasm differed by less than 2 Gy (RBE). The mean dose to the parotid glands differed by < 1 Gy (RBE). A number of outliers were however seen, which were mostly due to range degradation of fields stopping near organs-at-risk. For example, the maximum dose to the brain stem of patient 5 was 10 Gy (RBE) higher in the Monte Carlo dose calculation. The range degradation of one of the beams is shown in figure 6.6.

### 6.3.2. *In vivo* proton range verification

Although proton beams deliver dose mainly through electromagnetic interactions, detailed studies of proton-induced nuclear reactions during proton therapy are becoming of increasing importance because of the use of secondary $\gamma$-rays to facilitate *in vivo* verification of the beam delivery to the patient. As described in section §6.1.5, we implemented a method to evaluate the emission of secondary particles from specific reactions based on custom cross section libraries. As an example, shown in figure 6.7 is a comparison of the deposited dose and the emission of 6.13 MeV prompt $\gamma$-rays from the $^{16}O\,(p, p'\gamma_{6.13})\,^{16}O$ reaction at a 90 degree angle with the primary proton beam direction. This nuclear reaction simulation was based on differential

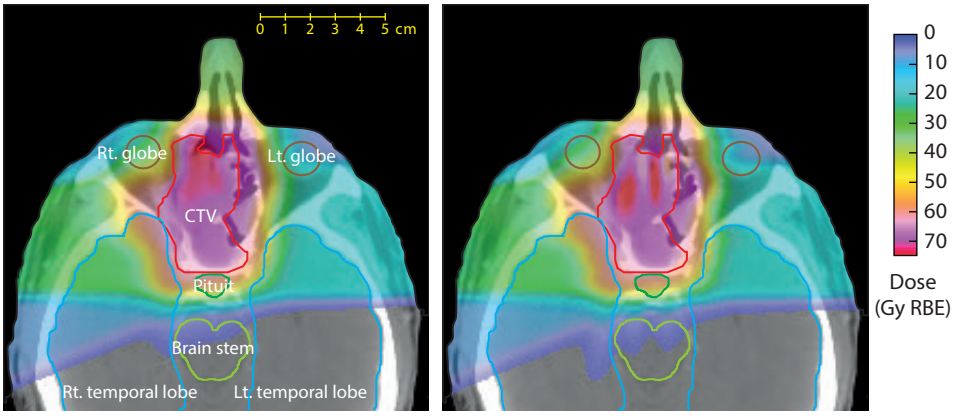



**Figure 6.4** | Axial slice of the pencil-beam and Monte Carlo dose calculation for patient 1. Doses above 1 Gy (RBE) are shown as a colour wash overlaid on the planning CT scan.



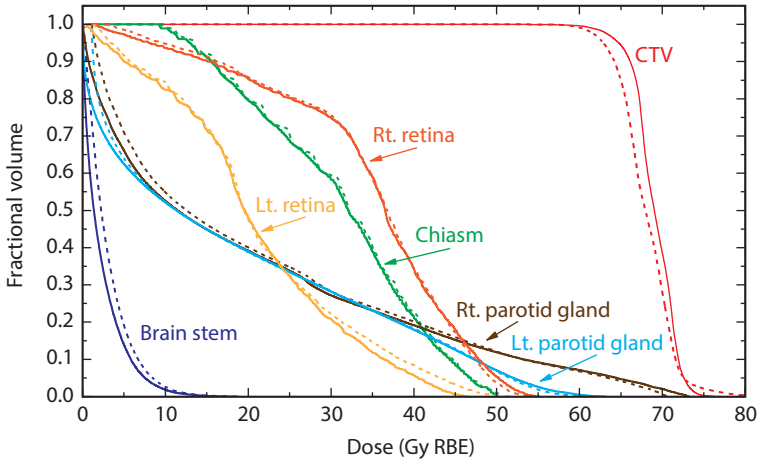



**Figure 6.5** | Dose-volume histograms of the treatment plan for patient 1, comparing the pencil-beam dose calculation (——) with Monte Carlo simulation (- - - -).

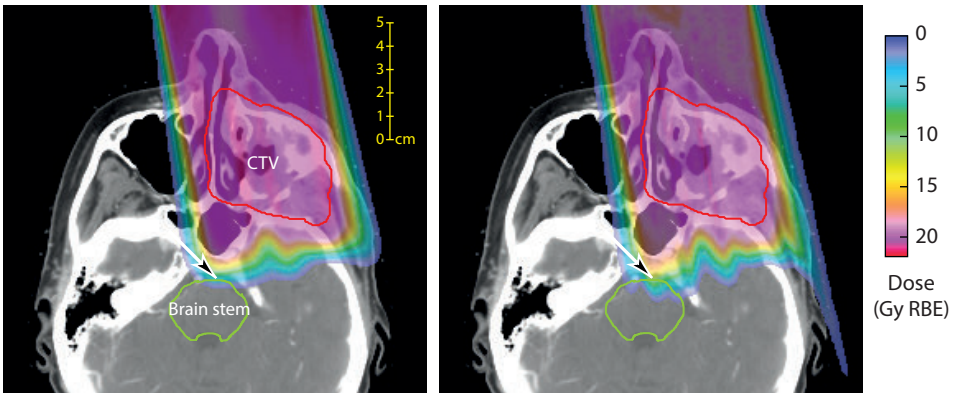



**Figure 6.6** | Axial slice of the pencil-beam and Monte Carlo dose calculation of a single anterior-posterior field for patient 5. Doses above 1 Gy (RBE) are shown as a colour wash overlaid on the planning CT scan.

cross sections which were experimentally determined [114]. These simulations can be further extended to include $\gamma$-ray attenuation and detection.

## 6.4. Discussion and conclusions



We have described a set of methods for routine Monte Carlo simulation of clinical proton therapy treatment plans. To take full advantage of the Monte Carlo method for accuracy of clinical proton therapy simulations, we found it is of prime importance to study carefully how uncertainties in the physics models relate to the main quantities of clinical relevance, such as the range of the proton beam, the dose deposited to water, and the prescription dose. Our methods were designed to ensure consistency in the approximations that are common to both Monte Carlo and clinical pencil-beam dose calculation methods, such as the stopping powers of patient tissues. We matched the clinical prescriptions with high accuracy, even if fundamental parameters in the physics models are not known with the same accuracy.

We developed a framework that automates simulations in a consistent manner. Treatment plans, CT scans and associated data were acquired from the treatment planning system and converted to input files for the TOPAS Monte Carlo application. Simulations were performed on a high performance computing cluster. The results were normalized to absolute units and can be exchanged through DICOM RT.

The primary goal of these methods and their implementation is to enable the routine use of clinical Monte Carlo simulations for proton therapy research. For example, for clinical trials it becomes feasible to retrospectively recalculate dose distributions for large groups of patients, to provide the best possible estimate of the delivered dose to correlate to outcome parameters. Improved dose calculation

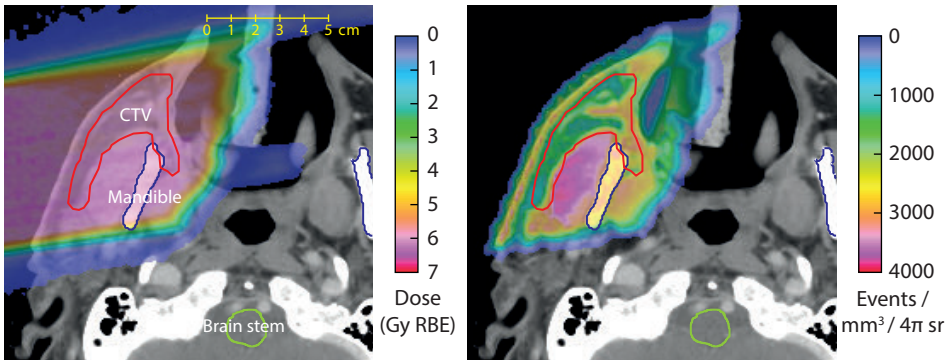



**Figure 6.7** | Monte Carlo dose calculation and simulation of a specific proton-induced nuclear interaction, resulting in 6.13 MeV prompt $\gamma$-ray emission, for a single proton field. The prompt $\gamma$-ray simulation is based on experimentally determined differential cross sections.

methods for clinical use can also be benchmarked against a validated Monte Carlo code. Patient-specific simulations of proton-nuclear interactions enable *in vivo* range verification methods to be evaluated in realistic clinical scenarios.

A comparison was shown between Monte Carlo dose calculations generated using MCAUTO and clinical calculations performed using a pencil-beam algorithm [42]. In this evaluation of 10 treatment plans for head and neck cancer patients, we found good agreement between pencil-beam and Monte Carlo dose calculations. Compared to the clinical dose calculation, the main differences we identified in the Monte Carlo simulation were a lower homogeneity of the dose delivered to the target, and increased range degradation around of the end-of-range of the beams. Both are a result of the more accurate simulation of proton scattering around interfaces between low- and high-density tissue in the beam path, which represents the most important benefit of Monte Carlo proton dose calculation. It is anticipated that clinical Monte Carlo dose calculations can facilitate a reduction of the currently used margins [24].

We have also shown a nuclear reaction simulation in which the general Monte Carlo models were combined with specifically optimized cross sections based on experimental data, which provide an accuracy than cannot be achieved with phenomenological nuclear reaction models. Large scale simulations of proton-nuclear interactions are expected to be of increasing importance as potentially clinically applicable *in vivo* range verification methods are being proposed and studied experimentally. For example, Monte Carlo methods are very suitable to evaluate the robustness of different methods in the presence of inhomogeneities in the patient anatomy and the uncertainties in tissue composition.

Our work also serves as an example of complete integration of the TOPAS Monte Carlo application, clinical treatment planning systems and computing infrastructure. While the modeling of the treatment head geometry and the interface to the treatment planning system is specific to our institute, TOPAS will facilitate development of such integrated infrastructure at other proton therapy centers.

# 7

# Computed tomography metal artifact reduction

We present and validate a CT metal artifact reduction method that is effective for a wide spectrum of clinical implant materials. Projections through low-$Z$ implants such as titanium were corrected using a novel physics correction algorithm that reduces beam hardening errors. In case of high-Z materials such as gold, platinum, and amalgam, projections through the implant were considered missing and regularized iterative reconstruction was performed. Both algorithms were combined if multiple implant materials were present. For comparison, a conventional projection interpolation method was implemented. In a blinded and randomized evaluation, 10 radiation oncologists ranked the quality of patient scans on which the different methods were applied. For scans that included low-$Z$ implants, the proposed method was ranked as best method in 90 % of the reviews. It was ranked superior to the original reconstruction, conventional projection interpolation and regularized limited data reconstruction. All reviewers ranked the method first for scans with high-$Z$ implants. We conclude that effective reduction of CT metal artifacts can be achieved by combining algorithms tailored to specific types of implant materials.

---

Based on published work: J.M. Verburg and J. Seco. CT metal artifact reduction method correcting for beam hardening and missing projections. *Phys. Med. Biol.* **57**, 2803–2818 (2012). Parts of this chapter are based on research performed during the master's thesis project of the author.

Metal artifacts in computed tomography (CT) appear as dark and bright streaks arising from implants such as orthopaedic hardware, dental fillings and coils. In diagnostic scans, they may obscure important information or can be difficult to distinguish from abnormal tissue. Quantitative errors in the CT numbers are an important concern in radiation therapy treatment planning [152].

The underlying causes of these artifacts are limitations in current CT technology. CT image reconstruction in clinical scanners is currently performed using the filtered-back projection (FBP) method, which assumes monochromatic x-ray attenuation and a complete set of projection measurements. If metallic implants are present, these approximations become inaccurate. Beam hardening is one of the issues: lower-energy photons are attenuated more as compared to higher-energy photons, causing the average energy of the beam to increase while passing through the patient. The polychromatic attenuation of implants differs significantly from normal tissue and is not accounted for in single-energy, FBP-based image reconstruction. Missing projections may occur if a highly attenuating object is present in the projection path, as insufficient photons are available for a sensible measurement of the projection. FBP does not provide a mechanism to deal with incomplete projections. The impact of partial volume effects, scatter and motion is also elevated by the presence of implants [153].

Several methods to reduce metal artifacts have been proposed previously [152–167] and can be categorized in two main groups. First, sinogram completion methods assume the projection measurements through metal are missing or incorrect. Accordingly, these projections are disregarded and image reconstruction is performed using limited projections, most commonly by computing surrogate data [152, 154, 155, 157–160, 165, 166], using iterative reconstruction [161, 162, 167], or a combination of both [156, 163]. The second group consists of methods attempting to reduce the mismatch between the monochromatic reconstruction model of FBP and the beam hardening effects that occur in reality [153, 164]. All information in the projections is retained and therefore the reduction of the artifacts depends on the accuracy of the model and the physical limitations of the detectors.

Metal artifact reduction on clinical scans has shown mixed results. Studies of the sinogram completion method indicate a trade-off in different aspects of image quality: while the main metal artifacts are mostly reduced, new artifacts may also be introduced and the removal of projections can lead to a loss of spatial resolution [163, 165]. Beam hardening correction methods are only effective when sufficient photons pass through the implant, and may require prior knowledge on the implant material or x-ray spectrum which is not always available in practise.

The aim of this chapter is to develop a method that provides a consistent reduction of metal artifacts for a broader spectrum of clinical cases. We hypothesize that a method combining corrections for beam hardening and missing projections may

overcome some of the drawbacks of existing methods. A method is proposed that applies either one of the corrections or a combination of both, depending on the implant materials present.

The chapter is organized as follows: section §1 describes the different parts of proposed metal artifact reduction method and their rationale, section §2 presents the implementation on a clinical CT scanner, and section §3 and section §4 show results of a phantom and patient study. Section §5 discusses the performance of the method and concludes the chapter.

## 7.1. Metal artifact reduction method

The main idea of the proposed CT metal artifact reduction method is to preserve information in projection measurements through lower-attenuating implants that do not fully attenuate the x-ray beam and only replace those projections through higher-attenuating implants that are unavailable because of complete attenuation. It consists of the following steps:

1. Identification of implants: determines the locations and types of implants in the patient.

2. Physics correction: reduces beam hardening errors in projections through lower-attenuating implants.

3. Iterative projection replacement: replaces missing projections through higher-attenuating implants.

4. Image reconstruction of the corrected projections.



In the following sections, these algorithms are discussed in detail. We consider an axial parallel-beam acquisition geometry with the projection domain $V$:

$$V = \{(\theta, r, z) | 0 \leq \theta < 2\pi, -1 \leq r \leq 1\}, \tag{7.1}$$

in which $\theta$ is the angle of the x-ray tube and detector, $r$ is the position along the detector, and $z$ is the slice location. The method can however also be generalized to directly reconstruct fan-beam or helical projections.

Reconstructed images are denoted as a vector $\mathbf{f} \in \mathbb{R}^3$ and measured projections as a vector $\mathbf{p}$. The discrete measurements are defined $\theta = 0, \Delta\beta, \ldots, 2\pi$, $r = -1, -1 + \Delta\gamma, \ldots, 1$, with $\Delta\beta$ the angle between the views, and $\Delta\gamma$ the distance between the projections bins. The slice thickness is denoted $\Delta z$. We also define a system matrix $\mathbf{A}$ that consists of the geometrical contributions of the projections to all voxels of the reconstructed volume: $\mathbf{p} = \mathbf{A}\mathbf{f}$.

### 7.1.1. Identification of implants

The locations of the implants in the patient were determined based on a filtered back-projection reconstruction of the original uncorrected scan data. Voxels with a CT number above a certain predefined threshold $H_l$ were considered implants. A second threshold $H_h$ was used to identify higher-attenuating implants, and CT numbers between the two thresholds were defined as lower-attenuating implants. The following masks are defined for the different regions:

$$\mathbf{m} = \begin{cases} 1 & \text{if } \mathbf{f} \geq H_l \\ 0 & \text{otherwise} \end{cases}$$

$$\mathbf{m}_l = \begin{cases} 1 & \text{if } \mathbf{f} \geq H_l \text{ and } \mathbf{f} < H_h \\ 0 & \text{otherwise} \end{cases} \tag{7.2}$$

$$\mathbf{m}_h = \begin{cases} 1 & \text{if } \mathbf{f} \geq H_h \\ 0 & \text{otherwise} \end{cases}.$$

Any regions in $\mathbf{m}_l$ that directly connect to regions in $\mathbf{m}_h$ were removed from the former, since these correspond to artifacts surrounding higher-attenuating implants. The subset of the projections $V$ that passes through the implants is denoted $Y$ and was determined through forward projection; these projections were also subdivided in $Y_l$ and $Y_h$:

$$Y = \begin{cases} 1 & \text{if } \mathbf{Am} > 0 \\ 0 & \text{otherwise} \end{cases}$$

$$Y_l = \begin{cases} 1 & \text{if } \mathbf{Am}_l > 0 \text{ and } \mathbf{Am}_h = 0 \\ 0 & \text{otherwise} \end{cases} \tag{7.3}$$

$$Y_h = \begin{cases} 1 & \text{if } \mathbf{Am}_h > 0 \\ 0 & \text{otherwise} \end{cases}.$$

Figure 7.1 shows an example of metallic implants in image space and the subsets of the projections passing through these implants. Note that all projections through the higher-attenuating implants are assigned to the same category, independent of the amount of metal in the projection path. Although the small subset of the projections that pass through the edges of these implants may not fully attenuate the beam, we did not include the values of these projections in the image reconstruction because they are affected by significant partial volume effects. Finally, we obtained for all projections the path length $\mathbf{t}$, which is the length of the intersection of the beam path with the implants: $\mathbf{t} = \mathbf{Am}$.

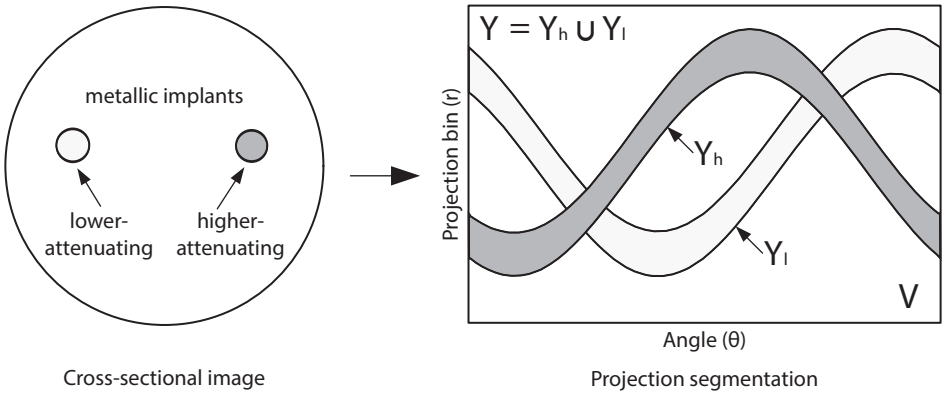



**Figure 7.1** | Metallic objects in a cross-sectional image and the corresponding subsets of the projections.

### 7.1.2. Physics correction

All current CT scanners employ some form of beam hardening correction in the pre-processing of the scan data. These methods are mostly designed to reduce beam hardening effects due to normal human tissue and have limited efficacy in reducing beam hardening introduced by implant materials. Our novel physics correction algorithms adds an additional correction to reduce the errors introduced by the implant. This error is estimated by comparing the projections through the implant with neighbouring projections that do not pass through the hardware.

An overview of this physics correction algorithm is given in figure 7.2. First, the original projection measurements through the implants (a) are interpolated form neighbouring projections (b). The differences between the original and interpolated projections (c) are then correlated with the path length through the implant (d), and used to fit a non-linear correction curve (e). This correction is applied to the original projections, which yields the corrected data (f). The next sections describe the implementation of these different parts of the algorithm.



#### Projection interpolation

The interpolation of projections through metal was performed on radiograph-like images which are reformatted from the projections measurements (as those shown in figure 7.2). These images consist of all projections acquired along the $z$-axis from a single angle. A smooth interpolation was performed in both the $\theta$ and $z$ directions using the Laplace equation, providing the interpolated projections $\mathbf{p}_{intp}$:

$$a\frac{\partial^2 \mathbf{p}_{intp}}{\partial \theta^2} + b\frac{\partial^2 \mathbf{p}_{intp}}{\partial z^2} + c\left(\mathbf{p}_{intp} - \mathbf{p}\right) = 0. \tag{7.4}$$

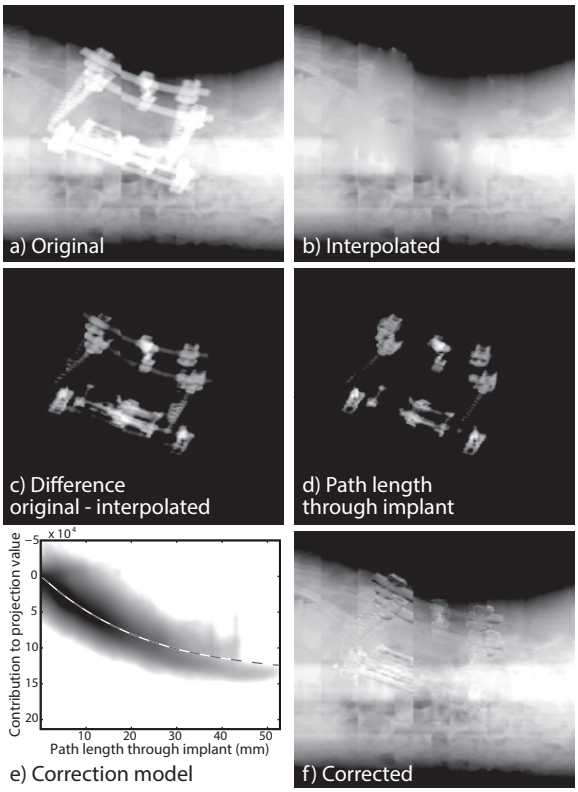



**Figure 7.2** | Schematic overview of the physics correction method. These projection images consist of the acquired CT projections from a single angle.



The parameters were set as $a = \frac{1}{\Delta z}$, $b = \frac{1}{\Delta \gamma}$, $c = 0$ for $(\theta, r) \in Y_l$ and $a = 0$, $b = 0$, $c = 1$ otherwise, i.e. the importance of smooth gradients is inversely related to the distance between the projections. The interpolation of the discrete projection data was implemented using a discrete finite difference approximation:

$$\begin{aligned} &\frac{1}{(2a+2b)}\left[a\mathbf{p}_{intp}(\gamma+\Delta\gamma,z)+a\mathbf{p}_{intp}(\gamma-\Delta\gamma,z)\right]+ \\ &\frac{1}{(2a+2b)}\left[b\mathbf{p}_{intp}(\gamma,z+\Delta z)+b\mathbf{p}_{intp}(\gamma,z-\Delta z)\right]+ \\ &+c\mathbf{p}=0. \end{aligned} \tag{7.5}$$

This linear system was solved directly for $\mathbf{p}_{intp}$. Because all projections through the implants are disregarded, the implant regions in the reconstructed images from the interpolated projections have no physical significance. Therefore, the contribution of this region was removed from $\mathbf{p}_{intp}$. A filtered-back projection reconstruction

was performed to obtain $\mathbf{f}_{intp}$, then the metal regions were extracted:

$$\mathbf{f}_{remove} = \begin{cases} \mathbf{f}_{intp} & if\ \mathbf{m} = 1 \\ 0 & \text{otherwise} \end{cases}, \tag{7.6}$$

and forward projected to be removed from the interpolated projections:

$$\mathbf{p}'_{intp} = \mathbf{p}_{intp} - \mathbf{A}\mathbf{f}_{remove}. \tag{7.7}$$

#### Correction of projections

A first-order model of the beam hardening error was determined by correlating the path length through the implant $\mathbf{t}$ to the difference between the original and interpolated projections, $\mathbf{d} = \mathbf{p}'_{intp} - \mathbf{p}$. This difference $\mathbf{d}$ can be considered an approximation of the contribution of the implant to the projection values. figure 7.2e shows an example of this correlation in the form of a density histogram. The correction curve $c(t)$ was fitted to the correlation using a least-square cubic spline fit which is positive, increasing and concave down. This corresponds to the expected beam hardening error: the contribution of additional implant material decreases with the path length because the harder x-rays are less attenuated. Finally, the correction curve was subtracted from the original projection to yield the corrected data $\mathbf{p}_{MAR}$:

$$\mathbf{p}_{MAR} = \mathbf{p} - c(\mathbf{t}). \tag{7.8}$$

In addition to beam hardening, other first-order errors in the projection values through implants are automatically incorporated in this model. This includes errors due to an increased level of background scatter, which mostly affects the aforementioned projections because of the low signal-to-noise ratio. Partial volume effects are not considered by the algorithm.



### 7.1.3. Iterative projection replacement

To replace missing projections, an iterative algorithm was developed to determine estimated values for those projections. In this process, the estimated projections were optimized to be consistent with the remaining data. This is however an ill-posed problem: many different sets of projections match the available data within measurement uncertainties. Therefore an additional total variation regularization was incorporated, which is a commonly used method in compressed sensing and limited data reconstruction [112, 162, 168].

#### Iterative reconstruction

The missing projections were obtained from images $\mathbf{f}_{opt}$ that were reconstructed according to the following optimization:

$$\underset{\mathbf{f}_{opt}\in\mathbb{R}^3, \mathbf{p}\in V\setminus Y_h}{\operatorname{argmin}} \left( \left| \mathbf{A}\mathbf{f}_{opt} - \mathbf{p} \right|_2^2 + \lambda \mathrm{TV}\left(\mathbf{f}_{opt}\right) \right). \tag{7.9}$$

The first term in this objective function describes the consistency of the image with the projections. The second regularization term penalizes high values of the total variation function:

$$\mathrm{TV}\left(\mathbf{f}_{opt}\right) = \int \left|\nabla \mathbf{f}_{opt}\right| dxdydz, \tag{7.10}$$

which represents the smoothness of the reconstructed images. The rationale for this form of regularization is the observation that human anatomy consists of relatively homogeneous tissue areas. If equally consistent with the measurements, an image with fewer gradients is considered more likely to be close to the actual anatomy.

The optimization was performed using the NESTA accelerated proximal gradient method [112, 169]. In the first iteration, reconstructed images from interpolated projections were used as initial $\mathbf{f}_{opt}$ in order to speed up convergence. Convergence of the iterations was determined by the relative change in the objective function averaged over 10 iterations; a change below $10^{-3}$ was used as stopping condition.

For this method to perform well, it is important to choose a good value for the parameter $\lambda$, which determines the trade-off between measurement consistency and smoothness. A low value for $\lambda$ results in non-physical solutions and unstable results; a small change of the input $\mathbf{p}$ may result in a very different image $\mathbf{f}_{opt}$. A too high value would result in overly smooth gradients in the estimated projections that do not follow the actual gradients in the patient anatomy. We found that a range of intermediate values yielded good results and selected a single value which was used for all scans.

7

### Replacement of projections

The images $\mathbf{f}_{opt}$ were not used directly as final images. Instead, a forward projection was performed from which the missing projection values were obtained:

$$\mathbf{p}_{MAR} = \begin{cases} \mathbf{p} & if\ (\theta, r) \notin Y_h \\ \mathbf{A}\mathbf{f}_{opt} & if\ (\theta, r) \in Y_h \end{cases}. \tag{7.11}$$

This facilitated a final image reconstruction using filtered back-projection, resulting in an appearance similar to the standard images from the scanner. Performance is also improved because the resolution of the full field-of-view iterative reconstruction does not need to match the resolution of the final images, which can be very high if a small area of the image is reconstructed.

### Extended field-of-view

Scans for radiation therapy treatment planning are performed on a flat couchtop, which often extends beyond the field-of-view (FOV) of the CT scanner. This is an issue that needs to be considered in the present algorithm. The objects outside the

FOV result in an inconsistency in the projection values; the subset of the projections passing through the parts of the couch outside the FOV are subject to additional attenuation which is not considered by the image reconstruction.

Normally, such errors average out against all other projections and do not significantly affect the reconstructed images. However, they cannot be ignored when dealing with incomplete projection data. Because the projections through the implants are estimated based on the available incomplete projections, the errors will propagate to those estimated values as the optimizer maximizes the consistency of the reconstructed images with the measurements. The result is an underestimation of density near the implants to compensate for the overestimated attenuation in nearby projections passing through the couchtop outside the FOV.

To correct this issue, the image reconstruction FOV was extended to a virtual FOV that included the entire couchtop. Using the couch position reported by the scanner and a model of the shape of the couch, the projections $C \in V$ passing through the couch outside the scanner FOV were determined. The total set $Y_h \cup C$ was then considered missing. Otherwise, the projections outside the FOV were assumed to pass only through air.

### 7.1.4. Image reconstruction of corrected projections

After all corrections are applied to the projection values, conventional filtered backprojection was used for the final image reconstruction. The images shown in this chapter were reconstructed with a Hamming filter with a frequency cut-off factor of 0.8. This filter was chosen as it approximately matches the standard reconstruction of commercial CT systems. Both the physics correction and iterative projection replacement methods do not yield any information on the parts of the images comprising the metallic implants. These areas were therefore copied from a reconstruction of the original projection data.



## 7.2. Implementation on clinical CT system

In this section we discuss the application of the method to scans obtained on a GE LightSpeed RT 16 CT scanner (GE Healthcare, Waukesha, WI).

### 7.2.1. Projection data

The projection data were already pre-processed by the manufacturer software, which applies several calibration steps. Also, they were rebinned from fan-beam to parallel-beam geometry. The resulting sinograms are identical to those normally used for reconstruction by the scanner. All scans were acquired in axial geometry with a 140 kV x-ray tube voltage. The parameters derived here therefore apply to this configuration, and may need to be adjusted if applied to a different CT scanner or using a different tube voltage.

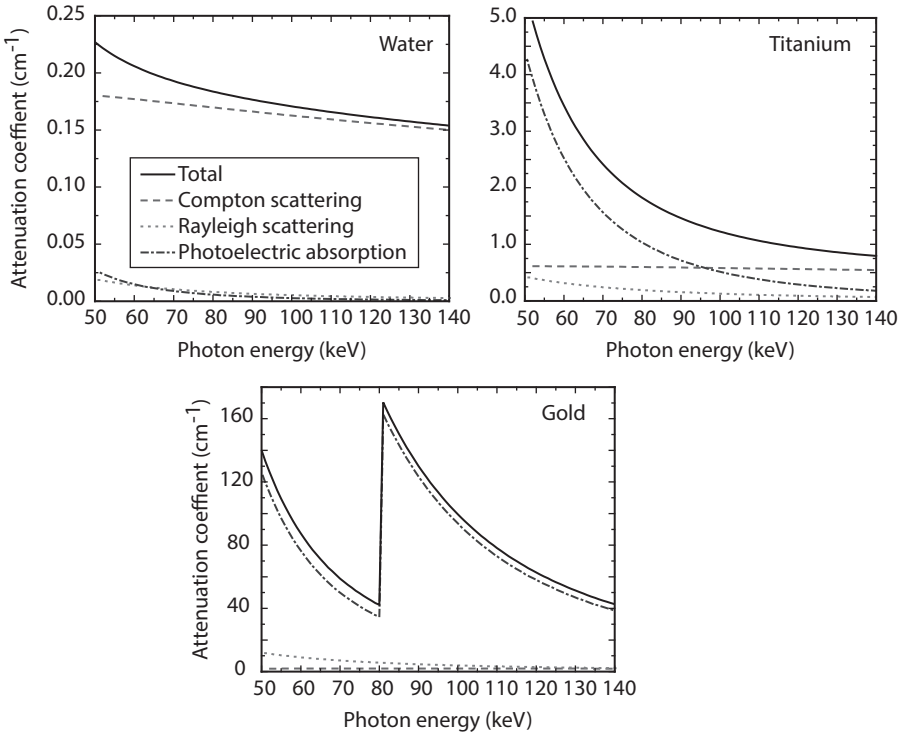



**Figure 7.3 |** Comparison of x-ray attenuation of water, titanium and gold. The attenuation coefficients have been obtained from the NIST XCOM database [170].

7

### 7.2.2. Attenuation of implants

Because the proposed method combines different correction methods, we need to determine which projections can be corrected and which implants are opaque to the x-rays. Figure 7.3 shows the attenuation of water, titanium and gold in the energy range of the x-ray beam (50 keV to 140 keV). The attenuation of the x-rays in this range is described by a combination of Rayleigh scattering, Compton scattering and photoelectric absorption.

Attenuation of water is dominated by Compton scatter. Titanium, which is commonly used for orthopedic implants, shows increased Compton scatter because of its higher electron density, and photoelectric absorption also becomes important. Because photoelectric absorption is more energy-dependent, beam hardening is increased. The total attenuation of titanium is about an order of magnitude higher than water. In case of high-$Z$ materials such as platinum, gold and amalgam, photoelectric absorption becomes the main effect, which has approximately a $Z^4$ dependence. The

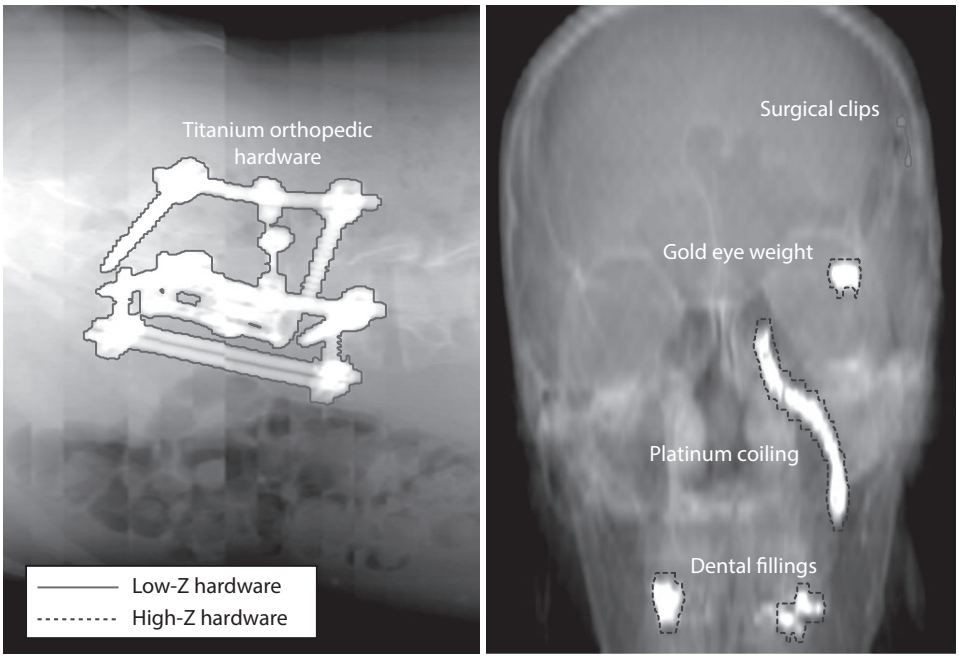



**Figure 7.4** | Projection images of two patients with various implants, with contours showing the projections determined as passing through low-Z and high-Z hardware.

attenuation of these materials is several orders of magnitude above the attenuation of water.

Inspection of the projection data of several patient and phantom scans showed clear contrast in projections through titanium hardware: projections that intersect a larger part of the implants have higher projection values as compared to projections that pass through smaller parts of the hardware. On the contrary, little contrast was visible in the projections through platinum and gold implants. Therefore, it is expected that physics correction is effective for low-$Z$ implants, and iterative projection replacement would be the appropriate method for high-$Z$ implants.



### 7.2.3. Projection segmentation

The segmentation thresholds have to be chosen to divide the projection space between normal tissue and implants, and to identify whether implants belong to the low-$Z$ or high-$Z$ category. These thresholds were determined by analysing projection images of different patient scans, which are not affected by image reconstruction artifacts. We found that $H_l$ = 2500 HU and $H_h$ = 8000 HU provided accurate results. Figure 7.4 shows projection images of two patients and the segmentation determined using these thresholds.

## 7.3. Phantom study

As an initial assessment of the proposed method, it was applied to two scans of a Gammex 467 tissue characterization phantom (Gammex Inc., Middleton, WI): one with two low-$Z$ inserts (titanium) and one with three high-$Z$ inserts (Wood's metal). A scan without metal inserts was obtained to serve as ground truth. figure 7.5(a) shows the arrangement of the different inserts in this phantom.

In order to verify the previously discussed results regarding the attenuation of implants, the physics correction algorithm and the iterative projection replacement method were applied to both the scans. We also show the conventional approach of only applying the projection interpolation as described in section §7.1.2.1.

Figure 7.5b and figure 7.5c depict slices of the phantom scans reconstructed from the original and corrected projection data. In case of low-$Z$ inserts, the physics correction method was successful in reducing the main artifact between the inserts. Some artifacts between the inserts and nearby bone remained, because the method does not take into account these higher-order beam hardening effects. Projection interpolation clearly performed worse as compared to the physics correction. Although the main artifact is somewhat reduced, many new artifacts appear throughout the image. This again indicates the projections through the implant still contain potentially usable information; the original projections were more accurate as compared to the interpolated ones. The iterative projection replacement method yielded excellent reduction of artifacts, even though the information in the projections through the inserts is not used. This result is however biased by the simple structure of the phantom, which enables accurate reconstruction from limited data.

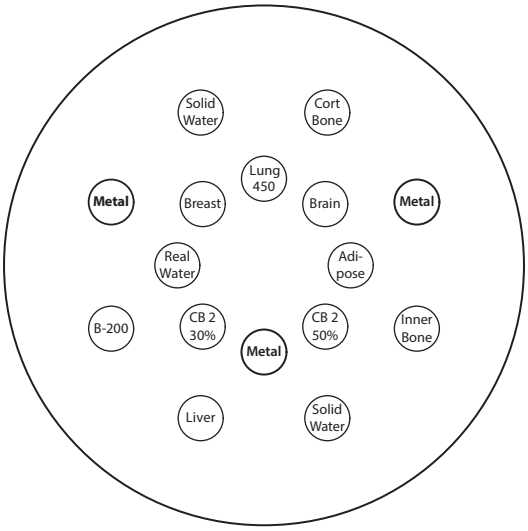
In case of high-$Z$ hardware (figure 7.5c), the iterative projection replacement method was able to eliminate almost all streak artifacts. The results are also superior as compared to projection interpolation, which leaves many remaining artifacts. Physics correction has little effect because of the near complete attenuation of the x-rays. These results confirm the rationale of applying physics correction to low-$Z$ implants and iterative projection replacement to high-$Z$ implants.



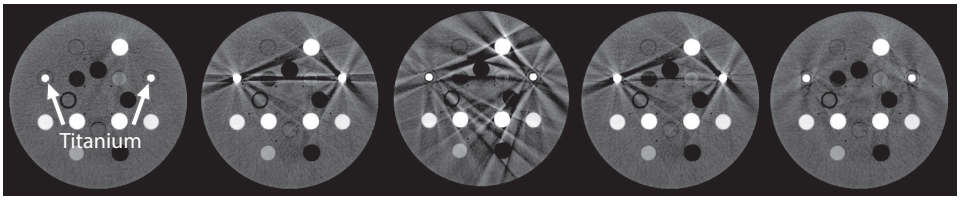
## 7.4. Patient study

The proposed artifact reduction method was applied in retrospect to four patient CT scans, which were obtained for radiotherapy treatment plans. This study was performed under an IRB-approved protocol. Similarly to the phantom study, we also compared two alternative metal artifact reduction methods: projection interpolation, and iterative projection replacement instead of physics correction in case of low-$Z$ implants. The types of implants and the imaged anatomy are listed in table 7.1.

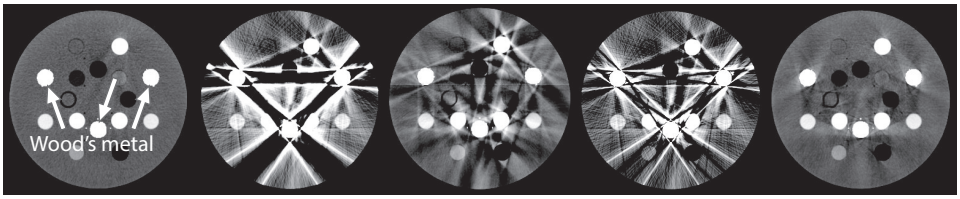
A group of 10 radiation oncologists ranked the image quality of the original and artifact reduced versions. They were blinded to the artifact reduction method and the images were shown in random order. For this comparison, four representative slices




(a) Tissue-equivalent and metal inserts in the phantom




(b) Low-$Z$: Two titanium inserts.




(c) High-$Z$: Three Wood's metal inserts

**Figure 7.5 |** Original and corrected images of tissue characterization phantom scans. Window level 40 HU, width 400 HU.

**Table 7.1** | List of patient scans used to evaluate the proposed metal artifact reduction method.

| | Implant material | Body site | Voltage (kV) | Current (mA) | Rotation (s) | Slice thickness (mm) |
|---|---|---|---|---|---|---|
| *Patient 1* | Titanium | Lumbar spine | 140 | 565 | 2.0 | 2.50 |
| *Patient 2* | Titanium and dental fillings | Cervical spine | 140 | 590 | 1.0 | 2.50 |
| *Patient 3* | Platinum | Base of skull | 140 | 600 | 2.0 | 1.25 |
| *Patient 4* | Gold | Brain | 140 | 600 | 1.0 | 1.25 |

of each patient scan were selected. To determine whether a significant difference existed between the methods, Friedman analysis was used to test the image quality ranks against a null hypothesis of random ranks. We then tested the differences in the mean image quality ranks for significance by post-hoc pairwise comparisons using the Wilcoxon-Nemenyi-McDonald-Thompson method; a single-step maxT procedure was used to adjust the $p$-values for multiple comparisons [171].

### 7.4.1. Patient 1: Low-Z titanium implant

Figure 7.6 shows a slice of the reconstructed images of the first patient with titanium orthopedic hardware, using the different metal artifact reduction methods. Projection interpolation resulted in an unsatisfactory result with new artifacts throughout the image. Iterative projection replacement seemed to perform somewhat better in regions far away from the implants, but still was not able to reconstruct accurately the areas surrounding the implant. The latter part of the image is mostly supported by projections through the implants, which are disregarded by these methods. The proposed method was able to reduce the artifacts around the implants without introducing new artifacts in other areas.



### 7.4.2. Patient 2: Low-Z titanium implant and high-Z dental fillings

The second case, as shown in figure 7.7, is a patient who has titanium hardware in the cervical spine and also has dental fillings. Therefore, the proposed method applied both physics correction for the low-$Z$ implant and iterative projection replacement for the high-$Z$ implant.

This case again shows the strength of the proposed method in preserving information in projections through the low-Z implant. The iterative projection replacement method applied to both implants reduced the artifacts around the high-Z implants but caused a significant loss of detail in the area around the low-Z implant. The proposed combination of algorithms yielded similar reduction of the high-Z artifacts, while reducing streaks between both artifacts and also retaining the detail around the low-Z implant.

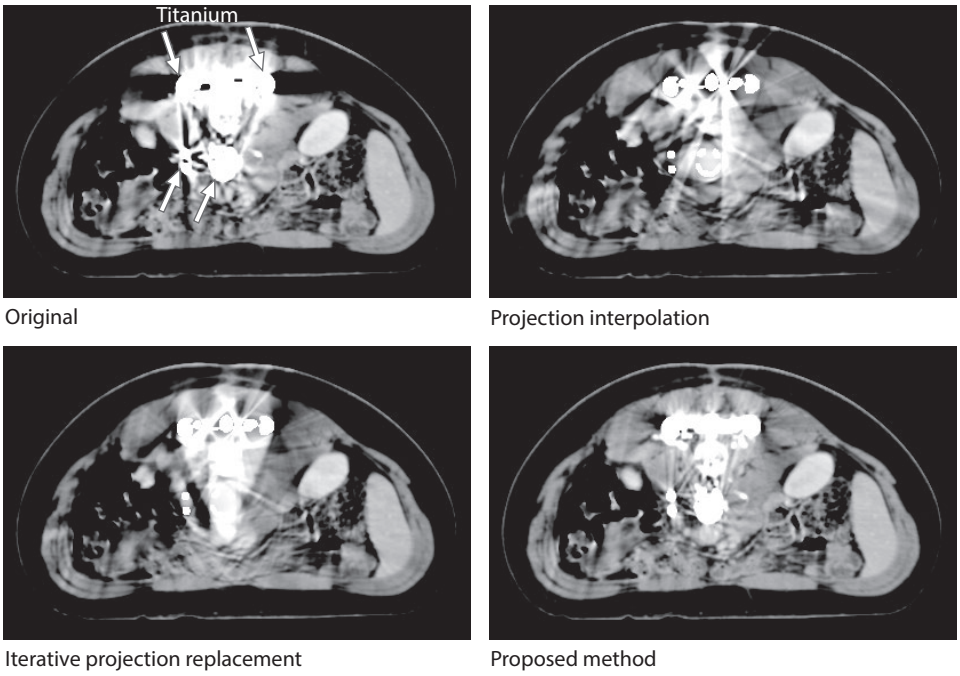



**Figure 7.6** | Titanium hardware in lumbar spine. Window level 40 HU, width 400 HU.

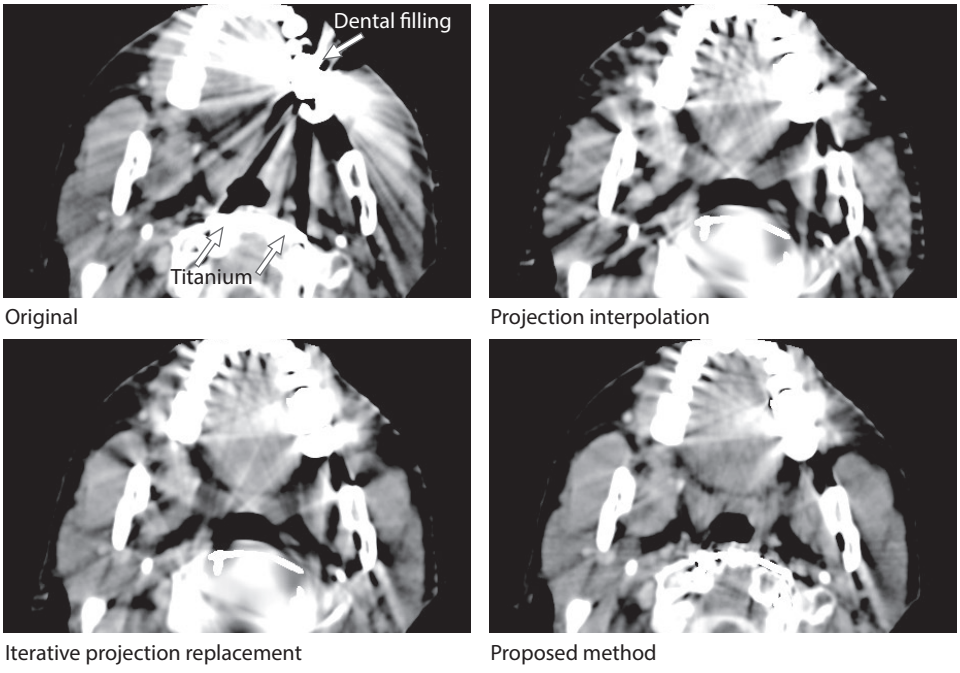



**Figure 7.7** | Titanium and dental fillings in skull base. Window level 40 HU, width 400 HU.

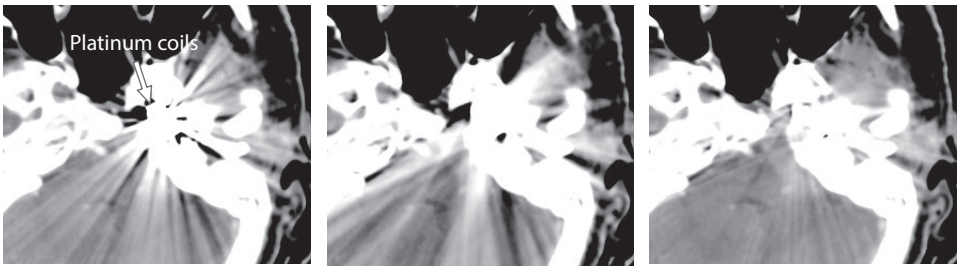



**Figure 7.8** | Platinum coiling in skull base. Window level 40 HU, width 400 HU.

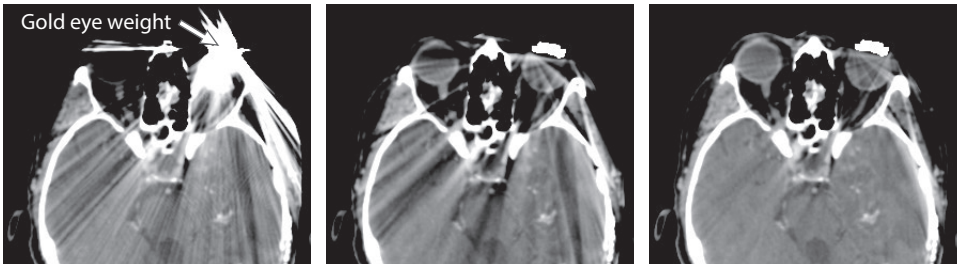



**Figure 7.9** | Gold eye weight. Window level 40 HU, width 400 HU.

### 7.4.3. Patient 3: High-Z platinum coiling

Patient 3 (figure 7.8) has only a high-Z implant. Because these implants are nearly opaque to the CT x-rays, it is reasonable to expect any sensible projection completion algorithm to somewhat reduce artifacts by replacing the invalid projections with estimated values. The proposed method appears to outperform projection completion, producing images with less remaining streak artifacts.

### 7.4.4. Patient 4: High-Z gold eye weight

The scan of patient 4, given in figure 7.9, included a gold eye weight which resulted in strong artifacts. Both projection interpolation and the proposed method reduced the main artifacts around the implant. Again, the proposed method resulted in images mostly free of streak artifacts, while some of these artifacts remained after projection interpolation.

### 7.4.5. Survey

Finally, we show in table 7.2 the results of the blinded survey in which radiation oncologists reviewed the image quality of the patient scans. Because the type of implant material has been shown to have a large impact on the artifact reduction, we analyzed low-$Z$ implants and high-$Z$ implants separately.

**Table 7.2** | Unique ranks assigned by the reviewers to the image quality of patient scans on which the different metal artifact reduction methods were applied. Rank 1 indicates best image quality.

| | *Reviewer* A | B | C | D | E | F | G | H | I | J | A | B | C | D | E | F | G | H | I | J | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| **Including low-Z implant** | *Patient 1* | | | | | | | | | | *Patient 2* | | | | | | | | | | *Mean* |
| *Original* | 2 | 4 | 4 | 4 | 1 | 2 | 2 | 1 | 4 | 3 | 2 | 2 | 3 | 2 | 2 | 2 | 2 | 4 | 3 | 4 | 2.65 |
| *Projection interpolation* | 3 | 2 | 3 | 3 | 4 | 4 | 3 | 4 | 3 | 2 | 4 | 4 | 4 | 4 | 4 | 4 | 4 | 3 | 4 | 3 | 3.45 |
| *Iterative proj. replacement* | 4 | 3 | 2 | 2 | 3 | 3 | 4 | 3 | 2 | 4 | 3 | 3 | 2 | 3 | 3 | 3 | 3 | 2 | 2 | 2 | 2.80 |
| *Proposed method* | 1 | 1 | 1 | 1 | 2 | 1 | 1 | 2 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1.10 |
| | Significance: Overall difference between methods $p < 0.0001$, rank of proposed method is higher as compared to original ($p = 0.0008$), projection interpolation ($p < 0.0001$), iterative projection replacement ($p = 0.0002$). | | | | | | | | | | | | | | | | | | | | |
| **High-Z implant only** | *Patient 3* | | | | | | | | | | *Patient 4* | | | | | | | | | | *Mean* |
| *Original* | 3 | 3 | 3 | 3 | 3 | 3 | 3 | 3 | 3 | 2 | 3 | 3 | 3 | 3 | 3 | 3 | 3 | 3 | 3 | 3 | 3.00 |
| *Projection interpolation* | 2 | 2 | 2 | 2 | 2 | 2 | 2 | 2 | 2 | 2 | 2 | 2 | 2 | 2 | 2 | 2 | 2 | 2 | 2 | 2 | 2.00 |
| *Proposed method* | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1.00 |
| | Significance: Overall difference between methods $p < 0.0001$, rank of proposed method is higher as compared to original ($p < 0.0001$), projection interpolation ($p = 0.004$). | | | | | | | | | | | | | | | | | | | | |

The reviewers were unanimous in their opinions about the image quality of the scans with high-$Z$ implants. All reviewers ranked the proposed method method as performing best, projection interpolation as second and the original FBP image as last. The better rank of the proposed method as compared to the original scan ($p < 0.0001$) and projection interpolation ($p = 0.004$) was highly significant.

In case of low-$Z$ implants, the proposed method was ranked first in 90 % of the cases. The other ranks however were more varied among the reviewers. This can be explained by the trade-off which is involved in applying projection interpolation or iterative projection replacement on scans with low-$Z$ implants. Although certain metal artifacts are reduced, spatial resolution decreases because projection measurements are removed which still hold usable information. The superior performance of the proposed method over the original scan ($p = 0.0008$) and projection interpolation ($p < 0.0001$) was again highly significant.

## 7.5. Discussion and conclusions

In this study, we demonstrated that significant reduction of metal artifacts in clinical CT can be achieved if the composition of the implant material is appreciated. We show it is essential to differentiate between low-$Z$ and high-$Z$ implants.

High-*Z* implants cause a complete or almost complete x-ray attenuation, resulting in incomplete projection data. We used an iterative projection replacement algorithm to estimate the missing projections. This algorithm uses the interrelations between projections and prior knowledge of the general structure of human anatomy in order to obtain sensible images from the incomplete data. A crucial part is the correction for other errors in the projection data such as the attenuation of the couch outside the FOV of the scanner, which prevents such errors form propagating to the estimated projections. A successful reduction of artifacts due to dental fillings, platinum coils and a gold eye weight was shown. The corrected images were free of major streak artifacts.

Low-*Z* implants such as titanium do not fully block the x-rays but result in strong beam hardening, which was corrected for in first order by the physics correction algorithm. Artifacts caused by titanium orthopedic hardware in the spine were significantly reduced by this method, which also preserved full anatomic detail of the area surrounding the implant. If higher-density implants are also present, we propose to use iterative projection replacement after physics correction to account for the missing projections. This combination is for example seen in head and neck cancer patients with spinal hardware and dental fillings.

The main advantage of the proposed method as compared to earlier work is the consistent improvement in image quality. Many previous methods did reduce metal artifacts but also removed detail from the image around the implants [163]. This is because all projections through implants were removed, even if they were not missing because of complete x-ray attenuation. Our method is also fully compatible with current CT technology and calibration methods.



These new methods and findings are important to improve diagnostic scans of patients with metallic implants and to reduce errors in CT-based radiotherapy treatment planning. Treatment plans for proton or heavy ion radiotherapy may particularly benefit from these improvements, as the finite range of the radiation beam requires a high quantitative accuracy of the CT numbers [172].

Some limitations of this study need to be acknowledged. Our physics correction algorithm is limited to first-order errors. Although the energy-integrating CT detectors limit these corrections in general, model-based iterative image reconstruction methods with higher-order models for beam hardening or scatter have the potential to further improve artifact reduction. Dual-energy CT technology can also provide additional beam hardening correction. Secondly, the scans that were used in this study were performed for radiotherapy treatment planning and thus obtained with a higher x-ray current as compared to diagnostic CT. For low-dose scans, the addition of a noise reduction technique might be beneficiary. Last, a few remarks on the computational complexity of our method are in order. The physics correction algorithm is fast and requires limited resources; the replacement of missing projections

however is an iterative method which requires greater computational time. This delay may be problematic in diagnostic CT but is likely to be resolved in the near future as faster computer hardware becomes available.

To summarize, we have shown that effective reduction of CT metal artifacts can be achieved by combining algorithms specific to low-Z or high-Z implants. Spatial resolution can be maintained by preserving projection information through low-Z implants, which makes it a viable approach for clinical use.

# 8

# Proton dose calculation for chordoma patients with titanium implants

We investigated dosimetric errors in proton therapy treatment planning due to titanium implants, and determined how these affect post-operative passively scattered proton therapy for chordoma patients with orthopaedic hardware. Artifacts in the computed tomography (CT) scan cause errors in the proton stopping powers used for dose calculation. Also, current clinical dose calculation algorithms are not designed to consider the effects of implants that are very different from human tissue. We first evaluated the impact of these two main issues. Dose errors introduced by metal artifacts were studied using phantoms with and without titanium inserts, and patient scans on which a metal artifact reduction method was applied. Pencil-beam dose calculations were compared to models of nuclear interactions in titanium and Monte Carlo simulations. Then, to assess the overall impact on treatment plans for chordoma, we compared the original clinical treatment plans to recalculated dose distributions employing both metal artifact reduction and Monte Carlo methods. Dose recalculations of clinical proton fields showed that metal artifacts cause range errors up to 6 mm distal to regions affected by CT artifacts. Monte Carlo simulations revealed dose differences > 10 % in the high-dose area, and range differences up to 10 mm. Since these errors are mostly local in nature, the large number of fields limits the impact on target coverage in the chordoma treatment plans to a small decrease of dose homogeneity.

---

Based on published work: J.M. Verburg and J. Seco. Dosimetric accuracy of proton therapy for chordoma patients with titanium implants. *Med. Phys.* **40**, 071727 (2013).

Proton radiotherapy is often the modality of choice to deliver post-operative radiotherapy to patients with head and neck and spinal tumours that require high doses for tumour control, such as chordoma. The required doses often cannot be delivered using conventional radiotherapy due to the dose limits on the spinal cord or brain stem.

These patients frequently have titanium orthopaedic hardware near the tumour site, which is needed to stabilize the vertebral column after surgery. The presence of implant materials can affect the accuracy of proton therapy treatment plans due to metal artifacts in the computed tomography (CT) scan [173, 174] and the approximations of pencil-beam dose calculation methods [34].

A recent clinical study of chordoma patients who received proton therapy, showed a significant association between the presence of titanium-based surgical stabilization and reduced tumour control [175]. Another study at our institution also reported a trend of local recurrence [176]. While many factors may be responsible, it warrants further investigation of dosimetric errors caused by this type of implant. Also, it is important to quantify the impact on clinical treatment plans, to establish if changes to these treatments are necessary and because concerns over the dosimetric errors have led to some institutions to decide not to treat these patients with proton therapy.

First, the impact of artifacts in the computed tomography (CT) scan for treatment planning is studied. Titanium causes hardening of the CT x-rays, which typically results in streak artifacts in the reconstructed CT images. CT numbers are used to determine the proton stopping power of tissues in the beam path; artifacts can therefore result in errors in the calculated proton range. In clinical treatment plans, proton stopping powers may be manually changed to attempt to reduce these errors [175]. The accuracy of such corrections is however uncertain, as it is difficult to manually determine anatomy and tissue density in images obscured by the artifacts.

Secondly, we investigate the accuracy of clinical pencil-beam dose calculations. Since these methods are based on water as a reference medium [42], they may not accurately simulate beam transport through titanium, which has characteristics very different from water and normal tissues. In particular, elastic multiple Coulomb interactions with the nuclei and non-elastic nuclear reactions are specific to each element. Multiple Coulomb scattering is relevant as it results in a local widening of the proton beam. Non-elastic interactions cause a loss of the primary proton fluence, which affects the downstream dose delivery.

Then, to assess the clinical impact of the dosimetric errors, we evaluate in retrospect two treatment plans of chordoma patients with titanium orthopedic hardware. The clinically planned dose distributions are compared to dose calculations that employ both CT metal artifact reduction and Monte Carlo methods to simulate the beam transport through the implants.

# 8.1. Methods and materials

## 8.1.1. Study design

The question we seek to answer is how CT metal artifacts and the approximations of the pencil-beam dose calculation method affect dose calculation accuracy in the presence of titanium implants. Also, we aim to quantify the impact of these potential dosimetric errors on clinical treatment plans for chordoma. The main issues under study are:

- CT metal artifacts due to the titanium implants.
- The simulation of beam transport through titanium, particularly:
  - Multiple Coulomb scattering of proton passing through implants.
  - Non-elastic nuclear interactions in implants.

### CT metal artifacts

To assess the impact of CT metal artifacts, we first studied a treatment plan designed on a phantom CT scan in which a severe metal artifact was created using titanium inserts. The purpose of this study was to determine the extent of possible proton range errors that can potentially be introduced by the CT artifacts.

The phantom was scanned with and without titanium inserts. Proton beams were planned to pass through different parts of the artifacts, with the spread-out Bragg peak (SOBP) targeted on a cylindrical volume. Using the pencil-beam dose calculation algorithm, the doses delivered by the beams were calculated on both the artifact-affected CT scans and ground truth scans without titanium inserts. We analysed the dose distributions of three proton beams passing through different parts of the metal artifact.

The effect of artifacts on clinical proton beams was studied by recalculating dose distributions on artifact-reduced CT images. Sinograms of patient treatment planning scans were acquired from the CT scanner and a previously developed CT metal artifact reduction method was applied [96]. Although no current method can completely eliminate all metal artifacts, our method significantly reduces the artifacts, providing reconstructed CT images that better reflect the actual patient anatomy.

### Beam transport through titanium

The second part of the study involves an analysis of the accuracy of pencil-beam dose calculations with respect to the calculation of beam transport through titanium implants. Pencil-beam dose calculation algorithms use measured depth-dose curves in water as input data, rather than using more fundamental physical principles. Therefore, although the energy loss of protons due to the titanium implant is taken

into account, most pencil-beam algorithms do not incorporate effects specific to the structure of the titanium nucleus. Essentially, it is assumed that titanium can be represented by a virtual high-density water medium.

To better understand the errors introduced by these approximations, we studied non-elastic nuclear reactions and multiple Coulomb scattering of protons passing through titanium. The fluence loss because of non-elastic nuclear reactions was calculated, and the root-mean-square (RMS) scattering angles due to multiple Coulomb interactions were determined. These results were compared to calculations using water with a stopping power equivalent to titanium.

To quantify the effects on dose calculation, a Monte Carlo simulation of a proton beam was performed on a computational phantom with a titanium cylinder, and compared to the results of a pencil-beam calculation. Last, Monte Carlo simulations of clinical proton beams were performed, using the artifact-reduced CT images.

#### Evaluation of chordoma treatment plans

To determine the impact of the implants on clinical treatment plans for chordoma, we combined the CT metal artifact reduction algorithm and Monte Carlo dose calculation methods to recalculate in retrospect the treatment plans of two patients. The original dose distributions were compared to the corrected dose calculations, and dose-volume histogram (DVH) analysis was performed.

### 8.1.2. CT scans

#### Phantom scan

A Gammex 467 tissue characterization phantom (Gammex Inc., Middleton, WI) was scanned on a GE LightSpeed RT 16 CT scanner (GE Healthcare, Waukesha, WI) with a 140 kV x-ray tube voltage and a current of 600 mA (see figure 8.1). These imaging parameters are similar to those used at our institution for patients with titanium implants. Two titanium cylinders with a diameter of 25.4 mm were placed in the phantom perpendicular to the transverse plane. A second scan without titanium inserts was performed to serve as ground truth. To the images of the ground truth scan, virtual titanium inserts were added with the exact size and location of the real inserts.



#### Patient scans

The CT scans of two chordoma patients were studied in retrospect. Both patients had titanium constructs that were used to stabilize the vertebral column, and were connected by cross-links. Details of the patient scans are listed in table 8.1. The patients were scanned on the same GE LightSpeed RT 16 scanner in axial mode. Sinograms were obtained from the scanner, and using the scanner system were calibrated and rebinned to parallel-bin geometry.

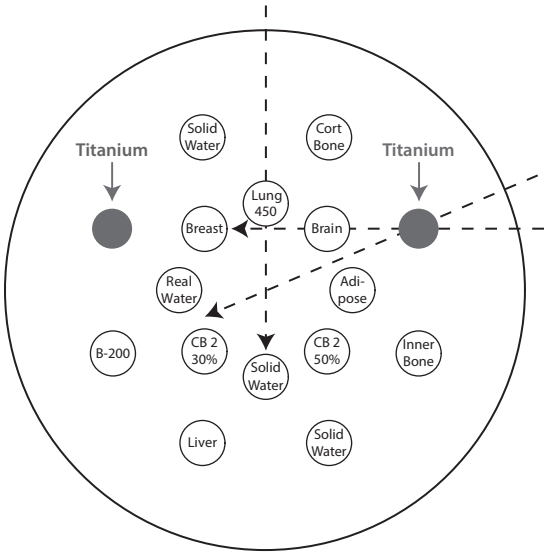



**Figure 8.1** | Cross-section of the tissue characterization phantom used to study CT metal artifacts. The phantom consists of a disk of solid water in which various tissue-equivalent plastics are inserted. One scan was obtained with two titanium cylinders in the positions indicated; a second ground truth scan was acquired without the titanium cylinders. The dashed lines indicate the beam angles used in the study.

The CT metal artifact reduction algorithm used is described in more detail in our previous work [96]. Metal artifacts due to the titanium stabilization hardware were reduced by a physics correction algorithm which employs a beam hardening correction. This method compares the projections through titanium with neighbouring projections not passing through the implants, creating a first-order model of beam hardening which was used to adjust the values of projections passing through the hardware. The locations of the titanium implants were determined using a threshold of 2500 Hounsfield units (HU) on the CT numbers in the original image reconstruction. Any voxels with a CT number above this threshold were considered to comprise the implants. This segmentation of the titanium implants was validated using CT projections rebinned into virtual radiographs, which are not affected by metal artifacts.

Artifacts due to high-Z implants, for example dental fillings, or a combination of high-Z implants and titanium hardware, were further reduced using an iterative optimization algorithm that replaces missing projections caused by near-complete x-ray attenuation. In this case, a higher threshold of 8000 HU was used to identify the high-Z implants.

### 8.1.3. Treatment planning

The treatment plans we studied were designed for passively scattered proton therapy delivery, and were created using the XiO (Elekta AB, Stockholm, Sweden) treatment planning system (TPS). All radiation doses are reported in Gy RBE (relative biological effectiveness), using the conventional constant proton RBE of 1.1.

#### Phantom for metal artifacts study

The treatment plan designed on the tissue characterization phantom, which was used to study metal artifacts, consisted of three coplanar proton beams passing through different parts of the main metal artifact between the two titanium cylinders. Two beams targeted a cylinder at the centre of the phantom, entering the phantom respectively from the top and at a right oblique 67° angle, as shown in figure 8.1. A third beam was oriented parallel to the two titanium inserts and targeted a cylinder in the middle of the inserts. All target volumes had a diameter of 5 cm.

#### Phantom for dose calculation study

To assess calculations of proton beam transport through titanium, a treatment plan was created on a virtual water phantom with a titanium cylinder. The plan featured a single beam designed to target a cylindrical region with a diameter of 5 cm, which was located 10 cm distal to a titanium cylinder with a diameter of 1 cm. The titanium cylinder was placed 5 cm from the surface of the phantom. The doses were calculated on a high-resolution grid with 0.65 $mm^3$ voxels.

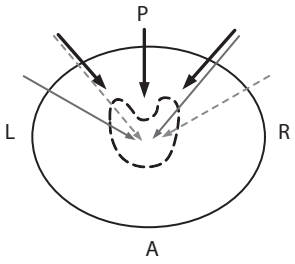
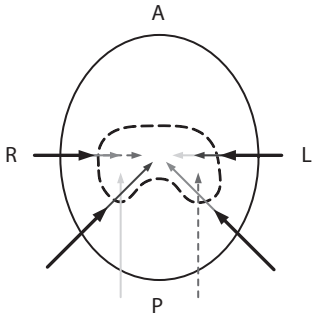
#### Patient treatment plans

The clinical treatment plans that we analysed were created for post-operative proton radiotherapy received by the chordoma patients. The patients also received pre-operative radiotherapy; this part of the treatment is however not relevant for the present study because the titanium implants were not yet present. The prescription doses and beam arrangements are provided in table 8.1.

The clinical treatment plans used several uniform beams combined with patch and through beams to spare the nearby spinal cord. This beam arrangement combines a through beam which delivers dose to the majority of the tumour with a patch field that is laterally conformed to the transverse edge of the through beam, and delivers dose the remaining part of the target next to the critical tissue [17]. Dose calculations were performed with a 2 $mm^3$ voxel size.

Treatment plans for passively scattered proton therapy at our institution are designed to be robust against a ±3.5 % variation of the water-equivalent range in the patient [17]. This margin is necessary because of various uncertainties, mainly related to the errors and degeneracy in the conversion of CT x-ray attenuation to proton stopping powers [177]. In addition, depending on the tumour site, a certain smearing of the range compensator is employed to ensure robustness against patient

**Table 8.1** | Details of the clinical CT scans and treatment plans.

| | **Patient 1** | **Patient 2** |
|---|---|---|
| *Body site* | Lumbar spine | Cervical spine |
| *Treatment position* | Head-first prone | Head-first supine |
| *Titanium hardware* | 4 rods, cage, cross-links | 3 rods, cage, cross-links |
| *Other hardware* | None | Dental fillings |
| *Prescribed dose* | 39.6 Gy (RBE) | 48.6 Gy (RBE) |
| *Total dose* | 70.2 Gy (RBE) | 75.6 Gy (RBE) |
| *Proton beams* | • 3 uniform fields<br>• 2 patch/through field combinations | • 4 uniform fields<br>• 4 patch/through field combinations |
| |  |  |
| | Thick lines: uniform fields; thin lines: patch/through field combinations. | |
| *CT scan* | • 140 kV tube voltage<br>• 565 mA current<br>• 2.5 mm slice thickness<br>• 2 s rotation | • 140 kV tube voltage<br>• 590 mA current<br>• 2.5 mm slice thickness<br>• 1 s rotation |

motion [172]. There is currently no standard to include additional margins when metallic implants are present.

### 8.1.4. Pencil-beam dose calculation algorithm

The pencil-beam dose calculation method that we used is an implementation of the algorithm by Hong et al. [42]. The broad proton beam is modelled as a combination of small computational pencil-beams for which the transport through the range compensator and patient is analytically calculated. Measured depth-dose curves in water for various energies and configurations of the treatment head provide the principal physical input parameters.

The energy loss of protons is simulated as a loss in residual proton range by considering the stopping power of the compensator and tissues relative to water, which is almost constant over the therapeutic energy range. Multiple Coulomb scattering is incorporated by increasing the radial emittance of the protons in a Gaussian approximation, using the generalized Highland [178] formula as described by Gottschalk et al. [16]. Material-specific properties are used to model multiple Coulomb scattering in the range compensator. In the patient, multiple Coulomb scattering is simulated assuming the tissue upstream in the beam path consists of a volume of water with the same integral proton stopping power as the tissue. Therefore, neither the scattering properties of titanium, nor the position of the titanium implants in the path of the pencil-beam is included.

### 8.1.5. Monte Carlo simulations

We performed Monte Carlo dose calculations with the TOPAS Monte Carlo code [144], using the GEANT4 9.5.p01 toolkit [64]. Our in-house MCAUTO code directly links the XiO TPS to this Monte Carlo simulation framework. All proton beam configurations and associated apertures and range compensators are automatically obtained from the planning system.

#### Treatment head

Unlike the pencil-beam algorithm, which uses measured depth-dose profiles in water as input parameters, the Monte Carlo code requires complete phase space information of the proton beams to perform patient dose calculations. This phase space data was generated by simulating the proton transport through the treatment head and the field-specific aperture and range compensator.

Proton beams in the XiO TPS are prescribed a beam weight, which is defined as the dose that would be delivered by said beam to water in the spread-out Bragg peak (SOBP) region, if the center of the SOBP is at the isocentre. For each field, the relationship between the beam weight and the number of protons at the entrance of the treatment head was determined by performing a separate Monte Carlo output

factor simulation, taking into account the effects of the air gap and the size of the aperture opening.

### Patient anatomy

Particle transport through the patient anatomy was simulated using the treatment planning CT scan. The range of CT numbers was divided in 24 bins, which were each assigned an elemental composition following the results of Schneider et al. [139]. In addition, CT numbers above 2500 HU were considered to comprise the titanium implants and were modelled as pure titanium in the Monte Carlo simulation.

A unique mass density was assigned to each CT number. These were determined by matching the relative proton stopping powers as determined by the Monte Carlo code with the relative stopping power used by the XiO TPS, which was calibrated to the CT scanners used at our department. This method ensures consistency between the assumptions in the pencil-beam dose calculation and the Monte Carlo simulations. In homogeneous media, the proton range as determined using the Monte Carlo code matches the clinical pencil-beam dose calculation within 1 mm. An on-the-fly conversion was performed to convert the Monte Carlo simulated dose to dose-to-water, which is the measure of dose reported by pencil-beam algorithms that is conventionally used in radiotherapy [140].

### Physics models

The GEANT4 physics models used have been validated for proton therapy simulation at our institution, and are also the default models used in TOPAS. Electro-magnetic interactions were simulated using the GEANT4 standard electromagnetic model with the high-precision option enabled, which increases the resolution of the stopping power tables. For protons, the UrbanMsc90 multiple Coulomb scattering model was used. This model uses a modified Highland [178] formula for the angular distribution of the scattering, combined with an empirical function for the non-Gaussian tail. Non-elastic proton-nucleus interactions were simulated with the binary cascade model at higher proton energies, followed by the pre-compound and evaporation model. The default Wellisch-Axen parametrisation of the total non-eleastic nuclear reaction cross-sections was employed [70].

### Physics model validation

We performed a number of validations of the GEANT4 physics models to ensure accurate simulation of proton interactions with titanium. The parametrization of the total non-elastic reaction cross-section used by GEANT4, which is the main parameter determining the loss of proton fluence distal to titanium implants, was compared to experimental data in the 9 MeV to 100 MeV incident proton energy range [179–182] as compiled by Bauhoff [89]. Comparisons were also made to optical model calculations using the TALYS 1.4 nuclear reaction code [58]. TALYS incorporates the

ECIS-06 optical model code, which was used with the Koning and Delaroche [74] global optical model potential.

No experimental data of proton multiple scattering in titanium was found in literature for the relevant energy range. We therefore compared Gaussian profiles fitted to the radial proton emittance in GEANT4 simulations with Gaussian fits to Molière theory [15]. Molière theory was shown to be accurate for other intermediate-Z materials [16]. The Molière theory evaluations were performed using the BGware code [183]. We used the method of Hanson et al. [184] to convert the characteristic scattering angle obtained using Molière theory to a Gaussian approximation.

Simulations of multiple Coulomb scattering using GEANT4 were performed by transporting protons through a slab of material. The incident protons were orientated perpendicular to this slab, and their angular momentum was registered after passing through the slab. A Gaussian distribution was fitted to the angular distribution of the outgoing protons by minimizing the least-square error. Simulations were performed with initial proton energies of 100 MeV, 150 MeV, and 200 MeV, and titanium slabs of 5 mm, 10 mm, 15 mm, and 20 mm thickness.

### 8.1.6. Proton-nuclear interactions with titanium

The difference between a pencil-beam dose calculation and a Monte Carlo simulation is a superposition of several effects that are better accounted for in the Monte Carlo models. To determine the magnitude of these effects, we performed several simulations in simple geometries, in which nuclear interactions with titanium were compared to a model in which the titanium is represented by water with an equivalent proton stopping power. Multiple Coulomb scattering was simulated in GEANT4 using the same approach as described in the previous section. In addition, fluence loss due to non-elastic nuclear interactions was calculated using the total non-elastic reactions cross-sections obtained from GEANT4.

## 8.2. Results

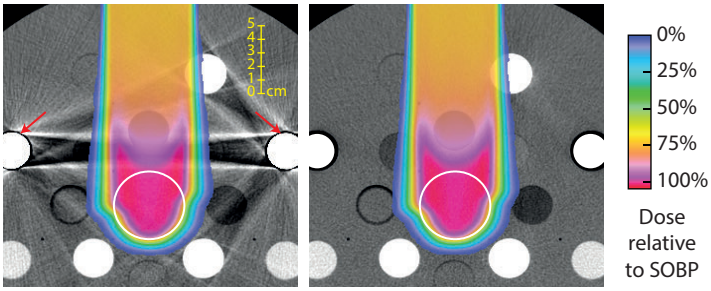


In this section, we first show the impact of CT metal artifact reduction and Monte Carlo simulation on the dose calculation in phantoms and patients. Then, we present the overall impact of these issues on treatment plans for chordoma patients.

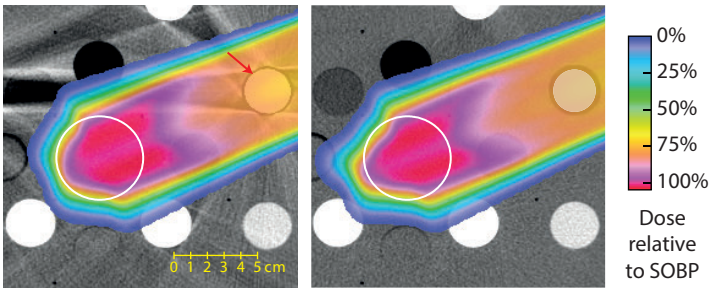
### 8.2.1. Impact of metal artifacts

#### Phantom study

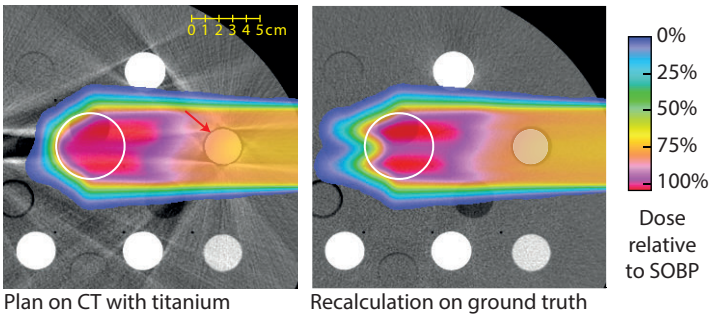
The impact of a severe metal artifact on the phantom dose calculation is shown in figure 8.2. The presence of two titanium cylinders results in the bright and dark streaks in the CT images, which mostly appear between the cylinders and between the titanium and other tissue equivalent materials with higher densities. As expected, the impact of the artifacts depended on the orientation of the proton beam. The proton




(a) Beam perpendicular to main artifact




(b) Beam through one titanium insert




(c) Beam parallel to main artifact

**Figure 8.2** | Dose distributions of three proton beams planned on a phantom CT scan with titanium inserts, which was affected by artifacts. The doses were recalculated on a ground truth scan without the metal. The ground truth scan includes a virtual titanium cylinder matching the physical one. Doses below 0.5% of the SOBP dose are not shown. CT window level: 40 HU, width: 400 HU.

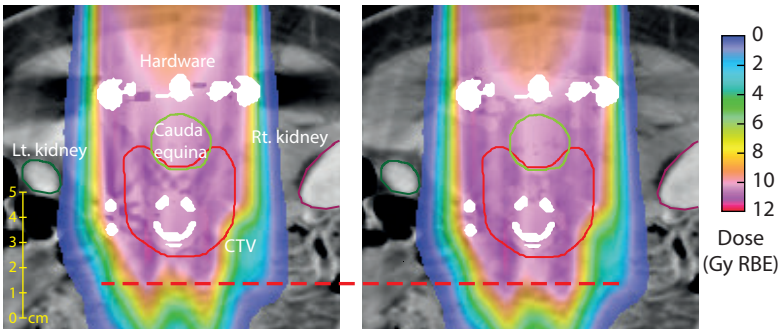



**Figure 8.3** | Dose calculations of a posterior-anterior beam traversing titanium hardware in the lumbar spine. The dashed red lines near the end-of-range of the beam are drawn to guide the eye. Doses below 0.1 Gy (RBE) are not shown. CT window level: 40 HU, width: 400 HU.

range errors varied between 1 mm to 10 mm. Because of the severe metal artifact in this phantom study, the impact on the clinical proton beams would normally be expected to be within this range.

The dose calculation of a beam oriented perpendicular to several bright and dark artifacts was almost unaffected by the artifacts, as seen in figure 8.2a. The small effect on the calculated range can be explained by the fact that errors due to several bright and dark streak artifacts cancel out. Shown in figure 8.2b is a beam coming from an oblique angle, which passed through the cylinder itself and the bright artifacts surrounding it. In this case, the recalculation of the dose on the ground truth scan showed an increase of the range in the central beam axis by about 5 mm. Finally, in figure 8.2c, the beam is oriented parallel to the main metal artifact. In this worst-case situation, range shifts on the order of 10 mm were observed in both directions.

8

## Patient study

The originally planned dose distributions for the two chordoma patient were recalculated on artifact-reduced CT images. In both cases, the recalculated dose distributions showed an increase of the proton range for the parts of the beams that passed through or near the titanium hardware. Target coverage was therefore not compromised. The maximum range overshoot was 6 mm in soft tissue at 80 % dose level.

Cross-section of the dose distribution for one of the proton beams of each treatment plan are shown in figure 8.3 and figure 8.4. The increase in range can clearly be seen in both cases. In the case of patient 2, the original beam was planned with the high-dose area just proximal to the parotid gland. The recalculated dose distribution shows th s area sh fted into the proximal side of the gland. This issue occurred to some extent in all left-right, right-left and oblique beams.

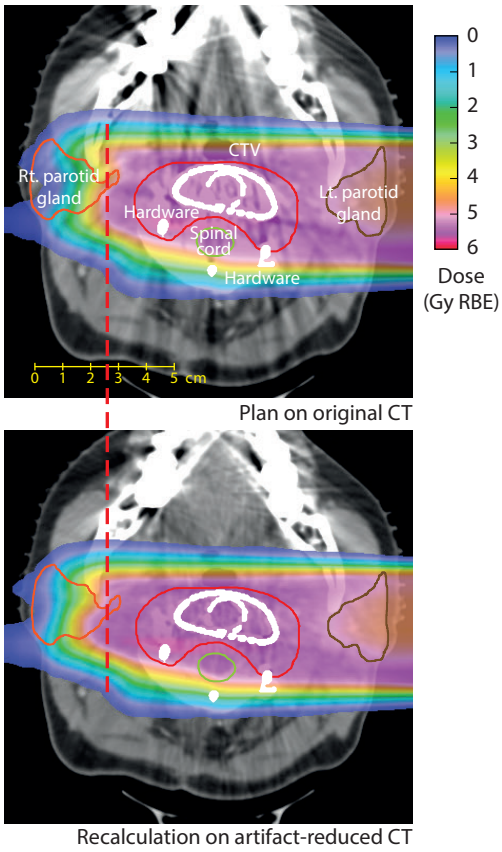



**Figure 8.4** | Dose calculations of a proton beam around titanium hardware in the cervical spine. The dashed red lines near the end-of-range of the beam are drawn to guide the eye. Doses below 0.1 Gy (RBE) are not shown. CT window level: 40 HU, width: 400 HU.

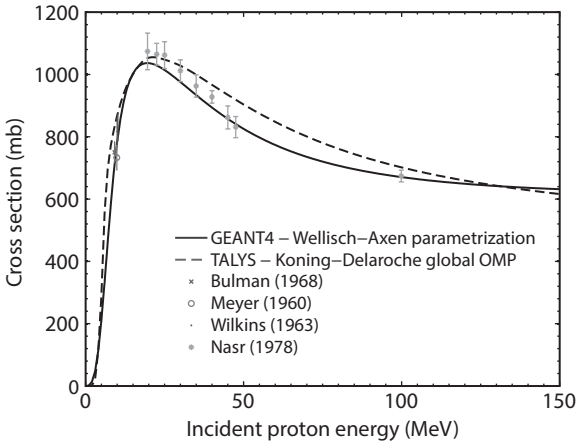



**Figure 8.5** | Total non-elastic nuclear reaction cross-section of proton-induced reactions on titanium with isotopes in natural abundance.

## 8.2.2. Impact of dose calculation method

### Geant4 physics model validation

The validation of the total non-elastic cross-section of proton-induced reactions on titanium are shown in figure 8.5. The GEANT4 9.5 cross-sections were compared to a TALYS optical model calculation and experimental data. A good agreement was found between the models and measurement data. The uncertainty in the cross-section is estimated to be below 10 % for the clinical proton energy range. Since only a small fraction of the protons passing through an implant will undergo a nuclear reaction, this uncertainly has a very small impact on dose calculation.

A comparison was also made between the GEANT4 simulations of multiple Coulomb scattering and Molière theory, the results of which are provided in table 8.2. The agreement of the root-mean-square (RMS) scattering angles was satisfactory with differences of only a few percent.



### Proton-nuclear interactions with titanium

Having validated the physics models, we analysed how the approximations of the pencil-beam dose algorithm differ from calculations specifically modelling the proton interactions with the titanium nuclei. Tabulated in table 8.3 is the RMS multiple Coulomb scattering angle for various combinations of proton energy and titanium thickness, compared to the RMS angle if the medium is assumed to be water with the same stopping power as titanium. In all the cases, the water-based simulation underestimated the RMS scattering angle by about 50 %. The error in this approximation therefore becomes greater with a larger amount of titanium in the beam path, or a lower proton energy, as both increase the multiple Coulomb scattering angle.

**Table 8.2** | Comparison of GEANT4 model predictions of proton multiple Coulomb scattering in titanium to Molière theory calculations. The RMS angle of Gaussian fits to the angular distributions is listed.

| Ingoing proton energy (MeV) | Titanium thickness (mm) | RMS angle Molière theory (mrad) | RMS angle GEANT4 (mrad) | Difference (%) |
|---|---|---|---|---|
| 100 | 5 | 27.19 | 25.99 | −4.41 |
| 100 | 10 | 42.05 | 41.03 | −2.43 |
| 100 | 15 | 58.86 | 56.74 | −3.60 |
| 100 | 20 | 80.24 | 77.96 | −2.84 |
| 150 | 5 | 17.82 | 17.15 | −3.76 |
| 150 | 10 | 26.90 | 25.91 | −3.68 |
| 150 | 15 | 33.50 | 33.51 | 0.03 |
| 150 | 20 | 41.45 | 40.80 | −1.57 |
| 200 | 5 | 13.28 | 12.98 | −2.26 |
| 200 | 10 | 19.24 | 19.36 | 0.62 |
| 200 | 15 | 24.42 | 24.67 | 1.02 |
| 200 | 20 | 30.04 | 29.50 | −1.80 |

**Table 8.3** | Comparison of GEANT4 simulations of multiple Coulomb scattering in titanium and a virtual high-density water medium with equivalent proton stopping power. The RMS angle of Gaussian fits to the angular distributions is listed.

| Ingoing proton energy (MeV) | Titanium thickness (mm) | RMS angle Titanium (mrad) | RMS angle Water-based (mrad) | Difference (%) |
|---|---|---|---|---|
| 100 | 5 | 25.99 | 13.31 | −48.79 |
| 100 | 10 | 41.03 | 20.93 | −48.99 |
| 100 | 15 | 56.74 | 28.89 | −49.08 |
| 100 | 20 | 77.96 | 39.79 | −48.96 |
| 150 | 5 | 17.15 | 8.80 | −48.69 |
| 150 | 10 | 25.91 | 13.23 | −48.94 |
| 150 | 15 | 33.51 | 17.06 | −49.09 |
| 150 | 20 | 40.80 | 20.73 | −49.19 |
| 200 | 5 | 12.98 | 6.67 | −48.61 |
| 200 | 10 | 19.36 | 9.91 | −48.81 |
| 200 | 15 | 24.67 | 12.58 | −49.01 |
| 200 | 20 | 29.50 | 15.00 | −49.15 |

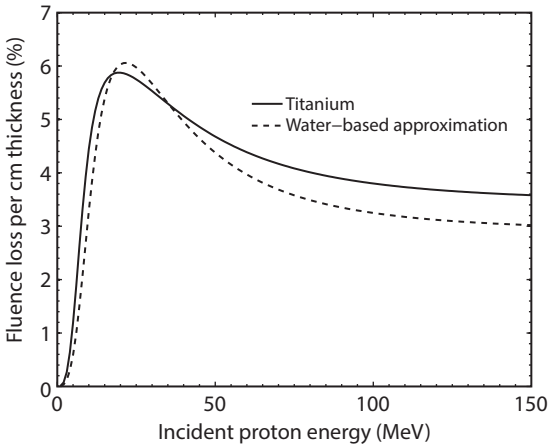



**Figure 8.6** | Calculation of the loss of primary proton fluence due to non-elastic nuclear reaction in titanium, using the Wellisch and Axen [70] parametrization. A calculation using the non-elastic reaction cross-section of titanium is compared to a calculation based on the cross-section for water scaled by the relative proton stopping power of titanium.

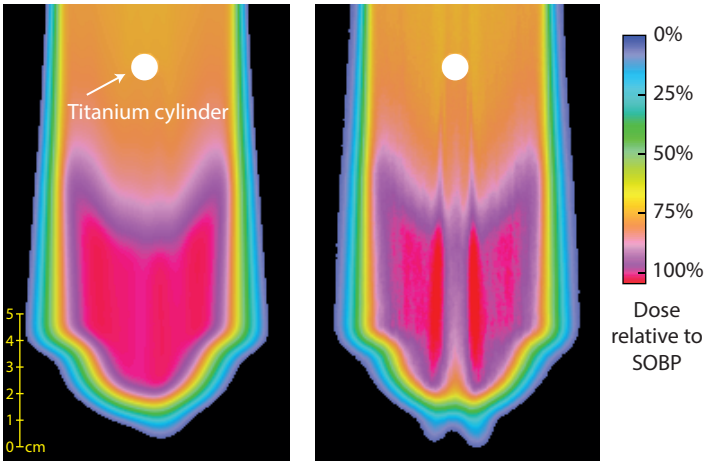



**Figure 8.7** | Dose calculations of a proton beam passing through a virtual uniform water phantom with a titanium cylinder. Doses below 0.5 % of the SOBP dose are not shown.

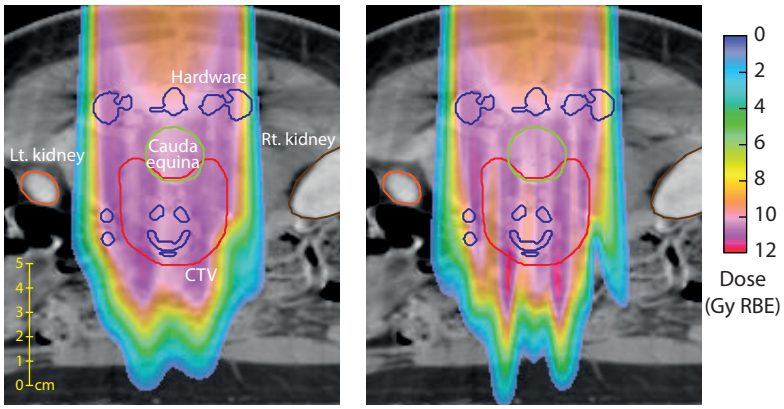



**Figure 8.8** | Dose calculations of a posterior-anterior beam passing through titanium orthopedic hardware. Doses below 0.5 Gy (RBE) are not shown. CT window level: 40 HU, window width: 400 HU.

Depicted in figure 8.6 is the effect of non-elastic nuclear interactions. The proton fluence loss per centimetre of titanium in the water-based approximation was compared to a calculation using the cross-sections for titanium as shown in figure 8.5. At low incident proton energies, the water-based approximation was almost equivalent to the titanium cross-section. The difference was somewhat larger in the case of higher energies, but remained limited.

### Phantom study

From the previous findings, it follows that the pencil-beam model of multiple Coulomb scattering is the main issue that may result in dosimetric errors when titanium implants are present. To assess the impact on dosimetry, a comparison of a pencil-beam dose calculation and a Monte Carlo simulation on the virtual voxelised water phantom was performed, wh ch is shown in figure 8.7. The Monte Carlo simulation showed a significantly larger reduction in dose delivered distal to the titanium cylinder. Also, elevated dose was seen distal to the edges of the titanium cylinder. For this particular geometry, the Monte Carlo simulation reported a dose approximately 50 % lower near the distal edge of the beam located behind the titanium cylinder. The degradation of the end-of-range of the proton beam, which is also due to multiple Coulomb scattering [185], was not reproduced by the pencil-beam method.

### Patient treatment plans

The analysis of the clinical proton beams revealed similar effects as seen in the phantom dose calculation. For example, shown in figure 8.8 are pencil-beam and Monte Carlo dose calculations of the posterior-anterior beam of the treatment plan

for patient 1. Again, the Monte Carlo simulations showed a large dose reduction distal to the implants and elevated dose distal to the edges of the hardware. The end-of-range of the beam was severely degraded; local range differences up to about 10 mm at 80 % dose level were observed.

### 8.2.3. Overall impact on chordoma treatment plans

Dose distributions for the treatment plans of both patients are shown in figure 8.9 and figure 8.10. We compared the original calculations, using the pencil-beam algorithm and original CT images, with recalculations utilizing Monte Carlo simulation and artifact-reduced CT images. Dose-volume histograms are plotted in figure 8.11.

The use of several uniform proton beams reduced the impact of the dose errors on the overall treatment plan. Although dose errors up to 10 % of the SOBP dose were found for the individual proton beams, as shown in figure 8.8, these errors were located at different positions for each beam, reducing the total impact. Since the treatment plans for patient 1 and patient 2 employed 7 and 12 beams respectively, the plans were quite robust against local dose errors in the individual fields.

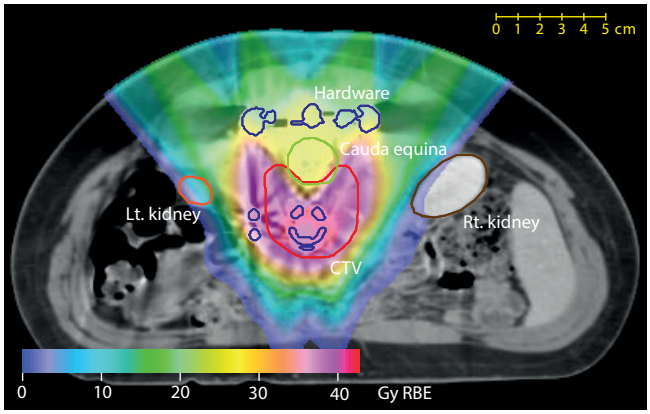
The recalculations of the treatment plans showed almost no difference in the mean dose to the clinical target volume (CTV). However, the homogeneity of the target dose was somewhat affected. A number of dose hot spots around the implants were observed within the CTV, which resulted in the high-dose tail in the DVHs shown in figure 8.11. In both cases, the maximum dose to the CTV was 4 Gy (RBE) higher in the recalculated dose distributions. In the case of patient 2, the mean dose to all salivary glands increased by 1 Gy (RBE).

## 8.3. Discussion and conclusions

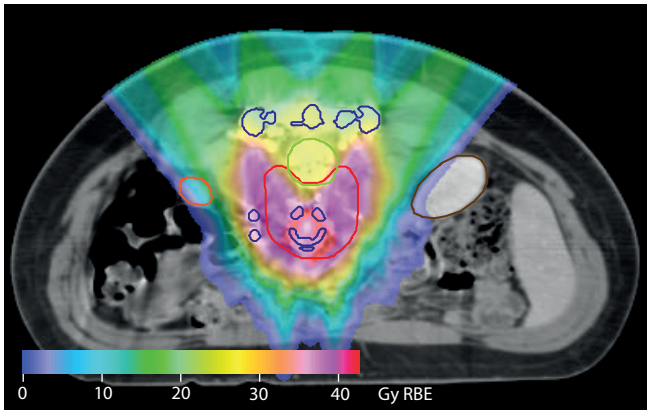
Our aim was to investigate the dosimetric errors in clinical proton therapy dose calculation due to the presence of titanium implants, and to determine how these errors affect treatment of chordoma patients with orthopaedic hardware. We found that CT metal artifacts mainly affect the range of the proton beam, while the pencil-beam dose calculation algorithm underestimates dose inhomogeneity and range degradation distal to the implants.

The magnitude of the errors introduced by CT metal artifacts is highly dependent on the geometry of the hardware and the orientation of the proton beam relative to the artifacts. A phantom study showed range errors due to a severe metal artifact can vary between 1 mm to 10 mm. The evaluation of two chordoma cases, with titanium hardware around respectively the lumbar and cervical spine, showed mostly an underestimation of the proton range due to metal artifacts. Recalculations of the dose distributions on artifact-reduced CT scans revealed range increases up to 6 mm.

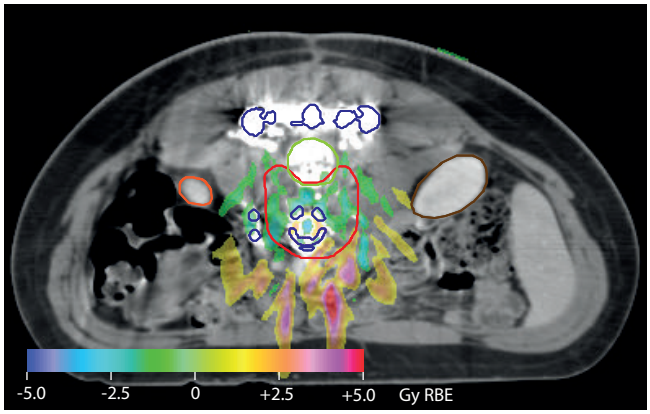
Second, the conventional pencil-beam dose calculation algorithms, which use water as a reference medium, do not consider effects specific to the nuclear struc-




Original dose distribution

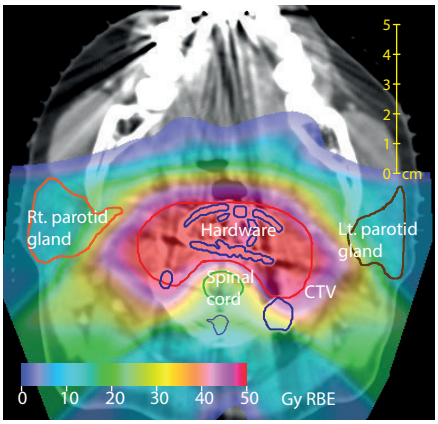



Corrected dose distribution

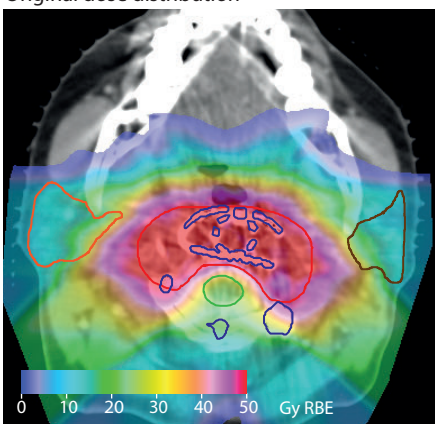



Dose difference: corrected – original

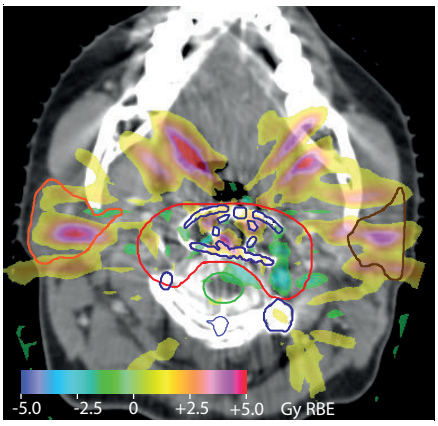
**Figure 8.9** | Original and corrected dose distributions of the treatment plan of patient 1. The difference between the two dose calculations in also shown. Doses and dose differences below 1 Gy (RBE) are not displayed. CT window level: 40 HU, width: 400 HU.




Original dose distribution

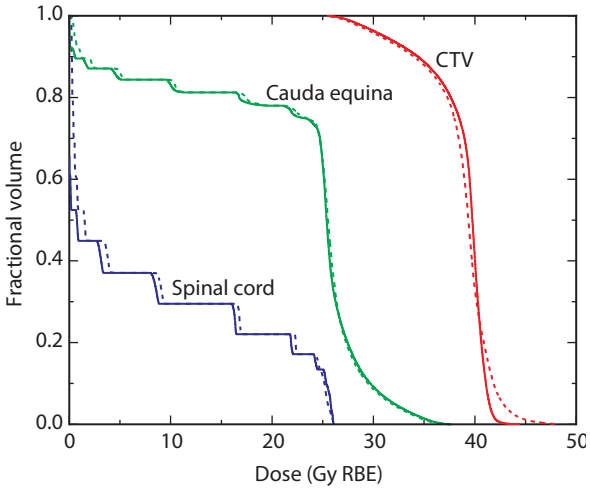



Corrected dose distribution

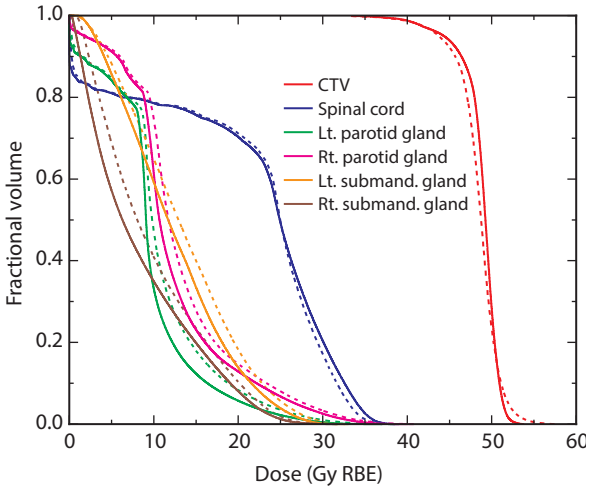



Dose difference: corrected – original

**Figure 8.10** | Original and corrected dose distributions of the treatment plan of patient 2. The difference between the two dose calculations in also shown. Doses and dose differences below 1 Gy (RBE) are not displayed. CT window level: 40 HU, width: 400 HU.




(a) Patient 1




(b) Patient 2

**Figure 8.11** | Dose-volume histograms of the patient dose distributions. Dose calculations on the original CT scans using the pencil-beam algorithm (——) and dose recalculations using metal artifact reduction and Monte Carlo simulations (- - - -).

ture of the irradiated materials. Also, scattering is calculated using the integrated radiological depth, which ignores the position of the implant in the beam path. We found that incomplete modelling of multiple Coulomb scattering results in the most significant errors, causing the dose delivered distal to typical titanium constructs to be overestimated by up to 10 %. Proton range errors up to 10 mm occur in addition. The loss of proton fluence due to non-elastic nuclear reactions in the implant is reasonably estimated by the pencil-beam dose calculation, with dose errors being limited to about 1 %. The result of these effects is that the characteristic 'dose shadow' distal to titanium implants, as observed experimentally by Parodi et al. [34] and Schneider et al. [186], is not well reproduced by the pencil-beam dose calculation.

The overall impact of dosimetric errors in the individual proton fields depends on the design of the treatment plan. The chordoma plans for passively scattered proton therapy were found to be quite robust against the errors introduced by the titanium implants. Dose recalculations which used artifact-reduced CT images and Monte Carlo simulations, showed only a small decrease in the homogeneity of the total dose delivered to the CTV. The mean dose to the CTV remained almost unchanged. This is mainly a result of the large number of proton beams used; since the errors due to the implants are generally at a different anatomical location in each proton field, the errors in the overall treatment plan tend to average out. Our results therefore do not show evidence for dosimetric inaccuracies in the treatment planning being a significant factor in the reduced tumour control of patients with titanium implants, which was reported by DeLaney et al. [176] and Staab et al. [175]. It should however be noted that the patient group studied by Staab et al. [175] received spot scanning proton therapy, and the treatment design is therefore not directly comparable with the cases that we studied. Also, analysis of a larger number of patients cases is needed to confirm these findings.

The dosimetric errors we identified apply in general to patients with titanium implants receiving proton therapy. Given the magnitude of the effects, we recommend to consider implementing methods to better estimate and mitigate the effects of the titanium implants.



Range errors due to CT metal artifacts can clearly be reduced using improved CT image reconstruction methods, which are starting to become commercially available. It is important to use a method that is designed to both provide an accurate segmentation of the high-density implants from the surrounding tissue, and to reduce artifacts throughout the reconstructed CT images. Methods designed to improve diagnostic image quality should be separately validated to ensure accurate segmentation of the implant which is essential for dose calculation. Unless beams are oriented parallel to a metal artifact, our results show a trend towards range overshoot due to the bright artifacts directly surrounding the implants. CT metal artifact reduction would therefore be expected to potentially yield some reduction in

dose to distal organs-at-risk such as the salivary glands. Because of range uncertainty in general, proton therapy treatment plans are normally designed with sufficient safety margins in the dose delivered to these organs.

The limitations of the pencil-beam dose calculation algorithm are a concern because dose inhomogeneity in the target is underestimated. While the physical interaction of the protons with the implants cannot be changed, an improved dose calculation algorithm that provides a good estimate of the effects can be helpful to determine how to mitigate them by changing the design of the treatment plan. Our results show that incorporating a better model of multiple Coulomb scattering would yield the most significant improvement. Potential methods to model this effect have previously been studied by Soukup et al. [187] and Szymanowski and Oelfke [188]. Until such methods are implemented, one could account for the expected errors by using a larger number of beams and conservative range margins.

We believe that the design of future treatments for patients with titanium implants using intensity modulated proton therapy (IMPT) needs to be studied carefully. The advantages of robust optimization of IMPT treatment plans to account for uncertainties in the patient position and CT numbers are well known [29], and are even more important in the case of patients with implants. However, if an IMPT plan is optimized without considering the limitations of the dose calculation algorithm, the possibility exists that a cold spot in the target volume as a result of multiple Coulomb scattering is not compensated for by other beams. Also, IMPT plans tend to require fewer beams to deliver highly conformal dose distributions, which would also increase the impact of the dosimetric errors on the overall treatment plan.

It should be mentioned that, in addition to the issues that we studied, the image quality of the planning CT scan may also affect the accuracy of fusion of MRI images to the CT images for target and organ delineation. Also, artifacts may introduce errors in image guided patient setup using the CT scan as a reference. The impact of these effects has not yet been quantified.

To summarize, we found that dose calculations of proton beams in the presence of titanium implants are subject to significant errors due to CT metal artifacts and the incomplete model of multiple Coulomb scattering used by pencil-beam dose calculations. The overall treatment plans of chordoma patients, which employed passively scattered proton therapy, were however not as significantly affected. This is due to the local nature of the errors combined with the large number of fields, causing the errors to average out to a large degree.

# 9

# Conclusions and perspective

In this chapter, we summarize the main conclusions of our research to reduce the range uncertainty of proton therapy beams. We also provide recommendations for the further clinical development of the proposed methods.

In current clinical practise, the range of proton therapy beams is determined by using population averages of tissue properties to convert the CT numbers of traversed tissue to proton stopping powers. To account for the uncertainty in the range of the beam in the patient, a margin is added to the range, and the sharp distal dose gradient is not positioned between the target and a nearby organ at risk.

We believe these current measures to make treatments robust against range uncertainty are not satisfactory, because they preclude the full use of the physical advantages of proton beams to create an optimal dose distribution. In this thesis, we investigated new technologies to reduce the proton range uncertainty. The ability to precisely position the end-of-range of the proton beam can remove the current limitations in treatment designs and further improve the benefit of proton therapy.

*In vivo* verification of the range of proton beams will be key to reduce the range uncertainty. Instead of using population averages of tissue properties to determine the proton range, it will enable the range to be fine-tuned for each beam delivered to an individual patient. Also, the range can be monitored continuously throughout the course of treatment. We propose to perform *in vivo* verification of the proton beam by performing quantitative spectroscopy of prompt $\gamma$-rays from proton-nuclear interactions, which are emitted virtually instantaneously and can escape the patient's body. These measurements are then compared to detailed models of the proton interactions, to determine the range of the beam. Accurate anatomical imaging and proton transport modelling should complement *in vivo* verification.

From our experimental results, we conclude that this approach is feasible within clinical beam delivery conditions. We therefore recommend further development towards clinical trials. In the next sections, we summarize the main conclusions. Recommendations for future research are also be given, and we discuss possibilities for clinical implementation and application of the proposed methods.

## 9.1. Prompt γ-ray detection

A small scale prototype detector, described in chapter 3, was developed to perform energy- and time-resolved measurements of proton-induced prompt $\gamma$-rays. The detector consisted of an actively shielded $LaBr_3$ (Ce) scintillator. We found that the energy resolution enabled discrete prompt $\gamma$-ray lines from specific nuclear reactions to be clearly resolved along proton beams stopped in matter. The timing resolution provided a separation of proton- and neutron-induced events within the cyclotron radiofrequency period.

A digital data acquisition system was developed for accurate quantification of the $\gamma$-ray counts and to handle high count rates. Continuous calibration and correction methods were developed to ensure stable absolute measurements for a wide range of proton energies, beam currents and radiation background levels.

## 9.2. Modelling prompt γ-ray emission

To facilitate the determination of the absolute range of the proton beam from prompt $\gamma$-ray measurements, we modelled discrete prompt $\gamma$-rays emissions from specific nuclear reactions. These discrete $\gamma$-ray lines were resolved from the background and measured quantitatively.

The existing phenomenological nuclear reaction models are generally accurate enough to simulate the impact of nuclear reactions on the dose distribution of proton beams. However, when quantitative information on prompt $\gamma$-rays is needed for range verification, the uncertainties in these models are too large, as shown in chapter 2. The angle-integrated total $\gamma$-ray emissions near the end-of-range differed by a factor of about two between different nuclear reaction models. Differences in the modelled emissions of specific discrete $\gamma$-rays were even larger. The anisotropic $\gamma$-ray emission further complicates the matter.

Therefore, experiments were needed to determine the cross sections of relevant nuclear reactions. Previous data exists for some $\gamma$-ray emissions, but these are not complete in the therapeutic proton energy range. We performed a study to optimize differential cross sections for 15 different $\gamma$-line excitations, which is described in chapter 4. It is recommended that additional thin target cross section measurements are performed to further improve the knowledge of the cross sections.

## 9.3. *In vivo* range verification

Using models of the prompt $\gamma$-rays that are expected to be detected in different scenarios, an optimisation procedure was developed to determine the range of the proton beam. The method developed in chapter 4 uses knowledge of specific nuclear reactions probabilities to directly determine the absolute residual range of the proton beam. Experiments were performed to assess the performance of the method.

### 9.3.1. Patient-specific prompt γ-ray models

To correlate the detected prompt $\gamma$-rays with the delivered treatment, patient-specific models of the $\gamma$-ray emission and detection are required. A practical approach to create such models, is to first simulate the distribution of proton energies, for example using a Monte Carlo simulation. Based on the proton energy distribution, prompt $\gamma$-ray emissions were be calculated using experimentally determined nuclear reaction cross sections, as shown in chapter 6.

In the case of proton pencil-beam scanning, prompt $\gamma$-ray models can be created for each individual pencil-beam, for which the $\gamma$-ray measurement can also be performed separately. For passively scattered beams, a model could be created for each energy layer. Using models of the detection system and an attenuation

correction for the $\gamma$-ray interactions with the patient, the expected detections can then be calculated from the $\gamma$-ray emissions.

Models of the prompt $\gamma$-ray emission and detection were created for different scenarios. This includes the case in which the beam delivery is fully consistent with the plan, but also different possible deviations. By matching the detected events to the models, it was then be determined how the beam was delivered in practise.

### 9.3.2. Optimization method

Instead of verifying the range of each proton pencil-beam individually, the optimization method was designed to verify the residual range or mix of ranges at different positions near the distal end of the proton beam. Data acquired during the delivery of multiple proton pencil-beams, delivered at the same position in the field but with different energies, were simultaneously considered in the optimization. Because these pencil-beams traverse the same tissue, they are affected by the same uncertainties. From the measurements, the error in the stopping power of the upstream tissue was determined, which relates directly to the proton range. By combining data from multiple pencil-beams, the method is more robust to statistical fluctuations.

The developed method incorporates the discrete $\gamma$-ray emissions. Based on the ability to identify nuclear reactions with different target elements, we also propose to include parameters other than the range in the optimization process. The precise elemental composition of the irradiated tissue, from which the prompt $\gamma$-rays are measured, is for example not generally known. Making assumptions about the elemental composition in the patient model can lead to systematic uncertainties. Therefore, we left the oxygen and carbon concentrations as free parameters. The values for these parameters were determined simultaneously with the range.

Another advantage of the method is the ability to verify the proton beam based on spectral measurements at a single point along the path of the beam. If a measurement is performed close to the end-of-range, the residual range can be determined accurately, since the total stopping power uncertainty downstream of this position will be small. Because the range determination does not depend on a spatial measurement of the $\gamma$-rays along the beam direction, the impact of inhomogeneities on the measurement is reduced. For certain treatment sites, some inhomogeneities will still be present within the width of the measurement plane. We recommend to study methods to incorporate additional prior knowledge from the patient's CT scan in the optimization procedure, to minimize the uncertainties in these cases.

We also noted that, after performing an absolute measurement of the proton range, changes relative to this measurement can be identified with a higher precision. The first measurement provides additional prior knowledge as it establishes a specific reference of all $\gamma$-ray emissions as well as the background radiation.

### 9.3.3. Phantom experiments

Using the small scale prototype detector, measurements were performed to verify the range of proton pencil-beams, as described in chapter 4. Within the area covered by a pencil-beam with a size of about 10 mm $\sigma$, it was shown that a standard deviation of 1.0 mm to 1.4 mm on the absolute range could be achieved in phantom measurements. This was based on measurements during the delivery of 5 proton pencil-beams with different energies, at a dose level about 5 times higher as compared to typical fractionated treatments. At a dose level comparable to clinical fields, the standard deviation was on the order of 3 mm. Range measurements relative to a previous one could be performed with a three times higher accuracy.

In chapter 5, we investigated the range verification of passively scattered proton beams. Because complete energy layers are delivered simultaneously by passively scattered beams, the location of range errors within the treatment field cannot be directly determined, unless a detection method is used that can establish the three-dimensional origin of the $\gamma$-ray emissions. Background radiation levels were found to be about an order of magnitude higher as compared to measurements with pencil-beams, since for typical fields a large fraction of the protons are stopped in the treatment head. Therefore, we focused on developing a setup to verify the consistency of the delivered range, using a collimation system with a large opening angle to improve the signal-to-background ratio.

With the same prototype detector and the collimation system that was optimized for passively scattered beams, uniform range errors affecting the entire treatment field area were detected with a standard deviation of 0.1 mm to 0.2 mm at clinical dose levels. We concluded that prompt $\gamma$-ray measurements using a simple detector can verify the consistency of the dose delivery by scattered beams. However, the potential for treatment adaptation is limited, because of the inability to determine the location of range errors, and because new patient-specific hardware would need to be machined.

## 9.4. Monte Carlo simulation of proton beam transport

Proton dose calculation algorithms in current clinical use are based on macroscopic models of beams in water. While such algorithms have a negligible uncertainty if the anatomy traversed by the beam is relatively homogeneous, they do have limitations when simulating proton scatter as the beam traverses complex tissue inhomogeneities, for example bony anatomy in the head and neck area.



Monte Carlo simulation methods, in which the microscopic interactions of many individual particles are simulated, can almost completely eliminate the range uncertainties resulting from the approximate calculations of proton scattering that are used in analytical dose calculation algorithms. Because individual protons are simulated, all interactions can be incorporated in the model.

We recommend that Monte Carlo methods are used when *in vivo* range verification is implemented. Although analytical methods could also be used, approximations in these methods may lead to a treatment to be based on a dose distribution that deviates to a some extend from one that physically deliverable. This affects *in vivo* range verification, because it relies on the possibility to model the delivered treatment and to make adaptations to make it consistent with the plan.

A framework was developed, described in chapter 6, that enables complete treatment plans to be automatically recalculated using the general purpose GEANT4 Monte Carlo code. This framework, which incorporates several methods and models previously developed by others, can serve as a reference to compare against optimized methods that are being developed for clinical treatment planning systems. We also incorporated features to simulate prompt $\gamma$-ray emissions.

## 9.5. Metal implants

For the specific case of patients having metal implants, we investigated CT imaging issues and the modelling of proton transport through the implant. These issues apply to patients who receive radiotherapy after surgical resection of a tumour around the spine, which often requires implants to be installed for stabilization.

The quantitative accuracy of the treatment planning CT scan for these patients was improved by correcting effects of different implant materials on the x-ray projections, that are not considered in regular CT image reconstruction models. In chapter 7, we concluded that these corrections significantly reduce the CT artifacts that lead to errors in the proton stopping powers used for treatment planning.

The impact of the implants on current chordoma treatment plans, delivered with passively scattered proton therapy, was evaluated in chapter 8. We used the improved CT reconstruction method and the developed Monte Carlo simulation framework for dose calculation. Local range overshoots up to +6 mm were found due to CT artifacts. Downstream of implants, Monte Carlo dose calculations of individual beams showed dose differences up to 10 % as compared to pencil-beam calculations, as well as range differences due to proton scattering around the implants.

The current treatment plans for chordoma patients at our institution were found be relatively robust against these uncertainties. These plans use a large number of beams and the dosimetric errors were located at different anatomical locations for each beam. The mean dose to the target remained almost unchanged after applying the corrections, but the dose homogeneity was somewhat reduced. The maximum target dose increased by about 4 Gy (RBE) as compared to the original plan.

It is recommended that the issues related to implants are taken into account when new treatment plan designs are developed, for example for intensity modulated proton therapy. For the prompt $\gamma$-ray model, it is also necessary to consider nuclear interactions with the implants, which will be different from interactions with tissue.

## 9.6. Outlook for clinical implementation

We believe proton range verification though spectroscopy of prompt $\gamma$-rays is promising for clinical use. The next recommended step is the development of a prototype prompt $\gamma$-ray detection system that is suitable for a first clinical trial. Because prompt $\gamma$-ray detection is fully non-invasive and does not affect or require changes in the treatment, a clinical trial is feasible in the near future. It will help to further evaluate the proposed method and to gain experience with the application in clinical practise. Further research is also recommended to determine the best methods to use *in vivo* range verification to adapt treatments.

### 9.6.1. Prototype development for clinical trials

We recommend the development of a clinical prototype prompt $\gamma$-ray detection system to be initially focused on pencil-beam scanning delivery, which is expected to become the main proton therapy delivery modality in the near future. For the first prototype for clinical trials, we believe it is necessary to clearly show that the range can be measured in patients with an accuracy that is significantly better than the current uncertainty of ±3.5 % ±1 mm of the range in water at a 1.5$\sigma$ confidence level, which corresponds to ±3 mm to ±12 mm for clinical beam ranges. The range verification method based on prompt $\gamma$-ray spectroscopy was specifically tailored to verify the absolute range with a high precision.

We consider a range precision requirement of ±2 mm water equivalent, combining systematic and statistical uncertainties, appropriate for a first clinical prototype. This would show a clear improvement as compared to current treatment plan designs. Extrapolating the experimental results from the small scale prompt $\gamma$-ray detector, *in vivo* range verification of pencil-beams with this precision appears feasible with a detector that has a total $\gamma$-ray detection efficiency per incident proton about 5 times higher as compared to the small scale prototype. We believe such a system can be developed within reasonable constraints on size and cost.

The stated precision should be achieved at a dose level for a proton beam in a fractionated treatment, which is about 1 Gy. For clinical trials, without making changes to the currently delivered treatments, the detector should also support a clinical pencil-beam current, which is 2 nA at our institution, and continuously monitor the proton treatment during the delivery of all treatment beams.

A complete functioning clinical prototype system will need to consist of a number of different components. In addition to the detector and associated data acquisition electronics, a mechanical system is required to reliably position the detector close to the patient and at a location near the distal end of each treatment beam. The range verification method that we developed is not sensitive to the exact position of the measurement plane relative to the patient, as long as the actual position is known accurately. A tolerance of a few millimetres on the positioning of the detector would

be acceptable for a clinical trial, if the detector setup is followed by a procedure to measure the position of the detector relative to isocenter.

Software needs to be developed to create a workflow to calculate the patient-specific prompt $\gamma$-ray models and to determine the absolute range of the beam. Monte Carlo methods will be suitable to develop these models, which can be calculated off-line for the first clinical trials. A sufficient level of integration with the clinical planning systems is needed to support the evaluation of different algorithms applied to the treatment plans and the acquired *in vivo* measurements.

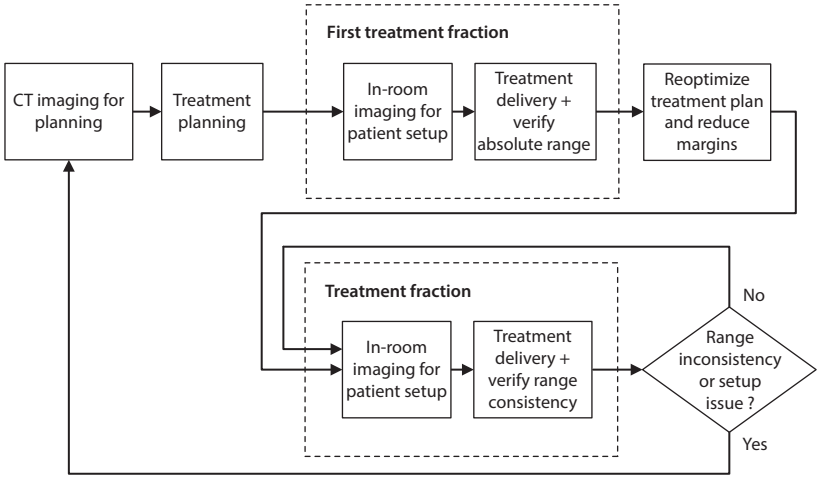
Further research is recommend to develop an end-to-end validation of a clinical prototype system, to assess the accuracy of the absolute proton range that is determined from the prompt $\gamma$-ray measurements. This validation will need to be based on a complex anthropomorphic or biological phantom in which direct dose measurements can be cross checked against the indirect verification through prompt $\gamma$-ray spectroscopy.

### 9.6.2. Range verification workflow

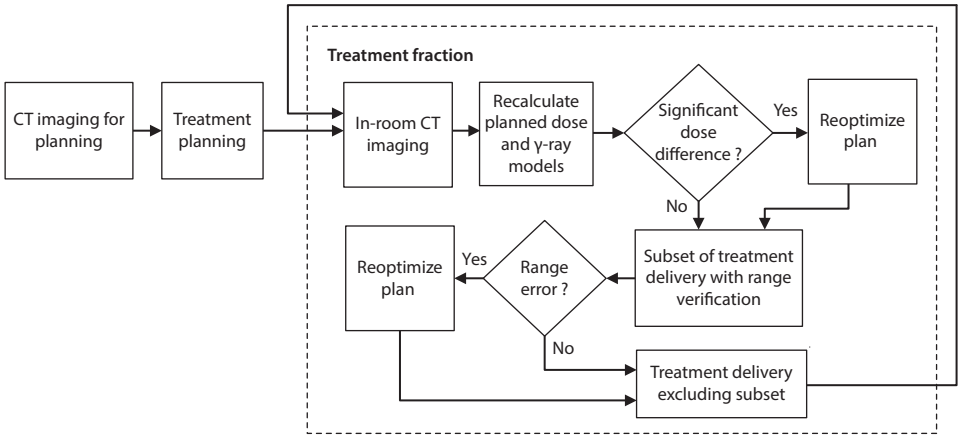
Based on the prompt $\gamma$-ray measurements that are performed during treatment, there are two main questions that need to be answered. First, is the delivered dose consistent with the treatment plan? And if not, which adaptations should be made to achieve the desired treatment?

Different potential methods to adapt proton therapy treatments based on *in vivo* prompt $\gamma$-ray measurements are envisioned, which vary in complexity and requirements. It is probably fair to say that the exact magnitude of the range uncertainty, including the systematic and random components, is not yet completely understood at this time. Even without adapting treatments, the ability to monitor the range of the beam *in vivo* will initially provide valuable data on the actual uncertainties that exist, which can be used to better evaluate different adaptation strategies.

A large component of the range uncertainty is likely systematic and affects all treatment fractions that are delivered. This is the uncertainty that relates to the estimation of the stopping power of the patient's anatomy in the beam path. The systematic uncertainty can be reduced based on *in vivo* measurements using an off-line process as shown in figure 9.1a. After treatment planning is performed, the first treatment fraction could be delivered using conventional range margins but with *in vivo* prompt $\gamma$-ray measurements for the verification of the absolute range of pencil-beams. After the treatment fraction has been delivered, the treatment plan can then be re-optimized and designed with smaller range margins. In this off-line process, anatomical imaging, using for example orthogonal x-rays, is performed to align the patient's anatomy with the proton beams. *In vivo* range verification can further be continuously used to monitor the consistency of the delivered range. When the patient's anatomy is found to be significantly different from the initial




(a) Off-line adaptation to reduce the systematic patient-specific range error.




(b) On-line adaptation using both in-room anatomical imaging and *in vivo* range monitoring.

**Figure 9.1 |** Incorporation of *in vivo* treatment verification in the proton therapy treatment process.

treatment plan, or when the range is found to be inconsistent, a new treatment plan is created for the remaining treatment fractions.

Range verification based on prompt $\gamma$-ray measurements generally provides a measurement of the range relative to the coordinate system of the treatment room. If anatomical changes are likely during the course of treatment, it is as important to ensure the treatment plan reflects the current anatomy. For this reason, CT scan equipment is nowadays being incorporated in proton therapy treatment rooms.

A potential workflow, in which anatomical imaging and *in vivo* range verification are integrated and used to adapt the treatment during each fraction, is shown in figure 9.1b. After an in-room CT scan, the treatment plan is first optimized based on the updated information on the patient's anatomy. Then, a number of proton pencil-beams is delivered, potentially somewhat proximal of the end-of-range such that these will certainly be on target. Based on prompt $\gamma$-ray measurements acquired during the delivery of these pencil-beams, the remaining set of pencil-beams to be delivered can be optimized, incorporating both the new anatomical information and *in vivo* range data. This treatment adaptation could also be performed at a regular interval instead of for every treatment fraction, depending on the changes that are expected. The workflow requires quick re-optimization of the treatment plan while the patient is in the treatment room.

# References


[1] A. Brahme, J.E. Roos, and I. Lax. Solution of an integral equation encountered in rotation therapy. *Phys. Med. Biol.* **27**, 1221 (1982).

[2] C.X. Yu. Intensity-modulated arc therapy with dynamic multileaf collimation: an alternative to tomotherapy. *Phys. Med. Biol.* **40**, 1435 (1995).

[3] W.H. Bragg and R. Kleeman. On the ionization curves of radium. *Phil. Mag. A* **37**, 581–587 (1904).

[4] R.R. Wilson. Radiological use of fast protons. *Radiology* **47**, 487–491 (1946).

[5] J.H. Lawrence. Proton irradiation of the pituitary. *Cancer* **10**, 795–798 (1957).

[6] S. Falkmer, F. B., B. Larsson, A. Lindell, J. Naeslund, and S. Stenson. Pilot study on proton irradiation of human carcinoma. *Acta Radiol.* **58**, 33–51 (1962).

[7] R.N. Kjellberg, A. Shintani, A.G. Frantz, and B. Kliman. Proton-beam therapy in acromegaly. *N. Engl. J. Med.* **278**, 689–695 (1968).

[8] R.N. Kjellberg, T. Hanamura, K.R. Davis, S.L. Lyons, and R.D. Adams. Bragg-peak proton-beam therapy for arteriovenous malformations of the brain. *N. Engl. J. Med.* **309**, 269–274 (1983).

[9] H. Suit et al. Evaluation of the clinical applicability of proton beams in definitive fractionated radiation therapy. *Int. J. Radiat. Oncol. Biol. Phys.* **8**, 2199–2205 (1982).

[10] M. Goitein, H.D. Suit, E. Gragoudas, A.M. Koehler, and R. Wilson. Potential for low-LET charged-particle radiation therapy in cancer. *Radiat. Res.* **8**, S297–S309 (1985).

[11] Particle therapy co-operative group. Patient statistics, 2013. http://ptcog.ch.

[12] H. Bethe. Bremsformel für Elektronen relativistischer Geschwindigkeit. *Z. Phys* **76**, 293–299 (1932).

[13] W.H. Barkas, J.N. Dyer, and H.H. Heckman. Resolution of the $\Sigma^-$-mass anomaly. *Phys. Rev. Lett.* **11**, 26–28 (1963).

[14] F. Bloch. Zur Bremsung rasch bewegter Teilchen beim Durchgang durch Materie. *Ann. Phys.* **408**, 285–320 (1933).

[15] G. Molière. Theorie der Streuung schneller geladener Teilchen II. Mehrfach- und Vielfachstreuung. *Z. Naturforsch. A* **3**, 78–96 (1948).

[16] B. Gottschalk, A.M. Koehler, R.J. Schneider, J.M. Sisterson, and M.S. Wagner. Multiple Coulomb scattering of 160 MeV protons. *Nucl. Instrum. Meth. B* **74**, 467–490 (1993).

[17] M.R. Bussiere and J.A. Adams. Treatment planning for conformal proton radiation therapy. *Technol. Cancer Res. Treat.* **2**, 389–399 (2003).

[18] A. J. Lomax et al. Intensity modulated proton therapy: a clinical example. *Med. Phys.* **28**, 317–324 (2001).

[19] U. Schneider, E. Pedroni, and A. Lomax. The calibration of CT Hounsfield units for radiotherapy treatment planning. *Phys. Med. Biol.* **41**, 111–124 (1996).

[20] A. M. Koehler, R. J. Schneider, and J. M. Sisterson. Range modulators for protons and heavy ions. *Nucl. Instrum. Methods* **131**, 437–440 (1975).

[21] H. Jiang, B. Wang, G. Xu, H. D. Suit, and H. Paganetti. Simulation of organ-specific patient effective dose due to secondary neutrons in proton radiation treatment. *Phys. Med Biol.* **50**, 4337–4353 (2005).

[22] T. Kanai, K. Kawachi, and Y. Kumamota. Spot scanning system for proton therapy. *Med. Phys.* **7**, 365–369 (1982).

[23] C. F. Von Essen, H. Blattmann, G. Bodendoerfer, J. Mizoe, E. Pedroni, E. Walder, and A. Zimmermann. The Piotron: II. methods and initial results of dynamic pion therapy in phase II studies. *Int. J. Radiat. Oncol. Biol. Phys.* **11**, 217–226 (1985).

[24] H. Paganetti. Range uncertainties in proton therapy and the role of Monte Carlo simulations. *Phys. Med. Biol.* **57**, R99–R117 (2012).

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

[26] A. V. Chvetsov and S. L. Paige. The influence of CT image noise on proton range calculation in radiotherapy planning. *Phys. Med. Biol.* **55**, N141–N149 (2010).

[27] S. España and H. Paganetti. Uncertainties in planned dose due to the limited voxel size of the planning CT when treating lung tumors with proton therapy. *Phys. Med. Biol.* **56**, 3843–3856 (2011).

[28] F. Albertini, A. Bolsi, A. J. Lomax, H. P. Rutz, B. Timmerman, and G. Goitein. Sensitivity of intensity modulated proton therapy plans to changes in patient weight. *Radiother. Oncol.* **86**, 187–194 (2007).

[29] J. Unkelbach, T. Bortfeld, B. C. Martin, and M. Soukup. Reducing the sensitivity of IMPT treatment plans to setup errors and range uncertainties via probabilistic treatment planning. *Med. Phys.* **36**, 149–163 (2009).

[30] A. Knopf and A. Lomax. *In vivo* proton range verification: a review. *Phys. Med. Biol.* **58**, R131–R160 (2013).

[31] J. Pawelke, L. Byars, W. Enghardt, W. D. Fromm, H. Geissel, B. G. Hasch, K. Lauckner, P. Manfrass, D. Schardt, and M. Sobiella. The investigation of different cameras for in-beam PET imaging. *Phys. Med. Biol.* **41**, 279–296 (1996).

[32] D. W. Litzenberg, D. A. Roberts, M. Y. Lee, K. Pham, A. M. Van der Molen, R. Ronningen, and F. D. Becchetti. On-line monitoring of radiotherapy beams: Experimental results with proton beams. *Med. Phys.* **26**, 992–1006 (1999).

[33] J. Llacer. Positron emission medical measurements with accelerated radioactive ion beams. *Nucl. Sci. Appl.* **3**, 111–131 (1988).

[34] K. Parodi et al. Patient study of *in vivo* verification of beam delivery and range, using positron emission tomography and computed tomography imaging after proton therapy. *Int. J. Radiat. Oncol. Biol. Phys.* **68**, 920–934 (2007).

[35] A. Knopf, K. Parodi, H. Paganetti, T. Bortfeld, J. Daartz, M. Engelsman, N.J. Liebsch, and H.A. Shih. Accuracy of proton beam range verification using post-treatment positron emission tomography/computed tomography as function of treatment site. *Int. J. Radiat. Oncol. Biol. Phys.* **79**, 297–304 (2011).

[36] C.H. Min, C.H. Kim, M.Y. Youn, and J.W. Kim. Prompt gamma measurements for locating the dose falloff region in the proton therapy. *Appl. Phys. Lett.* **89**, 183517 (2006).

[37] H.M. Lu. A potential method for *in vivo* range verification in proton therapy treatment. *Phys. Med. Biol.* **53**, 1413–1424 (2008).

[38] M.F. Gensheimer, T.I. Yock, N.J. Liebsch, G.C. Sharp, H. Paganetti, N. Madan, P.E. Grant, and T. Bortfeld. *In vivo* proton beam range verification using spine MRI changes. *Int. J. Radiat. Oncol. Biol. Phys.* **78**, 268–275 (2010).

[39] Y. Yuan, O.C. Andronesi, T.R. Bortfeld, C. Richter, R. Wolf, A.R. Guimaraes, T.S. Hong, and J. Seco. Feasibility study of *in vivo* MRI based dosimetric verification of proton end-of-range for liver cancer patients. *Radiother. Oncol.* **106**, 378–382 (2013).

[40] L. Sulak et al. Experimental studies of the acoustic signature of proton beams traversing fluid media. *Nucl. Instrum. Methods* **161**, 203–217 (1979).

[41] K.C. Jones, A. Witztum, C.M. Sehgal, and S. Avery. Proton beam characterization by proton-induced acoustic emission: simulation studies. *Phys. Med. Biol.* **59**, 6549–6563 (2014).

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

[43] X. Jia, J. Schuemann, H. Paganetti, and S.B. Jiang. GPU-based fast Monte Carlo dose calculation for proton therapy. *Phys. Med. Biol.* **57**, 7783–7797 (2012).

[44] M. Yang, G. Virshup, J. Clayton, X.R. Zhu, R. Mohan, and L. Dong. Theoretical variance analysis of single- and dual-energy computed tomography methods for calculating proton stopping power ratios of biological tissues. *Phys. Med. Biol.* **55**, 1343 (2010).

[45] A.M. Koehler. Proton radiography. *Science* **160**, 303–304 (1968).

[46] R. Schulte et al. Conceptual design of a proton computed tomography system for applications in proton radiation therapy. *IEEE Trans. Nucl. Sci.* **51**, 866–872 (2004).

[47] C. Talamonti et al. Proton radiography for clinical applications. *Nucl. Instrum. Meth. A* **612**, 571–575 (2010).

[48] M. Testa, J.M. Verburg, M. Rose, C.H. Min, S. Tang, E.H. Bentefour, H. Paganetti, and H.M. Lu. Proton radiography and proton computed tomography based on time-resolved dose measurements. *Phys. Med. Biol.* **58**, 8215–8233 (2013).

[49] T. Plautz et al. 200 MeV proton radiography studies with a hand phantom using a prototype proton CT scanner. *IEEE Trans. Med. Imag.* **33**, 875–881 (2014).

[50] S. España and H. Paganetti. The impact of uncertainties in the CT conversion algorithm when predicting proton beam ranges in patients from dose and PET-activity distributions. *Phys. Med. Biol.* **55**, 7557–7571 (2010).

[51] J.C. Polf, S. Peterson, G. Ciangaru, M. Gillin, and S. Beddar. Prompt gamma-ray emission from biological tissues during proton irradiation: a preliminary study. *Phys. Med. Biol.* **54**, 731–743 (2009).

[52] M. Testa et al. Real-time monitoring of the Bragg-peak position in ion therapy by means of single photon detection. *Radiat. Environ. Biophys.* **49**, 337–343 (2010).

[53] V. Bom, L. Joulaeizadeh, and F.J. Beekman. Real-time prompt gamma monitoring in spot-scanning proton therapy using imaging through a knife-edge-shaped slit. *Phys. Med. Biol.* **57**, 297–308 (2012).

[54] E. Seravalli et al. Monte Carlo calculations of positron emitter yields in proton radiotherapy. *Phys. Med. Biol.* **57**, 1659–1673 (2012).

[55] R. Capote et al. RIPL – Reference input parameter library for calculation of nuclear reactions and nuclear data evaluations. *Nucl. Data Sheets* **110**, 3107–3213 (2009).

[56] R.J. Murphy, B. Kozlovsky, J. Kiener, and G.H. Share. Nuclear gamma-ray de-excitation lines and continuum from accelerated-particle interactions in solar flares. *Astrophys. J. Suppl. Ser.* **183**, 142–155 (2009).

[57] M. Herman, R. Capote, B.V. Carlson, P. Oblozinsky, M. Sin, A. Trkov, H. Wienke, and V. Zerkin. EMPIRE: Nuclear reaction model code system for data evaluation. *Nucl. Data Sheets* **108**, 2655–2715 (2007).

[58] NRG Petten. TALYS 1.4, 2012. http://www.talys.eu.

[59] G. Folger, V.N. Ivanchenko, and J.P. Wellisch. The binary cascade – nucleon nuclear reactions. *Eur. Phys. J. A* **21**, 407–417 (2004).

[60] M.B. Chadwick. ENDF nuclear data in the physical, biological, and medical sciences. *Int. J. Radiat. Oncol. Biol. Phys.* **88**, 10–14 (2012).

[61] M.B. Chadwick, P.G. Young, R.E. MacFarlane, P. Moller, G.M. Hale, R.C. Little, A.J. Koning, and S. Chiba. LA150 documentation of cross sections, heating, and damage. Technical Report LA-UR-99-1222, Los Alamos National Laboratory, 1999.

[62] M.B. Chadwick and P.G. Young. Proton nuclear interactions up to 250 MeV for radiation transport simulations of particle therapy. *J. Brachyherapy Int.* **13**, 89–93 (1997).

[63] P.G. Young, E.D. Arthur, and M.B. Chadwick. Comprehensive nuclear model calculations: Introduction to the theory and use of the GNASH code. Technical Report LA-12343-MS, Los Alamos National Laboratory, 1992.

[64] J. Allison et al. GEANT4 developments and applications. *IEEE Trans. Nucl. Sci.* **53**, 270–278 (2006).

[65] Los Alamos National Laboratory. MCNP6 beta 2: Monte Carlo N-particle extended, 2012.

[66] J.M. Quesada, V. Ivantchenko, A. Ivanchenko, M.A. Cortes-Giraldo, G. Folger, A. Howard, and D. Wright. Recent developments in pre-equilibrium and de-excitation models in GEANT4. *Prog. Nucl. Sci. Techn.* **2**, 936–941 (2011).

[67] C.Z. Jarlskog and H. Paganetti. Physics settings for using the GEANT4 toolkit in proton therapy. *IEEE Trans. Nucl. Sci.* **55**, 1018–1025 (2008).

[68] V.F. Weisskopf and D.H. Ewing. On the yield of nuclear reactions with heavy elements. *Phys. Rev.* **57**, 472–485 (1940).

[69] J.K. Tuli. Evaluated nuclear structure data file. *Nucl. Instrum. Meth. A* **369**, 506–510 (1996).

[70] H.P. Wellisch and D. Axen. Total reaction cross section calculations in proton-nucleus scattering. *Phys. Rev. C* **54**, 1329–1332 (1996).

[71] R.K. Tripathi, F.A. Cucinotta, and J.W. Wilson. Universal parameterization of absorption cross section: Light systems. Technical Report TP-1999-209726, NASA, 1999.

[72] R.E. Prael. A new nuclear structure library for MCNPX and LAHET3. In *Proceedings of the Fourth International Topical Meeting on Nuclear Applications of Accelerator Technology*, pages 350–352, Washington DC, 2000.

[73] R.E. Prael and M. Bozoian. Adaptation of the multistage pre-equilibrium model for the Monte Carlo method. Technical Report LA-UR-88-3238, Los Alamos National Laboratory, 1988.

[74] A.J. Koning and J.P. Delaroche. Local and global nucleon optical models from 1 keV to 200 MeV. *Nucl. Phys. A* **713**, 231–310 (2003).

[75] P.G. Young. Status of optical model activities at Los Alamos National Laboratory. Technical Report LA-UR-95-3654, Los Alamos National Laboratory, 1995.

[76] D.G. Madland. Progress in the development of global medium-energy nucleon-nucleus optical model potentials. In *Proceedings of the OECD/NEA specialists' meeting on the nucleon-nucleus optical model up to* 200 MeV, page 129, Paris, 1997.

[77] R.K. Tripathi. Universal parametrization of absorption cross sections. Technical Report 3621, NASA, 1997.

[78] W. Hauser and H. Feshbach. The inelastic scattering of neutrons. *Phys. Rev.* **87**, 366–373 (1952).

[79] F. Ajzenbergselove. Energy-levels of light-nuclei A = 11–12. *Nucl. Phys. A* **506**, 1–158 (1990).

[80] D.R. Tilley, H.R. Weller, and C.M. Cheves. Energy-levels of light-nuclei A = 16–17. *Nucl. Phys. A* **564**, 1–183 (1993).

[81] P. Dyer, D. Bodansky, A.G. Seamster, E.B. Norman, and D.R. Maxson. Cross-sections relevant to gamma-ray astronomy – proton-induced reactions. *Phys. Rev. C* **23**, 1865–1882 (1981).

[82] J. Narayanaswamy, P. Dyer, S.R. Faber, and S.M. Austin. Production of 6.13 MeV gamma-rays from the $^{16}O(p, p'\gamma)^{16}O$ reaction at 23.7 and 44.6 MeV. *Phys. Rev. C* **24**, 2727–2730 (1981).

[83] F.L. Lang, C.W. Werntz, C.J. Crannell, J.I. Trombka, and C.C. Chang. Cross-sections for production of the 15.10 MeV and other astrophysically significant gamma-ray lines through excitation and spallation of $^{12}$C and $^{16}$O with protons. *Phys. Rev. C* **35**, 1214–1227 (1987).

[84] J. Kiener, M. Berheide, N.L. Achouri, A. Boughrara, A. Coc, A. Lefebvre, F.D.O. Santos, and C. Vieu. Gamma-ray production by inelastic proton scattering on $^{16}$O and $^{12}$C. *Phys. Rev. C* **58**, 2174–2179 (1998).

[85] A. Belhout et al. Gamma-ray production by proton and alpha-particle induced reactions on $^{12}$C, $^{16}$O, $^{24}$Mg, and Fe. *Phys. Rev. C* **76**, 034607 (2007).

[86] H. Benhabiles-Mezhoud et al. Measurements of nuclear gamma-ray line emission in interactions of protons and alpha particles with N, O, Ne, and Si. *Phys. Rev. C* **83**, 024603 (2011).

[87] K.T. Lesko, E.B. Norman, R.M. Larimer, S. Kuhn, D.M. Meekhof, S.G. Crane, and H.G. Bussell. Measurements of cross-sections relevant to gamma-ray line astronomy. *Phys. Rev. C* **37**, 1808–1817 (1988).

[88] ICRP. Reference man: Anatomical, physiological and metabolic characteristics. Technical Report 23, International Commission on Radiological Protection, 1975.

[89] W. Bauhoff. Tables of reaction and total cross-section for proton nucleus scattering below 1 GeV. *At. Data Nucl. Data Tables* **35**, 429–447 (1986).

[90] A. Knopf, K. Parodi, T. Bortfeld, H.A. Shih, and H. Paganetti. Systematic analysis of biological and physical limitations of proton beam range verification with offline PET/CT scans. *Phys. Med. Biol.* **54**, 4477–4495 (2009).

[91] M. Moteabbed, S. España, and H. Paganetti. Monte Carlo patient study on the comparison of prompt gamma and PET imaging for range verification in proton therapy. *Phys. Med. Biol.* **56**, 1063–1082 (2011).

[92] J. Smeets et al. Prompt gamma imaging with a slit camera for real-time range control in proton therapy. *Phys. Med. Biol.* **57**, 3371–3405 (2012).

[93] C.H. Min, J.G. Park, and C.H. Kim. Development of an array-type prompt gamma detection system for the online measurement of the range of the proton beam in a patient: A Monte Carlo feasibility study. *J. Korean Phys. Soc.* **52**, 888–891 (2008).

[94] S.W. Peterson, D. Robertson, and J. Polf. Optimizing a three-stage Compton camera for measuring prompt gamma rays emitted during proton radiotherapy. *Phys. Med. Biol.* **55**, 6841–6856 (2010).

[95] A.K. Biegun et al. Time-of-flight neutron rejection to improve prompt gamma imaging for proton range verification: A simulation study. *Phys. Med. Biol.* **57**, 6429–6444 (2012).

[96] J.M. Verburg, H.A. Shih, and J. Seco. Simulation of prompt gamma-ray emission during proton radiotherapy. *Phys. Med. Biol.* **57**, 5459–5472 (2012).

[97] J.C. Polf, R. Panthi, D.S. Mackin, M. McCleskey, A. Saastamoinen, B.T. Roeder, and S. Beddar. Measurement of characteristic prompt gamma rays emitted from oxygen and carbon in tissue-equivalent samples during proton beam irradiation. *Phys. Med. Biol.* **58**, 5821–5831 (2013).

[98] C. Kanzow, N. Yamashita, and M. Fukushima. Levenberg-Marquardt methods with strong local convergence properties for solving nonlinear equations with convex constraints. *J. Comput. Appl. Math.* **172**, 375–397 (2004).

[99] M.I.A. Lourakis. Levmar: Levenberg-Marquardt nonlinear least squares algorithms in C/C++, 2013.

[100] B. Clasie, N. Depauw, M. Fransen, C. Gomà, H.R. Panahandeh, J. Seco, J.B. Flanz, and H.M. Kooy. Golden beam data for proton pencil-beam scanning. *Phys. Med. Biol.* **57**, 1147–1158 (2012).

[101] P. Crespo, T. Barthel, H. Frais-Kolbl, E. Griesmayer, K. Heidel, K. Parodi, J. Pawelke, and W. Enghardt. Suppression of random coincidences during in-beam PET measurements at ion beam radiotherapy facilities. *IEEE Trans. Nucl. Sci.* **52**, 980–987 (2005).

[102] C. Robert et al. Distributions of secondary particles in proton and carbon-ion therapy: a comparison between GATE/GEANT4 and FLUKA Monte Carlo codes. *Phys. Med. Biol.* **58**, 2879–2899 (2013).

[103] U. Oelfke, G.K.Y. Lam, and M.S. Atkins. Proton dose monitoring with PET: quantitative studies in lucite. *Phys. Med. Biol.* **41**, 177–196 (1996).

[104] J.M. Verburg, K. Riley, T. Bortfeld, and J. Seco. Energy- and time-resolved detection of prompt gamma-rays for proton range verification. *Phys. Med. Biol.* **58**, L37–L49 (2013).

[105] K. Parodi, F. Ponisch, and W. Enghardt. Experimental study on the feasibility of in-beam PET for accurate monitoring of proton therapy. *IEEE Trans. Nucl. Sci.* **52**, 778–786 (2005).

[106] A. Miyatake, T. Nishio, T. Ogino, N. Saijo, H. Esumi, and M. Uesaka. Measurement and verification of positron emitter nuclei generated at each treatment site by target nuclear fragment reactions in proton therapy. *Med. Phys.* **37**, 4445–4455 (2010).

[107] A.B. Clegg, G.L. Salmon, R.E. Segel, and K.J. Foley. Gamma radiation from medium energy proton bombardment of lithium, beryllium, boron, carbon and nitrogen. *Proc. Phys. Soc.* **78**, 681–694 (1961).

[108] K.J. Foley, A.B. Clegg, and G.L. Salmon. Gamma radiation from bombardment of $^{16}$O and $^{19}$F nuclei with 150 MeV protons. *Nucl. Phys.* **31**, 43–52 (1962).

[109] J. Flanz, S. Durlacher, M. Goitein, A. Levine, P. Reardon, and A. Smith. Overview of the MGH–Northeast Proton Therapy Center plans and progress. *Nucl. Instrum. Meth. B* **99**, 830–834 (1995).

[110] S. Agarwal and K. Mierle. Ceres solver, 2014. http://ceres-solver.org.

[111] E. Sheldon and D.M. Van Patter. Compound inelastic nucleon and gamma-ray angular distributions for even- and odd-mass nuclei. *Rev. Mod. Phys.* **38**, 143–186 (1966).

[112] S. Becker, J. Bobin, and E.J. Candes. NESTA: A fast and accurate first-order method for sparse recovery. Technical report, California Institute of Technology, 2009.

[113] P.J. Huber. Robust estimation of a location parameter. *Ann. Math. Statist.* **35**, 73–101 (1964).

[114] J.M. Verburg and J. Seco. Proton range verification through prompt gamma-ray spectroscopy. *Phys. Med. Biol.* **59**, 7089–7106 (2014).

[115] M. Testa, C.H. Min., J.M. Verburg, J. Schuemann., H.M. Lu, and H. Paganetti. Range verification of passively scattered proton beams based on prompt gamma time patterns. *Phys. Med. Biol* **59**, 4181–4195 (2014).

[116] C.H. Min, H.R. Lee, C.H. Kim, and S.B. Lee. Development of array-type prompt gamma measurement system for *in vivo* range verification in proton therapy. *Med. Phys.* **39**, 2100–2107 (2012).

[117] D. Mackin, S. Peterson, S. Beddar, and J. Polf. Evaluation of a stochastic reconstruction algorithm for use in Compton camera imaging and beam range verification from secondary gamma emission during proton therapy. *Phys. Med. Biol* **57**, 3537–3553 (2012).

[118] M.H. Richard et al. Design guidelines for a double scattering compton camera for prompt-γ imaging during ion beam therapy: A Monte Carlo simulation study. *IEEE Trans. Nucl. Sci.* **58**, 87–94 (2011).

[119] F. Hueso-González et al. Test of Compton camera components for prompt gamma imaging at the ELBE bremsstrahlung beam. *J. Instrum.* **9**, P05002 (2014).

[120] J. Krimmer et al. Development of a Compton camera for medical applications based on silicon strip and scintillation detectors. *Nucl. Instr. Meth. Phys. Res. A* (2014). http://dx.doi.org/10.1016/j.nima.2014.11.042.

[121] M. Engelsman, H.M. Lu, D. Herrup, M. Bussiere, and H.M. Kooy. Commissioning a passive-scattering proton therapy nozzle for accurate SOBP delivery. *Med. Phys.* **36**, 2172–2180 (2009).

[122] 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. *Med. Phys.* **31**, 2107–2118 (2004).

[123] H. Paganetti, H. Jiang, K. Parodi, R. Slopsema, and M. Engelsman. Clinical implementation of full Monte Carlo dose calculation in proton beam therapy. *Phys. Med. Biol.* **53**, 4825–4853 (2008).

[124] M. Fippel and M. Soukup. A Monte Carlo dose calculation algorithm for proton therapy. *Med. Phys.* **31**, 2263–2273 (2004).

[125] A. Stankovskiy, S. Kerhoas-Cavata, R. Ferrand, C. Nauraye, and L. Demarzi. Monte Carlo modelling of the treatment line of the proton therapy center in Orsay. *Phys. Med. Biol.* **54**, 2377–2394 (2009).

[126] M.G. Pia, M. Begalli, A. Lechner, L. Quintieri, and P. Saracco. Physics-related epistemic uncertainties in proton depth dose simulation. *IEEE Trans. Nucl. Sci.* **57**, 2805–2830 (2010).

[127] J. Bauer et al. Integration and evaluation of automated Monte Carlo simulations in the clinical practice of scanned proton and carbon ion beam therapy. *Phys. Med. Biol.* **59**, 4635–4659 (2014).

[128] T. Akagi et al. Integration and evaluation of automated Monte Carlo simulations in the clinical practice of scanned proton and carbon ion beam therapy. *Prog. Nucl. Sci. Techn.* **4**, 896–900 (2014).

[129] J. Shin, J. Perl, J. Schuemann, H. Paganetti, and B.A. Faddegon. A modular method to handle multiple time-dependent quantities in Monte Carlo simulations. *Phys. Med. Biol.* **57**, 3295–3308 (2012).

[130] ICRU. Stopping powers and ranges for protons and alpha particles. Technical Report 49, International Commission on Radiation Units & Measurements, 1993.

[131] J. Ramos-Mendez, J. Perl, B. Faddegon, J. Schuemann, and H. Paganetti. Geometrical splitting technique to improve the computational efficiency in Monte Carlo calculations for proton therapy. *Med. Phys.* **40**, 041718 (2013).

[132] H. Paganetti. Monte Carlo calculations for absolute dosimetry to determine machine outputs for proton therapy fields. *Phys. Med. Biol.* **51**, 2801–2812 (2006).

[133] J.F. Janni. Energy loss, range, path length, time-of-flight, straggling, multiple scattering, and nuclear interaction probability. In two parts. Part 1. For 63 compounds. *At. Data and Nucl. Data Tables* **27**, 147–339 (1982).

[134] B. Bednarz, H.M. Lu, M. Engelsman, and H. Paganetti. Uncertainties and correction methods when modeling passive scattering proton therapy treatment heads with Monte Carlo. *Phys. Med. Biol.* **56**, 2837–2854 (2011).

[135] L. Grevillot, D. Bertrand, F. Dessy, N. Freud, and D. Sarrut. A Monte Carlo pencil beam scanning model for proton treatment plan simulation using GATE/GEANT4. *Phys. Med. Biol.* **56**, 5203–5219 (2011).

[136] S.J. Dowdell, B. Clasie, N. Depauw, P. Metcalfe, A.B. Rosenfeld, H.M. Kooy, J.B. Flanz, and H. Paganetti. Monte Carlo study of the potential reduction in out-of-field dose using a patient-specific aperture in pencil beam scanning protontherapy. *Phys. Med. Biol.* **57**, 2829–2842 (2012).

[137] K. Parodi, A. Mairani, S. Brons, B.G. Hasch, F. Sommerer, J. Naumann, O. Jakel, T. Haberer, and J. Debus. Monte Carlo simulations to support start-up and treatment planning of scanned proton and carbon ion therapy at a synchrotron-based facility. *Phys. Med. Biol.* **57**, 3759–3784 (2012).

[138] J. Schuemann, H. Paganetti, J. Shin, B. Faddegon, and J. Perl. Efficient voxel navigation for proton therapy dose calculation in TOPAS and GEANT4. *Phys. Med. Biol.* **57**, 3281–3293 (2012).

[139] W. Schneider, T. Bortfeld, and W. Schlegel. Correlation between CT numbers and tissue parameters needed for Monte Carlo simulations of clinical dose distributions. *Phys. Med. Biol.* **45**, 459–478 (2000).

[140] H. Paganetti. Dose to water versus dose to medium in proton beam therapy. *Phys. Med. Biol.* **54**, 4399–4421 (2009).

[141] S. Agosteo, C. Birattari, M. Caravaggio, M. Silari, and G. Tosi. Secondary neutron and photon dose in proton therapy. *Radiother. Oncol.* **48**, 293–305 (1998).

[142] G. Dedes, M. Pinto, D. Dauvergne, N. Freud, J. Krimmer, J.M. Létang, C. Ray, and E. Testa. Assessment and improvements of GEANT4 hadronic models in the context of prompt-gamma hadrontherapy monitoring. *Phys. Med. Biol.* **59**, 1747–1772 (2014).

[143] H.M. Kooy, B.M. Clasie, H.M. Lu, T.M. Madden, H. Bentefour, N. Depauw, J.A. Adams, A.V. Trofimov, D.D.T.F. Delaney, and J.B. Flanz. A case study in proton pencil-beam scanning delivery. *Int. J. Radiat. Oncol. Biol. Phys.* **76**, 624–630 (2010).

[144] J. Perl, J. Shin, J. Schuemann, B. Faddegon, and H. Paganetti. TOPAS: an innovative proton Monte Carlo platform for research and clinical applications. *Med. Phys.* **39**, 6818–6837 (2012).

[145] M. Testa, J. Schuemann, H.M. Lu, J. Shin, B. Faddegon, J. Perl, and H. Paganetti. Experimental validation of the TOPAS Monte Carlo system for passive scattering proton therapy. *Med. Phys.* **40**, 121719 (2013).

[146] J. Schuemann, S. Dowdell, C. Grassberger, C.H. Min, and H. Paganetti. Site-specific range uncertainties caused by dose calculation algorithms for proton therapy. *Phys. Med. Biol.* **59**, 4007–4031 (2014).

[147] D. Giantsoudi, C. Grassberger, D. Craft, A. Niemierko, A. Trofimov, and H. Paganetti. Linear energy transfer-guided optimization in intensity modulated proton therapy: feasibility study and clinical potential. *Int. J. Radiat. Oncol. Biol. Phys.* **87**, 216–222 (2013).

[148] S. Dowdell, C. Grassberger, G.C. Sharp, and H. Paganetti. Interplay effects in proton scanning for lung: a 4D Monte Carlo study assessing the impact of tumor and beam delivery parameters. *Phys. Med. Biol.* **58**, 4137–4156 (2013).

[149] Y. Jia, C. Beltran, D.J. Indelicato, S. Flampouri, Z. Li, and T.E. Merchant. Proton therapy dose distribution comparison between Monte Carlo and a treatment planning system for pediatric patients with ependymoma. *Med. Phys.* **39**, 4742–4747 (2012).

[150] J.M. Verburg and J. Seco. Dosimetric accuracy of proton therapy for chordoma patients with titanium implants. *Med. Phys.* **40**, 071727 (2013).

[151] C. Grassberger, J. Daartz, S. Dowdell, T. Ruggieri, G. Sharp, and H. Paganetti. Quantification of proton dose calculation accuracy in the lung. *Int. J. Radiat. Oncol. Biol. Phys.* **89**, 424–430 (2014).

[152] M. Bazalova, L. Beaulieu, S. Palefsky, and F. Verhaegen. Correction of CT artifacts and its influence on Monte Carlo dose calculations. *Med. Phys.* **34**, 2119–2132 (2007).

[153] B. de Man, J. Nuyts, P. Dupont, G. Marchal, and P. Suetens. Metal streak artifacts in x-ray computed tomography: A simulation study. *IEEE Trans. Nucl. Sci.* **46**, 691–696 (1999).

[154] W.A. Kalender, R. Hebel, and J. Ebersberger. Reduction of CT artifacts caused by metallic implants. *Radiology* **164**, 576–577 (1987).

[155] E. Meyer, R. Raupach, M. Lell, B. Schmidt, and M. Kachelriess. Normalized metal artifact reduction (NMAR) in computed tomography. *Med. Phys.* **37**, 5482–5493 (2010).

[156] O. Watzke and W.A. Kalender. A pragmatic approach to metal artifact reduction in CT: merging of metal artifact reduced images. *Eur. Radiol.* **14**, 849–856 (2004).

[157] S.Y. Zhao, D.D. Roberston, G. Wang, B. Whiting, and K.T. Bae. X-ray CT metal artifact reduction using wavelets: an application for imaging total hip prostheses. *IEEE Trans. Med. Imag.* **19**, 1238–1247 (2000).

[158] G.H. Glover and N.J. Pelc. An algorithm for the reduction of metal clip artifacts in CT reconstructions. *Med. Phys.* **8**, 799–807 (1981).

[159] C. Xu, F. Verhaegen, D. Laurendeau, S.A. Enger, and L. Beaulieu. An algorithm for efficient metal artifact reductions in permanent seed implants. *Med. Phys.* **38**, 47–56 (2011).

[160] W.J.H. Veldkamp, R.M.S. Joemai, A.J. van der Molen, and J. Geleijns. Development and validation of segmentation and interpolation techniques in sinograms for metal artifact suppression in CT. *Med. Phys.* **37**, 620–628 (2010).

[161] J. Prince and A.S. Willsky. Constrained sinogram restoration for limited-angle tomography. *Opt. Eng.* **29**, 535–544 (1990).

[162] Z. Zhang, J. Wang, and L. Xing. Metal artifact reduction in x-ray computed tomography (CT) by constrained optimization. *Med. Phys.* **38**, 701–711 (2011).

[163] F.E. Boas and D. Fleischmann. Evaluation of two iterative techniques for reducing metal artifacts in computed tomography. *Radiology* **259**, 894–902 (2011).

[164] J.F. Williamson, B.R. Whiting, J. Benac, R.J. Murphy, G.J. Blaine, J.A. O'Sullivan, D.G. Politte, and D.L. Snyder. Prospects for quantitative computed tomography imaging in the presence of foreign metal bodies using statistical image reconstruction. *Med. Phys.* **29**, 2404–2418 (2002).

[165] M. Yazdi, M.A. Lari, G. Bernier, and L. Beaulieu. An opposite view data replacement approach for reducing artifacts due to metallic dental objects. *Med. Phys.* **38**, 2275–2281 (2011).

[166] M. Oehler, B. Kratz, T. Knopp, J. Müller, and T. Buzug. Evaluation of surrogate data quality in sinogram-based CT metal-artifact reduction. In *Proceedings of SPIE* volume 7076, page 707607, 2008.

[167] C. Lemmens, D. Faul, and J. Nuyts. Suppression of metal artifacts in CT using a reconstruction procedure that combines MAP and projection completion. *IEEE Trans. Med. Imag.* **28**, 250–260 (2009).

[168] L.I. Rudin, S. Osher, and E. Fatemi. Nonlinear total variation based noise removal algorithms. *Physica D* **60**, 259–268 (1992).

[169] Y. Nesterov. Smooth minimization of non-smooth functions. *Math. Program.* **103**, 127–152 (2005).

[170] L. Gerward, N. Guilbert, K.B. Jensen, and H. Levring. X-ray absorption in matter. Reengineering XCOM. *Radiat. Phys. Chem.* **60**, 23–24 (2001).

[171] M. Hollander and D. Wolfe. *Nonparametric statistical methods.* Wiley-Interscience, 2nd edition, 1999.

[172] M. Urie, M. Goitein, and M. Wagner. Compensating for hetrogeneities in proton radiation-therapy. *Phys. Med. Biol* **29**, 553–566 (1984).

[173] O. Jakel and P. Reiss. The influence of metal artefacts on the range of ion beams. *Phys. Med. Biol.* **52**, 635–644 (2007).

[174] W.D. Newhauser, A. Giebeler, K.M. Langen, D.M. Mirkovic, and R. Mohan. Can megavoltage computed tomography reduce proton range uncertainties in treatment plans for patients with large metal implants? *Phys. Med. Biol.* **53**, 2327–2344 (2008).

[175] A. Staab, H.P. Rutz, C. Ares, B. Timmermann, R. Schneider, A. Bolsi, F. Albertini, A. Lomax, G. Goitein, and E. Hug. Spot-scanning-based proton therapy for extracranial chordoma. *Int. J. Radiat. Oncol. Biol. Phys.* **81**, e489–e496 (2011).

[176] T.F. DeLaney et al. Phase II study of high-dose photon/proton radiotherapy in the management of spine sarcomas. *Int. J. Radiat. Oncol. Biol. Phys.* **74**, 732–739 (2009).

[177] M. Yang, X.R. Zhu, P.C. Park, U. Titt, R. Mohan, G. Virshup, J.E. Clayton, and L. Dong. Comprehensive analysis of proton range uncertainties related to patient stopping-power-ratio estimation using the stoichiometric calibration. *Phys. Med. Biol.* **57**, 4095–4115 (2012).

[178] V.L. Highland. Some practical remarks on multiple scattering. *Nucl. Instrum. Methods* **129**, 497–499 (1975).

[179] P.J. Bulman and J.A.R. Griffith. Proton total reaction cross sections at 9.1 MeV. *Nucl. Phys. A* **111**, 315–3120 (1968).

[180] V. Meyer and N.M. Hintz. Charged particle and total reaction cross sections for protons at 9.85 MeV. *Phys. Rev. Lett.* **5**, 207–209 (1960).

[181] B.D. Wilkins and G. Igo. 10 MeV proton reaction cross sections for several elements. *Phys. Rev.* **129**, 2198–2206 (1963).

[182] T.N. Nasr, A.M. Sourkes, D.J. Margaziotis, R.F. Carlson, and A.J. Cox. Measurements of the total reaction cross section for protons on Ti and B between 20 and 50 MeV. *Can. J. Phys.* **56**, 56–62 (1978).

[183] B. Gottschalk. BGware version 04aug09, 2009.

[184] A.O. Hanson, L.H. Lanzl, E.M. Lyman, and M.B. Scott. Measurement of multiple scattering of 15.7 MeV electrons. *Phys. Rev.* **84**, 634–637 (1951).

[185] G.O. Sawakuchi, U. Titt, D. Mirkovic, and R. Mohan. Density heterogeneities and the influence of multiple Coulomb and nuclear scatterings on the Bragg peak distal edge of proton therapy beams. *Phys. Med. Biol.* **53**, 4605–4619 (2008).

[186] U. Schneider, P. Pemler, J. Besserer, M. Dellert, M. Moosburger, J. de Boer, E. Pedroni, and T. Boehringer. The water equivalence of solid materials used for dosimetry with small proton beams. *Med. Phys.* **29**, 2946–2951 (2002).

[187] 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).

[188] H. Szymanowski and U. Oelfke. Two-dimensional pencil beam scaling: An improved proton dose algorithm for heterogeneous media. *Phys. Med. Biol.* **47**, 3313–3330 (2002).

# Summary

**Reducing Range Uncertainty in Proton Therapy**

Proton therapy as a modality for cancer radiotherapy offers important advantages over conventional radiotherapy using x-rays. X-ray beams aimed at a tumour also deliver dose to tissue proximal and distal to the tumour. Protons, on the other hand, have a finite range and deliver the highest dose near the end of their range. By using protons with different energies, the high-dose region can be shaped to the tumour not only laterally but also in depth, thereby reducing dose proximal to the tumour and sparing the uninvolved distal tissue from harmful radiation. This enables a higher dose to be delivered to the tumour if needed to improve local control, or to lower the integral dose the patient, thereby reducing the chance of side effects.

Because the range of the proton beam in the patient is subject to uncertainties, the finite range and the resulting sharp distal dose gradient also pose challenges. An important source of range uncertainty is the conversion from computed tomography images to proton stopping powers for treatment planning. No exact relation exists for this conversion, because of degeneracy and differences in the tissue compositions of patients. In addition, the dose calculation and imaging methods currently in clinical use have limitations in the presence of inhomogeneous tissue or implants.

Since the proton beam delivers no dose beyond the end-of-range, it is essential that the range covers the distal end of the tumour. Therefore, in clinical practice, safety margins are applied of about 3.5 % of the range is water, which results in a margin up to 1 cm for deeper seated tumours. Also, to avoid delivering high dose to organs-at-risk, the sharp distal edge of the beam is not normally used to spare organs in close proximity to the tumour. The requirement to use the lateral edge leads to a shallower dose gradient and precludes the use of certain beam angles which would otherwise be favourable for sparing of normal tissue.

In this thesis, we investigated the reduction of range uncertainty, the goal of which is to fully utilize the physical advantages of proton beams, to enable better treatments with smaller margins. *In vivo* verification of the proton range will be key to reduce the uncertainties and to safely deliver treatments with small margins. We developed a method to verify the range *in vivo* by performing spectroscopy of prompt $\gamma$-rays from proton-nuclear reactions. Although the primary proton beam stops inside the patient, these $\gamma$-rays can escape the body and therefore can be exploited to monitor the beam. They are emitted virtually instantaneously during treatment as the excited residual nuclei decay to their ground state, enabling real-time verification.

We performed experimental and theoretical studies to investigate the quantitative detection of discrete prompt $\gamma$-rays from specific proton reactions with the main elements in human tissue. A small scale prototype prompt $\gamma$-ray detection system was developed, consisting of an actively shielded scintillator with tungsten collimation and a digital data acquisition system. The $\gamma$-rays were resolved in dimensions of energy and of time relative to the accelerated proton bunches. The energy resolution allowed discrete $\gamma$-ray lines to be resolved, while the time resolution provided separation between proton- and neutron-induced events. Fifteen $\gamma$-ray line excitations from proton reactions with $^{12}$C and $^{16}$O could be resolved, for which differential cross sections were optimized.

A Monte Carlo simulation framework was developed to model the transport of the proton beam in phantoms and based on clinical treatment plans. Using the experimentally determined cross sections, these simulations can predict the detected prompt $\gamma$-rays in different scenarios. Detailed models of the proton interactions can also improve treatment planning, because they provide a more accurate dose calculation as compared to analytical algorithms, in particular when lateral heterogeneities within the area of a proton field result in a complex mix of different ranges.

To verify the absolute range of delivered proton beams, we developed an optimization method that fits the models of the proton interactions to measured prompt $\gamma$-ray spectra. The optimization method determined the absolute range of the proton beam without prior knowledge of the elemental composition of the irradiated tissue. The performance of the range verification method was assessed in different phantom experiments, using the small scale prototype detector. The absolute range determined from prompt $\gamma$-ray measurements agreed with dose measurements within a few tenths of a millimetre. A statistical accuracy on the order of one millimetre was achieved with typical proton pencil-beams, at a dose level about 5 times higher as compared to standard fractionated treatments.

For the specific case of patients with metallic implants, which are common when radiotherapy is used after surgical resection of spinal tumours, we also improved the image reconstruction of the computed tomography images used for treatment planning. The developed correction methods were shown to significantly reduce the artifacts which normally occur, enhancing the accuracy of dose calculation. Monte Carlo calculations also improved the dose accuracy distal to implants.

We conclude that using the developed technologies, a 1 mm to 2 mm accuracy of the absolute proton range at a $1.5\sigma$ confidence level, appears feasible for clinical proton fields delivered using proton pencil-beams. Further developments to facilitate a clinical trial are recommended, which will require a larger-scale prompt $\gamma$-ray detection system and an efficient workflow for patient-specific modelling of prompt $\gamma$-ray emissions.

# Acknowledgements

I am grateful for the experiences I gained during the last four years and would like to thank the people who have helped to make this thesis possible.

I would like to thank Thomas Bortfeld and Joao Seco for giving me the opportunity to join the Radiation Oncology physics research group at the MGH. I enjoyed the freedom I had to pursue my research interests and I am grateful for your input and for your enthusiasm about the experimental prompt $\gamma$-ray detection work. I appreciated your continued help to obtain funding for the experiments. I look forward to continuing to develop a system for clinical trials.

Thanks to Ward Cottaar and Pieter Wijn for facilitating my PhD defence in Eindhoven. I would also like to extend my gratitude to the other members of the Doctorate Committee: Frank Verhaegen, Martijn Engelsman, Jom Luiten, and Gerrit Kroesen.

A number of people have helped with the experiments. A special thanks to Mauro Testa, for helping out on several night shifts, and for brainstroming about the experiments. Ben Clasie, thanks for assisting with the pencil-beam scanning system. Thanks to Ethan Cascio and Bob Brett for their help to link the data acquisition with the dosimetry equipment. Tom Ruggieri, thanks for machining the hardware for detector alignment. I would like to thank Jay Flanz for proving beam time and other resources. Thanks to the operators at the FHBPTC for running the cyclotron during the night shifts. Kent Riley, who sadly passed away last year, provided valuable input when I was designing the experimental setup. Helen Shih and Hsiao-Ming Lu, thanks for helping with the funding of detector components.

I would like to acknowledge Jan Schümann, Clemens Grassberger, Harald Paganetti, and the TOPAS developers, for collaborating on the Monte Carlo dose calculation developments. Also thanks to Hanne Kooy and Tom Madden for providing the implementation of the pencil-beam dose calculation algorithm.

My colleagues in Boston have made my time as a PhD student very enjoyable. Thanks to everyone for the scientific and non-scientific discussions, the fun events, and the Friday drinks!

Finally, I want to thank my parents, Lidy and Chris, and my brother, Paul, for their support and encouragement.

# List of publications

**J.M. Verburg**, M. Testa, and J. Seco. Range verification of passively scattered proton beams using prompt gamma-ray detection. *Phys. Med. Biol.* **60**, 1019–1029 (2015).

**J.M. Verburg** and J. Seco. Proton range verification through prompt gamma-ray spectroscopy. *Phys. Med. Biol.* **59**, 7089–7106 (2014).

**J.M. Verburg**, K. Riley, T. Bortfeld, and J. Seco. Energy- and time-resolved detection of prompt gamma-rays for proton range verification. *Phys. Med. Biol.* **58**, L37–L49 (2013).

**J.M. Verburg** and J. Seco. Dosimetric accuracy of proton therapy for chordoma patients with titanium implants. *Med. Phys.* **40**, 071727 (2013).

**J.M. Verburg**, H.A. Shih, and J. Seco. Simulation of prompt gamma-ray emission during proton radiotherapy. *Phys. Med. Biol.* **57**, 5459–5472 (2012).

**J.M. Verburg** and J. Seco. CT metal artifact reduction method correcting for beam hardening and missing projections. *Phys. Med. Biol.* **57**, 2803–2818 (2012).

M. Testa, C.H. Min, **J.M. Verburg**, J. Schuemann, H.-M. Lu, and H. Paganetti. Range verification of passively scattered proton beams based on prompt gamma time patterns. *Phys. Med. Biol.* **59**, 4181–4196 (2014).

M. Testa, **J.M. Verburg**, M. Rose, C.H. Min, S. Tang, E.H. Bentefour, H. Paganetti, and H.-M. Lu. Proton radiography and proton computed tomography based on time-resolved dose measurements. *Phys. Med. Biol.* **58**, 8215–8233 (2013).

M.F. Spadea, **J.M. Verburg**, G. Baroni, and J. Seco. The impact of low-Z and high-Z metal implants in IMRT: A Monte Carlo study of dose inaccuracies in commercial dose algorithms. *Med. Phys.* **41**, 011702 (2013).

M.F. Spadea, **J.M. Verburg**, G. Baroni, and J. Seco. Dosimetric assessment of a novel metal artifact reduction method in CT images. *J. Appl. Clin. Med. Phys.* **41**, 011702 (2013).

# Curriculum vitae

Joost Mathijs Verburg was born on November 26th, 1985 in Nijmegen, The Netherlands. He received pre-university education (VWO) at Pantarijn in Wageningen, obtaining his diploma in 2004. In the same year, he started his studies at the Eindhoven University of Technology, where he received his B.Sc. degree in Applied Physics in 2009. For his bachelor's thesis, he performed a research internship at the Molecular Biosensing for Medical Diagnostics group at the same university, investigating magnetic nanoparticles for use in biosensors. During his studies, he was a member of the departmental education committee. He continued his graduate studies at the Eindhoven University of Technology, receiving a M.Sc. degree in Applied Physics in 2011, with a specialization in medical physics. During 2009 and 2010, he spent 15 months at the Department of Radiation Oncology at Harvard Medical School and Massachusetts General Hospital, to research computed tomography image reconstruction for radiotherapy treatment planning, which was the topic of his master's thesis. From 2011 to 2015, he was a Visiting Fellow at the Graduate School of Arts and Sciences at Harvard University. At the Francis H. Burr Proton Therapy Center at Massachusetts General Hospital, he performed research on the reduction of range uncertainty in proton radiotherapy, which is described in this thesis. He presented his work in peer-reviewed publications and at several international conferences and workshops. Two of his publications were selected as featured article in Physics in Medicine and Biology. At the 2014 IEEE Medical Imaging Conference in Seattle, WA, he received the second place student paper award.

Protons are increasingly used for cancer radiotherapy. A proton beam has a finite range and deposits maximum dose at a specific depth. High radiation doses can therefore be delivered to tumours, while better sparing surrounding healthy tissue. To fully utilize these advantages, the range of the proton beam in the patient needs to be established with a high accuracy, such that the dose distribution can be precisely shaped to the tumour.

In this work, technologies are developed to reduce the uncertainty in the range of proton beams. A new method is proposed to monitor the range *in vivo* and in real-time, by performing spectroscopy of prompt γ-rays from proton-nuclear reactions. CT image reconstruction algorithms and models of the proton interactions with tissue are also improved. The better knowledge of the proton range can enable more optimal and precise treatments.