HollandPTC Protonen Therapie Centrum

TU Delft

Erasmus MC Universitair Medisch Centrum Rotterdam

Delft University of Technology

Multi-dimensional Uncertainty Analysis for Proton FLASH Radiotherapy

Including machine-intrinsic uncertainties in pencil beam placement and cyclotron proton beam current

BM51035: BME MSc Thesis

Odin Kruithof

Multi-dimensional Uncertainty Analysis for Proton FLASH Radiotherapy

Including machine-intrinsic uncertainties in pencil beam placement and cyclotron proton beam current

by

Odin Kruithof

to obtain the degree of Master of Science
at the Delft University of Technology,
to be defended publicly on Wednesday August 23, 2023 at 3:00 PM.

Student number:	481768	
Project duration:	December 1, 2022 – August 23, 2023	
Thesis committee:	dr. ir. D. Lathouwers,	TUD, daily supervisor
	dr. S. Habraken,	HPTC & LUMC, daily supervisor
	prof. dr. M. Hoogeman,	EMC, HPTC, & TUD, committee member

This thesis is confidential and cannot be made public until August 16, 2025.

Cover:	An impressionist version of Girl with a Pearl Earring painted by Johannes Vermeer. The precision found in his paintings reflects the care offered to patients at the Holland Proton Therapy Centre.

Acknowledgements

I devoted the past nine months to research. This would not have been possible without the support of the people around me. First, I want to thank my daily supervisors Steven Habraken, Danny Lathouwers and Kees Spruijt. Steven, thank you for your expertise and guidance. Your suggestions and critical questions in our weekly meetings steered me in the right direction while keeping me in charge of the project. I also enjoyed our non-scientific-related discussions. Danny, thanks for your help with the simulations and the ODE. Without your expertise, the simulations would still be running and the ODE would be non-differentiable. Kees, thank you for providing me with beam measurement data from the HPTC. The measurements were essential for setting up the analytical dose model. Our collaboration for the abstract submission was also effective and pleasant. Next to my daily supervisors, I would also like to thank others who contributed to my project. Tiberiu Burlacu (a.k.a. Tibi), thanks for your help with the 3D dose engine. Thanks for making the engine useful for my project and for all the bug fixes. Sebastiaan Breedveld, thanks for the help with the installation of iCycle in FLASH mode.

I was lucky enough to do my project at three different institutes, the EMC, the HPTC and the TUD. During the weeks, I worked at different locations which created a vibrant environment I enjoyed and can recommend to others. Because I did my project at three different places, I met many people with whom I could spend the days. Thanks to my fellow MSc thesis students for their support and the fun. Thanks to the PhD students from the EMC. I enjoyed our lunches together and you showed me the ins and outs of doing a PhD. If you enjoy drinking coffee just like me, I would recommend spending most days at the HPTC. Here, the coffee is good (and free) in contrast to the EMC and TUD.

Lastly, I want to thank the people that stand close to me for their endless support. Thanks to my girlfriend Josephine for providing energy, motivation and structure. This project wouldn't have been the same without your help, especially in dealing with my sometimes not-so-good planning skills and chaotic thinking. I also want to thank my parents for their support over the past nine months. I enjoyed our trips to Rome and Drenthe. They gave me the energy to continue my work. I want to end with a paraphrase which I really like. It catches the essence of the first law of thermodynamics: ”Energy cannot be created or destroyed, it can only be changed from one form to another.”. All of you gave me the energy to complete this project.

Odin Kruithof
Delft, August 2023

Abstract

With FLASH proton radiotherapy, healthy tissue is spared more compared to conventional proton therapy. The FLASH effect is present at high fractional doses of more than 8 Gy, at ultra-high mean dose rates of more than 40 Gy per second and at dose delivery times of less than 200 milliseconds. However, particle accelerator intrinsic uncertainties can influence the FLASH effect negatively. FLASH proton therapy is affected by pencil beam positioning errors and by proton beam current fluctuations. In this work, these two types of uncertainties were added to a Gaussian two-dimensional analytical dose model. The impact of pencil beam placement errors was evaluated on dose deposition and related to clinically used metrics. These are the V95, V107, D2 and D98. The constraint on the V95 was violated for 5% of patients first. For all evaluated FLASH dose fields, the allowed Gaussian pencil beam placement error is less than 1.0-mm in standard deviation. A trade-off between target coverage and the level of FLASH effect was found. Proton beam current fluctuations were coupled to the pencil beam scanning dose rate (PBSDR98), the minimum dose rate for 98% of the target volume. For all evaluated FLASH fields the permitted beam current fluctuation was 6.8%. This is considered more achievable than the limit on pencil beam placement errors. Proton FLASH radiotherapy is feasible, but proton beam characterisation is required to quantify for which patients this is the case.

The PBSDR is sensitive to added uncertainties and contradicts our current understanding of the FLASH effect. In this thesis, a more robust metric is developed which depends on the radiolytic oxygen depletion hypothesis for the FLASH effect. This metric is called the FER_{eff}. The oxygen concentration over time in a cell is modelled with an ordinary differential equation (ODE) involving an oxygen depletion and re-oxygenation term. The FER_{eff} can be calculated with the oxygen concentration over time for a treatment plan. Currently, The onset of the FLASH effect lies between a prescribed dose of 8 Gy and 16 Gy. This is in line with the FLASH dose threshold. The shortcomings of the PBSDR are dealt with.

A three-dimensional uncertainty analysis can be done with a semi-analytical dose engine from the TUD. With this engine, changes in dose deposition are coupled to Hounsfield unit (HU) perturbations without the necessity of a full dose re-computation. To study the machine uncertainties, an EMC FLASH treatment plan formed with iCycle was recreated with the TUD dose engine. The iCycle optimisation results contain all the required information for a dose calculation. The gamma pass rates between the iCycle dose and engine dose were satisfactory for the target but not good enough for an organ-at-risk (OAR) yet. Similarity can be improved by adjusting the dose recreation procedure. When the iCycle dose and engine dose are comparable, a trustworthy clinical uncertainty analysis can be done. The idea is to simulate field and individual pencil beam placement errors by shifting the original patient CT to a virtual CT. First, it should be studied if the computed dose response is valid for a combination of CT shifts and HU perturbations. In a later stage, pencil beam placement errors should be coupled to the clinically used ICRU metrics and constraints. The impact of beam current fluctuations can also be studied in the future.

Contents

Nomenclature

Abbreviations

Abbreviation	Definition
3D-CRT	Three-dimensional conformal radiation therapy
CT	Computed tomography
DADR	Dose averaged dose rate [Gy/s]
EMC	Erasmus Medical Centre
FED	FLASH enhanced dose [Gy]
FER	FLASH enhancement ratio
FER_{eff}	Effective FLASH enhancement ratio
GTV	Gross target volume
HPTC	Holland Proton Therapy Centre
HU	Hounsfield unit
ICRU	International Commission on Radiation Units and Measurements
IMRT	Intensity-modulated radiation therapy
MC	Monte Carlo
MGH	Massachusetts General Hospital
MU	Monitoring unit
NTCP	Normal tissue complication probability
OAR	Organ at risk
ODE	Ordinary differential equation
OER	Oxygen enhancement ratio
PBSDR	Pencil beam scanning dose rate [Gy/s]
PTV	Planning target volume
QA	Quality assessment
TCP	Tumour control probability
TUD	Technical University of Delft
-VH	-Volume histogram
VMAT	Volumetric modulated arc therapy

Symbols

Symbol	Definition	Unit
C	Concentration	[mol m^{-3}]
C_{ss}	Steady-state concentration	[mol m^{-3}]
D	Dose	[Gy]
$\dot{D}$	Dose rate	[Gy s^{-1}]
DR	Dose rate	[Gy s^{-1}]
Dx	Dose-at-volume	[Gy]
d	Dose-deposition matrix	[Gy protons^{-1}]
E	Scientific notation	[-]
G	Gaussian function	[-]
H	Henry’s constant	[mol m^{-3} mmHg]
p	Partial pressure	[mmHg]
T	Irradiation time	[s]
Vx	Volume-at-dose	[m^{3}]

Symbol	Definition	Unit
w	Pencil beam weight	[protons]
α	Relative weight of pure pencil beam against halo	[-]
β	Halo fitting parameters	[-]
Γ	Low's Gamma index	[-]
γ	Gamma passing rate	[%]
ΔD_M	Gamma index dose threshold	[Gy]
Δd_M	Gamma index distance threshold	[m]
δ	Dose difference	[Gy]
η	Halo fitting parameters	[-]
θ	Gantry angle	[degrees]
κ	Radiolytic oxygen depletion model parameter	[mol m^{-3} Gy^{-1}]
μ	Mean of Gaussian function	[m]
σ	Standard deviation	[m]
τ	Re-oxygenation model parameter	[s]
ϕ	OER model parameter	[mmHg^{-1}]
ψ_O/ψ_D	Weight of oxygen-induced over direct cell death	[-]

1

Introduction

1.1. Background

The incidence of cancer in the European Union is expected to increase by more than 50% in a few decades. This increase is driven by several demographic factors like the overall ageing of the population and the increase in life expectancy. [48] Especially in high-income countries, the curing rates of various types of cancer are steadily increasing due to earlier detection and better treatments. [3] Nonetheless, this increase in overall survival is not big enough to tackle the growing incidence of cancer. As a result, cancer mortality will increase over the coming decades. Therefore, the long-term strategy is to constantly improve our cancer treatments. [4]

Treatment of cancer is done via three primary modalities. Surgery is used for the resection of non-metastasised tumours, radiotherapy is applied to achieve local and regional control of cancer with ionising radiation and chemotherapy is used to fight systemic cancers. In a curative treatment setting, these modalities are often combined. A typical treatment consists of local resection of the tumour and irradiation of potential micro-metastases for regional control. Furthermore, ongoing research is aiming for the improvement of curative cancer treatment with these modalities. In this work, the focus lies on the field of radiotherapy.

Radiotherapy can be given to the patient with both external and internal ionising radiation. The biological basis of these types of radiation is the same, namely, they induce DNA damage. High-energy radiation has ionising properties which causes double-strand DNA breaks. This results in cell death. The goal of radiotherapy is to kill tumour cells while limiting cell death of healthy tissue. This is possible because tumour cells are more sensitive to ionising radiation than healthy cells. DNA damage can be repaired less consistently. With curative cancer treatment, all tumour cells should be killed to prevent recurrence and healthy tissue should be spared as much as possible for reduction of harmful side-effects. This trade-off characterises the challenge of applying any type of radiotherapy. Two essential quantities in the field of radiotherapy are coupled to this trade-off. These are the tumour control probability (TCP) and the normal tissue complication probability (NTCP). Both metrics describe the relation between dose and tissue reactions. The area between the two metrics is known as the therapeutic window. The area is computed as TCP*(1-NTCP). The highest probability of treatment success is found at the dose value where the TCP and NTCP are separated the most. The type of radiotherapy treatment is chosen such that the therapeutic window is as large as possible. This yields the highest treatment success rate. Research has always focused on increasing the size of the therapeutic window. The development of fractionation, i.e. delivering the total radiation dose in multiple fractions, increased the therapeutic window. Fractionation works because healthy cells can repair radiation damage more consistently than tumour cells.

The type of ionising radiation applied is patient-specific. Generally, most patients receive external photon irradiation, i.e. X-rays, generated by a linear accelerator. In such a machine, electrons are accelerated between a cathode and anode onto a target. The energy of the X-rays is proportional to the degree of acceleration. Electrons are accelerated with radiofrequency waves. The number of released electrons at the cathode depends on the temperature. The target is often a piece of tungsten. The high-energy photons (several MVs) are mainly formed by slowing down of electrons because of interaction with the tungsten nuclear core (bremsstrahlung). The generated X-rays are mounted with the help of collimators on the tumour area of the patient during treatment. By using multiple beam angles, healthy tissue is spared whereas tumour coverage is maintained.

There are three well-known techniques for dose delivery with high-energy photons. These are three-dimensional conformal radiation therapy (3D-CRT) [18], intensity-modulated radiation therapy (IMRT) [22] and volumetric modulated arc therapy (VMAT) [36]. With 3D-CRT high-energy photon beams are mounted on the tumour from different angles. The incoming beams are shaped with a multileaf collimator such that they match the tumour outline. 3D-CRT is an example of forward treatment planning in which dose deposition is linked with beam directions. With forward treatment planning, an optimal treatment plan for the patient is not guaranteed. IMRT is an advanced form of 3D-CRT in which dose delivery is more conformal. Again, multiple beam directions are used and the incoming beams are shaped with a multileaf collimator. However, the intensity of each beam is modulated such that healthy tissue is spared more. IMRT is based on inverse treatment planning. This implies that the found beam intensities are optimised to match a predefined dose distribution. The optimal dose is often expressed in a list of constraints and objectives. IMRT is further improved with VMAT. In this case, the dose is delivered with a continuous arc consisting of many more beam directions. The dose can be delivered to the patient quicker than IMRT and is more conformal. However, the optimisation of the beam intensities is more complex for so many beam angles. As the dose is delivered with a single rotation, the multileaf collimator constantly changes shape according to the beam angle.

The downside of irradiation with high-energy photons relates to their main physical characteristic. The linear attenuation coefficient of high-energy photons in water is in the order of 0.01 cm^{-1}. [24] This implies that these high-energy photons can traverse the human body with ease. In other words, energy is deposited along the whole path of the photons and healthy tissue will always be targeted with photon radiotherapy. For some difficult tumour cases, e.g. certain brain tumours, photon radiotherapy is not an option because too much healthy tissue will be targeted. This may result in severe brain damage and decreased quality of life. In such cases, proton radiotherapy is more suited as the dose can be delivered to the tumour accurately while limiting dose delivery to healthy tissue by 50% up to 70% compared to photon radiotherapy. [49]

The treatment success rate of any radiotherapy treatment also depends on the uncertainties present. Setup errors are an essential source of uncertainty. Setup errors can be separated into a random part and a systematic part. Typically they both follow a Gaussian distribution with varying mean and standard deviation. The mean of the random error distribution is zero whereas the mean of the systematic error distribution is non-zero. Additionally, the standard deviation of the random setup error distribution is smaller than the standard deviation of the systematic distribution. Systematic errors should be measured but cannot be avoided. With daily measurements, correction for time trends is possible. Random errors are more machine intrinsic and should be measured as well. Correction for random errors is not possible. Another benefit of fractionation is the fact that both random and systematic errors are averaged out throughout the treatment. This yields better treatment success rates. Contrarily, fractionation can be challenging because changes to the irradiated anatomy occur from day to day (inter-fractional uncertainties). Within a clinical setting, the goal is to minimise all possible uncertainties to achieve the highest possible curing rate.

1.1.1. Proton Radiotherapy

The benefit of proton radiotherapy over photon radiotherapy for a variety of treatment sites has been shown. [33] Because the benefit is substantial, clinical interest in proton therapy increased worldwide over the years. Ever since the first hospital-based proton therapy facility has been built in the late 1980s, the number of available facilities and the total number of patients treated per year with proton therapy shows an exponential increase. [49] In Fig. 1.1 this exponential increase is shown together with important milestones in the field of clinical proton radiotherapy. The most important technological improvement of proton therapy is the introduction of pencil-beam-scanning (PBS). This made proton therapy more consistent and more user-friendly compared to the passive proton scattering technique applied before. [38]

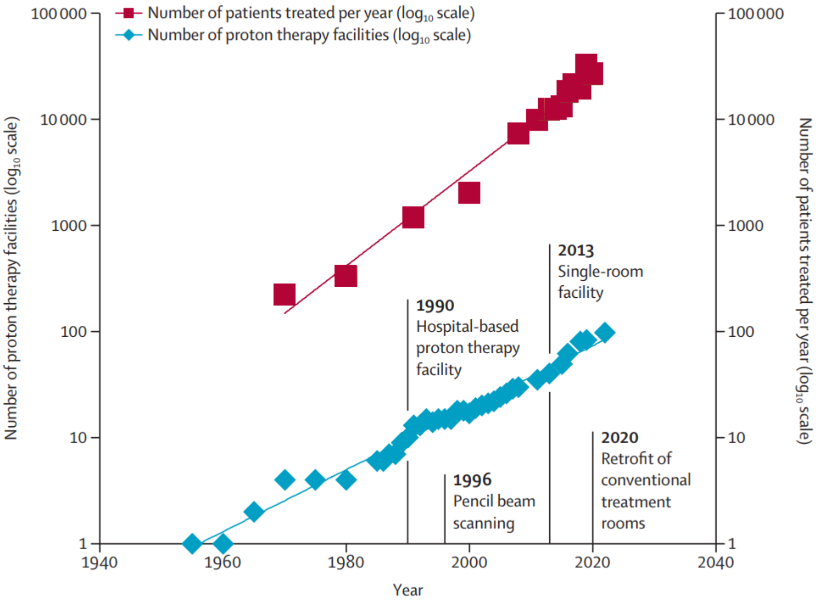


Figure 1.1: Representation of the number of proton therapy facilities (left y-axis) and number of treated patients per year with proton radiotherapy (right y-axis) over the last years (x-axis). Note that both y-axes are on a $\log_{10}$ scale. This figure is adapted from [49].

High-energy protons can be used for cancer treatment because they also have ionising properties, similar to photons. The main difference between the two is that a certain energy range of protons does not traverse the patient completely. Instead, the protons stop inside the patient and deposit their energy locally. This allows for more healthy tissue sparing compared to photon therapy. The range of a proton directly relates to its energy and the physical stopping power in MeV per cm. For clinically used proton energies (70-250 MeV), the stopping power in water is around 10 MeV per cm for protons of 70 Mev and 5 MeV per cm for protons of 250 MeV. [5] From these stopping powers it is clear that the range of protons is in the order of centimetres to decimetres. A well-validated fit to the stopping power data for the proton range in water is the following formula [6, 37]:

$$R(E) \approx 0.00244E^{1.75}[cm]$$

Filling in this formula for 70 MeV and 250 MeV protons returns the mentioned proton ranges in water.

Another advantage of protons over photons relates to their energy deposition profile. The deposited energy along the path of a proton beam is inversely proportional to its velocity. This results in a sharp increase in dose delivery at the end of the beam range because of the constant slowing down of the protons inside the volume. For a proton beam, this sharp increase is known as the Bragg peak and this peak typically is a few millimetres in width. [37] With multiple proton beam energies and intensities, a spread-out Bragg peak can be shaped such that the tumour is fully covered. Additionally, multiple modulated beam angles can be used such that healthy tissue is spared even more. In Fig. 1.2 a simplification of the difference between photon and proton dose deposition is shown. From the figure, it is clear that with photons more integral dose is given to healthy tissue to achieve similar tumour coverage.

Based on the energy deposition profile of protons in comparison to photons, proton radiotherapy outperforms photon radiotherapy in an ideal scenario. However, throughout treatment, the efficiency of proton radiotherapy is influenced by uncertainties. Similar to photon radiotherapy, patient-specific setup errors are again an important class of uncertainties. Moreover, proton range errors affect the

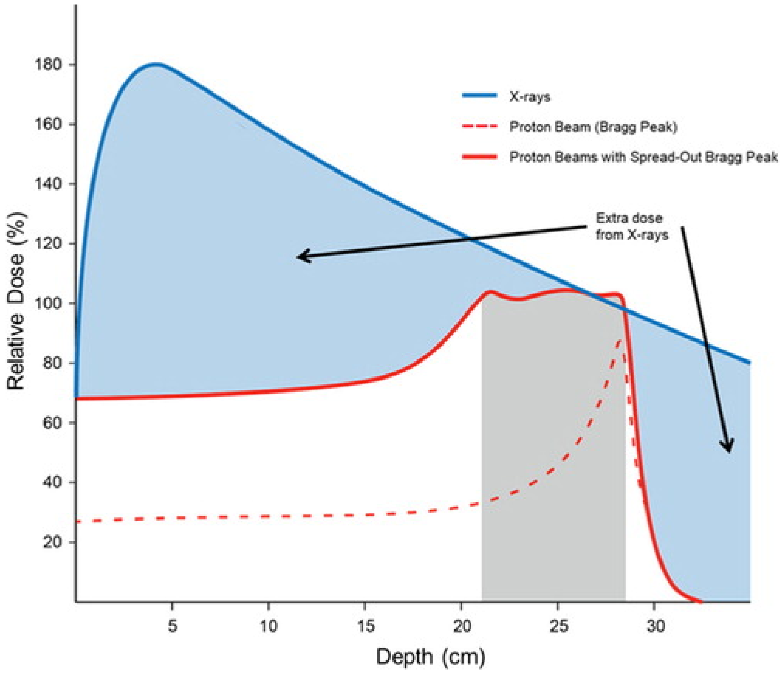


Figure 1.2: A comparison between photon radiotherapy with high energy photons (X-rays) and proton radiotherapy for an arbitrary tumour shown with the grey area. For protons, a single proton beam Bragg peak of 200 MeV is shown together with the modulated spread-out Bragg peak. On the x-axis, the depth in cm is shown. The y-axis corresponds to the relative dose. The tumour has to receive approximately 100% of the relative dose. The relative dose difference between photon and proton radiotherapy is displayed in the light blue area. This figure is adapted from [13].

dose deposition. Because a proton Bragg peak has a width of several millimetres, the impact of any type of added uncertainty is bigger compared to photon therapy. In short, proton therapy is an example of high-precision radiotherapy but is limited by its precise delivery because of all possible uncertainties. To further boost the applicability of proton radiotherapy, the focus lies on more precise tumour targeting together with better treatment delivery quality assessment.

As mentioned before, PBS is one of the milestones in the field of proton radiotherapy which has led to an exponential increase in treated patients per year with protons (Fig. 1.1). The introduction of spot scanning with proton beams, i.e. pencil beams, has shaped the current clinically used type of proton radiotherapy, namely, intensity-modulated proton therapy (IMPT). Before the introduction of PBS, passive scattered proton therapy was used. Every treatment required patient-specific materials to reshape an incoming proton beam to the tumour shape. With a scattering material, the proton beam of limited lateral width was broadened. PBS allowed for the direct placement of the pencil beams on the tumour and required less patient-specific materials. For a patient treatment with IMPT, individual pencil beam positions together with their intensity are optimised such that an ideal dose to the patient can be delivered. To reach all pencil beam scanning positions, quadruple magnets are used to steer the incoming proton beam laterally. Multiple depths are reached with modulation of the proton energy by an energy degrader. The scanning technique and intensity optimisation make IMPT a high-precision type of radiotherapy.

The clinical range of proton energies is 70 MeV to 250 MeV. These high-energy protons have to be generated with a particle accelerator. There are two main types available, a synchrotron and a cyclotron. Both make use of electromagnetic fields to accelerate particles and keep them in focus. In clinical institutes, cyclotrons are most frequently used as they have a more compact design. The typical diameter of a clinical cyclotron is about 3.5 to 5 metres, whereas a clinical synchrotron has a diameter of 6 to 8 metres. [37] Additionally, a synchrotron produces proton pulses with a time interval of several seconds whereas cyclotrons produce a continuous proton beam. In the centre of the proton accelerator, ionised hydrogen is located such that plasma is formed. The plasma is accelerated which yields the high-energy protons. The production of high-energy protons is more expensive than the production of high-energy photons. Moreover, the size of a particle accelerator together with the rotating gantry makes it challenging to implement on a large scale compared to linear accelerators used in photon radiotherapy. More work is needed to close the gap between the accessibility of photon and proton radiotherapy. Especially in low and middle-income countries, the potential of proton therapy is large. [49] Additionally, the efficacy of proton therapy can be exploited further. For instance, FLASH proton therapy has the potential to further increase healthy tissue sparing while maintaining tumour coverage.

1.1.2. FLASH Radiotherapy

The beneficial effect of FLASH radiotherapy over conventional radiotherapy is called the FLASH effect. This biological effect describes healthy tissue sparing achieved at certain dose delivery thresholds. In FLASH research, multiple thresholds have been proposed. The general, more conservative, threshold definition is as follows: the FLASH effect is guaranteed at (1) fractional doses of more than 8 Gy, (2) a delivery time of less than 200 milliseconds and (3) a mean dose rate of 40 Gy per second. [7] Numerous animal studies have shown the presence of this FLASH effect (mice [16], zebrafish [34], pigs [46] and cats [46]). More recently, the first patients have been treated with electron and proton FLASH radiotherapy. [8, 32]

For clinical applications, a relatively small FLASH effect may have a significant impact. The FLASH effect corresponds to an increase in the size of the therapeutic window. In other words, the difference between the NTCP and TCP with FLASH radiotherapy is larger than compared to conventional radiotherapy. In Fig. 1.3 the FLASH effect is clear from experimental data. The data is acquired from mice with implanted human brain tumours. [7] For conventional radiotherapy the therapeutic window is smaller than for FLASH radiotherapy. From the figure, it is also clear that the FLASH effect is mainly driven by normal tissue sparing. When comparing conventional with FLASH radiotherapy, the NTCP fitted lines (blue) show a more pronounced response compared to the TCP fitted lines (red). For 10 Gy dose delivery, the FLASH NTCP is near 0% against 90% for conventional radiotherapy.

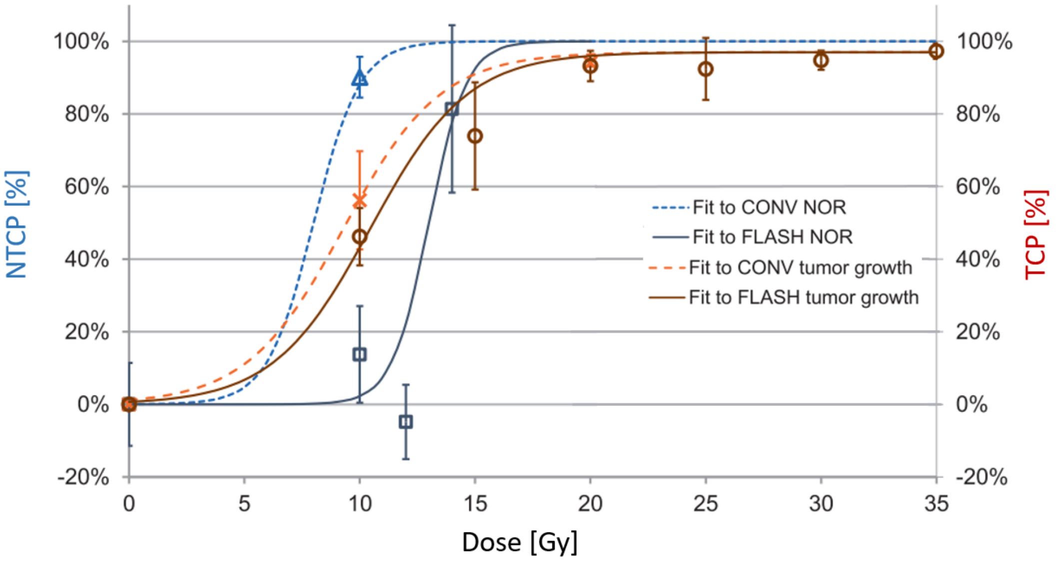


Figure 1.3: NTCP and TCP correlated with the dose on the x-axis for conventional and FLASH electron radiotherapy. FLASH radiotherapy is shown with solid lines. The curves for conventional radiotherapy are shown with dashed lines. The NTCP is shown on the left y-axis with blue lines. With a novel object recognition test (NOR) for mice, the NTCP was determined. The TCP is shown on the right y-axis with red lines. The TCP is directly related to relative tumour growth after irradiation. For conventional radiotherapy dose rates of 0.01 Gy/s are reported. For FLASH radiotherapy, dose rates of 125 Gy/s are achieved with electron pulses. The curves are fitted to the found mean NTCP and the mean TCP of at least five measurements at various dose levels. The error bars at each mean correspond to the standard deviation in acquired data. This figure is adapted from [7].

The reason why more healthy tissue is spared during FLASH radiotherapy is not fully understood yet. Multiple biological processes are linked to the FLASH effect such as radiolytic oxygen depletion in cells, the formation of reactive oxygen species (ROS) and the level of immune response. [25] A full explanation of the FLASH effect probably involves a variety of biological processes altered with the FLASH irradiation compared to the conventional mode. Especially the first factor, the relation with oxygen depletion, has been studied extensively. Work from the 1960s already correlated oxygen concentration with bacterial survival at FLASH dose rates. [15] In hypoxic conditions, bacterial survival was bigger than in normoxic conditions.

Due to the usage of high dose rates, oxygen in cells depletes rapidly creating a hypoxic environment. Within this environment, cells are less sensitive to ionising radiation. The FLASH effect is noticeable because tumour cells are responding less to oxygen concentration changes than healthy cells. This is driven by the fact that tumours are oxygenated poorly because of rapid and uncontrolled growth. In a more recent study done on prostate cancer cells, the FLASH effect was again linked to oxygen concentration and depletion. [2] There is a clear correlation between the degree of hypoxia and the FLASH effect in tissue. More research is needed to understand all biological processes contributing to the extra healthy tissue sparing under FLASH conditions.

Most research in the FLASH field is focusing on the clinical translation of FLASH radiotherapy with electrons. Mainly because electron pulses adhering to the mentioned FLASH thresholds can be made with a relatively simple linear accelerator. During electron pulses dose rates of thousands to millions of Gy per second can be achieved. [7] However, the potential clinical impact of electron FLASH radiotherapy is limited. The most frequently used linear accelerated electrons have an energy of several MeVs (5 MeV to 10 MeV) implying that the dose cannot be delivered deep inside a patient. In this range of energies, the stopping power of electrons in water is around 2 MeV per cm. [5] From this physical property it is clear that electrons can only be used to irradiate superficially located tumours. Ultra-high electron energy (typically more than 100 MeV) devices for FLASH radiotherapy are also in development. [42, 43] These ultra-high energy electrons can reach deep-seated tumours. However, the complex design of the device together with high costs limits the applicability of ultra-high electron energy beams for FLASH radiotherapy. It is unlikely that this technique will become available soon. [42]. For FLASH proton therapy, there is more clinical potential as all types of tumours can be reached with current accelerators.

FLASH Proton Radiotherapy

The FLASH threshold can be met with the currently used cyclotrons for proton acceleration. Such high mean dose rates are achieved by operating at high beam currents (>200 nA) and high energies (>200 MeV). However, up to now, energy switching for the formation of the spread-out Bragg-peak is not possible as the treatment time and dose rate thresholds will not be achieved. [28] Right now, clinical FLASH proton therapy is envisioned to work without energy switching. There are two options to deliver FLASH proton therapy: (1) transmission beam proton therapy making it similar to photon radiotherapy and (2) proton therapy with ridge filters. With transmission FLASH, the characteristic Bragg-peak is disregarded and dose to tissue is delivered with the plateau region of the deposition curve (see dose deposition plateau region in Fig. 1.2). This might seem counter-intuitive, as the Bragg-peak is considered the functional basis of proton therapy, but a sufficient FLASH effect makes the transmission technique viable in some treatment cases. [23] Transmission beams are not suited for all clinical scenarios. With transmission FLASH, PBS can still be applied for lateral coverage. Another option to apply FLASH proton therapy is to work with ridge filters, somewhat similar to clinical proton therapy before the introduction of PBS. With the help of a scattering source and a ridge filter placed in the path of the generated proton beam, the dose can be delivered to the whole target quickly. Research has shown that creating a ridge filter for FLASH proton therapy is possible. [50] At the Holland Proton Therapy Centre (HPTC), PBS is used clinically, which makes transmission FLASH the proposed type of FLASH proton therapy. Both types of FLASH proton therapy are possible, but in this thesis, the emphasis will lie on transmission FLASH mode.

Next to high beam currents and high proton energies, other constraints are required to adhere to all the FLASH thresholds. Only relatively small tumours can be irradiated with transmission FLASH. Otherwise, PBS takes too long due to the total number of spots. The constraint on tumour size can be linked to the operational beam current. The cyclotron can create a proton beam current of approximately 100 nA to 200 nA at the HPTC. This implies that for big tumours, the FLASH thresholds of 40 Gy per second and 200 millisecond delivery time will not be reached. Additionally, hypofractionation is required for FLASH radiotherapy. In conventional radiotherapy, fractions of approximately 2 Gy are given to patients, whereas the fractional FLASH dose threshold is 8 Gy. All these constraints make the room for error small. To guarantee the FLASH effect for patients, excellent quality assessment is required. FLASH proton therapy is sensitive to uncertainties, even more than other types of radiotherapy. Hypofractionation is one of the reasons why FLASH proton therapy is more sensitive to uncertainties. Throughout the treatment, the placement error is averaged out less because of hypofractionation.

Moreover, for FLASH treatments not only uncertainties influencing the dose but factors affecting the dose rate should be considered as well. In that light, two important machine intrinsic uncertainties should be taken into account. The random spot positioning error induced by the scanning magnets and fluctuations in the beam-current of the cyclotron. The machine spot positioning errors influence the dose and dose rate, whereas the change in beam current only affects the dose rate. Dose delivery is not changed as for every spot the total monitoring units (MUs) are predefined and are measured during delivery. The magnets steer to the next spot when the planned MU, and thus dose, is given to the corresponding spot. With this scanning procedure, only the total beam time spent at each spot varies with beam-current fluctuations and not the dose delivered per spot. Both types of machine uncertainties are substantial. At the HPTC, spot position changes in the order of millimetres are registered and up to 20% fluctuation in beam current is possible from day to day.

1.2. Purpose

The presence of the FLASH effect has been shown numerous times in various animal studies. [7, 16, 34, 46] Furthermore, in a clinical trial the technical feasibility of transmission proton FLASH radiotherapy for humans has been proven. [12, 32] It has been shown that particle accelerators can be used to produce proton beams adhering to the FLASH thresholds in delivery time, dose and dose rate. The next step of FLASH proton therapy is further clinical translation. An important point that has to be addressed before clinical implementation, is the relation between FLASH parameters and machine intrinsic uncertainties. By analysing this relationship, an answer to the following question can be found: 'Is FLASH proton therapy viable when taking into account machine intrinsic uncertainties?' The goal of the project is to quantify the impact of machine uncertainties on the efficacy of FLASH proton therapy. Finding the exact relation between the uncertainties and FLASH dose delivery will contribute to treatment quality assurance and shape future proton FLASH radiotherapy.

The question is answered by simulating two types of uncertainties. These are pencil beam placement errors and proton beam current fluctuations. Dose and dose rate changes are simulated corresponding to spot position and beam current changes. This is done for a one-, two- and three-dimensional setup. The one- and two-dimensional setup is based on an arbitrary field with a target. For the three-dimensional setup, a FLASH treatment plan from the Erasmus Medical Centre (EMC) was used. Dose distributions were calculated millions of times to achieve high power statistics. In particular, for the three-dimensional setup, dose calculations are time-consuming. This makes an uncertainty analysis with sampling not possible. With a semi-analytic dose engine created at the Technical University of Delft (TUD), the three-dimensional uncertainty analysis is possible with an adjoint analysis. [10] The uncertainty in pencil beam position can be analysed without a full dose re-computation based on HU perturbations. More information about fast dose-calculating algorithms and models can be found in my literature study included in the appendix (see appendix part D).

1.3. Overview

The methods used throughout the one-, two- and three-dimensional uncertainty analyses are described in chapter 2. Thereafter, in chapter 3 the one- and two-dimensional uncertainty analysis is provided for a fixed machine uncertainty scenario. The dose rate metrics used in this analysis are sensitive and not robust. In chapter 4, a heuristic, more robust, biologically inspired metric is used for the uncertainty analysis. From there, the two-dimensional analysis is extended to translate various scales of machine uncertainties to clinical constraints in chapter 5. Forward dose calculations and concepts for the uncertainty analysis in 3D are given in chapter 6. Afterwards, the results and implications of the work are reviewed in chapter 7. In that chapter, future research directions are also given. Lastly, a conclusion is drawn in chapter 8.

2

Methods

In this chapter, all methods used throughout the project are discussed. The methods are sorted in order of appearance in this thesis.

2.1. 1D & 2D uncertainty analysis

The addition of spot placement errors and beam current fluctuations into the analytical model yields changes in dose and dose rate per voxel. The dose in each voxel is given by the summed dose deposition per weighted pencil beam and is given in Gy. However, it is not clear how the dose rate should be defined. In the field of FLASH proton therapy, a dose rate in Gy per second can be calculated in many ways. Two options are the dose-averaged dose rate (DADR) and the pencil-beam scanning dose rate (PBSDR). The two types of dose-rate definitions will be used throughout the paper (chapter 3 and chapter 5).

In chapter 3 the setup of the 1D and 2D analytical model is described in detail.

2.1.1. DADR

The DADR has been first introduced by Van de Water et al. [45] This dose-rate definition takes into account the instantaneous dose rate per pencil beam and weights it according to the relative dose deposition per pencil beam. In Eq. 2.1 the expression for the DADR per voxel (i) is given for a given number of pencil beams (n). Weighting is done with the pencil beam weight (w [protons]), the pencil beam intensity (BI [protons/s]) and the dose-deposition matrix per voxel and pencil beam (d [Gy/proton]). The weighting with dose is better visible when rewriting the equation to the second form. The multiplication of the dose-deposition matrix with every pencil beam weight returns the dose deposition in a voxel. The multiplication of the beam intensity in a specific voxel with the corresponding dose-deposition matrix gives the instantaneous dose rate of the pencil beam. In an arbitrary case where all pencil beam weights are equal, the DADR will be the mean instantaneous dose rate.

$$DADR_i = \sum_{j=1}^{n} \frac{(d_{i,j}w_j)(d_{i,j}BIj)}{\sum_{j=1}^{n} d_{i,j}w_j} = \sum_{j=1}^{n} \frac{(d_{i,j}w_j)(DR_{i,j})}{D_i} \tag{2.1}$$

From the expression of the DADR, it is clear that the time between two consecutive pencil beams is disregarded as the instantaneous dose rate is used. This may lead to an overestimation of the dose rate. [45] Furthermore, the pencil beam scanning pattern is not included in this dose-rate metric which makes it less ideal for implementation into the envisioned clinical transmission of FLASH proton therapy.

2.1.2. PBSDR

The PBSDR has been developed by Folkert et al., specifically for FLASH proton radiotherapy. [17] This formulation of the dose rate depends on the scanning pattern. The expression for the PBSDR per voxel is given in Eq. 2.2. In this equation, the total dose (D) in a voxel minus two times the predefined dose threshold (d_t) is divided by the effective irradiation time (T_i). The dose threshold is defined as a fraction of the prescribed dose. The effective time can be rewritten as ($T_1 - T_0$), where T_0 equals the irradiation time at which the dose threshold is exceeded and T_1 the irradiation time at which the maximum dose minus the threshold is reached. With a dose threshold of 0 Gy, the PBSDR will equal the dose maximum in a voxel divided by the total irradiation time. As the FLASH effect is a local and time-dependent effect, the zero threshold scenario does not correspond to reality. With a proper dose threshold, dose rates in line with the FLASH effect are found.

$$PBSDR_i = \frac{D_i - 2d_t}{T_i} \tag{2.2}$$

In Fig. 2.1, the dose for an arbitrary field is given with the corresponding PBSDR. From the figure, it is clear that the PBSDR is highest at the first and last column of the pencil beam scanning pattern. This results from the fact that the dose at these high PBSDR locations is given with relatively few pencil beams. Consequently, the effective irradiation time will be relatively low. In the middle of the field, where many pencil beams deposit dose, the PBSDR is the lowest. The dose-rate profile of the PBSDR is in line with the idea behind FLASH proton radiotherapy. With this definition, the FLASH effect will be dominant at the edges as the highest dose rates are achieved there. This is beneficial as healthy tissue can be found near the edges.

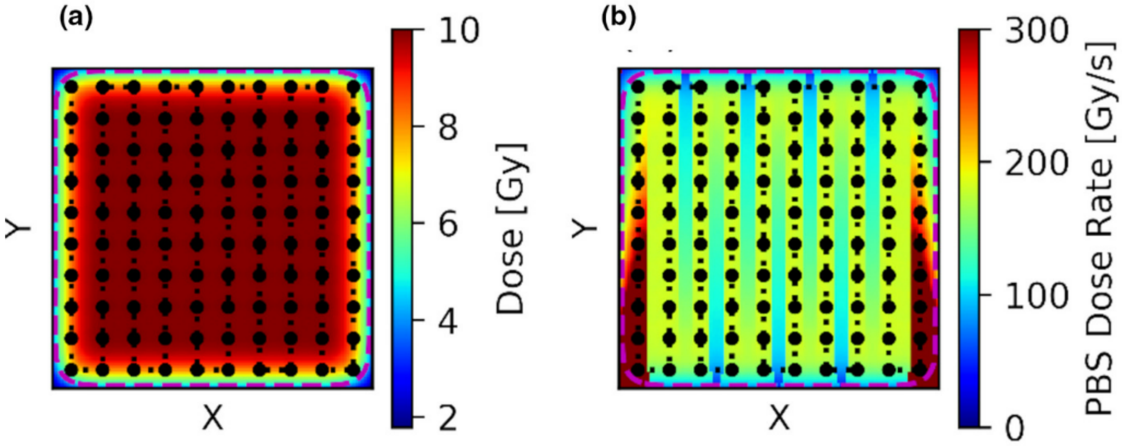


Figure 2.1: A dose profile (a) and a PBSDR profile (b) for an arbitrary 2D field with 10x10 pencil beams. In both sub-figures, the pink dotted line represents the 50% isodose line which is considered the target. The dose threshold for the PBSDR is 1% of the prescribed dose. This figure is adapted from [17].

From both Eq. 2.2 and Fig. 2.1 it is evident that the PBSDR depends on the scanning pattern because the effective irradiation time is related to the predefined dose threshold. This dependency is a flaw of the PBSDR as the selection of the dose threshold is rather subjective. With too high and too low dose thresholds, unrealistic dose rates will be found. It is also not known which dose threshold matches best with the FLASH effect. On the other hand, as the dose rate depends on the scanning pattern, optimisation of the scanning pattern to maximise the dose rate at the edges and thus the FLASH effect based on the PBSDR is possible. [41]

2.2. Biological FLASH metric

One of the contributing factors to the FLASH effect is the concentration of oxygen in irradiated cells. [2] If the FLASH thresholds are met, oxygen depletes rapidly in healthy cells, increasing the radioresistance of the cells. The constructed FLASH model in this thesis depends on the oxygen concentration and depletion behaviour. The metric created is called the effective FLASH enhancement ratio (FER$_{eff}$). In the quantification of the FER$_{eff}$, the oxygen concentration in cells over time and important FLASH metrics are involved. The FLASH quantities involved are the oxygen enhancement ratio (OER), the FLASH enhancement ratio (FER) and the FLASH enhanced dose (FED). All used metrics in the model are time traces, implying that the outcome is scanning pattern dependent. The model is locally defined without any oxygen diffusion and metabolism. The FLASH effect depends on oxygen diffusion and metabolism as well. However, these processes are difficult to parameterise. [39]

The oxygen concentration over time has been modelled with an ordinary differential equation consisting of the dose rate and the concentration difference from normoxic conditions. Oxygen concentration linearly depletes with dose rate and exponentially recovers with oxygen concentration difference from the steady state. From the oxygen concentration time trace, the OER and FER over time can be calculated. The FED over time can be calculated by division of a non-FLASH dose (under normoxic conditions) with the FER. Lastly, the FER$_{eff}$ is defined for one position by the ratio of the non-FLASH dose over FED at full completion of the scanning pattern. The structure is shown in Fig. 2.2. For each voxel position (r) the FER$_{eff}$ depends on the time traces (t) of the other used metrics.

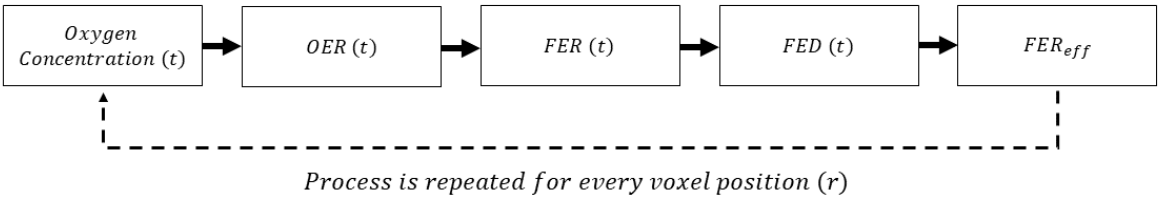


Figure 2.2: Structure of the biologically inspired FLASH model. The first four quantities are time traces (t). The outcome is the effective FLASH enhancement ratio which is defined per voxel position (r). The process is repeated for the whole field to get the FER_{eff} distribution.

2.2.1. Oxygen concentration

The oxygen concentration over time is modelled with an ordinary differential equation. In this model, oxygen depletes linearly and re-oxygenation is modelled as an exponential process. In Eq. 2.3 the equation is given. The two used fit parameters are based on work from M. Diepeveen (his MSc thesis [14] and an article under review). The κ equals 5.0 E-4 $mol\ m^{-3}\ Gy^{-1}$ and τ equals 0.001 s. The dose rate ($\dot{D}$) over time is calculated as the forward finite difference between consecutive time points: $\dot{D} = \frac{D(t+dt)-D(t)}{dt}$. The steady-state oxygen concentration (C_{ss}) represent the oxygen concentration in normoxic conditions. It has been determined by bounding the FER and reversely calculating the corresponding OER and steady-state oxygen concentration. The max FER was bounded at 2.0 as found FER values in animal studies seem to approach this enhancement ratio. [7]

$$\frac{dC(t)}{dt} = -\kappa\dot{D}(t) + \frac{1}{\tau}\Big(C_{ss} - C(t)\Big) \tag{2.3}$$

The concentration over time is found by numerically solving this ordinary differential equation with the backward Euler approach. This implies that the oxygen concentration at the next timepoint (t+dt) is related to $\frac{dC(t+dt)}{d(t+dt)}$ instead of $\frac{dC(t)}{dt}$ for the forward Euler solution. The backward Euler method is needed to force the solution for the oxygen concentration to stay above zero and remain differentiable. Oxygen depletion is modelled with $-\kappa\dot{D}$ which implies that with forward Euler, the depletion term can drive the oxygen concentration below zero. The backward Euler solution for the concentration over time is given in Eq. 2.4. The initial solution at the start of the scanning pattern (t=0) is the steady-state oxygen concentration. In appendix part A the derivation of this equation can be found.

$$C(t+dt) = \frac{\left(\frac{C_{ss}}{\tau} + \frac{C(t)}{dt}\right)}{\left(\frac{1}{dt} + \frac{1}{\tau} + \frac{\kappa \dot{D}}{C(t)}\right)} \tag{2.4}$$

The two fit parameters are included in the model for fine-tuning purposes. Both the linear depletion term and the exponential re-oxygenation term can be adjusted with model parameters. However, in essence, this ordinary differential equation only depends on a linear combination of κ and τ. The proof for this is given in the appendix part B.

2.2.2. OER

The OER is defined as the ratio between the dose given under normoxic conditions and hypoxic conditions yielding the same treatment outcome. [20] As mentioned before, under hypoxic conditions tissue is more resistant to ionising radiation implying that more dose is needed to achieve the same result relative to the normoxic scenario.

Grimes and Partridge formed an expression for the OER related to oxygen partial pressure. [20] This equation is based on multiple experimental studies and is given in Eq. 2.5. The fraction of ψ_O over ψ_D relates to the relative weight of oxygen-caused cell death (ψ_O) against direct dose cell death (ψ_D). The best fit for this fraction was 1.63. The other parameter ϕ is calculated according to the thermal velocity, the mean-free path and the collision probability with the DNA of oxygen molecules. [19] The ϕ is measured to be 0.026 [$mmHg^{-1}$]

$$\text{OER}(p(t)) = 1 + \frac{\psi_O}{\psi_D}(1 - e^{-\phi p(t)}) \tag{2.5}$$

In the theoretical limit, the OER approaches a maximum of 2.63 with partial oxygen pressure increasing to infinity. However, as the steady-state concentration of oxygen was fixed in the model, the theoretical limit of the OER will never be reached. Similar to the FER, the OER is bound between 1 and 2.

Grimes and Partridge chose the partial pressure for oxygen in mmHg as input for Eq. 2.5. The biological FLASH metric is based on the depletion of oxygen concentration in cells, meaning that rewriting the equation is necessary. The partial pressure of oxygen molecules is proportional to concentration according to Henry's law: $H = \frac{p}{C}$. With Henry's constant (H) for oxygen molecules at body temperature, the partial pressure (p) can be converted to concentration (C). H equals 7.29 E-5 [$mol\ m^{-3}$ $mmHg$] for O_2. [21] Note that parameter ϕ should be corrected with Henry's constant as well such that ϕ' is expressed in [$m^3\ mol^{-1}$]. The functioning of the biological FLASH model is not affected by the conversion and gives the same output for either partial pressure or concentration. The formula for the OER used in the model depending on oxygen concentration instead of partial pressure is given in Eq. 2.6.

$$\text{OER}(C(t)) = 1 + \frac{\psi_O}{\psi_D}(1 - e^{-\phi' \frac{p(t)}{H}}) \tag{2.6}$$

2.2.3. FER

The FER is a ratio describing the dose under hypoxic against normoxic conditions to establish the same effect. Note that the FER is inversely related to the OER and thus is linked to the OER according to Eq. 2.7. When setting up this model, it was decided to allow FER and OER values between 1 and 2, which justifies this bound relation. From the equation it is clear that an OER of 2 corresponds with a FER of

1, showing the inverse relation between the two metrics.

$$\text{FER}[\text{OER}(C(t))] = \frac{2}{\text{OER}(C(t))} \tag{2.7}$$

This inverse relation between OER and FER might seem counterintuitive, but it is not. The main discrepancy between the OER and FER is that they are defined for other types of tissue. The OER is defined for tumour tissue. A high OER relates to more tumour radio-sensitivity because of high oxygen concentrations and results in higher tumour coverage at lower doses. However, the FER is defined for healthy tissue, in which increased radio-sensitivity would be considered as bad for the patient. More healthy tissue is spared when the concentration of oxygen is as low as possible inside the healthy tissue. The extra tissue sparing in these hypoxic conditions is described with the FER. In short, the OER is of main importance for tumour tissue, whereas the FER is for healthy tissue.

2.2.4. FED

To relate the found FER over time with the dose delivered, the FED can be used. The FED represent a biologically plausible dose. At a max FER of 2.0, the dose to the respective tissue is lowered by a factor of 2. The FED is defined as the nominal normoxic dose (D_n) delivery divided by the FER over time (Eq. 2.8). Keep in mind that the FER is calculated according to Eq. 2.7 and depends on oxygen concentration and the resulting OER over time.

$$\text{FED}(t) = \frac{D_n(t)}{\text{FER}(t)} \tag{2.8}$$

The FLASH effect is present if the biologically delivered dose (FED) at the time (t) in healthy tissue is lower than in tumour tissue.

2.2.5. FER$_{eff}$

The outcome of the biological FLASH model is the FER$_{eff}$. This metric brings back the FED during the dose delivery to one effective value. The FER$_{eff}$ is a quantity for the total FLASH effect during the dose delivery. To guarantee the FLASH effect, this effective variable should be higher in healthy tissue than in tumour tissue. Again, as the OER and FER are bound between 1 and 2, the FER$_{eff}$ is between the same boundaries. The FER$_{eff}$ is calculated by division of the nominal normoxic dose (D_n) with the FED after full completion of dose delivery. Full completion of the dose delivery is at timepoint T_{end}. In Eq. 2.9 the expression for the FER$_{eff}$ is given for one voxel position (r).

$$\text{FER}_{eff}(r) = \frac{D_n(t = T_{end})}{\text{FED}(t = T_{end})} \tag{2.9}$$

After calculating the FER$_{eff}$ for every voxel position, a distribution of the novel FLASH metric for the field or target can be set up. With this metric various properties related to clinical FLASH proton therapy are studied. For instance, the relation between the scanning pattern and this metric. Furthermore, the robustness can be tested by implementing machine-specific uncertainties into the model. In chapter 4 the results of this model are given.

2.3. 2D clinical model

The 2D analytical uncertainty analysis is extended to a clinical perspective in chapter 5. The size of random machine intrinsic uncertainties in spot position and beam current are related to clinical outcomes.

In the end, this extension helps with further translation of FLASH proton radiotherapy into the clinic. The evaluation of the machine uncertainties is done with several clinically used parameters. These are the volume-at-dose (V95, V107), dose-at-volume (D2,D98) and the dose-rate-at-volume (PBSDR98).

By sampling machine uncertainties 10.000 times per random error, good statistics for these clinical metrics are found. The clinical metrics are reported with their median, 5th percentile and 95th percentile value. The random error in beam current and spot position is varied for multiple target sizes and inter-pencil beam spacing. This results in millions of dose calculations. For high accuracy in the dose calculations, a high-resolution grid is needed to compute the dose deposition of the pencil beams. This is computationally expensive. Therefore, Gaussian quadrature sampling was implemented to speed up the dose calculations while maintaining computation accuracy.

2.3.1. Clinical evaluation quantities

Volume-at-dose (Vx)

The volume-at-dose is a clinical quantity describing the relative volume receiving at least 'x'% of the prescribed dose. In this thesis, the V95 and V107 are used. These are the clinical metrics recommended by the ICRU. [26, 27] For clarity, the V95 is the relative volume that receives at least 95% of the prescribed treatment dose. The V95 is a metric used to quantify target coverage. The V107 gives insight into hot spots in the dose distribution.

Dose-at-volume (Dx)

The dose-at-volume is a metric for the dose at 'x'% of a defined volume. The uncertainty analysis is done with the D2 and D98. These are again ICRU recommended metrics. [26, 27] E.g. the D2 represent the minimum dose which 2% of the defined volume receives. The D2 can be considered as the dose given at certain hot spots. The D98 represents the overall target dose. The relation between the D98 and D2 is a measure of the homogeneity of dose deposition.

Dose-rate-at-volume (DRx)

The dose rate-at-volume is similar to the dose-at-volume. It is a quantity relating the dose rate to a defined volume 'x'. In the uncertainty analysis, the pencil-beam scanning dose rate was used at 98% of the volume as a metric (PBSDR98). For FLASH therapy this overall dose rate value is more important than the dose rate given in a small area of the target.

2.3.2. Gaussian quadrature sampling

It is possible to implement Gaussian quadrature because the dose deposition is computed analytically. In short, the dose is a summation of probability density functions of Gaussian pencil beams (more information in chapter 3). The summation of all probability functions is equal to taking an integral for all pencil beams. With Gaussian quadrature, the probability density functions on a fixed grid are better approximated. Especially on low-resolution grids, where the analytical Gaussian pencil beam is sampled poorly, the increase in accuracy because of Gaussian quadrature is substantial.

For second-order Gaussian quadrature, sampling of the probability density function per pencil beam at position R is equal to weighted samples at R$\pm\frac{1}{\sqrt{3}}$. The weights for second-order Gaussian quadrature are 1. [35] For implementation in the 2D analytical model, the position shifts with $\pm\frac{1}{\sqrt{3}}$ are in two directions (x and y). Additionally, the shifts are scaled to match the resolution of the 2D grid. This results in position shifts to R_x of $\pm\frac{dx}{2\sqrt{3}}$ and R_y of $\pm\frac{dy}{2\sqrt{3}}$, where dx and dy are the stepsize in the two directions. With second-order Gaussian quadrature in 2D, four dose calculations per pencil beam are done to achieve a more accurate output. The total sampling time can be reduced as calculating the dose four times on a lower-resolution grid is quicker than on a high-resolution grid. In Fig. 2.3 a schematic representation of Gaussian quadrature sampling for one voxel and one shift is shown. The four dotted red lines show the shifts (S_x, S_y) towards the red crosses. For the positive shift in both x and y, the new voxel is visible. The middle of the new voxel lies on the shift applied to the original position $[R_x, R_y]$

Implementation in the 2D analytical model was achieved with the following steps. (1) Create the four shifted grids respective to the original grid with the mentioned shift size. The middle of each voxel in the grid should now be located at the origin plus shift. (2) Compute the Gaussian probability density function for all pencil beams on the four created grids. (3) Take the sum of all probability densities of the four grids and compute the dose accordingly. (4) With the weight factors (all are 1 in the second-order Gaussian quadrature) take the weighted average such that the dose is computed accurately.

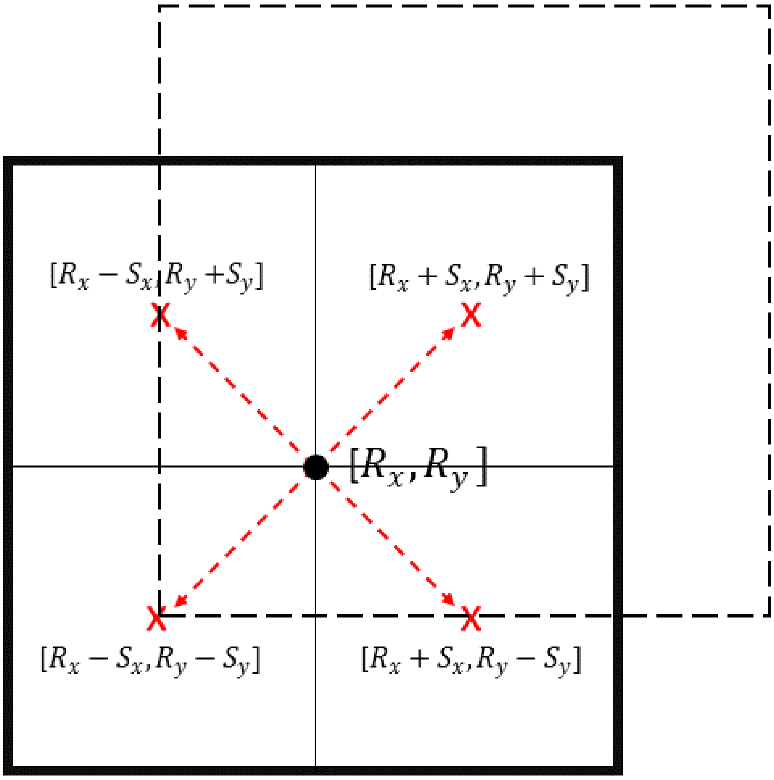


Figure 2.3: Schematic overview of second-order Gaussian quadrature for one voxel represented by the solid black square. With shifts in two directions (x and y), the middle of the original voxel is shifted to four new locations shown with the red crosses. The shifts are drawn with red striped arrows. S_x is $\frac{dx}{2\sqrt{3}}$, S_y equals $\frac{dy}{2\sqrt{3}}$. The shifted voxel corresponding to the positive S_x and S_y is shown with the striped black square. With Gaussian quadrature, the integral in the original middle position $[R_x, R_y]$ is calculated more accurately based on the four shifted voxels.

2.4. 3D uncertainty analysis

The three-dimensional FLASH proton therapy-related uncertainty analysis is done for one treatment plan. This plan is generated within iCycle; the developed treatment planning system at the EMC. To quantify the effect of spot placement machine uncertainties on the dose distribution, a semi-analytical dose engine from the TUD is used. With this engine, the effect of HU perturbations can be calculated directly, without full dose re-computation. Beforehand, the dose distribution from iCycle and the semi-analytical engine are compared with a gamma analysis.

2.4.1. Treatment plan generation & characteristics

The treatment plan used for the uncertainty analysis is generated with iCycle. Treatment plans are optimised with a multicriterial optimisation procedure, based on a predefined wish-list of constraints and objectives. In the first phase of the optimisation, all objectives are optimised sequentially from highest to lowest priority. When an objective is met, an extra constraint is added to the optimisation such that this objective will not be negatively affected by the continuing optimisation. After optimisation of all objectives, a treatment plan is generated adhering to all set constraints. In the next phase, objectives are successively optimised to their bare minimum, again according to their priority. [9]

The optimisation result is a pencil-beam weight for various positions and proton energies. The dose in a patient is then calculated with a matrix multiplication of the weights with the dose-influence matrix. The optimised weights in iCycle are in Gy. One iCycle weight corresponds with 1 Gy in the Bragg peak.

The single optimised FLASH proton therapy treatment plan for a relatively small lung tumour consists of 74 pencil beam positions. The proton energy per pencil beam is fixed at 244 MeV, such that a transmission FLASH plan is generated. The transmission beams enter the patient with a gantry angle such that the path to the tumour is minimised. In this plan, the gantry angle is 110 degrees. For simplicity, no table rotations are allowed. The predefined wish-list is patient-specific. The prescribed dose is 54 Gy in 3 fractions. The organs at risk (OARs) are the left lung, right lung, spinal cord, oesophagus, trachea and left bronchus. Initial constraints are set on the dose delivered to the GTV, PTV-GTV and the OARs. Objectives are prioritised as follows: PTV-GTV (1), GTV (2), PTV shells (3-10), the region outside PTV (11-12), healthy tissue in the left lung (13-14), right lung (15-16), spinal cord (17-18), oesophagus (19-20), trachea (21-22), left bronchus (23-24) and the delivered MUs (25). The PTV shells are implemented with margins in millimetres around the PTV. The full wish-list can be found in appendix part C.

2.4.2. Semi-analytical dose engine

A semi-analytical dose engine was used to perform the three-dimensional uncertainty analysis. This dose engine was developed at the Department of Medical Physics & Technology from the TUD by Burlacu et al. [10] It is analytical as its fundamental basis is the Boltzmann transport equation. With approximations to the transport equation, energy deposition and dose in tissue can be calculated forward. Additionally, the first-order effect of HU perturbations on dose delivery is calculated backwards with an adjoint analysis. The outcome is a response function linking the forward and backward results. The purpose of this analysis is to allow fast re-computations of the dose when anatomy changes during treatment occur (online adaptive radiotherapy). Anatomy changes are linked to HU perturbations. For the uncertainty analysis, pencil-beam placement errors are of interest. By shifting the input CT virtually, the effect of these displacements on the dose can be quantified with the adjoint analysis. More information about this dose engine can be found in the work from Burlacu et al. [10] and in my literature study (appendix part D).

2.4.3. Gamma analysis

The dose distributions resulting from iCycle and the semi-analytical dose engine are compared with a gamma analysis. The gamma analysis is considered the golden standard for comparison between two dose distributions. The gamma index was originally introduced by Low et al. (1998) [30] The gamma index for a given simulated voxel position (r_s) compared with a reference position (r_c) is given by Eq. 2.10

$$\gamma(r_s) = \min\{\Gamma(r_s, r_c)\}\forall\{r_c\} \tag{2.10}$$

In this equation, the $\Gamma(r_s, r_c)$ for a voxel is given by Eq. 2.11. The $\Gamma(r_s, r_c)$ consists of two parts. The first term relates to the distance-to-agreement and the second is the dose difference. In this equation, the r is the distance between the reference and the simulated voxel. The δ is the dose difference. Both terms are normalised with a user-selected margin in distance and dose difference (respectively Δd_M and ΔD_M). After the full gamma analysis of two dose-deposition matrices, every voxel has a value for $\gamma(r_s)$. If $\gamma(r_s)$ of a voxel is smaller than one, the two compared dose-deposition matrices lie within the selected margins. If $\gamma(r_s)$ is bigger than one, the reference and simulated matrices cannot be matched within the margins. The gamma pass rate is defined as the fraction of the simulated dose matrix with a $\gamma(r_s)$ below one. A graphical representation of the gamma index for one position is shown in Fig. 2.4. There are two types of gamma pass rates used, the local and global gamma pass rates.

With a local gamma analysis, two different dose-deposition matrices are compared with each other and analysed according to the set margins. A global analysis varies. Here, the dose-deposition matrices are compared to a single dose value, e.g. the prescribed dose. Additionally, a lower dose threshold is often used for the gamma analysis. With the threshold, low-dose areas, which might disturb the gamma analysis are excluded.

$$\Gamma(r_s, r_c) = \sqrt{\frac{r^2(r_s, r_c)}{\Delta d_M^2} + \frac{\delta^2(r_s, r_c)}{\Delta D_M^2}} \tag{2.11}$$

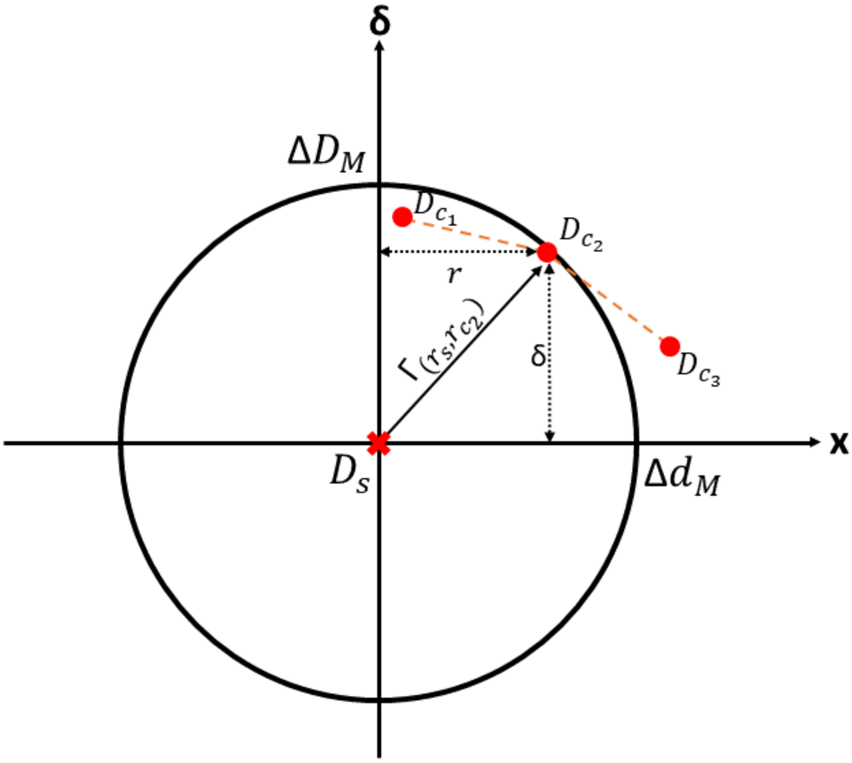


Figure 2.4: 2D representation of the gamma index between a simulated (D_s) and three reference dose points (D_{c_i}). Note that D_{c_1} is located within the margins in dose and distance $(\Delta D_M, \Delta d_M)$, D_{c_2} on the margin boundary and D_{c_3} outside the margin space. This implies that the $\Gamma < 1$ for point D_{c_1}, $\Gamma = 1$ for D_{c_2} and $\Gamma > 1$ for D_{c_3}. The minimal Γ value equals the value for γ. In this example, the gamma index is passed as at least one dose point is located within the margin space. This figure is adapted from [30].

3

1D & 2D Analytical Simulations

In this chapter, the full sensitivity and uncertainty analysis for a one and two-dimensional setup is given. First, the general procedure of the analysis is described for the one-dimensional setup, thereafter, it is extended toward the two-dimensional analysis.

3.1. One-dimensional analysis

In this section, the one-dimensional analysis is explained. First, the simulation of individual pencil beams is described in 3.1.1. The optimisation procedure of the pencil beam weights is stated in 3.1.2. Then the uncertainty in spot position and beam current is explained in respectively 3.1.3 and 3.1.4. In section 3.1.5 the two types of uncertainties are combined and analysed.

As all the pencil beams are on a single line in the 1D setup, the PBSDR is not used in this analysis. There is no variation in scan pattern possible for the 1D situation. Consequently, the reported dose rates are DADRs.

3.1.1. Pencil beam simulation

The individual pencil beams are simulated according to the Gaussian dose distribution with means μ_i and standard deviation σ_{pb}. The pencil beam standard deviation is fixed based on beam measurements and is approximately equal to 3.2 mm. The inter-pencil beam distance equals 0.5 cm. This is in line with the first-in-human clinical FLASH trial. [29] The means are linearly spaced between -2.5 cm and 2.5 cm. This altogether results in a total of eleven individual Gaussian beams. To convert the probability density function to a dose value [Gy] instead of a probability, the normalised Gaussian is multiplied with a factor: $D_{prescribed} * \sqrt{2\pi\sigma}$. The total dose along the x-axis is given by the summation of the individual Gaussian pencil beams. In Fig 3.1 an example dose profile for the eleven Gaussian pencil beams is given together with the summed dose of the pencil beams.

The simulated pencil beams with σ_{pb} at a fixed depth do not include any lateral spreading due to atomic, electronic and nuclear interactions. Whereas, in reality, the energy deposited at a certain position consists of some lateral spreading. This is the so-called halo effect. [1] The dose can be modelled as the summation of two Gaussian pencil beams, one for the 'pure' pencil beams and the other one for the possible interactions at a fixed depth. The two Gaussian functions have equal means but varying standard deviations σ_{pb} and σ_h. In Eq 3.1 the relation between the two is given. Where the parameter α denotes a relative weight between the two Gaussian functions and C_i a Gaussian-specific constant to convert the distribution to dose. [11]

$$D = \alpha C_{pb} G(\mu, \sigma pb) + (1 - \alpha) C_h G(\mu, \sigma_h) \tag{3.1}$$

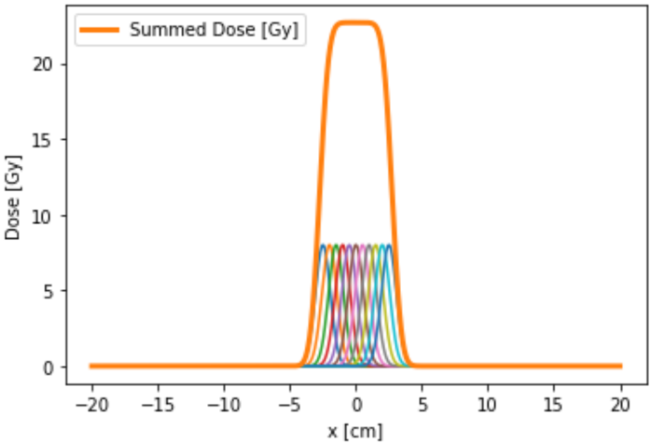


Figure 3.1: A dose profile for a one-dimensional situation with the underlying eleven pencil beams. The spacing between pencil beams is 0.5 cm, placed between -2.5 cm and 2.5 cm. The σ_{pb} is equal to 3.2 mm. The prescribed dose $(D_{prescribed})$ is 8 Gy for all pencil beams.

To be able to compute the dose including the halo effect, the parameter α and σ_h have to be determined. Weight factor α is range dependent and can be modelled according to Eq 3.2 where t equals the traversed depth over total proton range $t = (\frac{d}{R})$. The range of a proton is energy-dependent and is approximately equal to $R \approx 0.00244E^{1.75}$. [6, 37]

$$\alpha(R) = \text{min}(1, \beta_0(R) + \beta_1(R) \cdot t + \beta_2(R) \cdot t^2) \tag{3.2}$$

The values for the betas can be computed with the equation below. The fitting of the eta values is done on data acquired from an IBA cyclotron at the Massachusetts General Hospital (MGH). The determined eta values are given in table 3.1. [37]

$$\beta_i(R) = \eta_{0,i} + \eta_{1,i} \cdot R + \eta_{2,i} \cdot R^2 \tag{3.3}$$

Table 3.1: Fit values in order to compute α (R) and β_i (R) from Eq 3.2 & 3.3

	β_0	β_1	β_2
$\eta_{0,i}$	1.002	2.128E-3	-2.549E-3
$\eta_{1,i}$	-5.900E-4	-2.044E-2	2.125E-2
$\eta_{2,i}$	0	3.178E-4	-3.788E-4

For our one and two-dimensional analysis, the depth fraction (t) was fixed for a proton of 244 MeV and a depth of 5 cm. Solving both Eq 3.2 & Eq 3.3 returns an α of 0.942

The remaining unknown to compute the dose, σ_h, can be calculated according to Eq 3.4. This relation is also found with proton beam data from the MGH. [37]

$$\sigma_h(R) = 6.50 - 0.34 \cdot R + 0.0078 \cdot R^2 \tag{3.4}$$

For a beam of 244 MeV, the proton range (R) is approximately 36.8 cm. This yields a σ_h of 4.5 cm, around 14 times the standard deviation of the 'pure' pencil beam (σ_{pb}).

The new dose deposition, including the halo effect, is shown in Fig 3.2. By visual inspection, it is evident that the dose deposition along a one-dimensional axis changes as compared to Fig 3.1. The dose plateau of Fig 3.2 is less homogeneous, the summed dose reaches a higher max dose and the dose fall-off region is larger due to the lateral spreading of the energy.

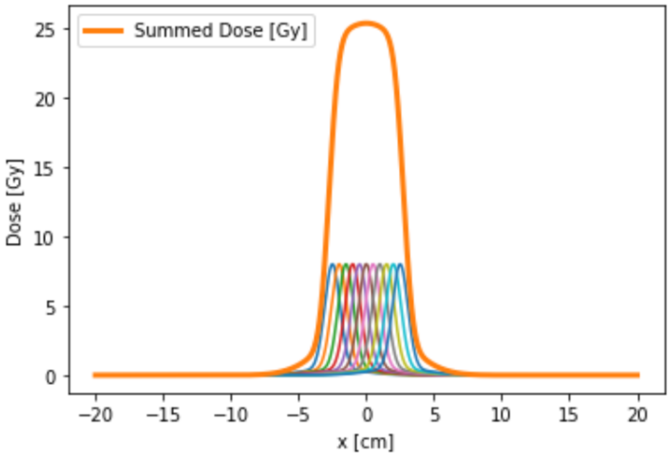


Figure 3.2: A dose profile for a one-dimensional situation including the halo effect with the underlying eleven pencil beams. The spacing between pencil beams is 0.5 cm, placed between -2.5 cm and 2.5 cm. The σ_{pb} equals 3.2 mm and the σ_h equals 4.541 cm. The prescribed dose $(D_{prescribed})$ is 8 Gy for all pencil beams.

To be able to calculate a dose rate profile related to the summed dose profile, the irradiation time per spot position is needed. The time per spot was calculated such that a single spot yields 100% $D_{prescribed}$ in the Bragg peak. This gives a theoretical max dose rate per spot. In practice, multiple spots contribute to the prescribed dose, yielding lower dose rates than the theoretical maximum. For the one and two-dimensional analysis, this irradiation time per spot was kept constant and calculated based on beam data from the HPTC. For a beam of 244 MeV, it was determined that 0.368 Giga protons give 1 Gy. [40] With this beam characteristic, the irradiation time per spot can be calculated according to Eq 3.5, where EC is the elementary charge $(1.602E-19[C/pr])$. Solving the equation for a beam current of 250 nA yields an irradiation time of 0.0019 seconds. The instantaneous maximal dose rate in FLASH mode thus equals: $8[Gy]/0.0019[s] \simeq 4241[Gy/s]$

$$T_{irr}[s] = \frac{0.368E9[pr/Gy] \cdot D_{prescribed}[Gy]}{\frac{I[C/s]}{EC[C/pr]}} \tag{3.5}$$

The DADR corresponding to the dose profile of Fig 3.2 is shown in Fig 3.3. Note that each pencil beam is at 100% of the max dose rate as the pencil beams all contribute 8 Gy to the dose profile. Furthermore, take notice that the highest dose rate is achieved at the first and last pencil beams. At these positions, the dose is given with fewer individual pencil beams as compared to the middle of the one-dimensional field.

3.1.2. Pencil beam weight optimisation

A pencil beam weight optimisation is added to the analytical model to improve the dose deposition profile. The weights of the pencil beams were selected such that the max of the summed dose profile equals the prescribed dose. The cost function used for optimisation includes a term for the dose difference with the prescribed dose, D98 and D2. These three terms have the same relative weight. With optimisation on the prescribed dose for the dose-at-volume metrics, the resulting dose is as homogeneous as possible. The cost function takes into account dose values from the mean of the first until the last pencil beam. For the pencil beams displayed in Figs 3.1, 3.2 and 3.3 this implies dose values between -2.5 cm and 2.5 cm. With this selection of dose values, solely the dose plateau is optimised and the dose fall-off region is disregarded. In Eq 3.6 the mathematical form of the cost function (CF)

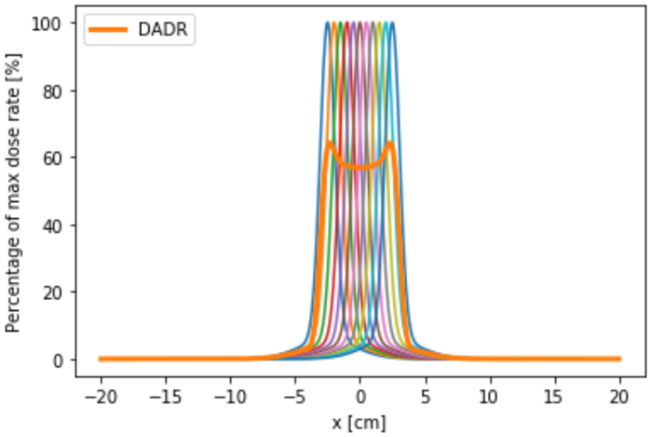
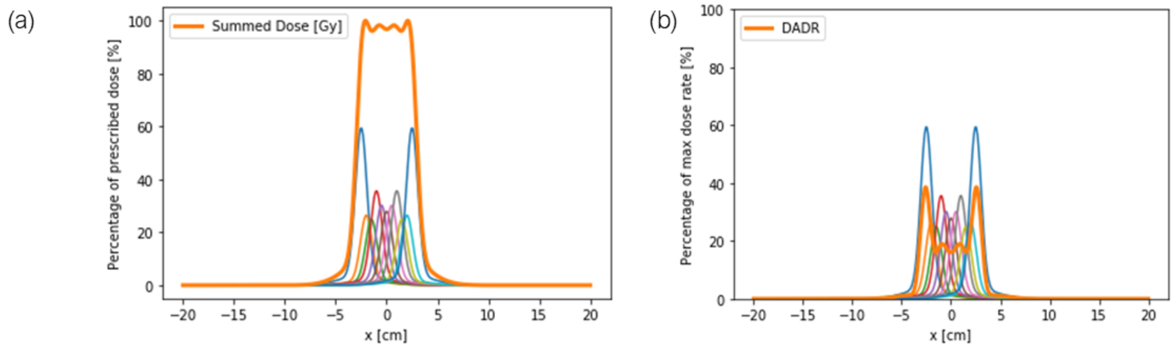


Figure 3.3: A relative DADR profile for a one-dimensional situation including the halo effect with the underlying eleven pencil beams. The spacing between pencil beams is 0.5 cm, placed between -2.5 cm and 2.5 cm. The σ_{pb} equals 3.2 mm and the σ_h equals 4.541 cm. The prescribed dose $(D_{prescribed})$ is 8 Gy for all pencil beams.

is given for (j) pencil beams and (i) voxels. This function is iteratively minimised for an input weight matrix (w). After full optimisation, the last weight matrix is stored and used for dose and dose rate computations. To simulate a clinical treatment optimisation, the weight per pencil beam was bound between 0 and infinity.

$$CF = \sum_{i=\mu_1}^{i=\mu_N} \sum_{j=1}^{j=N} + \Big(D_{i,j}(w_j) - D_{pres}\Big)^2 + \Big(D_{i,j}(w_j) - D_{98}\Big)^2 + \Big(D_{i,j}(w_j) - D_2\Big)^2 \tag{3.6}$$

The calculated dose and dose rates and the found weight factors are shown in Fig 3.4. The max of the summed dose profile equals the prescribed dose and the profile is as smooth as possible in the region between the first and last pencil beams. The dose rate profile shows similar behaviour as shown in Fig 3.3, namely, the dose rate is highest at the first and last pencil beams. The individual pencil beams do not reach the maximal theoretical dose rate as the weights are below 1.0. The underlying individual pencil beams with their corresponding weights are symmetrical around the middle pencil beam with weight (w_6).



(c)

w_1	w_2	w_3	w_4	w_5	w_6	w_7	w_8	w_9	w_{10}	w_{11}
0.59	0.26	0.25	0.36	0.30	0.28	0.30	0.36	0.25	0.26	0.59

Figure 3.4: The dose (a) and dose rate (b) profiles with the underlying eleven pencil beams after pencil beam weight optimisation with Eq 3.6. The found weights are listed in (c). The inter-pencil beam distance is 0.5 cm and the beams are placed between -2.5 cm and 2.5 cm. The σ_{pb} equals 3.2 mm and the σ_h equals 4.541 cm. The weights are selected such that the max of the summed dose equals the prescribed dose.

3.1.3. Spot position uncertainties

Proton therapy requires the positioning of pencil beams with sub-millimetre precision in the patient. Small deviations in spot position can influence dose conformity. Therefore, the spot position uncertainty quantification and minimisation remain one of the major research directions within the proton therapy field.

After optimisation of the dose profile, spot uncertainties can be added to our analytical model. There are two possible ways to introduce spot position uncertainties. The first option is to add all possible spot position errors with their associated probability at once with a convolution. With this approach, the computed dose and dose rate profiles will resemble the average of infinite fractions. The convolutional approach provides a good insight into the sensitivity of dose and dose rate due to small errors in spot positioning. In Eq 3.7 the analytical description of the dose including spot uncertainties is given. The summed dose of N pencil beams is convolved by a Gaussian distribution with mean zero and standard deviation (σ_u). The size of the standard deviation was kept constant at 0.5 mm for the one and two-dimensional analysis and is based on beam positioning data from the HPTC. After the convolution of the dose profile, the dose rate profile also includes spot uncertainty as it is calculated from the dose profile.

$$\bar{D} = \Big[\sum_{i=\mu_1}^{i=\mu_N} D_0(\mu_i, \sigma_{pb}, \sigma_h) \Big] \circledast G(\mu = 0, \sigma_u) \tag{3.7}$$

In Fig 3.5 the convolved summed dose and dose rate profile are shown together with the profiles without including spot position uncertainties. When comparing the two outcomes with each other, it is hard to observe differences. The only visual difference between the two situations is at the max of the dose rate profiles. Here the dose rates with spot uncertainties yield lower max dose rates due to smearing out of the dose because of the Gaussian uncertainty distribution.

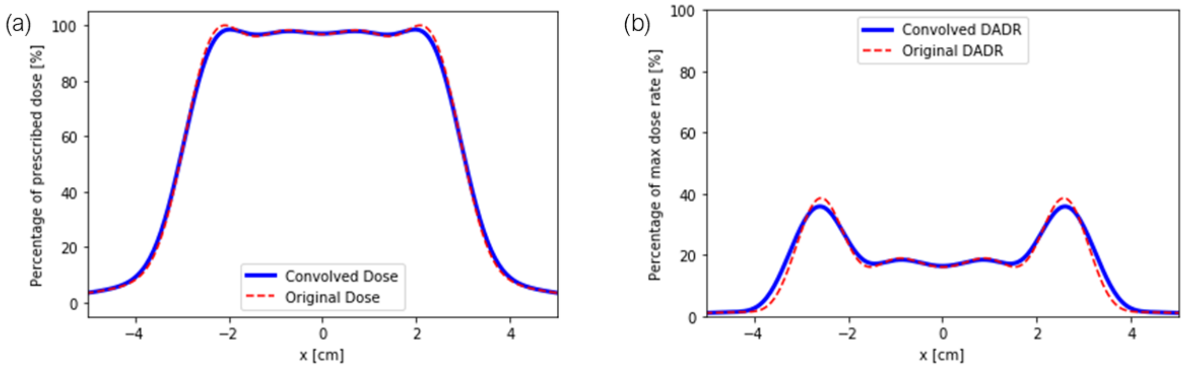


Figure 3.5: The summed dose (a) and dose rate (b) profiles with spot uncertainties (blue solid line) and without spot uncertainties (red dashed line). The σ_{pb} equals 3.2 mm, the σ_h 4.541 cm and the σ_u 0.5 mm.

The second possibility to analyse spot uncertainties is by sampling dose and dose rate profiles with randomly generated spot errors from the Gaussian distribution with σ_u. Per pencil beam location a random spot deviation is drawn. This deviation is added to the 'true' spot position such that the dose and dose rate profiles can be sampled with uncertainties. To make this sampling procedure relevant for the uncertainty and sensitivity analysis, many dose and dose rate samples are needed (MC approach). With sampling, not the average dose and dose rate deviation of infinite fractions is found, but rather the possible deviation on a single fraction basis per position. As sampling with a Gaussian distribution can be skewed by outliers, the maximum and minimal dose and dose rate profiles are not a realistic representation on an individual treatment basis. To analyse the maximal and minimal deviations due to spot position errors, the 5th percentile and 95th percentile are used for the dose and dose rate profiles. For the 5th percentile, this implies that at least 95% of the fractions given, reach that minimal dose and dose rate. In Fig 3.6 the 95th percentile maximal and 5th percentile minimal dose and dose rates are given together with their relative difference as compared to the profiles without the inclusion of uncertainties. From sub-figure (c) it can be concluded that the dose profile changes up to approximately 20%. The deviation in dose rate is bigger and is up to around 40% as shown in sub-figure (d). From the results, it is also evident that the biggest relative difference is situated at the dose fall-off region of the field. Note that the relative difference is not constant with dose, nor with dose rate. This results from the fact that at the end of the field, relatively few pencil beams are in play to deliver the dose. In this situation, the effect of the spot position error on dose and dose rate is higher as the spot error is less averaged out. Also, observe from both sub-figure (c) and (d) that the relative difference is higher for the 95th percentile maximal dose and dose rate than the 5th percentile minimal dose and dose rate at the end of the field. This implies that the relation between the dose per spot and DADR is non-linear. In other words, the DADR is sensitive to the rate of dose change at the low spot-density region.

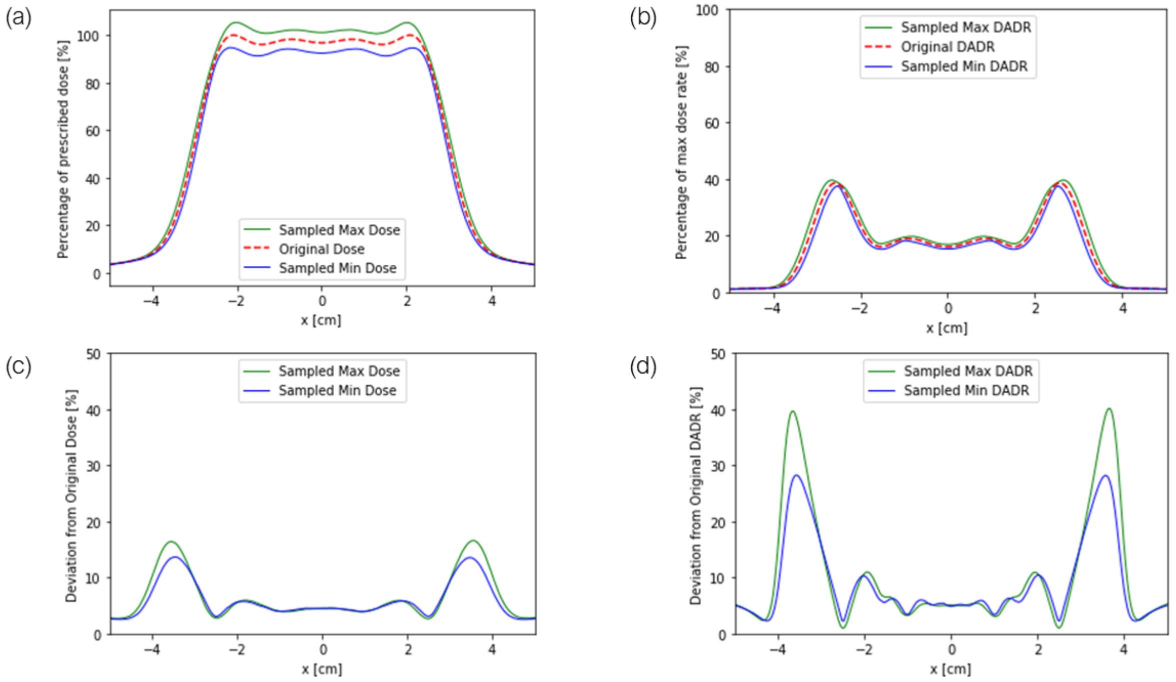


Figure 3.6: The summed dose profile (a) and dose rate (b) profiles after including spot position uncertainties for the 95th percentile maximum (solid green line), 5th percentile minimum (solid blue line) and without spot uncertainties (red dashed line). The relative difference between the maximum and minimum with the dose (c) and dose rate (d) without spot errors is also depicted. 10,000 samples were taken for this figure.

3.1.4. Proton beam-current uncertainties

The cyclotron beam current is another machine parameter influencing the dose rate profile. Especially in FLASH radiotherapy, where high doses and dose rates are being used, deviations in beam current can influence the total FLASH effect.

At the HollandPTC up to ± 10% day-to-day variation in beam current is present in general. At the daily QA of the beam, the current is measured and treatment plans are corrected accordingly. This implies that the dose is still given as planned. However, due to beam current fluctuations, the dose rate is affected. Recall from Eq 3.5 that the irradiation time per spot is inversely proportional to the beam current. When the current changes ± 10%, the irradiation time and the dose rate are altered by ± 10% as well. This creates the following ranges for respectively the irradiation time per spot and corresponding maximal instantaneous dose rate: $T_{irr} = [0.0017s, 0.0021s]$, $\dot{D}_{max} = [3854Gy/s, 4711Gy/s]$.

The uncertainty analysis on dose rate due to beam current fluctuations is done with sampling. Random samples are made by drawing random deviations from a Gaussian distribution with a relative standard deviation of 5%. This standard deviation is based on the beam current database of HollandPTC. In

Fig 3.7 the 95th percentile maximal and 5th percentile minimal relative dose rates are depicted together with the deviation from the analytical dose rate without beam current errors. Sub-figure (b) confirms that the beam current affects the dose rate linearly as the deviation of both the max and min sampled dose rate is constant over the whole field. In other words, the beam current error is equal for each delivered pencil beam. The 95th percentile max and 5th percentile min differences correspond with the standard deviation of the randomly drawn Gaussian current fluctuation.

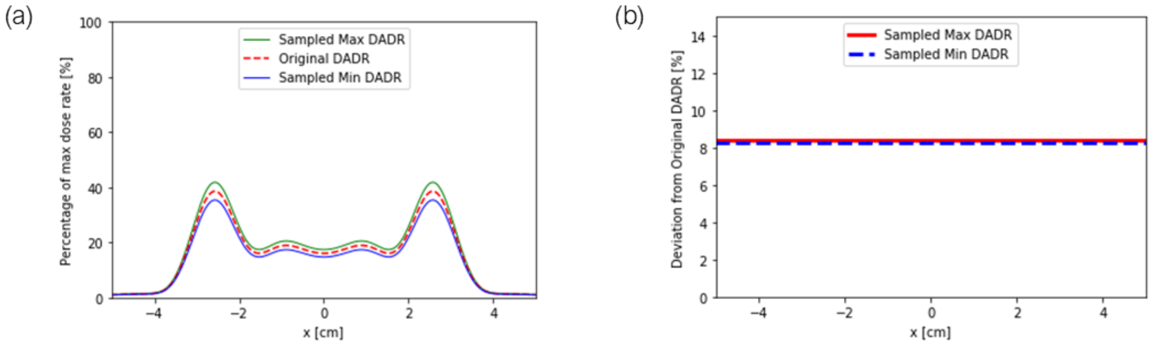


Figure 3.7: The summed dose rate profile (a) after including beam current uncertainties for the 95th percentile max sample (solid green line), 5th percentile min sample (solid blue line) and without spot uncertainties (red dashed line). The relative difference between the maximum and minimum with the dose rate (b) without beam current uncertainties is also depicted. 10,000 samples were taken for this figure.

3.1.5. Combination of spot position and proton beam current uncertainties

The effect of both spot and beam current uncertainties on dose and dose rate were also investigated for the one-dimensional situation. With sampling, the two types of errors were added to the analytical model. The random errors were drawn from the same Gaussian distributions as described in 3.1.3 and 3.1.4. As mentioned before, a change in beam current does only affect the dose rate and not the dose delivered. Therefore, the dose profile with the inclusion of both types of uncertainty is equal to the sampled dose profiles shown in Fig 3.6. The dose rate profile of the field after including both uncertainty in spot position and beam current is shown in Fig 3.8. In essence, the profile is a combination of the profiles shown in Fig 3.6(b) and 3.7(a). At the fall-off region of the field, the error is due to spot position uncertainties. In the field, the dose rate is sensitive to both the beam current and spot position uncertainty.

This relation is proven by Fig 3.9. Here the relative difference between both the 95th percentile maximum and 5th percentile minimum dose rate and the error-free dose rate is depicted. The difference is given for three situations: 1) only spot uncertainties, 2) only beam current uncertainties and 3) the combination of both spot position and beam current uncertainties. From both sub-figures, it is evident that the constant error induced by the beam current uncertainty acts as a plateau. The combination of both uncertainties is never below the constant blue line. From the figure, it is also clear that the combined relative error is a combination of both types of uncertainties. The peaks at $\pm 3.8cm$ correspond to the fall-off region of the field. Here the relative error is driven by the uncertainty in the spot position. In the field itself, ranging from -2.5 cm to 2.5 cm, the relative dose rate deviation is dominated by the uncertainty in beam current. From the data presented here, one can conclude that the fluctuation in beam current is the main factor influencing the dose rate in the field. However, in the dose and dose rate fall-off region, the relative difference is driven by the error in the spot position. However, the relevance of this deviation is questionable as the dose and dose rate are lower than the dose and dose rate in the field.

3.2. Two-dimensional analysis

In this section, the one-dimensional analysis is extended to two-dimensional. The pencil beam creation is similar to the procedure described in 3.1.1. Minor differences are that in the two-dimensional

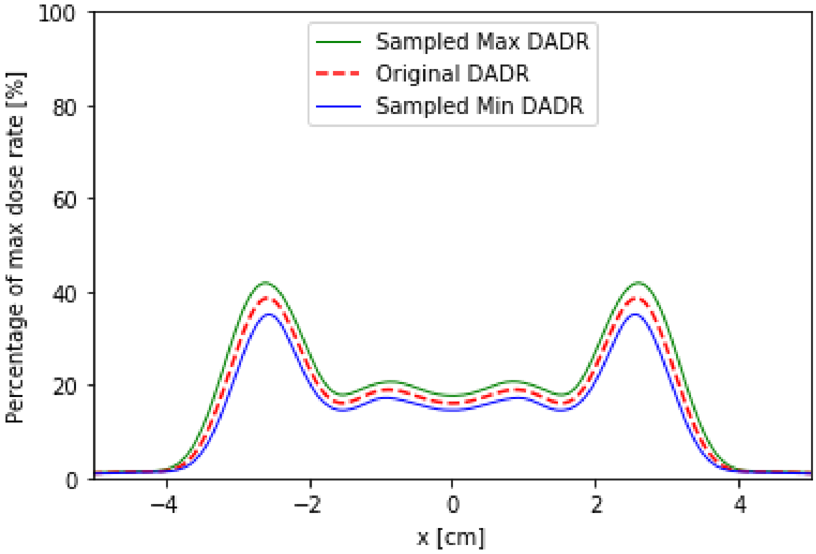


Figure 3.8: The summed dose rate profile after including spot position and beam current uncertainties for the 95th percentile max sample (solid green line), 5th percentile min sample (solid blue line) and without spot uncertainties (red dashed line). 10,000 samples were taken for this figure.

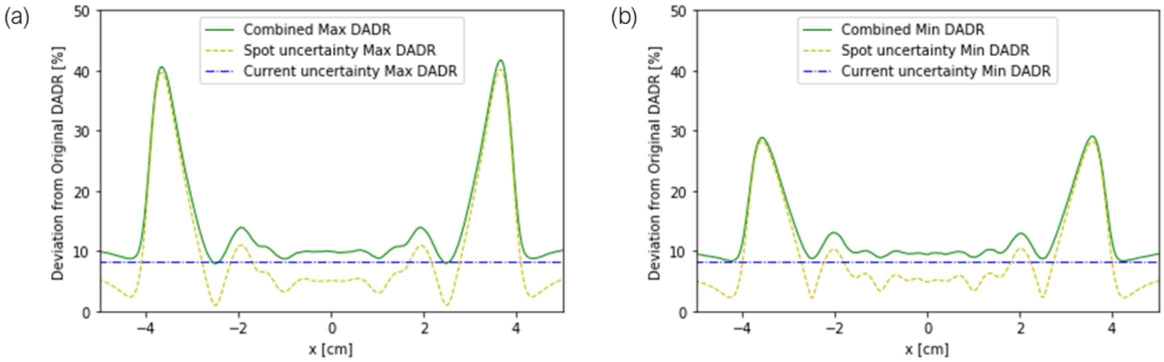


Figure 3.9: The relative difference for each type of uncertainty between the 95th percentile maximum (a) and 5th percentile minimum (b) sampled DADR compared with the error-free dose rate profiles. 10,000 samples were taken for this figure.

setup, 13x13 pencil beams are positioned on a two-dimensional grid between x and y of [17.0 cm, 23.0 cm] instead of the eleven pencil beams along the one-dimensional x-axis. The six-by-six centimetre scanning pattern is chosen based on field sizes used in the first-in-human FLASH study. [12, 29] Furthermore, the weights of the individual pencil beams are scaled such that the median dose in the field equals the prescribed dose. In the 1D calculations, they were weighted according to the maximum dose value. The inter-pencil beam distance still equals 0.5 cm. The individual pencil beams are two-dimensional (multivariate normal distribution) Gaussian distributions with σ_{pb} and σ_h. In order to convert the probability density functions to dose, multiplication with a pencil beam-specific constant is needed. The multiplication factor varies from the one-dimensional setup and equals $D_{prescribed} * \sqrt{4\pi^2 \det(\sigma_i)}$, where the determinant of the two-dimensional standard deviation is needed. For all pencil beams simulated, the two dimensions are assumed independent. In other words, the determinant of the standard deviation (σ_i) is the standard deviation squared:

$$\det(\sigma_i) = \det \begin{bmatrix} \sigma_i & 0 \\ 0 & \sigma_i \end{bmatrix} = \sigma_i^2$$

For the two-dimensional setup, σ_{pb}, σ_h, α, T_{irr} and $\dot{D}_{max}$ are equal to the values used in the one-dimensional analysis. However, as the dose rate depends on the pencil beam scan pattern, dose rates are not computed as DADRs but rather as PBSDRs. The PBSDR is the leading metric used in all Varian proton therapy centres. [17] The results acquired for the two-dimensional analysis are all computed with a PSBDR threshold of 10% $D_{prescribed}$ unless stated otherwise. Pencil beam weight optimisation is again done with the minimisation of the cost function given in Eq 3.6.

In the following section, the dose and dose rate profiles without the inclusion of uncertainties are given. Thereafter, the uncertainties in spot position and beam current are added and described in respectively sections 3.2.2 and 3.2.3. Then in section 3.2.4, both types of uncertainties are included in the model and analysed. Particularly of interest is the sensitivity of the PBSDR for the two types of uncertainties as this will be the leading dose rate metric for FLASH proton therapy at HollandPTC.

3.2.1. Error-free situation

The optimised pencil beam weights are shown in Fig 3.10 together with the resulting dose in the field. Note that the weights found are again symmetrical around the middle pencil beam positioned at 20 cm, 20 cm. Higher weights are required for pencil beams on the first and last rows of the field to make the dose profile delivered as homogeneous as possible. Similar to the one-dimensional setup, weight optimisation is not done for the dose fall-off region.

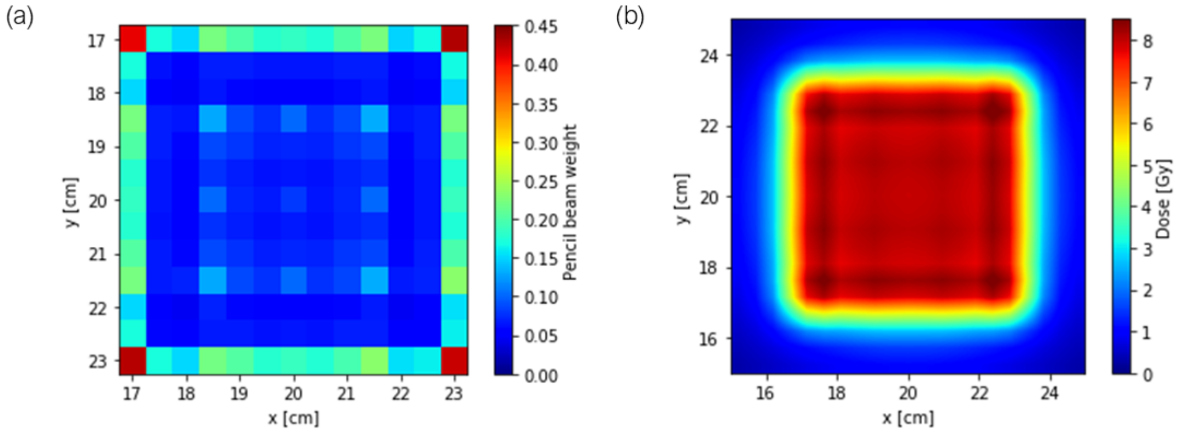


Figure 3.10: The optimised pencil beam weights (a) with the corresponding dose 2D dose profile (b). The inter-pencil beam distance is 0.5 cm and the beams are placed between x,y = 17.0 cm and 23.0 cm. The σ_{pb} equals 3.2 mm and the σ_h equals 4.541 cm. The weights are selected such that the max of the summed dose equals the prescribed dose. The dose values shown are bi-linearly interpolated.

The pencil beam scanning pattern used in current FLASH research is the so-called snake pattern. This is a line-by-line scanning pattern. The primary direction of scanning is in the y direction, the second direction is the x direction. This is because the scanning velocity of the proton beam is four times quicker in y than in x (20 m/s against 5 m/s). The underlying cause is the rotational velocity of the treatment table and nozzle. The travel time between the two spots is disregarded in this analysis. In Fig 3.11 the line-by-line pattern is given for the simulated FLASH field together with the corresponding PBSDR. The dose rate displayed in sub-figure (b) is symmetrical around the middle of the field. The reason for this is twofold. Firstly, the six-by-six-centimetre scanning pattern results in symmetrical weight factors for the chosen cost function (see Fig 3.10). Secondly, in the definition of the PBSDR, the scanning direction of the pattern is not taken into account. In other words, scanning the pattern in reverse yields equal dose rates. When switching from one-dimensional to two-dimensional, the dose rate in the field drops. Keep in mind that 10% of the maximal dose rate corresponds with 424.1 Gy/s which is several orders higher than dose rates seen in conventional proton therapy. From sub-figure (b) it is also evident that the dose rates are highest on the first and last row of the pattern. In particular, the dose rate around the first and last spot of the pattern is the highest. The dose rate is highest on these rows and locations because here the dose given is established with relatively few pencil beams with higher weights. In the middle of the field, the lowest and more constant dose rate values of around 10 to 15% of the theoretical maximal dose rate are found as here many spots (with lower weights as well) contribute to the dose.

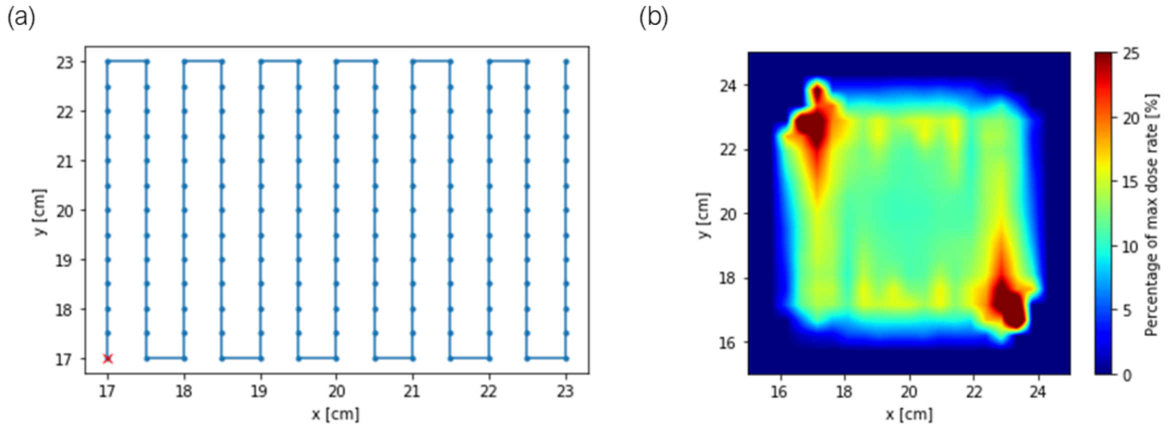


Figure 3.11: The pencil beam scanning snake pattern (a) together with the resulting PBSDR distribution (b). The start of the scanning pattern is located at the red cross. The dose rate profile shown is a bi-linearly interpolation of the found PBSDRs. The threshold of the PBSDR is 10% of the prescribed dose.

Another way to visualise the PBSDR for a given field is with a PBSDR volume histogram (PBSDR-VH). In such a representation, a relative volume (y-axis) is given which receives at least a certain percentage of the theoretical max dose rate (x-axis). With PBSDR-VHs, various dose rate profiles can be compared with each other. In Fig 3.12 the PBSDR-VH of dose rates with varying PBSDR threshold values is given for a relative volume, V95 [%]. The V95 is defined as the area which receives at least 95% of the prescribed dose in the field defined with the individual pencil beams (Fig 3.10).

From the figure, it is clear that the PBSDR metric is sensitive to the dose input threshold. For low dose thresholds (0.1%, 1.0%) of the prescribed dose, the PBSDR is in the tens of percentage region of the theoretical maximal dose rate. For the 5.0%, 10% dose threshold of the prescribed dose, the PBSDR increases for the whole field. For even higher dose thresholds (20 % and 40%), the dose rate becomes unrealistically high for the field. Also, note that the volume histograms of the 20% and 40% of the prescribed dose cross each other. From a clinical perspective, this should never happen. The figure indicates a drawback of the PBSDR metric as the computed dose rates depend on the selected dose threshold.

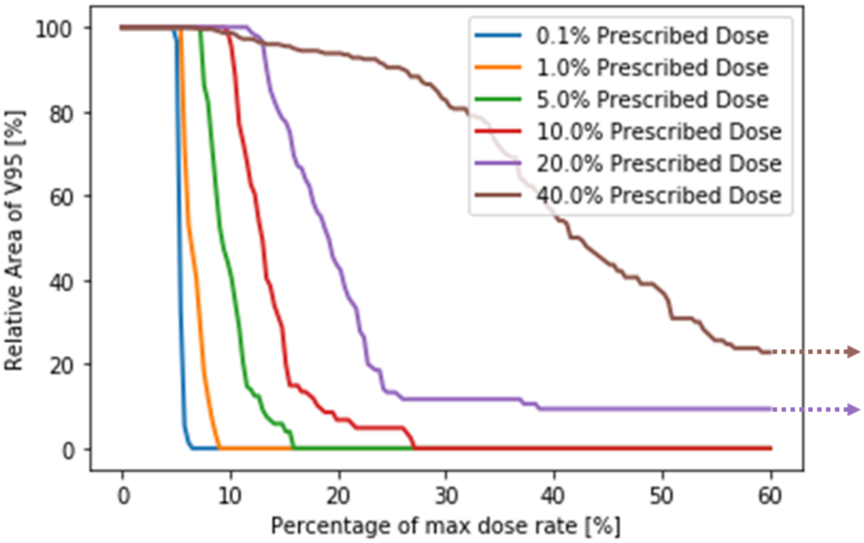


Figure 3.12: Pencil beam scanning dose rate volume histogram (PBSDR-VH) of dose rate profiles with varying dose thresholds. All PBSDRs are computed with the snake pattern shown in Fig 3.11.

3.2.2. Spot position uncertainties

Uncertainties in spot position are again added with two approaches: (1) the convolution approach and (2) the sampling approach. They are described in detail in section 3.1.3. In Fig 3.13 the relative difference between the convoluted and error-free dose and dose rate profile are depicted. Sub-figure (a) shows the relative difference in dose. For the spots on the first and last rows and columns of the scan pattern, relatively big differences are found between the two profiles. The differences in dose are the highest here because these spots are situated on the edge of the pattern. The change in spot position will not be compensated by other pencil beams resulting in a lower dose at the edge of the pattern.

The change in dose rate is more substantial (sub-figure (b)). On the spots on the edge of the pattern, the addition of spot position uncertainties results in a lower dose rate. Surrounding the field, some hot spots of increased dose rate can be found due to spot position errors. The hot spots reside around pencil beams with relatively high weights (See Fig 3.10(a)). The figure proves that the PBSDR is sensitive to dose changes at low-density pencil beam regions with relatively high weights. For dose changes in the middle of the field, where many spots with relatively low weights contribute to the dose, the PBSDR remains the same.

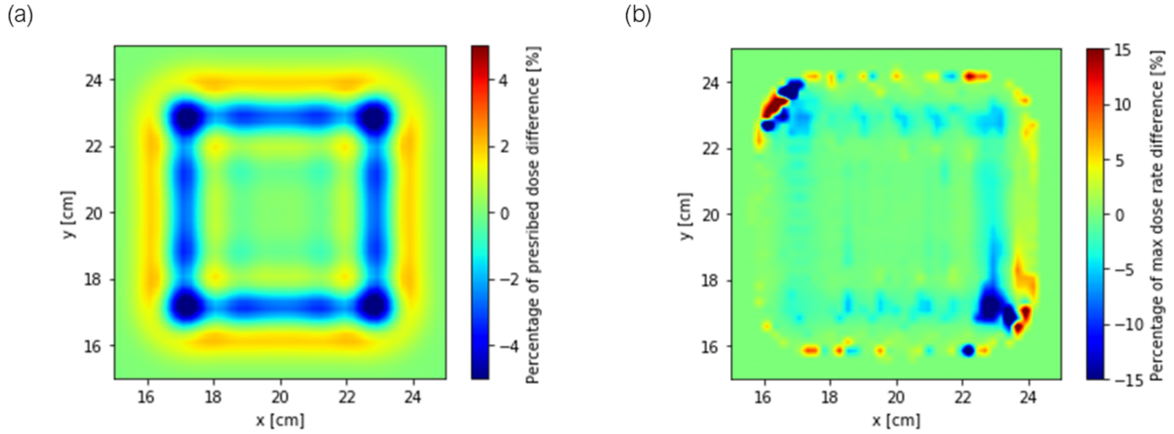


Figure 3.13: The convolution minus error-free dose (a) and dose rate (b) relative difference profiles for the snake scanning pattern. Pencil beams are placed between x, y = 17.0 cm, 23.0 cm. The σ_{pb} equals 3.2 mm, the σ_h 4.541 cm and the σ_u 0.5 mm. The relative dose and dose rate difference values are shown bi-linearly interpolated. The threshold for the PBSDR is 10% of the prescribed dose.

The same uncertainty and sensitivity analysis can be done with Gaussian spot position error sampling with σ_u = 0.5 mm in both the x and y directions. It is assumed that the weight between the error in both directions is equal. In Fig 3.14 the relative difference between the sampled 95th percentile max and 5th percentile min dose and dose rate and the error-free situation is shown. For the difference in dose shown in sub-figure (a), it is evident that the biggest changes in dose are located in the surrounding of the field. Only in the corners and first and last row of the field, deviations in dose can be observed. Everywhere else in the field, the difference is dose is relatively low. Sub-figure (b) confirms this behaviour. Only surrounding the field, big relative differences in dose rate are found. The biggest relative differences are again situated around the four corners of the field, where the spots have the highest weight. Note that this is similar to the results shown in Fig 3.13.

The correlation of both PBSDR and pencil beam weight with the relative sampled error due to spot position changes is shown in Fig. 3.15. The relative error correlates with PBSDR (sub-figure (a)) and with pencil beam weights (sub-figure (b)). Even though the dose rate does not linearly depend on small changes in spot position (recall Fig. 3.6(d)), a linear fit for the relative error is shown as well. This gives an insight into the behaviour of the error with increasing PBSDR and pencil beam weight. The figure proves that the PBSDR is more sensitive to position changes at high-weighted pencil beams at relatively high PBSDRs. When considering an error-free PBSDR profile, this should be taken into account.

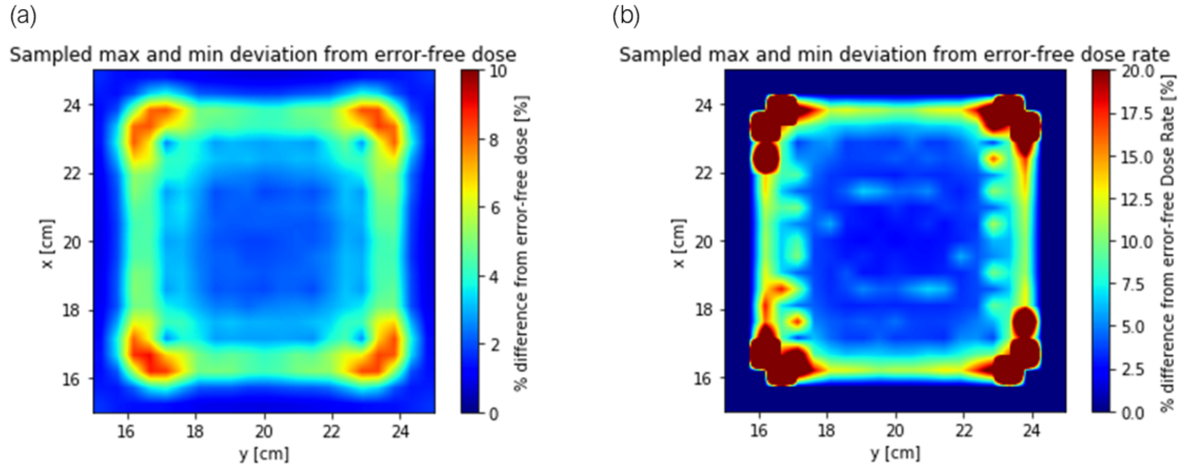


Figure 3.14: The absolute relative difference between the sampled spot position deviation max/min and error-free dose (a) and dose rate (b). The random spot position deviations are drawn from a Gaussian distribution with a 0.5 mm standard deviation. The PBSDRs are calculated according to the snake scanning pattern. Pencil beams are placed between x, y = 17.0 cm, 23.0 cm. The σ_{pb} equals 3.2 mm and the σ_h 4.541 cm. The relative dose and dose rate difference values are shown bi-linearly interpolated. The threshold for the PBSDR is 10% of the prescribed dose. The max and min are the 95th percentile and 5th percentile dose and dose rate values of 500 samples.

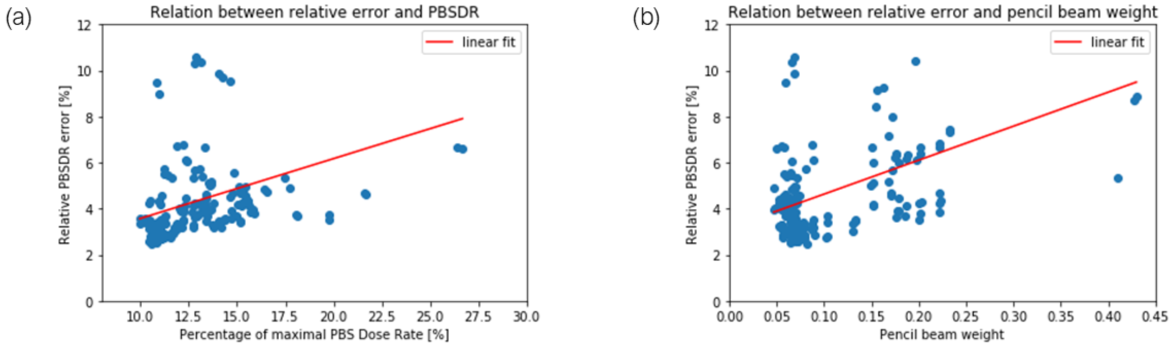


Figure 3.15: The correlation between the PBSDR on the x-axis and the relative max and min sampled error on the y-axis (a) and the correlation between the pencil beam weights on the x-axis and the relative error on the y-axis (b). The 95th percentile maximal and 5th percentile minimal values correspond to the relative error. 500 samples were taken for this figure. For both sub-figures, a first-order polynomial is fitted to the data and shown with the red line. The threshold for the PBSDR is 10% of the prescribed dose.

3.2.3. Proton beam current uncertainties

For the variation in beam current, sampling of the fluctuation is done with a 5% relative standard deviation based on HPTC proton beam data. Recall from one-dimensional results described in 3.1.4 that fluctuations in beam current only influence the dose rate. The variation in beam current is a day-to-day phenomenon, which makes the beam current equal for each delivered spot. Therefore, the relative error over the field is constant due to the current fluctuations (see Fig. 3.7). This also holds for the two-dimensional setup. In Fig. 3.16 this is depicted twice. In sub-figure (a) the relative difference between the 95th percentile max and 5th percentile min sampled dose rate and the error-free dose rate is shown. Note that for the whole field, the deviation is around 8%, similar to the deviation in the one-dimensional setup. In sub-figure (b) the PBSDR-VH is given for three profiles: the sampled 95th percentile max dose rate, the error-free dose rate and the sampled 5th percentile min dose rate. At every relative volume of V95, the difference between the max/min and the error-free case is around 8%. The effect of fluctuations in beam current on the dose rate is independent of both the absolute PBSDR and pencil beam weight. In general, the PBSDR is sensitive to current changes, however, it does not depend on field-specific factors. This is the main difference between the spot position and the beam current uncertainty added.

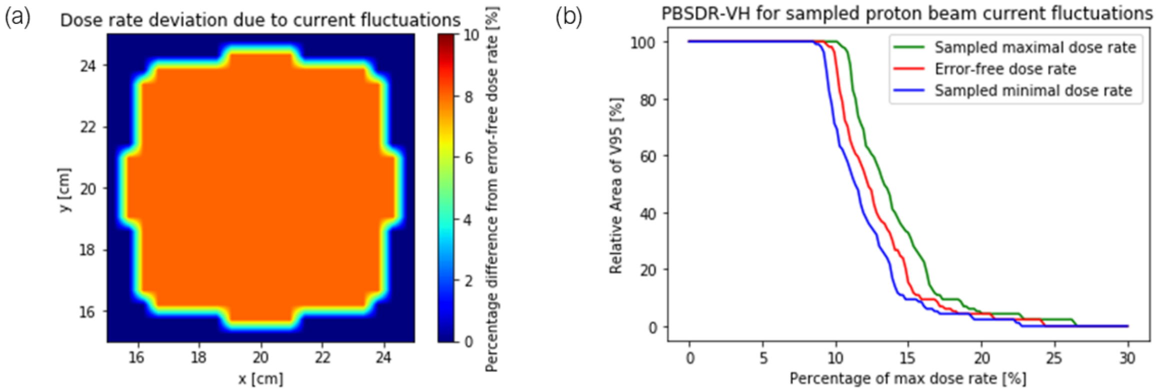


Figure 3.16: (a): The difference between the 95th percentile maximal and 5th percentile minimal sampled dose rate and the error-free dose rate for randomly drawn beam current fluctuations. The drawn fluctuations are drawn from a Gaussian with a 5% relative standard deviation. The PBSDR values are bi-linearly interpolated. (b): The PBSDR-VH with the percentage of the theoretical max dose rate on the x-axis and the relative area of V95 on the y-axis. For both sub-figures, the threshold for the PBSDR is 10% of the prescribed dose. 500 current fluctuation samples were taken for this figure. All PBSDRs are calculated according to the snake scanning pattern.

3.2.4. Combination of spot position and proton beam current uncertainties

To quantify the impact of uncertainties on the PBSDR metric, they have to be added to the model simultaneously. With sampling, the behaviour of the combination of the uncertainties is analysed. In Fig. 3.17 the relative difference between the sampled maximal and minimal dose rates and the error-free dose rate is shown. The relative deviations show aspects of both individual relative dose rate difference profiles (See Fig. 3.16(a) and Fig. 3.14(b)). Inside the field, the relative error is constant and around 10%. This is approximately the summation of both individual error types ($\sim$8% due to beam current fluctuations, $\sim$2% due to spot position errors). Surrounding the field, the total relative error resembles the deviation induced by the spot position uncertainties. Especially around the spots in the corners, which have a high individual pencil beam weight, relatively high total differences are found. This again gives evidence of the non-linearity of the deviation concerning dose rate and pencil beam weight.

This correlation between relative error and the PBSDR and pencil beam weight is confirmed by the two plots in Fig. 3.18. For both sub-figures, the correlation for the three separate scenarios of uncertainty is given. The red and green dotted lines respectively correspond to the sampled relative errors induced by only spot position and only beam current fluctuations. As mentioned before in this chapter, the error due to spot position changes is non-constant whereas the error due to beam current deviations is constant for the whole field. The data points included in blue are from the combined error scenario. Similar to the scenario of only spot position uncertainties, the relative error is increasing with both PBSDR and pencil beam weight. However, the slope of the fitted line for the data is lower than the slope of only spot position uncertainties. This implies that when adding the two types of uncertainties together in the PBSDR metric, the resulting error is smoothed. Thus, the two types of errors are not independent. Also, note that relative error in dose rate induced by current uncertainties serves as a plateau. For almost the whole field, the relative error is close to the constant $\sim$ 8% beam current error. Only for high dose rates (>25% of maximal theoretical dose rate) and high pencil beam weights (>0.30), the error added by the spot position uncertainty becomes dominant over the beam current uncertainty. Encountering such high values of pencil beam weight is rare. Again, the PBSDR shows high sensitivity for dose rates and pencil beam weights.

In conclusion, beam current fluctuations in the order of 5% standard deviation have a bigger impact on the PBSDR than spot position changes of 0.5-millimetre standard deviation.

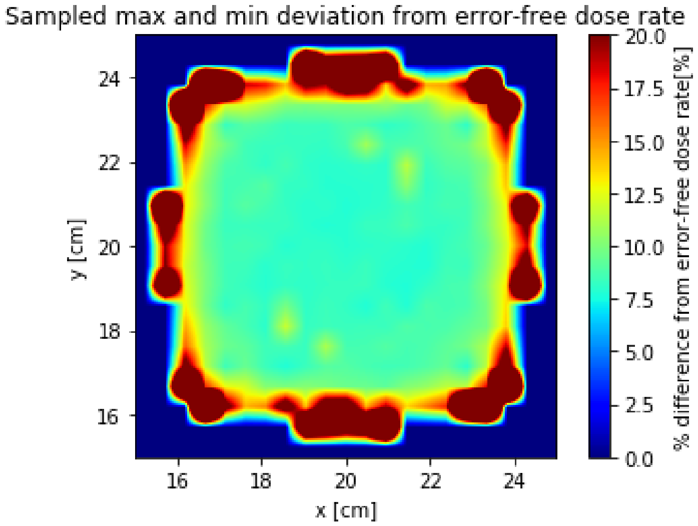


Figure 3.17: The absolute relative difference for the two types of uncertainties between the sampled max/min and error-free dose rate. The random spot position deviations are drawn from a Gaussian distribution with a 0.5 mm standard deviation and the random current fluctuations with a 5% relative standard deviation. The PBSDRs are calculated according to the snake scanning pattern. Pencil beams are placed between x, y = 17.0 cm, 23.0 cm. The σ_{pb} equals 3.2 mm and the σ_h equals 4.541 cm. The relative dose rate difference values are shown bi-linearly interpolated. The threshold for the PBSDR is 10% of the prescribed dose. The max and min are the 95th percentile and 5th percentile dose and dose rate values of 500 samples

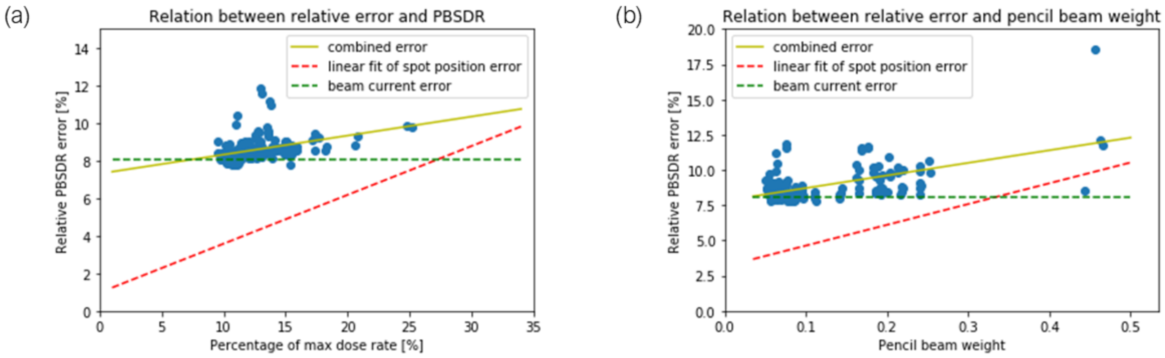


Figure 3.18: The correlation between the PBSDR on the x-axis and the relative max and min sampled error on the y-axis (a) and the correlation between the pencil beam weights on the x-axis and the relative error on the y-axis (b). The correlation is plotted for the three scenarios: only spot position errors, only beam current fluctuations and the combination of spot and bundle uncertainties. The 95th percentile maximal and 5th percentile minimal values correspond to the relative error. 500 samples were taken for this figure. In both sub-figures, a first-order polynomial is fitted to the data points and shown with the yellow line. The threshold for the PBSDR is 10% of the prescribed dose.

4

Biological FLASH metric

For the clinical implementation of FLASH proton radiotherapy, a robust metric related to the FLASH effect is needed. With a robust FLASH-related metric, the FLASH effect can be quantified consistently. The level of robustness can also help shape FLASH patient quality assessment. Additionally, treatment optimisation based on such a metric will guarantee a maximal FLASH effect for the patient. From the uncertainty analysis in chapter 3 it is clear that the currently used metrics to compute a dose rate for FLASH proton therapy display shortcomings. The DADR is not suitable for transmission FLASH proton therapy as it is independent of the scanning pattern. The FLASH effect has been linked to oxygen depletion in healthy cells, which implies that the scanning pattern should correlate with the FLASH effect. [2]

The other dose-rate metric, the PBSDR, developed by Folkert et al. does take into account the pencil-beam scanning pattern. [17] However, the PBSDR is not perfect as well. First of all, it is sensitive to changes in spot position and beam current (Fig.3.18). Secondly, crucial characteristics of the PBSDR are not in line with the current understanding of the FLASH effect. The fluctuations in proton beam current show a constant effect on the PBSDR (see Fig.3.16(a)) and the direction of irradiation of a pencil beam scanning pattern does not influence the resulting dose rate (note the x-axis symmetry in Fig.3.11(b)). The FLASH effect should depend on the scanning direction and should not be linearly affected by the beam current. Thirdly, the PBSDR is not differentiable in extreme cases of short irradiation times (observe the dose-rate hot spots at the corners in Fig.3.11(b)). In other words, the PBSDR approaches infinity for irradiation times going to zero. Lastly, the PBSDR depends on the predefined dose threshold (see Fig.3.12). It is unclear which dose threshold can be used for quantifying the dose rate in line with the FLASH effect.

The results of the biologically inspired FLASH metric will be given in this chapter. The mathematical functions and the structure of the model are described in 2.2. The mentioned shortcomings of the DADR and PBSDR are tackled with this new metric. The output is defined as the effective FLASH enhancement factor (FER_{eff}) and is coupled to the FLASH effect. This is more intuitive than a dose-rate definition for the FLASH effect.

The model is applied to the same field as described in 3.2. A six-by-six centimetre scanning pattern with 0.5 cm pencil-beam spacing is used to deliver an as homogeneous as possible dose of 8 Gy. The pencil-beam weights are optimised with the same cost function. The structure shown in Fig. 2.3 will be followed in 4.1. The sensitivity of the FER_{eff} for machine intrinsic uncertainties is described in 4.2.

4.1. Model structure

The time traces of the oxygen concentration, OER, FER and FED are given for the voxel in the middle of the six-by-six centimetre scanning pattern. This voxel was selected because, in the middle of the field, multiple pencil beams contribute to the dose. In this case, the effect of dose deposition of multiple pencil beams on the time traces is visible.

4.1.1. Oxygen concentration

Recall that the oxygen concentration over time is calculated with the following ordinary differential equation.

$$\frac{dC(t)}{dt} = -\kappa \dot{D}(t) + \frac{1}{\tau}\Big(C_{ss} - C(t)\Big) \tag{4.1}$$

With κ the radiolytic oxygen depletion is modelled induced by the dose rate. Re-oxygenation to the steady state is exponentially modelled with weight τ. The combination of the two model parameters describes the local oxygen concentration coupled with the instantaneous dose rate. The numerical solution is found with the backward Euler method. The delivery of every pencil beam is divided into 100 time steps. The total irradiation time is in the order of milliseconds for the 8 Gy median prescribed dose. The model parameters κ and τ are fitted according to work from M. Diepeveen (his MSc thesis [14] and an article under revision). To solve the oxygen concentration over time, the time traces for the cumulative dose and related instantaneous dose rate are required. The dose and dose-rate profiles for a voxel in the middle of the scanning pattern are shown in Fig. 4.1. The x-axis is given in the percentage of pattern completion. This is proportional to the irradiation time. The cumulative dose in sub-figure (a) shows that the prescribed dose of 8 Gy is given in this voxel. From the dose-rate profile in sub-figure (b), it is clear that many pencil beams contribute to the dose as the dose rate is mostly non-zero. This wide contribution yields from the implemented Gaussian halo with a standard deviation of 4.541 cm.

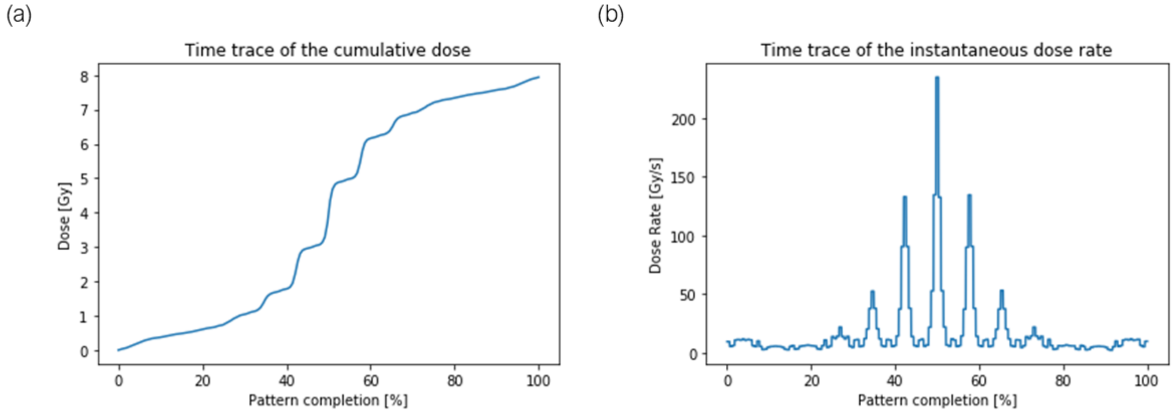


Figure 4.1: Time traces of the cumulative dose (a) and instantaneous dose rate (b) for the voxel in the middle of the scanning pattern of six-by-six centimetres. The prescribed median dose is 8 Gy. The instantaneous dose rate per position is scaled to the size of the time steps and calculated with the forward difference equation. The time is given in percentage of pattern completion.

The dose-rate profile is used to solve Eq. 4.1. The oxygen concentration over time for the middle voxel in the field is shown in Fig. 4.2 with the blue line. As the voxel is located in the middle of the field, most oxygen is depleted around 50% total pattern completion. Re-oxygenation occurs at most other points in time. With the selected system parameters the oxygen concentration never fully depletes. Both the depletion term of Eq. 4.1 and the re-oxygenation term are visible for this voxel with the selected system parameters.

4.1.2. OER

From the oxygen concentration, the OER can be calculated according to Eq. 2.6. As mentioned before in 2.2, the OER is based on tumour tissue. A low OER relates to hypoxic conditions and yields more tumour radioresistance. The OER for the voxel in the middle of the scanning pattern is shown in Fig. 4.2 with the red line. At 50% pattern completion, the OER goes down because of oxygen depletion. The similarity with the oxygen concentration time trace is noticeable.

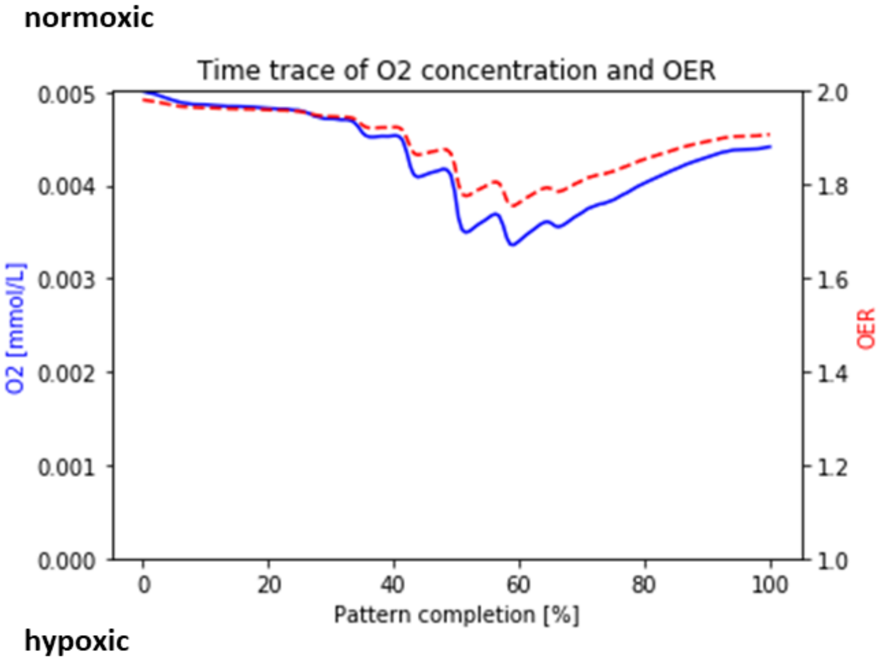


Figure 4.2: Time trace of the oxygen concentration and OER for the voxel in the middle of the scanning pattern of six-by-six centimetres. The left y-axis corresponds to the oxygen concentration and the right y-axis to the OER. The equation for the oxygen concentration is solved numerically with the backward Euler method. The prescribed median dose is 8 Gy, κ equals 0.0005 $mol\ m^{-3}\ Gy^{-1}$ and τ 0.001 s. The steady-state concentration is the normoxic concentration. Hypoxic conditions relate to depleted oxygen concentrations. For the OER, normoxic conditions equal high OERs and hypoxic conditions relate to low OERs. The OER is bound between 1 and 2 due to the fitting of the steady-state oxygen concentration. The time is given in percentage of pattern completion.

4.1.3. FER

The FER per voxel can be calculated by taking the bound inverse of the OER ($FER = \frac{2}{OER}$). The FER has been developed for healthy tissue instead of tumour tissue. A high FER is seen under hypoxic conditions and causes more healthy tissue sparing because of less radiolytic-induced damage. In Fig. 4.3 the time trace of the FER for the middle voxel of the scanning pattern is depicted. Again, the increase in FER is located around the time of 50% pattern completion.

4.1.4. FED

The FED is defined as the cumulative dose divided by the FER per voxel. In essence, the FED is a scaled version of the cumulative dose. In the target, the resulting FED will be lower than the planned dose deposition. The decrease in dose delivery is caused by oxygen depletion in cells. At lower oxygen concentrations, healthy tissue and tumour tissue are less sensitive to dose. The FED can be seen as a biologically delivered dose instead of a physically planned dose. The time trace of the cumulative FED together with the planned dose for the voxel in the middle of the scanning pattern is shown in Fig. 4.4. The physically prescribed median dose of 8 Gy is not reached. At 50% pattern completion, the difference between the biological FED and physical dose increases the most. Here the rate of oxygen depletion is the highest (Fig. 4.2).

4.1.5. FER$_{eff}$

With the cumulative FED and planned dose, the FER$_{eff}$ can be computed. It is defined as the ratio between the planned and biological dose delivery at 100% scanning pattern completion. For example, the FER$_{eff}$ for the voxel in the middle of the scanning pattern is 1.08 (FER$_{eff}$ = $\frac{7.94}{7.37}$). This calculation is repeated for every voxel and yields the FER$_{eff}$ distribution for the whole field. For 8 Gy prescribed median dose, the distribution is shown in Fig. 4.5. The FER$_{eff}$ distribution is homogeneous. Only for the outermost pencil beams of the pattern, the FER$_{eff}$ value is higher than in the rest of the field. The

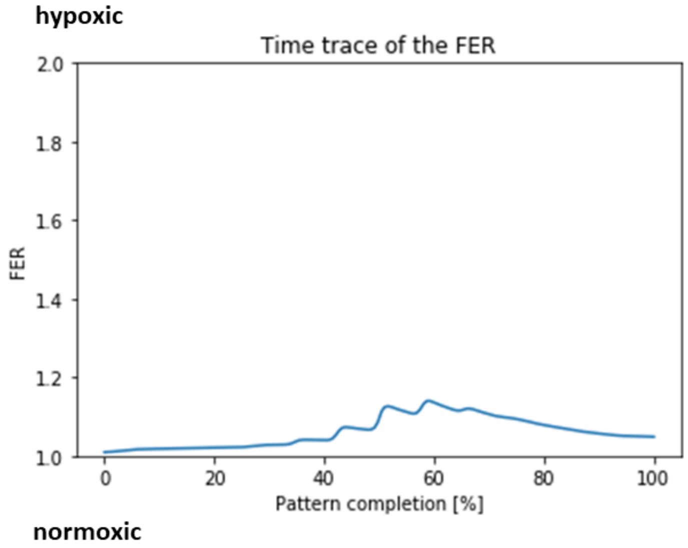


Figure 4.3: Time trace of the FLASH Enhancement Ratio (FER) for the voxel in the middle of the scanning pattern of six-by-six centimetres. Normoxic conditions equal low FERs and hypoxic conditions relate to high FERs. The time is given in percentage of pattern completion. The FER is bound between 1 and 2 due to the fitting of the steady-state oxygen concentration.

FLASH effect is limited for this 8 Gy median prescribed case as the FER$_{eff}$ is similar for tumour and healthy tissue.

With the selected model parameters, the FLASH effect is limited at the conservative threshold of 8 Gy. The FER$_{eff}$ distribution for a higher dose of 54 Gy together with various FER$_{eff}$ volume histograms are shown in Fig. 4.6. Note that the volume histograms are for relative volumes of V50, the volume receiving 50% of the prescribed dose. The V50 is the field size area instead of the target area. From sub-figure (a) it is clear that the FER$_{eff}$ values at the edges of the pattern are higher than in the middle. Now, the FLASH effect is guaranteed. The increase of the FER$_{eff}$ at the edges of the scanning pattern is also visible in its respective volume histogram shown in sub-figure (b). The effective FER volume curves are not a straight line down anymore. The span width of the volume curves of the effective FER correlates with the FLASH effect. Based on the various FER$_{eff}$ volume histograms depicted, the prescribed dose threshold lies around 16 Gy right now for the FLASH effect. In theory, tweaking the model parameters is possible such that the FLASH effect is noticeable for the set FLASH effect threshold of 8 Gy as well.

4.2. Sensitivity analysis

The same uncertainty analysis as in chapter 3 can be done to quantify the robustness of the effective FER compared to the PBSDR. The same Gaussian machine uncertainties of 0.5 mm standard deviation in spot position placement and 5% standard deviation in proton beam current were added into the biological FLASH model. With these added uncertainties, the relative error per voxel in the target could again be quantified. The relative error is defined the same as before. It is the relative difference between the error-free effective FER and its 5th and 95th percentile value. In Fig. 4.7 the relative error of the FER$_{eff}$ and PBSDR are depicted together. In sub-figure (a) the sensitivity for spot position changes is shown and in sub-figure (b) the sensitivity for beam current fluctuations is shown. For the FER$_{eff}$ both added types of uncertainties have a limited effect of $\approx 0.5\%$ relative error inside the target area (V95). Compared to the PBSDR, the FER$_{eff}$ is less sensitive to the same machine uncertainties. The sensitivity of the FER$_{eff}$ also does not increase with pencil beam weight. From both sub-figures, it is clear that the FER$_{eff}$ is way more robust than the PBSDR.

The fact that the relative error is the same for the whole target is not surprising. Recall from Fig. 4.5 that for the 8 Gy prescribed median dose, the FER$_{eff}$ distribution for this six-by-six centimetre scanning

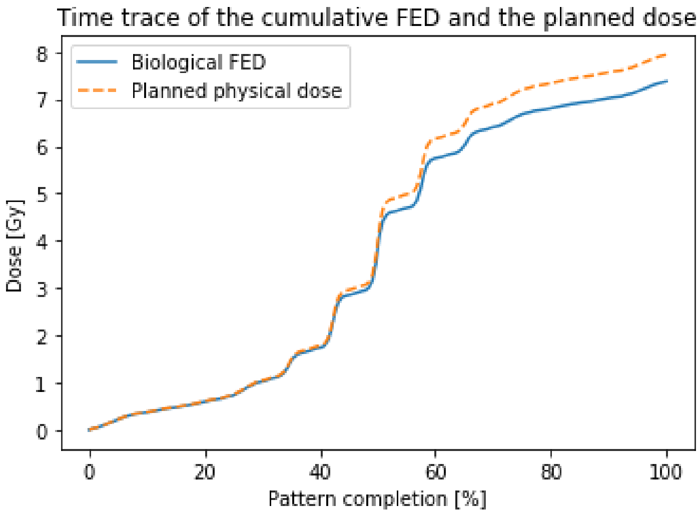


Figure 4.4: Time trace of the FLASH Enhanced Dose (FED) and the planned physical dose for the voxel in the middle of the scanning pattern of six-by-six centimetres. The physical dose is optimised such that 8 Gy median prescribed dose is achieved. The biologically delivered dose is lower than 8 Gy because of the oxygen depletion added into the model.

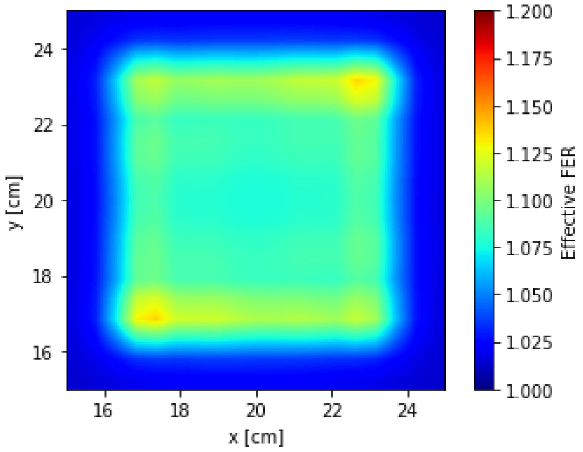


Figure 4.5: The effective FLASH Enhancement Ratio (FER$_{eff}$) distribution for 8 Gy median prescribed dose. A pencil-beam scanning pattern of six-by-six centimetres is used between 17 cm and 23 cm. The exact scanning pattern is shown in Fig. 3.11(a)

pattern is almost completely homogeneous. This implies that the underlying time traces of the FER$_{eff}$ follow the same depletion and re-oxygenation behaviour. Sampling small uncertainties will therefore always generate an average constant error over the whole target. To study the sensitivity of the FER$_{eff}$ further, it is helpful to implement bigger uncertainties in inhomogeneous FER$_{eff}$ distribution cases.

Fixed beam current fluctuations of $\pm$ 10% and $\pm$ 20% were added to the 54 Gy prescribed median dose case of Fig. 4.6(a). The effect of these large beam current uncertainties are depicted in the FER$_{eff}$ distribution plots of Fig. 4.8. Sub-figures (a) and (b) depict the -10% and -20% beam current. This implies that irradiation of the plan takes longer than for the error-free scenario. The other two sub-figures, (c) and (d), correspond to the +10% and +20% proton beam current fluctuations. From the sub-figures, it is clear that the relative error is not constant for the whole field anymore. At the edges of the scanning pattern, the induced relative error is smaller than in the middle of the pattern. This is caused by the fact that at the edges only a few pencil beams are affected by the beam current fluctuation. The depletion and re-oxygenation are not hindered by the change. In the middle of the field, where many pencil beams are involved, the FER$_{eff}$ is affected.

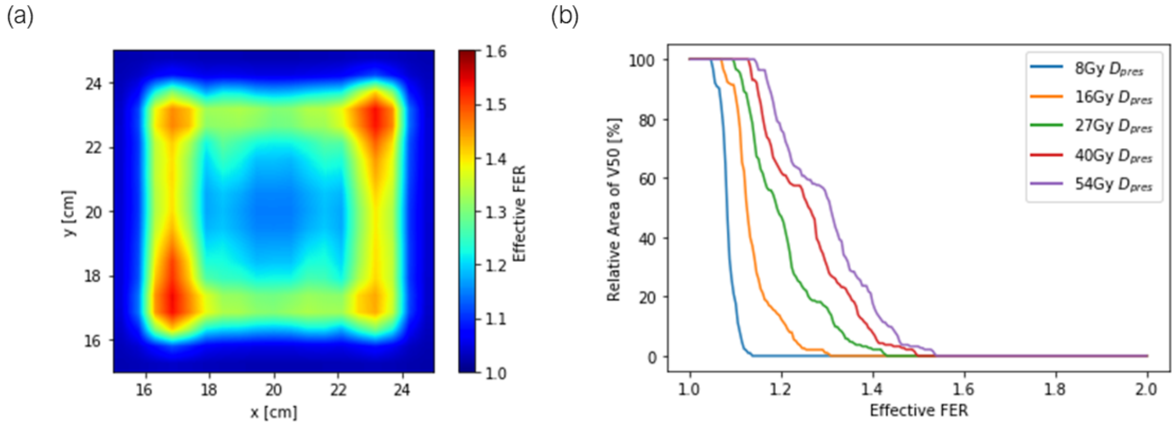


Figure 4.6: The effective FLASH Enhancement Ratio (FER$_{eff}$) distribution for 54 Gy median prescribed dose (a) and FER$_{eff}$ volume histograms for varying prescribed dose values (b). A snake pencil-beam scanning pattern of six-by-six centimetres is used between 17 cm and 23 cm. The exact scanning pattern is shown in Fig. 3.11(a)

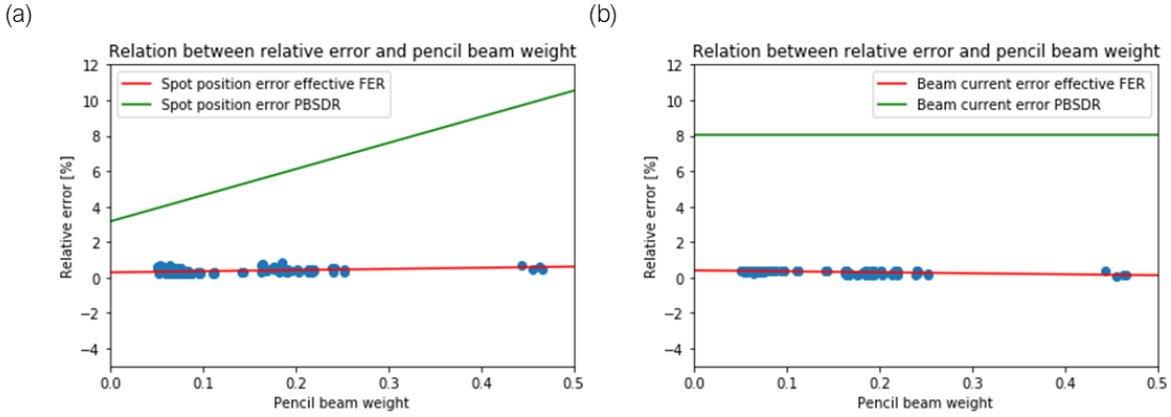


Figure 4.7: The correlation between the pencil beam weights on the x-axis and the relative error on the y-axis of spot position uncertainties of the FER$_{eff}$ and the PBSDR (a) and the correlation of pencil beam weights and the relative error for beam current fluctuations (b). The linear fitted PBSDR relative errors are shown with the green lines. The linear fits to the error of FER$_{eff}$ are red. The difference between the error-free scenario and the 95th percentile maximal and 5th percentile minimal values correspond to the relative error. 500 samples were taken for this figure. In both sub-figures, a first-order polynomial is fitted to the blue data points. The threshold for the PBSDR is 10% of the prescribed dose. The FER$_{eff}$ model parameters are κ = 0.0005 $mol\ m^{-3}\ Gy^{-1}$ and τ = 0.001 s. The prescribed median dose is 8 Gy.

From Fig. 4.8 it is also made clear that an increase in beam current has a bigger impact on the FER$_{eff}$ than an equal decrease. Additionally, the relative error is proportional to the size of the beam current fluctuation. The volume histograms of the relative errors are shown in Fig. 4.9. With -20% beam current, the relative error is approximately 2.5% for the target whereas the relative error is 3% for the +20% beam current deviation. This indicates that the oxygen depletion term related to dose rate is dominant over the re-oxygenation term with the current model parameters. The relative error for the $\pm$ 10% fluctuation is lower than the $\pm$ 20% cases.

With this biologically inspired FLASH metric, the shortcomings of the PBSDR are successfully dealt with. Firstly, the sensitivity for pencil beam placement and beam current errors is way lower. Secondly, fluctuations in proton beam current do not affect the FER$_{eff}$ with a constant factor as two separate time scales are into play, the depletion and re-oxygenation, instead of a single time scale for the PBSDR. Also because of this reason, the direction of irradiation does matter for the FER$_{eff}$ in contrast to the PBSDR. Thirdly, the set-up ordinary differential equation is solved in such a way that it is always differentiable. Lastly, the user-defined dose threshold for the PBSDR is omitted. The model has two important parameters, κ and τ. These can be fitted to match the found experimental values of the FLASH enhancement ratio in healthy tissue.

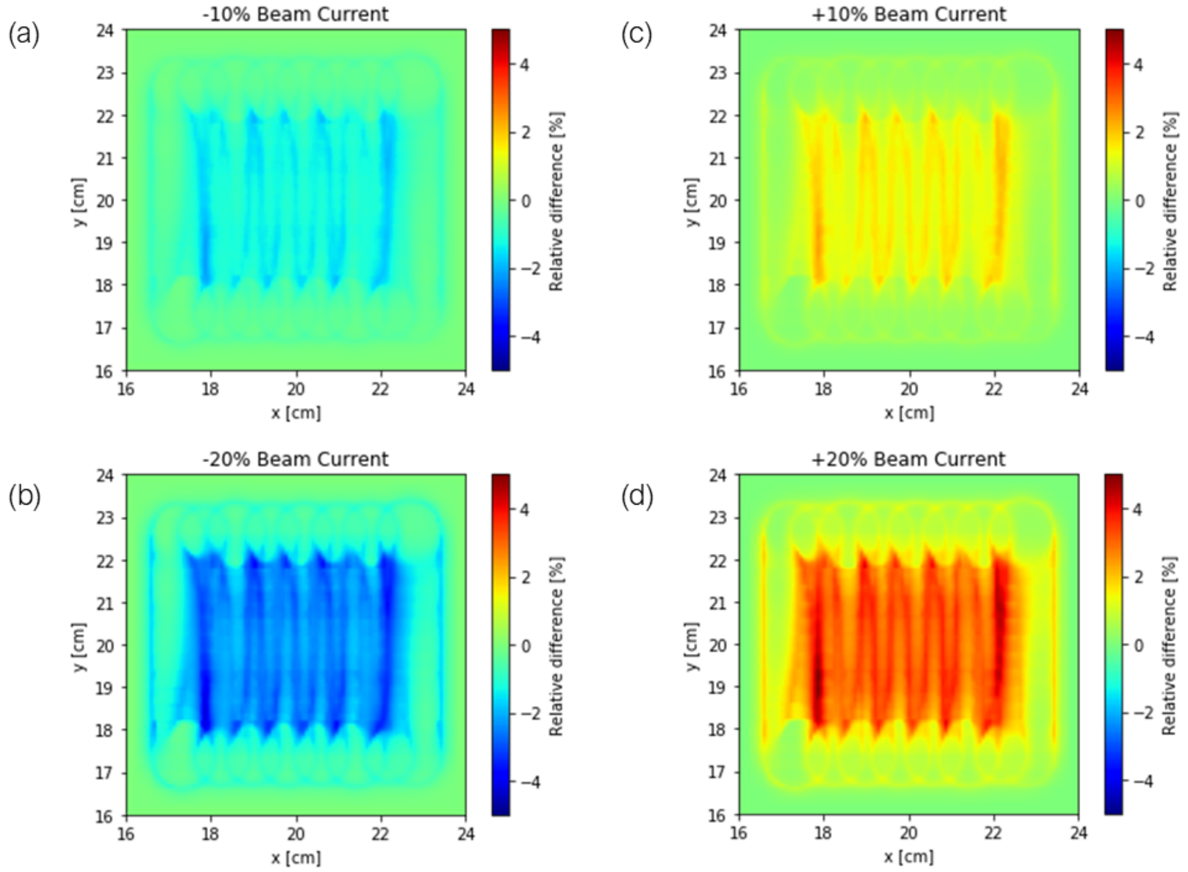


Figure 4.8: FER$_{eff}$ relative error distribution plots for various fixed beam current fluctuations. The relative error is defined as the percentage difference with the error-free FER$_{eff}$ distribution. The scanning pattern is six-by-six centimetres and is placed from 17 cm to 23 cm with 0.5 cm spacing. A 20% increase in beam current implies that the dose is delivered 1.2 times quicker than in the error-free scenario.

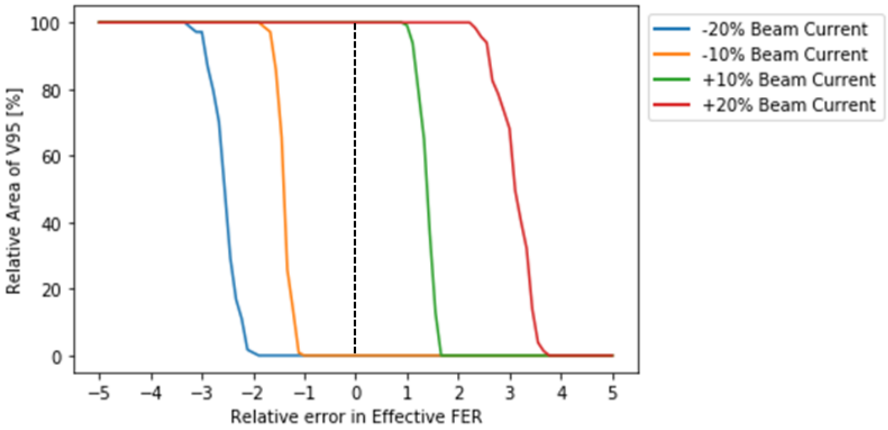


Figure 4.9: Relative error of FER$_{eff}$ volume histograms for the ± 10% and ± 20% beam current fluctuations. The striped black line represents the error-free scenario.

5

2D clinical model

In chapter 3 the 2D analytical dose model has been introduced. The convolutional and sampling approach added pencil beam position and beam current uncertainties to the model. These uncertainties had a fixed Gaussian standard deviation of respectively 0.5 mm and 5%. Moreover, the uncertainty analysis was solely done for a six-by-six centimetre scanning grid with a fixed 0.5 cm pencil beam spacing. In the clinic, various field sizes and pencil beam spacings are used.

The clinical relevance of the 2D analytical model is boosted by sampling uncertainties of different sizes for multiple field sizes and pencil beam spacings. Pencil beam position random placement errors were generated as 0 mm up to 2 mm standard deviation. The beam current fluctuations range from 0% to 20% standard deviation. These uncertainties were added to 9 different field sizes with 3 pencil beam spacings (so 27 fields in total). The field sizes are selected in correspondence with the FAST-01 clinical trial. [12]. In the FAST-01 trial, 7 field sizes are reported (7.5 cm × 7.5 cm, 7.5 cm × 10 cm, 7.5 cm × 12 cm, 7.5 cm × 14 cm, 7.5 cm × 16 cm, 7.5 cm × 18 cm and 7.5 cm × 20 cm). As transmission proton FLASH is mainly suitable for small tumours, fields of approximately 3 cm x 3 cm and 5 cm x 5 cm were added to the analysis as well. The 3 selected pencil beam spacings are 0.3 cm, 0.4 cm and 0.5 cm. The pencil beam scanning patterns were optimised with the same cost function as before (Eq. 3.6). The pencil beams were chosen such that they come as close to the reported field sizes as possible with a clinically viable scanning pattern. This implies that the scanning pattern is symmetrical around the pencil beam in the isocentre. The field size is determined as the area receiving at least 50% of the prescribed dose. All field sizes for the various pencil beam spacings with the underlying number of pencil beams are given in table 5.1. In general, the generated field sizes approximate the field size goals pretty well. For the larger FAST-01 field sizes and 0.3 cm pencil beam spacing, the total number of pencil beams increases beyond 1000. With so many pencil beams, adhering to the irradiation time threshold for the FLASH effect of 200 milliseconds is challenging.

The 27 different fields are sampled with increasing uncertainties in pencil beam placement and proton beam current. As mentioned before, pencil beam placement errors affect the dose and dose rate, whereas beam current fluctuations only impact the dose rate. Here, the impact of pencil beam placement errors concerning the added uncertainty is only evaluated on the dose deposition. The effect of beam current uncertainties is evaluated with the PBSDR. Every individual added error is sampled 10,000 times for each field. This implies that for both the impact on dose deposition as well as the PBSDR, 2.7 million calculations are required (27 fields, 10 errors and 10,000 samples). Gaussian quadrature sampling is used for high-accuracy calculations while limiting the computation time. To translate the impact of the added uncertainties to the clinic, they were linked to clinically used metrics. Dose deposition changes were evaluated on the V95, V107, D2 and D98. The impact on dose rate was evaluated on the PBSDR-98.

The same 2D analytical model as in chapter 3 with the snake pencil beam scanning pattern is used for sampling. However, it is used in its more basic form, without the second Gaussian for the halo effect. Moreover, there is one fundamental difference between the two sampling procedures. In chapter 3 the

Table 5.1: Overview of the 27 generated field sizes together with the number of pencil beams (PBs). The field size is determined as the area receiving at least 50% of the prescribed dose. The pattern is selected such that the resulting field size is resembling the FAST-01 fields and the 3 cm and 5 cm square fields.

	Pencil Beam Spacing		
	0.3 cm	0.4 cm	0.5 cm
	Created Field Size		
Field Size Goal			
3.0 cm x 3.0 cm	3.4 cm x 3.4 cm 81 PBs (9 by 9)	3.4 cm x 3.4 cm 49 PBs (7 by 7)	3.0 cm x 3.0 cm 25 PBs (5 by 5)
5.0 cm x 5.0 cm	5.0 cm x 5.0 cm 225 PBs (15 by 15)	5.0 cm x 5.0 cm 121 PBs (11 by 11)	5.0 cm x 5.0 cm 81 PBs (9 by 9)
7.5 cm x 7.5 cm	7.4 cm x 7.4 cm 529 PBs (23 by 23)	8.2 cm x 8.2 cm 361 PBs (19 by 19)	8.0 cm x 8.0 cm 225 PBs (15 by 15)
7.5 cm x 10.0 cm	7.4 cm x 9.8 cm 713 PBs (23 by 31)	8.2 cm x 10.6 cm 475 PBs (19 by 25)	8.0 cm x 10.0 cm 285 PBs (15 by 19)
7.5 cm x 12.0 cm	7.4 cm x 12.4 cm 897 PBs (23 by 39)	8.2 cm x 12.2 cm 551 PBs (19 by 29)	8.0 cm x 12.0 cm 345 PBs (15 by 23)
7.5 cm x 14.0 cm	7.4 cm x 14.0 cm 1035 PBs (23 by 45)	8.2 cm x 14.6 cm 665 PBs (19 by 35)	8.0 cm x 14.0 cm 405 PBs (15 by 27)
7.5 cm x 16.0 cm	7.4 cm x 15.8 cm 1173 PBs (23 by 51)	8.2 cm x 16.2 cm 741 PBs (19 by 39)	8.0 cm x 16.0 cm 465 PBs (15 by 31)
7.5 cm x 18.0 cm	7.4 cm x 18.4 cm 1357 PBs (23 by 59)	8.2 cm x 18.6 cm 855 PBs (19 by 45)	8.0 cm x 18.0 cm 525 PBs (15 by 35)
7.5 cm x 20.0 cm	7.4 cm x 20.2 cm 1495 PBs (23 by 65)	8.2 cm x 20.2 cm 931 PBs (19 by 49)	8.0 cm x 20.0 cm 585 PBs (15 by 39)

impact of the uncertainties was evaluated as the minimal and maximal voxelwise difference. Here, it is quantified as the difference occurring per treatment. This implies that real fractional doses are sampled which is more clinically relevant.

In this chapter, the Gaussian quadrature implementation is verified in section 5.1. After that in section 5.2, the impact of pencil beam displacement and beam current fluctuation is evaluated for one specific field size and pencil beam spacing. In the next section (5.3), the clinically relevant analysis is extended to multiple field sizes. The evaluation is extended to multiple pencil beam spacings in the last section (5.4).

5.1. Gaussian quadrature verification

The performance of Gaussian quadrature sampling was evaluated according to the clinical metrics and constraints. The impact of pencil beam positioning errors on dose deposition was analysed for a high-resolution grid (HR), a low-resolution grid (LR no Quad) and a low-resolution grid with Gaussian quadrature (LR Quad). High resolution corresponds with 0.02 cm and low resolution with 0.1 cm voxel width. Note that low resolution still matches the resolution of a typical CT scan. For the analysis, 10,000 samples per Gaussian standard deviation error were taken. The scanning pattern used for the analysis is equal to the one described in chapter 3. It is six-by-six centimetres with 0.5-centimetre pencil beam spacing. In Fig. 5.1 the median results of the samples in the target are shown. Sub-figure (a) shows the full analysis for the V95, V107, D98 and D2 with increasing Gaussian errors. The clinically used constraints are also visualised with the dotted lines. These correspond to 95% of V95 for the V95, 1.5 cm^2 of V95 for the V107, $0.90D_{pres}$ for the D98 and $1.12D_{pres}$ for the D2. This is in line with the ICRU recommendations and current clinical practice. [26, 27] In sub-figure (b) a close-up is shown for a Gaussian error between 0.350 and 0.475 for the V95. The V95 is selected as the clinical constraint is violated at the smallest Gaussian error size. From the figure, it is clear that the difference between the low-resolution results with and without Gaussian quadrature is negligible. The difference between the high-resolution and low-resolution results is visible. The intersection with the clinical constraint is 0.40 mm for the low-resolution grid and 0.44 mm for the high-resolution grid.

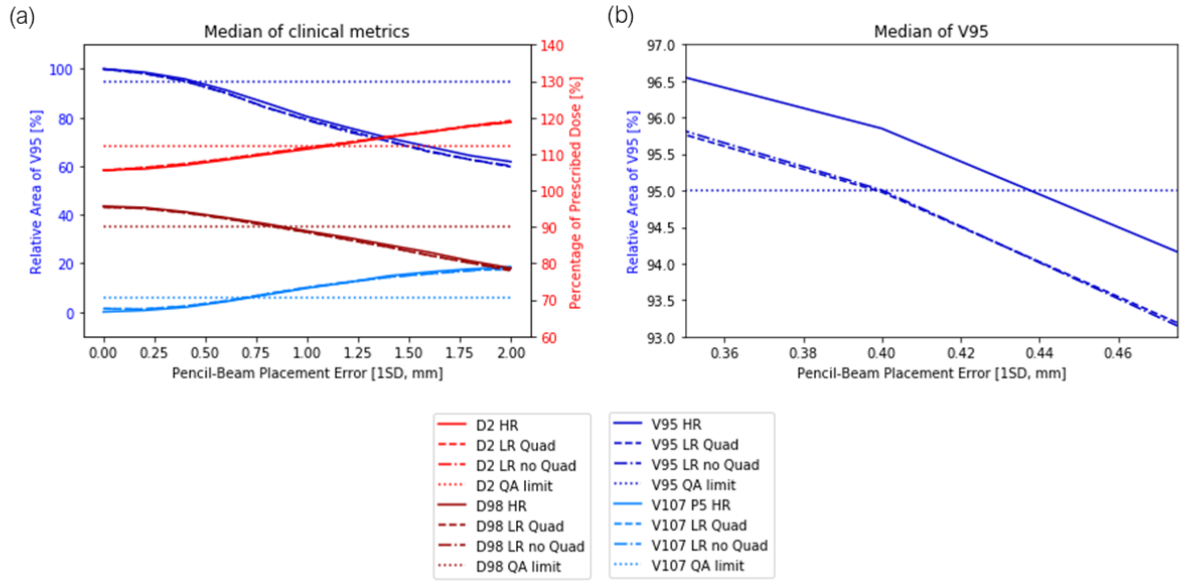


Figure 5.1: The correlation between Gaussian pencil beam displacement standard deviation on the x-axis and four clinical metrics (a) on both y-axes and a close-up for the V95 (b). 10,000 samples were taken per standard deviation and the median value for each metric is shown. The left y-axis corresponds to the V95 (dark blue) and V107 (light blue). The right y-axis is used for the D2 (light red) and D98 (dark red). The horizontal dotted lines indicate the clinically used constraints. These are 95% of V95 for the V95, 1.5 cm^2 of V95 for the V107, $0.90D_{pres}$ for the D98 and $1.12D_{pres}$ for the D2. High resolution (HR) are the solid lines, low resolution with quadrature sampling (LR Quad) are the striped lines and low resolution without quadrature (LR no Quad) are the other lines. High-resolution is 0.02 cm and low-resolution is 0.1 cm. The prescribed dose is 8 Gy.

The impact of Gaussian quadrature sampling is limited as the difference between the low-resolution with and without Gaussian quadrature is negligible. The lack of improvement can be linked to the pencil beam weight optimisation. The optimisation is affected by the resolution of the grid. On the high-resolution grid of 0.02 cm voxel width, the Gaussian pencil beams of 0.32 cm standard deviation are way better sampled in contrast to the 0.1 cm voxel width grid. This leads to differences in pencil beam weights with the least square optimisation. In Fig. 5.2 the relative difference between the high-resolution and low-resolution pencil beam weights is shown. Especially at the edges of the scanning pattern, where relatively few pencil beams are used for dose deposition, the relative difference in pencil beam weight is present. The true power of Gaussian quadrature will probably come to light when the underlying pencil beam weights are equal.

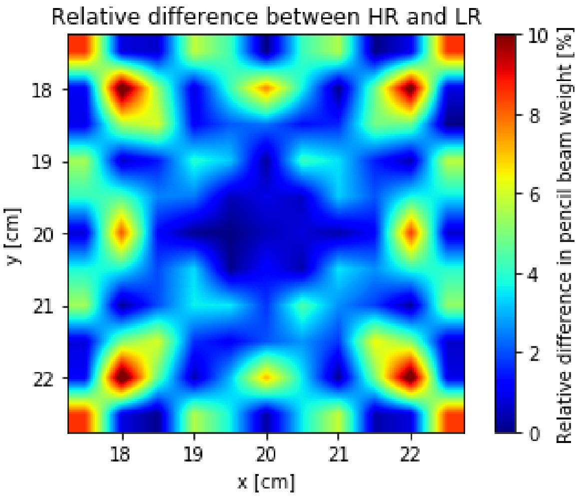


Figure 5.2: Relative difference in pencil beam weights for the low-resolution and high-resolution grid. The six-by-six centimetre scanning pattern is used. The relative differences are bilinearly interpolated for better visualisation.

Sampling of pencil beam placement and beam current errors with the high-resolution grid was not feasible. The time estimate for the sampling procedure for the largest field was more than a year. From Fig. 5.1 it is found that quadrature sampling is not affecting the results. Consequently, the results shown in the remainder of this chapter are accurate but are bound to the 0.1 cm resolution.

5.2. Single-field analysis

Here, the results for the 8.0 cm by 8.0 cm field with 0.5 cm pencil beam spacing are given. The relation between error size and the clinical evaluation metrics is given for the median, the 5th percentile and the 95th percentile value. The point of intersection thus represents the error size at which 95% of patients receive clinically acceptable treatment. The effect of pencil beam uncertainties on dose deposition for this field is shown in Fig. 5.3 The first clinical constraint that is violated is the V95. At 0.38 mm pencil beam Gaussian uncertainty, the tumour coverage is not good enough anymore for 5% of patients according to quality assessment standards. The other first intersections are at 0.39 mm standard deviation for the V107, 0.82 mm for the D2 and 0.74 mm for the D98.

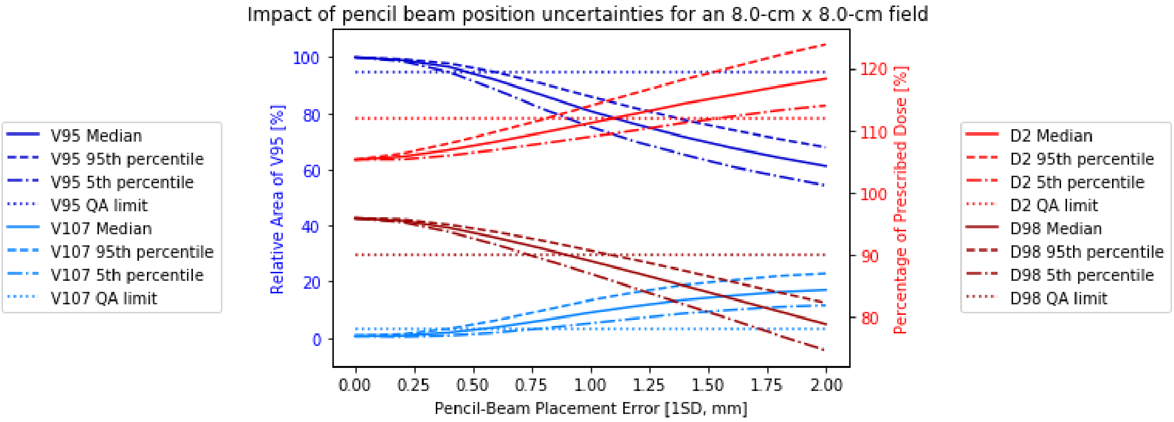


Figure 5.3: The correlation between Gaussian pencil beam displacement standard deviation on the x-axis and four clinical metrics on both y-axes for an 8.0 cm by 8.0 cm field with 0.5 cm pencil beam spacing. 10,000 samples were taken per standard deviation. The median (solid line), 5th percentile (dotted/solid line) and 95th percentile values (striped line) are given for each clinical evaluation metric. The left y-axis corresponds to the V95 (dark blue) and V107 (light blue). The right y-axis is used for the D2 (light red) and D98 (dark red). The horizontal dotted lines indicate the clinically used constraints. These are 95% of V95 for the V95, 1.5 cm^2 of V95 for the V107, $0.90D_{pres}$ for the D98 and $1.12D_{pres}$ for the D2. The prescribed dose is 8 Gy.

The effect of beam current fluctuations on the PBSDR is shown in Fig. 5.4. At 6.8% standard deviation beam current fluctuation the constraint is breached. With such fluctuations, the dose rate for 5% of patients decreases too much for this specific field. It can also be observed that the PBSDR is more robust for a beam current decrease than an increase. At 20% Gaussian standard deviation, the relative error is -25% for the 5th percentile against 52% for the 95th percentile.

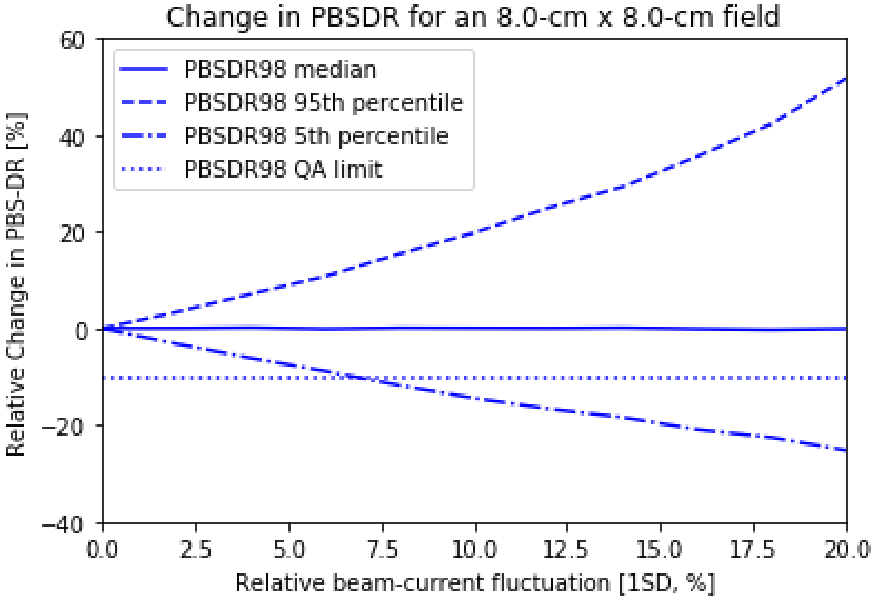


Figure 5.4: The correlation between Gaussian beam current fluctuation on the x-axis and the relative change in PBSDR on the y-axis for an 8.0 cm by 8.0 cm field with 0.5 cm pencil beam spacing. 10,000 samples were taken per standard deviation. The median (solid line), 5th percentile (dotted/solid line) and 95th percentile values (striped line) are given. The horizontal dotted line indicates the clinically relevant constraint. This is a -10% relative change for the PBSDR98. For the PBSDR a dose threshold of 10% prescribed dose is used.

From the single-field analysis, it is clear that pencil beam positioning errors have the biggest impact on the target coverage. Whether or not the 0.38 mm standard deviation for FLASH proton therapy is achievable has to be validated with proton beam measurements. The fluctuation in beam current of 6.8% standard deviation seems easier to adhere to.

5.3. Multi-field analysis

The median of the V95, V107, D2 and D98 for all fields with 0.5 cm pencil beam spacing is depicted in Fig. 5.5. All sub-figures are close-ups of the relevant area for the analysis. The V95 is shown in sub-figure (a). There is a correlation between the field size and the pencil-beam Gaussian error size for constraint violation visible. The constraint is breached for relatively small fields more easily. This indicates that with increasing field size, and thus an increasing number of pencil beams, the dose delivery is more robust. The same holds for the V107 in sub-figure (b), the D98 in sub-figure (c) and the D2 in sub-figure (d). For the V107 there is no constraint depicted. The reason behind this is that the V107 constraint varies for each field size. It is defined as 1.5 cm^2 of the field area. For the larger fields, this constraint is too strict for the generated fields optimised on the prescribed dose, D98 and D2. Smooth lines are fitted to the data with (cubic) spline interpolation. With this approach, the curvature of the fitted lines is minimised. The sample points are depicted with the crosses.

From Fig. 5.5 it is clear that target coverage is affected by pencil beam displacements the most. The median values of the V95 seem to approach the same random error size for constraint violation with increasing field sizes. For the generated FAST-01 fields, the intersecting values lie close together. More clinically relevant are the 5th percentile curves for the V95 shown in Fig. 5.6. At the intersection with the QA limit, 5% of patients do not receive treatment in accordance with the set ICRU metrics and

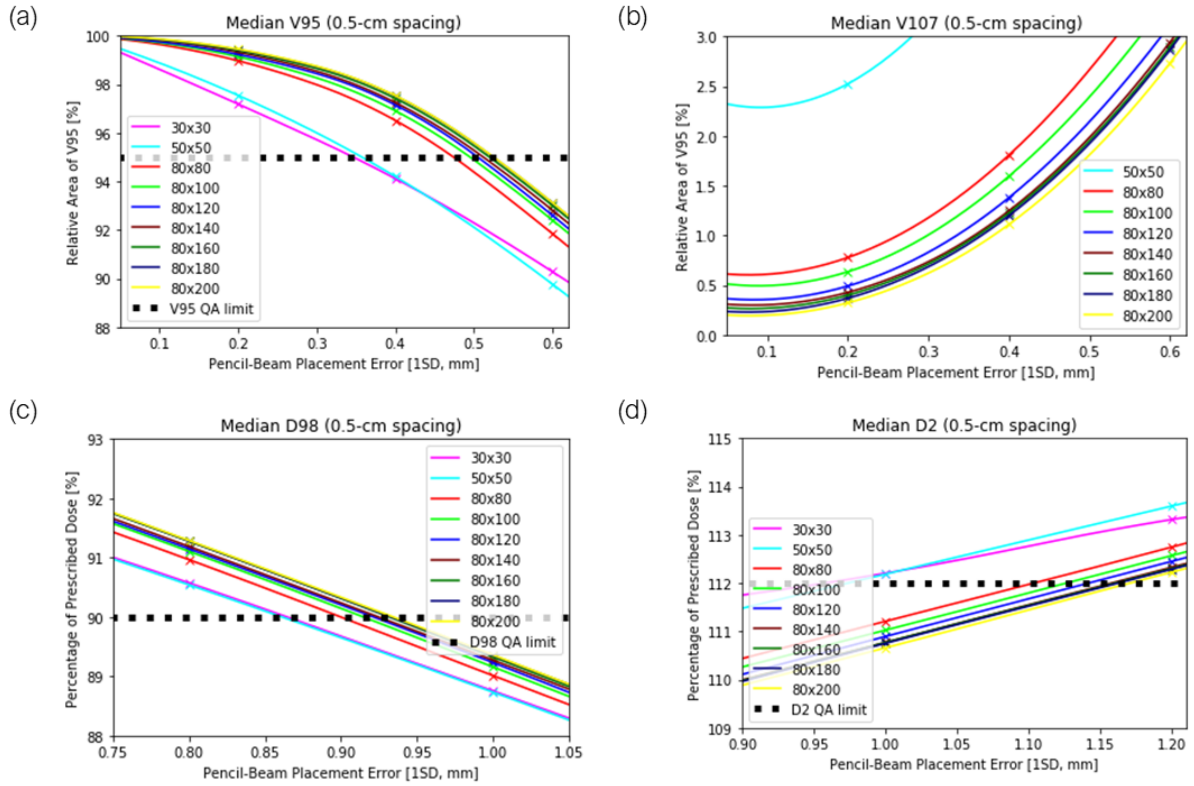
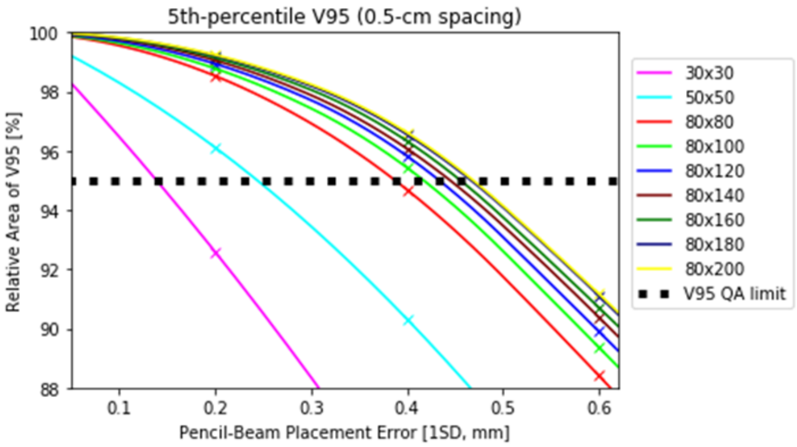


Figure 5.5: The correlation between Gaussian pencil beam displacement standard deviation on the x-axis and the median V95 (a), the V107 (b), the D98 (c) and the D2 (d) for various field sizes in mm with 0.5 cm pencil beam spacing. 10,000 samples were taken per standard deviation. Curves are fitted to the sample points with cubic B-splines. The horizontal black dotted lines indicate the clinically used constraints. These are 95% of V95 for the V95, $0.90D_{pres}$ for the D98 and $1.12D_{pres}$ for the D2. The prescribed dose is 8 Gy for each field.

constraints. The correlation with field size is again visible. In sub-figure (b) the intersection points are given. With increasing field sizes, it becomes easier to adhere to the constraint. In other words, the dose is delivered more robustly. The intersection values are more separated compared to Fig. 5.5(a). The median value of the clinical metrics is affected less than its 5th percentile and 95th percentile.

The impact of beam current fluctuations on the PBSDR is equal for all field sizes and identical to the representation in Fig. 5.4. This is because the PBSDR has only one time scale. The decrease or increase in beam current imposes a direct effect on the PBSDR related to the standard deviation of the added uncertainty. Doing this analysis with the biological FLASH metric (chapter 4) will probably return a field size dependency as two time scales are used.

(a)



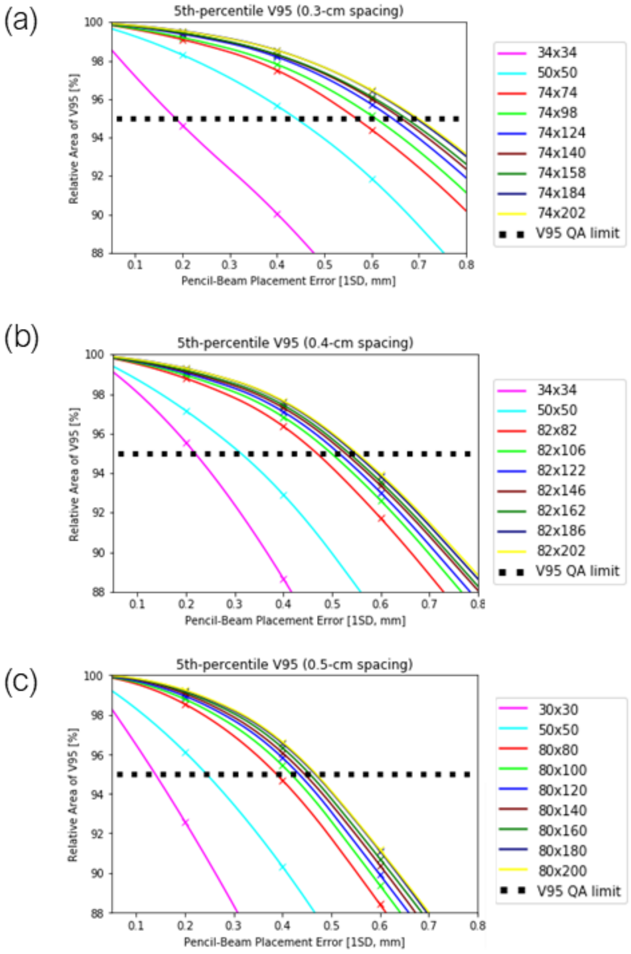
(b)

I_{30x30}	I_{50x50}	I_{80x80}	I_{80x100}	I_{80x120}	I_{80x140}	I_{80x160}	I_{80x180}	I_{80x200}
0.14 mm	0.24 mm	0.38 mm	0.42 mm	0.43 mm	0.44 mm	0.46 mm	0.47 mm	0.47 mm

Figure 5.6: The correlation between Gaussian pencil beam displacement standard deviation on the x-axis and the 5th percentile V95 (a) for various field sizes in mm with 0.5 cm pencil beam spacing. The points of intersection ($I_{..x..}$) with the QA limit are also given (b). 10,000 samples were taken per standard deviation. Curves are fitted to the sample points with cubic B-splines. The horizontal black dotted lines indicate the clinically used constraint. This is 95% of V95 for the V95. The prescribed dose is 8 Gy for each field.

5.4. Full clinical analysis

Dose delivery is more robust with increasing field sizes and the number of pencil beams. To guarantee the FLASH effect for small fields a smaller pencil beam spacing might be suitable. In Fig. 5.7 the 5th percentile of the V95 for the 27 fields together with the quality assessment constraint intersection point is shown. Sub-figure (a) corresponds to 0.3 cm pencil beam spacing, sub-figure (b) to 0.4 cm spacing and sub-figure (c) to 0.5 cm spacing. The V95 is analysed as it is the most critical clinical evaluation metric (Fig. 5.3). From the figure, it is clear that the pencil beam spacing indeed correlates with the Gaussian error intersection value. If more pencil beams are used to deliver similar doses, the individual allowed pencil beam displacement size increases. The allowed error size for the 5 cm square field jumps from 0.24 mm to 0.44 mm for 0.5 cm spacing against 0.3 cm spacing. Delivering FLASH proton therapy for small fields becomes more viable with smaller spacings. For the larger FAST-01 fields, the intersection value also increases with decreasing pencil beam spacing.



I_{34x34}	0.19 mm	I_{74x140}	0.66 mm
I_{50x50}	0.44 mm	I_{74x158}	0.67 mm
I_{74x74}	0.57 mm	I_{74x184}	0.68 mm
I_{74x98}	0.61 mm	I_{74x202}	0.70 mm
I_{74x124}	0.64 mm		

I_{34x34}	0.22 mm	I_{82x146}	0.53 mm
I_{50x50}	0.31 mm	I_{82x162}	0.54 mm
I_{82x82}	0.47 mm	I_{82x186}	0.55 mm
I_{82x106}	0.50 mm	I_{82x202}	0.55 mm
I_{82x122}	0.51 mm		

I_{30x30}	0.14 mm	I_{80x140}	0.44 mm
I_{50x50}	0.24 mm	I_{80x160}	0.46 mm
I_{80x80}	0.38 mm	I_{80x180}	0.47 mm
I_{80x100}	0.42 mm	I_{80x200}	0.47 mm
I_{80x120}	0.43 mm		

Figure 5.7: The correlation between Gaussian pencil beam displacement standard deviation on the x-axis and the 5th percentile V95 for various field sizes in mm with 0.3 cm pencil beam spacing (a), 0.4 cm pencil beam spacing (b) and 0.5 cm pencil beam spacing (c). The points of intersection ($I_{..x..}$) with the QA limit are shown next to the plots. 10,000 samples were taken per standard deviation. Curves are fitted to the sample points with cubic B-splines. The horizontal black dotted lines indicate the clinically used constraint. This is 95% of V95 for the V95. The prescribed dose is 8 Gy for each field.

6

3D uncertainty analysis

The 3D uncertainty analysis can be done with the dose engine described in 2.4.2. The adjoint analysis of the engine allows the quantification of dose deposition differences caused by anatomy changes throughout a proton therapy treatment. Perturbations in HU are directly coupled to the change in dose deposition. To quantify the dose response when pencil placement errors occur instead of anatomy changes, smart usage of the engine combined with the adjoint analysis is required. Dose rates cannot be analysed with the dose engine as there are no time scales in play. The uncertainty of beam current fluctuations can therefore not be analysed for this transmission FLASH treatment plan.

The results of the 3D uncertainty analysis are given in this chapter. First, the procedure of recreating the iCycle dose with the TUD dose engine based on the iCycle optimisation results is described in section 6.1. This step is essential as the FLASH treatment plan is made within iCycle. When the iCycle dose is recreated with the dose engine, the adjoint analysis can be applied. In this section, the iCycle and engine calculated dose matrices are shown and compared for the transmission FLASH patient as well. In the next section (6.2), various ideas for a pencil beam placement uncertainty analysis with the TUD dose engine are described.

6.1. From iCycle to dose engine

6.1.1. Dose engine input

The following treatment plan-specific inputs are required for a dose calculation with the TUD engine: (1) (x,y,z) coordinates relative to the CT scan of the start and end points of every pencil beam, (2) the energy per pencil beam, (3) the energy spread per pencil beam, (4) the number of protons per pencil beam and (5) the spatial spread (standard deviation) of a pencil beam. Inputs (2), (3) and (5) could easily be determined. The energy per pencil beam is fixed at 244 MeV with a 5 MeV standard deviation as the iCycle FLASH plan consists only of 244 MeV transmission beams. A 5 MeV energy spread is sufficient because it barely affects the dose calculation of transmission beams. The energy spread is needed for accurate dose calculations at Bragg-peak depths. The spatial spreading of a pencil beam is fixed at 0.32 cm, equal to the pencil beams used in the 2D model.

The procedure of finding the (x,y,z) start and end positions of individual pencil beams on the CT scan is less trivial. These positions have to be found with trigonometry based on the iCycle optimisation results of the treatment plan. The pencil beam locations in the plane of the scanning pattern have to be translated to the CT dimensions. The pencil beam scanning isocentre and the individual pencil beam shifts to the isocentre are utilised for this task. These shifts to the isocentre give rise to the scanning pattern. The isocentre is the CT (x,y,z) position of the middle pencil beam. The shifts in each direction are given per pencil beam in millimetres. The FLASH treatment plan is created with only a gantry rotation which slightly simplifies the problem. When first assuming a xy-slice at a fixed depth, the x and y start and end positions of each pencil beam can be found with only the isocentre location. For transmission beams the full CT extent in the x-direction is traversed. For the patient CT used, this implies that the starting position of each pencil beam is 25 cm and the end position is -25 cm. The y start and end position depends on the isocentre and the gantry angle. The y start and end location($Y_{s,e}$) are calculated with Eq. 6.1. In this equation, the gantry angle is denoted as θ in degrees and the isocentre locations in x and y as I_x and I_y in centimetres. The CT extent in x (x_{extent}) is constant and equals 50 cm in our case.

$$Y_{s,e} = I_y \mp \text{arctan}(\theta)(\frac{x_{extent}}{2} \pm I_x) \tag{6.1}$$

For the start and end locations in the z-direction, the isocentre z-position together with the individual pencil beam shifts is required. In the plane of the scanning pattern, there is no variation in depth because only a gantry angle is used for irradiation. However, when translating it to the CT domain, slight variations in start and end depth are present. These variations are induced by the pencil beam shifts in x and y. This means that the start and end z positions can be computed with the 2D rotation matrix. In Eq. 6.2 the expression for the start and end z-locations ($Z_{s,e}$) is given. The isocentre position in z is denoted as I_z and the pencil beam shifts in x and y as S_x and S_y

$$Z_{s,e} = I_z + \begin{pmatrix} S_x \\ S_y \end{pmatrix} \begin{pmatrix} \text{cos}(\theta) & \text{sin}(\theta) \\ -\text{sin}(\theta) & \text{cos}(\theta) \end{pmatrix} \tag{6.2}$$

The input number of protons per pencil beam for the dose engine have to be matched to the iCycle optimisation results as well. For the right scaling, the weight per pencil beam from the iCycle plan was used. In the iCycle version used, unity weight equals 1 Gy dose delivery in the Bragg peak. Recall from chapter 3 that for 244 MeV proton beam, 0.368 giga-proton deposit 1 Gy in the Bragg peak. Based on this beam characteristic, the iCycle weights could be converted to the number of protons. The decision was made to work with 0.3 giga-proton per Gy and afterwards match the iCycle and engine dose matrices with a least square optimisation. The inverse relation is valid as well and gives 3.33 Gy per giga-proton. The expression for the conversion of iCycle pencil beam weights to the number of protons is given in Eq. 6.3. The number of protons for a pencil beam (N_i) is related to its weight (w_i) and the assumed proton beam characteristic ($0.3\ GprGy^{-1}$).

$$N_i[pr] = w_i[Gy] * 0.3[GprGy^{-1}] * 10^9[pr/Gpr] \tag{6.3}$$

6.1.2. Dose computation and comparison

After determining both the start/end positions and the number of protons per pencil beam, a dose can be calculated with the engine. This dose should be similar to the iCycle-generated dose as the underlying pencil beams are equal. However, they will not be identical because iCycle and the TUD dose engine each use a different method for dose calculations. A dose within iCycle is calculated with a pencil beam algorithm involving multiple approximated kernels, [44] whereas the TUD engine is based on approximations to the Boltzmann particle transport equation. [10] Both algorithms have their strengths and limitations but will never return identical dose matrices. In my literature study (see appendix part D) more information about pencil beam and Boltzmann-based dose algorithms can be found.

The dose engine has difficulties with accounting for lateral inhomogeneities in the CT. Dose calculations for lung patients are affected relatively the most. To better account for lateral inhomogeneities, every pencil beam is split into smaller Gaussian beamlets. This improves the accuracy of the dose computations for the used FLASH lung patient. The splitting scheme applied is equal to [1, 6, 6, 12, 12], where each number matches the points on consecutive equidistant rings. This does imply that every input pencil beam is split into 37 beamlets. The dose calculations now have to be done for 2738 beamlets instead of 74 input pencil beams. When splitting the pencil beams into beamlets, the start/stop beam position and number of protons are adjusted accordingly. The gain in accuracy is paired with a loss in computation speed. By limiting the dose calculations to the patient CT outline, the whole calculation process is sped up.

With everything discussed above, a dose calculation is realised with the dose engine. The resulting dose matrix from the engine is on the same resolution as the CT scan. For the FLASH patient, it equals 512x512x259 with a resolution of 0.098 cm x 0.098 cm x 0.15 cm (x,y,z). The iCycle dose matrix is also calculated based on the CT scan extent but removed rows and columns where the dose deposition is 0 Gy. For this specific patient, it means that the iCycle dose is 334x504x259 with the same resolution as the engine dose. Before a valid comparison between the two can be made, the removed rows and columns have to be added back to the iCycle dose matrix such that the dimensions match. This is achieved with an implemented fitting function. Every possible positioning of the 334x504x259 iCycle dose inside the 512x512x259 engine dose is tested and related to a cost function. The best fit is used to obtain the 512x512x259 iCycle dose. The proton beam characteristic of 0.3 giga-proton per Gy is not accurate. With a least square optimisation, a calibration factor is found for the engine dose to better represent the iCycle dose. The found factor equals 1.1741, meaning that approximately 17% more protons per pencil beam are required. In other words, the beam characteristic should be 0.352 giga-proton per Gy instead of 0.3 giga-proton per Gy. This lies within 5% of the measured 0.368 giga-proton per Gy for 244 MeV proton beams at the HPTC.

The dose matrices for one xy-slice are shown in Fig. 6.1. The iCycle dose distribution is depicted in sub-figure (a) and the TUD engine dose in sub-figure (b). The target is the PTV, i.e. the red contour in the lung. Other organs-at-risk contours are shown as well. From these dose matrices, it is clear that the oesophagus receives a relatively high dose. The main difference between the two dose matrices is the smoothness. The iCycle dose is more smooth due to coarse approximation kernels involved in the pencil beam dose algorithm. The engine dose distribution is affected by density heterogeneity.

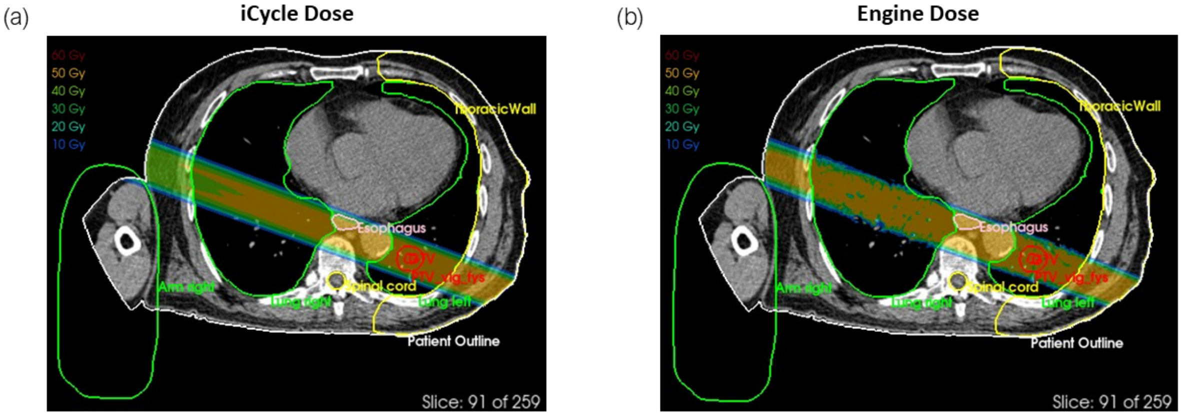


Figure 6.1: Transverse view of the iCycle dose distribution (a) and TUD engine dose distribution (b). The red contours represent the GTV and PTV. All other relevant contours are visible as well. The dose is shown for slice 91 of 259 from the CT scan. The dose values in Gy correspond to the colour scale in the upper left corner of the sub-figures. The plots of the dose distributions are created in RTstudio (at EMC).

Various 3D gamma pass rates were also computed for the dose matrices. In Fig. 6.2 the local gamma index distribution for the PTV is plotted. The dose margins of 1%, 2% and 3% together with distance margins of 1 mm, 2 mm and 3 mm are used. The dotted line corresponds to the 1%/1 mm gamma distribution, the striped line to the 2%/2 mm distribution and the solid line to the 3%/3 mm distribution. With these gamma curves, the gamma pass rate for the PTV can be found. It is defined as the relative volume with a Γ of less or equal to 1. For the PTV the following pass rates are found: (1) $\gamma_{1\%/1mm}$ = 62.2%, (2) $\gamma_{2\%/2mm}$ = 94.8% and (3) $\gamma_{3\%/3mm}$ = 99.7%. Because both algorithms are fundamentally different, the 1%/1 mm pass rate is not close to the other pass rates.

The oesophagus is an organ at risk for this lung patient so the local gamma pass rates are computed for this organ as well. The $\gamma_{1\%/1mm}$ = 75.3%, the $\gamma_{2\%/2mm}$ = 80.5% and the $\gamma_{3\%/3mm}$ = 85.5%. From these pass rates, it can be concluded that the dose distributions computed are not completely comparable for the oesophagus. At the edges of the structure, the dose varies relatively the most. This might be caused by a discrepancy in the spatial spreading of the pencil beams. For the dose engine, a spreading of 0.32 cm is used, whereas iCycle might use wider pencil beams.

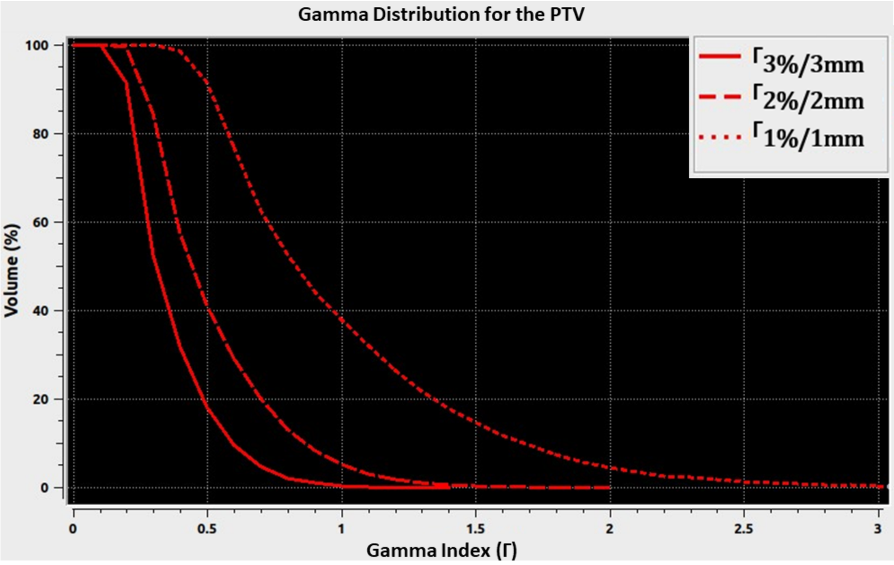


Figure 6.2: The local Gamma index volume histograms for the PTV. The relative area for a gamma index (Γ) of 1 is the gamma pass rate. The gamma evaluation is done within RTstudio (at EMC).

6.2. Concepts for the uncertainty analysis

The main strength of the TUD dose engine is its adjoint analysis part. With a forward dose calculation, such as the dose distribution shown in Fig. 6.1(b) and HU perturbations, the dose difference can be computed caused by the density shifts. Two types of pencil beam placement errors can be simulated with limited changes to the dose engine. The first and easiest error to implement is an equal displacement for every pencil beam. The second error is an individual pencil beam placement error. This can be simulated by rewriting the engine code. The procedures to simulate these errors are described below. They were not implemented and analysed yet because the project came to an end. Before performing the uncertainty analysis, it should also be studied if CT shifts do not disrupt the adjoint analysis.

6.2.1. Field error

The procedure to quantify the impact of a single pencil beam displacement with the adjoint analysis is straightforward. Two CTs are required. One original CT for the forward dose calculation and one 'virtual' CT for the pencil beam shift. First, the dose is calculated by tracking the input pencil beams through the original CT. Thereafter, the dose response is formulated based on the second virtual CT. In short, a second CT has to be created which relates the pencil beam shift to a HU shift.

With the following steps, such a virtual CT can be constructed. Firstly, import the original CT in matrix form. This can be achieved within various programming environments. MATLAB has built-in DICOM functions for importing CT scans. The Pydicom module for Python should work as well. Secondly, a random single pencil beam shift has to be generated in x, y and z. The pencil beam error can be generated by drawing a number from a Gaussian distribution with a relevant standard deviation. For instance, the 0.5 mm standard deviation used in chapter 3. Thirdly, the three randomly drawn errors for the pencil beams have to be related to the imported CT. This step is essential, as the resolution of the imported CT is limited and not constant in x, y and z. As an example, the CT of the lung patient used in this study has a resolution of 0.098 cm x 0.098 cm x 0.15 cm. The drawn random errors have to be matched with the voxel resolutions. Additionally, (linear) interpolation of voxel HU is required as the drawn random error will most likely not be an exact multiple of the respective resolution. Another crucial aspect to realise is that a generated pencil beam displacement relates to the inverse virtual CT shift. For instance, when a pencil beam is shifted 1 mm in an upwards direction, the virtual CT scan has to be shifted 1 mm downwards. This guarantees that the originally put-in beam directions match the random displacement. Lastly, the new virtual CT has to be saved with DICOM standards. The most logical approach would be to copy all metadata of the first original CT to the second virtual CT. The only difference would be the 3D HU matrix.

To clarify the procedure, an example scenario comes in handy. Assume that a random error of -0.098 cm x 0.245 cm x 0.10 cm (x,y,z) is drawn for the CT resolution of the described lung patient. The shifts can be implemented per direction. The CT shift in the x-direction is the inverse of the drawn error, so here it equals 0.098 cm. This implies that the imported CT is shifted 1 voxel in the positive x-direction. The y-direction CT shift is -0.245 cm and is not an exact multiple of the y-resolution. 0.245 cm is 2.5 times the resolution in the y-direction. This means that interpolation of voxel values is required. To implement this, first shift the current CT 2 negative y-direction voxels. The remaining 0.5 voxel shift has to be achieved by interpolation. With linear interpolation, every voxel should be changed to the weighted average of two consecutive voxels in the negative y-direction. The weight is equal (0.5) for this drawn shift. The z-direction shift is -0.10 cm and requires interpolation as well. Not a single full voxel can be shifted and direct interpolation is required. 0.10 cm is $\frac{2}{3}$ of the z-resolution. This implies that unequal weights are required for the voxel interpolation. Here, they should be $\frac{1}{3}$ and $\frac{2}{3}$. After applying these shifts, a new virtual CT is formed which is a shifted version of the original CT. The adjoint analysis can be done with the virtual CT without any rewriting of code.

With this procedure, one single pencil beam error is simulated and implemented in the system. More clinically relevant would be to quantify the impact of many various possible placement errors. This should be possible with the adjoint analysis as well. First-order changes can correctly be mapped to the CT displacement a priori. The first step would be to unravel the mathematical relation between the response function and the displacement.

6.2.2. Individual pencil beam error

For even more clinical relevance, an uncertainty analysis of the displacement error per pencil beam is needed. This cannot be achieved with the procedure described above. When shifting the original CT to a virtual CT for multiple pencil beams, the shifts get intertwined in the virtual CT. In other words, the relation between displacement and CT voxel shifts is lost. One possible solution would be to create an adjoint response function for every pencil beam. This is a tedious process but does not require rewriting the engine code. The treatment plan for the lung patient with a relatively small PTV already has 74 pencil beams, implying that 74 consecutive dose and adjoint calculations are required for the system. The procedure would go as follows: (1) draw a random displacement in x,y and z per pencil beam. (2) Create a virtual CT for every displacement. (3) Calculate the response function per pencil beam. (4) Take the sum of all response functions for the dose difference.

It is also possible to implement individual pencil beam errors by rewriting the code. The creation of multiple virtual CTs is not needed with this procedure. The general idea is to track the traversed voxels per shifted pencil beam and use these voxel values in the adjoint analysis. The following procedure has to be introduced in the dose engine: (1) save all voxel values and positions the original pencil beams traverses. This is possible within the model as there is already a function which is capable of pencil beam tracking. (2) Draw a random error per pencil beam. (3) Change the start and end position of the beam according to the drawn error. (4) Save all voxel values and positions that the displaced beam traverses. (5) Create a virtual CT with the found voxel differences. The original encountered voxel values should be changed to the values encountered after the beam shifts. After completion of this procedure, there should be one virtual CT which can be used for the adjoint analysis.

Again, to increase clinical relevance, the exact relation between the response function and the pencil beam displacement should be found.

7

Discussion

This project aimed to quantify the impact of machine uncertainties on the efficacy of FLASH proton therapy. Further clinical translation of FLASH proton therapy is only possible when machine-intrinsic uncertainties do not affect the FLASH effect. In this chapter, the results are put into the project purpose perspective. Limitations and future research directions are mentioned as well. The results related to the analytical model are discussed in section 7.1. In section 7.2 the implications of the clinical 3D analysis are described.

7.1. Analytical dose model

In this thesis, an analytical dose model was introduced and used to analyse machine uncertainties in pencil beam placement and beam current. The general setup of this analytical model is discussed in sub-section 7.1.1. The findings related to the first uncertainty analysis are described here as well. In the following sub-section (7.1.2), the implications of the new biological FLASH metric are given. Thereafter, the results from the implemented clinical model are discussed in sub-section 7.1.3.

7.1.1. Model setup and first uncertainty analysis

In chapter 3, an one- and two-dimensional analytical model was formed with Gaussian pencil beams in line with the set FLASH thresholds. With a least-square optimisation, relative weights of the pencil beams were chosen such that the 8 Gy prescribed dose was delivered as homogeneous as possible. After weight optimisation, machine uncertainties of 0.5 mm standard deviation in pencil beam position and 5% standard deviation in beam current were added to the model. The impact on dose and dose rate delivery was evaluated per voxel with a convolutional and sampling approach.

Figures 3.4 and 3.10 indicate that the analytical model with weight optimisation functions as expected. The one- and two-dimensional delivered dose is homogeneous and equals the prescribed dose of 8 Gy. The pencil beam placement is of importance for dose and dose rate delivery. Especially at the edge of the scanning pattern, the error due to pencil beam placement is large (Fig. 3.14). At these positions, relatively few pencil beams deliver the dose. In other words, the error is not compensated by other pencil beams. This is confirmed by Fig. 3.15(b). The relative error in PBSDR correlates positively with the pencil beam weight. The proton beam current fluctuation only affects the dose rate. The main finding is that the fluctuation in beam current imposes a direct effect on the PBSDR. With a 5% standard deviation, approximately 8% relative error in the PBSDR is found for the whole field (Fig. 3.16(a)). Figure 3.18 shows the results for the combination of both types of uncertainties. The PBSDR is affected the most by the beam current fluctuation. Only at high pencil beam weights, the pencil beam placement is more important.

The results in chapter 3 also made clear that the PBSDR is a sensitive metric and not in line with the current knowledge about the FLASH effect. The impact of pencil beam placement and beam current errors on the PBSDR is high and correlates with the pencil beam weight (Fig. 3.18). The direction of the pencil beam scanning pattern does not affect the PBSDR. The PBSDR only has one characteristic time scale, which implies that a constant error is added with beam current fluctuations (3.16(a)). The PBSDR requires a dose threshold as input and the computed dose rate depends on this threshold (Fig. 3.12). This dose threshold cannot directly be linked to the FLASH effect.

The analytical dose model has some limitations that should be mentioned. First of all, the modelling of the halo effect and its corresponding alpha value is based on data from the MGH. It would be better to incorporate the halo and alpha values following HPTC beam data. Additionally, the 'pure' Gaussian pencil beam standard deviation was fixed at 0.32 cm. This is an approximation to the found values for proton pencil beams at the HPTC but not exactly fitted to experimental data. The pencil beam weights are found with a simple least-square optimisation. The optimisation of a clinical treatment plan is more advanced. Moreover, the optimisation is done for the peak location of the first pencil beam until the last pencil beam. This implies that dose delivery is not optimised for the fall-off region. Including this region will yield better target coverage. All computed dose rates are coupled to the maximum theoretical dose rate at the Bragg peak. However, this dose rate will never be reached with transmission proton beams. The proton fluence at the plateau is approximately 0.3 times the fluence in the Bragg peak. [6]

The uncertainty analysis for 0.5 mm standard deviation in pencil beam position and 5% standard deviation in beam current consists of some limiting factors as well. The clinical importance of this analysis is limited as the induced relative error is evaluated per voxel and not per treatment plan. The research question can partially be answered with this analysis. The uncertainty analysis will be strengthened by an analysis on a treatment basis instead of a voxel basis. Furthermore, only the random individual pencil beam placement error and systematic beam current fluctuation are included. The analysis becomes more realistic when also systematic whole pencil beam field displacements and random individual pencil beam current fluctuations are added to the model. Lastly, the uncertainty analysis is only done for one predefined error size in pencil beam position and beam current fluctuation. It is unclear if this predefined error is completely in line with the HPTC machinery. Measurements are needed to characterise the errors at HPTC. These uncertainties can be added to the model and help with patient QA.

In the future, the analytical dose model should be fine-tuned to match measurements from the HPTC. My suggestion would be to remove the double Gaussian pencil beams and replace them with a single measured Gaussian at a certain depth. This measured Gaussian should already include the energy dispersion because of particle interactions. In this case, the model will also not depend on beam data from an IBA cyclotron at MGH. The pencil beam weight optimisation can be improved further such that dose delivery is more robust. I would suggest starting with adjusting the set optimisation boundaries. The components of the uncertainty analysis can also be adjusted in future work. Adding a whole-field pencil beam displacement and individual pencil beam current fluctuation will strengthen the uncertainty analysis. Additionally, the characterisation of the machine-intrinsic errors at HPTC is necessary to improve the relevance of the uncertainty analysis.

In conclusion, the analytical dose model is suited to simulate dose delivery and can be used for the uncertainty analysis in pencil beam position and beam current fluctuation. The error in pencil beam position affects the dose and dose rate, whereas the error in beam current only affects the dose rate. When combining both types of machine uncertainties, the beam current fluctuation affects the dose rate more than pencil beam displacements. The PBSDR consists of several shortcomings which limit its relevance for FLASH proton therapy. The analytical model and uncertainty analysis should be improved further to increase clinical impact.

7.1.2. Biological FLASH metric

The new FLASH metric described in chapter 4 was developed to account for the mentioned limitations of the PBSDR. A more robust and more FLASH effect-related metric is essential for the clinical translation of FLASH proton therapy. The created metric is the effective FLASH enhancement ratio (FER_{eff}) and is based on the radiolytic oxygen depletion hypothesis for the FLASH effect. Healthy tissue is spared more compared to conventional proton radiotherapy because the ultra-high dose rates induce rapid oxygen depletion in the healthy cells. Cells in hypoxic conditions are more protected against indirect radiation damage. With time traces of the oxygen concentration per voxel, the FER_{eff} is quantified. The oxygen concentration over time is modelled with an ordinary differential equation which includes a term for oxygen depletion caused by the dose rate and re-oxygenation related to current oxygen concentration. There are two model parameters, κ and τ, which are related to depletion and re-oxygenation respectively. Re-oxygenation also depends on the steady-state oxygen concentration which is predefined in accordance with other research.

Again, the 0.5 mm pencil beam position and 5% beam current fluctuation uncertainties were added to the model. From Fig. 4.7 it is clear that the FLASH metric is more robust for the added uncertainties than the PBSDR. For 8 Gy prescribed dose the FER_{eff} is almost constant for the field (Fig. 4.5), which implies that the FLASH effect is limited. At higher prescribed dose values, the FER_{eff} changes and the differential FLASH effect becomes visible (Fig. 4.6). The onset of the FLASH effect lies somewhere between 8 Gy and 16 Gy, which is in agreement with the set FLASH dose threshold of 8 Gy. In contrast to the PBSDR, the impact of beam current fluctuations is not constant for the whole field (Fig. 4.8). This is induced by the fact that two separate time scales are used for the FER_{eff}, the oxygen depletion and re-oxygenation. From the figure, it is also clear that the FER_{eff} at the edges of the scanning pattern is less sensitive to beam fluctuations than at the middle of the pattern. The reason for this is that at the edges of the pattern, relatively few pencil beams deliver the dose. This is a key result, as the FER_{eff} should be high and robust at the edges of the scanning pattern. Here the healthy tissue is located. In the middle of the scanning field, where the tumourous tissue is situated, the FER_{eff} is of less importance. This also indicates that scanning pattern optimisation is beneficial for achieving the FLASH effect. The variable dose threshold of the PBSDR is not required for the calculation of the FER_{eff}. Additionally, as a FLASH enhancement ratio is calculated instead of a dose rate, the relation with the FLASH effect is more pronounced.

The current findings are limited in several ways. The main limitation of the FLASH metric relates to the model parameters. The calculated FER_{eff} depends on the chosen κ, τ and C_{ss}. The κ and τ were fitted following the research of M. Diepeveen. (his MSc thesis [14] and an article under revision). The selected values might not be correct for our developed model. The steady-state oxygen concentration was reverse-engineered such that the FER_{eff} is bound between 1 and 2. The lower bound of 1 corresponds to the normoxic condition and the upper bound of 2 to the hypoxic condition. The upper bound was fixed at 2 in agreement with found FERs in animal studies. However, whether this truly is the upper bound is not clear and should be studied further.

More work is needed to increase the reliability of the FLASH metric. The FER_{eff} shows a dependency on the prescribed dose. This implies that the model parameters should be fitted such that the computed FLASH effect matches reality. Instead of currently using one source, multiple studies should be used for fitting the model parameters. Another interesting direction that could be explored is scanning pattern shape optimisation. The results show that the FLASH effect depends on the scanning pattern so optimisation of the shape is promising. Evaluating the FER_{eff} for the snow-flake-shaped scanning patterns found by Santo et al. would be a good start. [41]

The FER_{eff} is a robust metric which directly relates the FLASH effect to oxygen depletion induced by the instantaneous dose rate. All the shortcomings of the PBSDR are dealt with. The FER_{eff} is suited for quantification of the FLASH effect under the influence of pencil beam placement errors and beam current fluctuations. The next step would be to fit the model parameters to more clinical data.

7.1.3. Clinical model evaluation

The uncertainty analysis of the analytical model is extended for various standard deviation sizes in pencil beam placement error and beam current fluctuation in chapter 5. Instead of one field with a fixed 0.5 cm pencil beam spacing, the impact of the errors was evaluated on 9 typical FLASH fields with 0.3 cm, 0.4 cm and 0.5 cm pencil beam spacing. This results in a total of 27 fields. Moreover, the effect of the added uncertainties was not evaluated voxel-wise but on an individual treatment base. The found errors directly relate to a potential FLASH proton therapy treatment. In contrast to quantifying the relative error in dose and dose rate, the error was linked to clinically used ICRU metrics and respective constraints. [26, 27] These are the V95, V107, D2, D98 and PBSDR98. The decision was made to relate pencil beam placement errors to the V95, V107, D2 and D98. Beam current fluctuation-induced errors were only evaluated according to the PBSDR98. The impact of machine-intrinsic uncertainties on FLASH proton therapy can be evaluated. The found results will shape future QA.

Gaussian quadrature sampling was implemented to speed up dose calculations while maintaining accuracy. The difference between a high-resolution and low-resolution dose calculation is still visible in Fig. 5.1. The threshold for the high-resolution median V95 is at 0.44 mm whereas the quadrature sampled median V95 threshold is at around 0.40 mm. Moreover, the results from the quadrature sampling are identical to the low-resolution results. The effect of quadrature sampling is limited. Fig. 5.2 shows the difference. The pencil beam weight optimisation depends on the voxel resolution. With the high-resolution, the Gaussian pencil beams are better sampled resulting in more conformal dose distributions. Consequently, the constraint violation value for the V95 is somewhat higher.

From the single field uncertainty analysis for the 8.0 cm by 8.0 cm field with 0.5 cm pencil beam spacing shown in Fig. 5.3 it is clear that target coverage is the limiting factor. This sensitivity can be explained by the optimisation procedure. The dose delivery is not optimised for the fall-off region, which limits target coverage under the influence of pencil beam displacements. At 0.38 mm standard deviation pencil beam placement error, the V95 clinical constraint is exceeded for 5% of patients treated. The PBSDR98 constraint for this field is breached at 6.8% beam current fluctuation (Fig. 5.4). Adhering to the allowed pencil beam placement error seems more challenging than the beam current fluctuation at this point. In particular, because patient set-up uncertainty is not included in the analysis. Patient positioning does not affect the total beam current uncertainty. In Fig. 5.6 the decrease in tumour coverage is correlated with the pencil beam position uncertainty for the 9 field sizes with 0.5 cm pencil beam spacing. There is a clear correlation between V95 constraint violation and field size visible. The bigger the field, the bigger the allowed pencil beam positioning error. This relation can directly be linked to the number of pencil beams per field. With relatively few pencil beams, the added error in pencil beam placement cannot be compensated by other pencil beams. The V95 curves also seem to converge to a pencil beam placement error with increasing field size as the change in gain in constraint violation diminishes. For the biggest evaluated field of 8 cm by 20 cm, a pencil beam placement uncertainty of 0.47 mm is allowed to deliver a FLASH dose to at least 95% of patients. This constraint is again strict as patient positioning is not taken into account. By using smaller pencil beam spacings, a higher pencil beam positioning error is possible. In Fig. 5.7 the V95 curves for the 0.3 cm, 0.4 cm and 0.5 cm pencil beam spacings are shown. The constraint of the 7.4 cm by 20.2 cm field lies at the 0.7 mm pencil beam placing uncertainty. In general, the allowed pencil beam uncertainty increases for every field when the pencil beam spacing is decreased. Again, the number of pencil beams used for every field explains this fact. For all pencil beam spacings, the constraint violation pencil beam positioning error value converges with field size.

Some of the 0.3 cm spacing plans consist of more than 1000 pencil beams. Delivery of all these pencil beams while adhering to the FLASH threshold in treatment time (200 ms) is problematic. When a dose rate of 4000 Gy per second per pencil beam is achieved to deliver 8 Gy, the total treatment time is 2 seconds for 1000 pencil beams. This is an order bigger than the FLASH threshold. Scanning pattern shape optimisation might be a solution to still guarantee a local FLASH effect in such scenarios.

Whether FLASH transmission proton therapy is viable when taking into account machine-intrinsic uncertainties can be partially answered by the discussed findings. The constraint in beam current fluctuation of 6.8% standard deviation seems manageable. This value is constant for every field and every pencil beam spacing as the used PBSDR for this analysis only consists of one time scale. The impact of pencil beam positioning on accurate FLASH dose delivery is the limiting factor. The characteristic pencil beam placement error at HPTC should be found to select an optimal pencil beam spacing. The FLASH effect is delivery time-dependent and with fewer pencil beams the FLASH effect will be more guaranteed. However, irradiating a field with 0.5 cm pencil beam spacing is not always possible because of the constraint on target coverage. There is a trade-off between target coverage and the FLASH effect. This also explains the results of the first 10 patients treated within the FAST-01 clinical trial. There was no significant difference between conventional proton therapy and FLASH transmission proton therapy found. [12, 32] This indicates that the trade-off between target coverage and the FLASH effect came into play. Target coverage most likely was maintained but at the cost of the FLASH effect. The 0.5 cm pencil beam spacing fields used in the clinical trial, the 7 biggest fields in table 5.1, consist of hundreds of pencil beams. The FLASH effect is most likely affected because of the total irradiation time.

The uncertainty analysis with the clinical model with relevant dosimetric parameters consists of some limiting factors as well. Firstly, the analytical model used for the uncertainty analysis was used without the second Gaussian for the halo effect. This removes the dependency on data from the MGH. Only the pure pencil beam width of 0.32 cm was used. However, this is probably not exactly the width of the pencil beams at HPTC of transmission proton beams. Secondly, the impact of pencil beam positioning was only evaluated on the dose delivery and not on the dose rate. The dose rate is also affected by pencil beam position and might tighten the trade-off between target coverage and the FLASH effect further. Thirdly, the loss of V95 was dominant over all other evaluation constraints. This indicates that target coverage is not perfect. As mentioned before, the pencil beam weights are optimised from the peak of the first pencil beam to the last one. The dose fall-off region is disregarded. Better target coverage might be achieved when adjusting the optimisation area, however, homogeneity will likely go down. There is a trade-off between homogeneity and conformity. Another limitation is the used voxel resolution. The uncertainty analysis was done for a grid with a 0.1 cm resolution. At finer resolutions, the constraint violation on dose delivery improves. Optimisation of pencil beam weights at a higher resolution might return a more conformal dose distribution and strengthen the uncertainty analysis. Lastly, the evaluation of the impact of the beam current fluctuation on the dose rate is done with the PBSDR. The FER_{eff} will return other, more clinically relevant results. Because two separate time scales are used in this metric, there will possibly be a field size and pencil beam spacing correlation with beam current fluctuation size visible.

In the near future, the pencil beam placement and beam current fluctuations have to be measured for various scanning patterns and pencil beam spacings. Ideally, a database is created for these Gaussian error sizes. With such a database, a clinical FLASH treatment plan can be created and delivered. Monitoring the total treatment time is also required to quantify the FLASH effect. It is also interesting to adjust the optimisation procedure of the pencil beam weights. A more clinically relevant treatment plan will be found when optimising the weights on homogeneity, conformity and a higher grid resolution. Currently, it is unlikely that big field sizes can be irradiated while adhering to the FLASH thresholds. Therefore, novel dose-delivery techniques should also be explored in future research. Treatment of an 8.0 cm by 20.0 cm field can also be achieved by irradiating four consecutive local fields of 8.0 cm by 5.0 cm. A study has shown that the average dose rate increases when splitting a treatment plan into multiple sub-plans. [31] The relation between the local and full-field FLASH effect can be studied further with the analytical model. With the same reasoning, the effect of fractionation on the uncertainty analysis should be studied. The found constraints on target coverage might be slightly loosened because of fractionation. Again the relation between the fractional FLASH effect and the full-treatment FLASH effect has to be established.

Based on the findings presented in this chapter, FLASH proton therapy seems viable when taking into account machine-intrinsic uncertainties. However, the field sizes and pencil beam spacings for which FLASH proton therapy is possible have to be determined at HPTC. Both the constraints on the ICRU metrics and FLASH thresholds have to be fulfilled for FLASH treatments. The analytical model with the uncertainty analysis will help shape future FLASH proton therapy and can be used for QA purposes.

7.2. 3D analysis

In this section, the results of chapter 6 are discussed. The goal was to quantify the impact of pencil beam placement errors on a FLASH treatment plan from the EMC. The uncertainties can be added to the treatment plan by using a dose engine from the TUD. The first step was to recreate the dose distribution from the EMC (iCycle) with the TUD dose engine. Thereafter, pencil beam placement errors can be evaluated with an adjoint analysis. The uncertainty analysis has not yet been done. However, several concept ideas are given to perform the analysis. In sub-section 7.2.1 the results from the dose recreation are discussed. Limitations and future steps are described as well. In the last sub-section, the concepts are discussed together with possible pitfalls.

7.2.1. Dose calculation and comparison

The iCycle dose distribution was recreated with the TUD dose engine with the help of the iCycle optimisation results and the HPTC proton beam data. The start and end positions of pencil beams on the CT were found by the iCycle isocentre position together with pencil beam shifts and the gantry angle of 110°. The number of protons per pencil beam was calculated by relating the beam characteristic of 0.3 $GprGy^{-1}$ to iCycle pencil beam weights. An iCycle weight of 1 corresponds to 1 Gy in the Bragg peak. The energy per pencil beam was fixed at 244 MeV, similar to the iCycle treatment plan. A 5 MeV energy spread was assumed together with a 0.32 cm pencil beam width (equal to the width used in the 2D model). The dose computed with the TUD engine was compared with the iCycle dose with a gamma evaluation. A local gamma pass rate is determined for the PTV and the oesophagus. The gamma pass rates indicate that the dose matrices are not completely similar. This is intuitive as the iCycle and TUD dose engines use different types of dose calculation models with varying levels of computation accuracy. The analysed treatment plan is from a lung patient with a relatively small PTV. The dose calculation for lung patients is challenging because of the low mass density of air and the presence of large density gradients. Because of the approximations present in both types of dose algorithms, the computed dose will lack accuracy compared to full-scale MC simulations. Consequently, the difference between the iCycle dose and the TUD engine dose is not surprising.

The local gamma index curves for the PTV are shown in Fig. 6.2. The found gamma pass rates are: $\gamma_{1\%/1mm}$ = 62.2%, $\gamma_{2\%/2mm}$ = 94.8% and $\gamma_{3\%/3mm}$ = 99.7%. From these pass rates, it is clear that the iCycle and TUD engine doses are not completely identical. Especially the 1%/1 mm pass rate is relatively low. The gamma pass rates in the oesophagus were also determined. They were $\gamma_{1\%/1mm}$ = 75.3%, $\gamma_{2\%/2mm}$ = 80.5% and $\gamma_{3\%/3mm}$ = 85.5%. For this OAR, the 2%/2 mm and the 3%/3 mm pass rates are relatively low compared to the PTV pass rates. At the z-directional edges of the oesophagus, the iCycle dose and engine dose differ. To quantify the effect of pencil beam positioning errors on the FLASH effect, the dose should be correct for OARs. The FLASH effect is defined for healthy tissue. If the computed dose with the dose engine is not completely correct, the uncertainty analysis will lose significance.

The lack of similarity can be caused by multiple limitations of the dose computation procedure. First of all, the iCycle dose is reshaped to match the dimensions of the TUD engine dose. The reshaping might induce minor errors in voxel-wise dose values. However, this step is inevitable because otherwise the dose distributions cannot be compared voxel-wise. Secondly, the similarity between the iCycle dose and the engine dose is improved with a least-square optimisation. With this procedure, a calibration factor is found for the beam characteristic. The adjusted beam characteristic lies within 5% of beam measurements at HPTC. The calibration factor might be unnecessary and omitting it potentially yields better results for OARs. Thirdly, the used 0.32 cm pencil beam width might not be identical to the pencil beam width used within iCycle. Slightly wider pencil beams might further improve the gamma pass rates for all structures. Lastly, the stopping power values are linearly interpolated for 244 MeV proton pencil beams with the dose engine. The dose computation will be more accurate when polynomial interpolation is applied. This is not possible yet because the stopping power values are not determined for protons of more than 250 MeV. The stopping power difference will not be major as in this energy regime the stopping power values are almost constant. [5]

More research should be done to verify if the iCycle dose and engine dose are as comparable as possible. A good follow-up would be to benchmark the engine dose against a full-scale MC simulated dose. This can be done with FLUKA or GEANT4. Small changes to the current iCycle to dose engine conversion can be introduced to further increase the similarity between the iCycle dose and the dose engine dose. The calibration factor should be omitted and the pencil beam width should be identical. The gamma pass rate should be as high as possible such that the uncertainty analysis is trustworthy.

In general, the procedure to mimic an iCycle dose with the TUD dose engine functions. However, the gamma pass rates are not optimal yet. The significance of the uncertainty analysis depends on the gamma pass rate. Multiple strategies should be explored to increase similarity.

7.2.2. Uncertainty analysis

As mentioned before, the impact of pencil beam placement errors has not been evaluated with the adjoint analysis yet. Postulations for a single field error and individual pencil beam errors are conceptualised. They involve HU perturbations related to CT translations. Pencil beam placement errors are mapped to a virtual CT from which the impact of pencil beam displacements can be quantified. More work is needed to implement the concept ideas.

There is one crucial aspect that should be investigated before doing the uncertainty analysis. Based on the MSc thesis of S. Wendt [47], the adjoint calculated dose response is accurate for small HU perturbations of several percentages and small detector transformations of approximately $\pm$ 1 millimetre. However, greater random pencil beam errors will be drawn with the uncertainty analysis implying that the adjoint analysis might not be accurate anymore. Consequently, the HU perturbations created by the pencil beam shifts can become large at high mass density gradients, for instance, a voxel changing from lung to bone tissue. The envisioned uncertainty analysis for FLASH proton therapy is limited by these factors. First, it should be investigated if CT transformations are not compromising the adjoint analysis. The accuracy of the adjoint analysis can be quantified by comparing it with another forward dose calculation for the shifted scenario. If CT transformations are not possible, other concepts have to be developed to simulate pencil beam placement errors.

When a procedure is validated for the use of the adjoint analysis for pencil beam displacements, the most logical first step would be to simulate a field error. The field error corresponds to a position change of all pencil beams with the same shift. This random error is drawn according to a Gaussian distribution with a certain standard deviation. It is more clinically relevant to simulate the change in dose deposition for infinite randomly drawn errors. The focus should be on determining the mathematical relation between the response function and the drawn pencil beam error. This relation should be well validated and generalised for multiple Gaussian standard deviations. After the implementation of a field error, the uncertainty analysis can be extended for individual pencil beam errors.

With the current adjoint analysis, the response function is determined and validated for the dose, the NTCP and the TCP. In future research, the adjoint system should also be developed for the clinical metrics used in the 2D uncertainty analysis. These are the V95, V107, D2 and D98. When these adjoint systems are found, the 3D uncertainty analysis can be done following the ICRU metrics and constraints. [26, 27]

Currently, it is not possible to include beam current fluctuations in the analysis as only a dose response is evaluated. Dose deposition is not affected by the change in beam current. To include this in the analysis, an adjoint model for the FER_{eff} or PBSDR should be created. This will strengthen the uncertainty analysis. These models should be created to analyse the impact of machine uncertainties on FLASH proton therapy.

In theory, the adjoint analysis is suited for an accurate and fast uncertainty analysis. However, several points have to be addressed before the analysis is possible for iCycle FLASH treatment plans. When all limitations are tackled, the impact of pencil beam positioning on FLASH dose delivery can be quantified.

8

Conclusion

The aim of the project is to quantify the impact of machine uncertainties on the efficacy of FLASH proton therapy. Two machine-intrinsic uncertainties affecting FLASH proton therapy are implemented in an analytical dose model. The impact of different sizes of individual pencil beam placement errors is evaluated on the V95, V107, D2 and D98. For the generated fields, the constraint for the V95 is violated for 5% of patients at the smallest random pencil beam displacement size. Two important relations for dose delivery are found. The allowed pencil beam error scales with increasing field size and decreasing pencil beam spacing. However, under these conditions, the FLASH effect diminishes. A trade-off between the FLASH effect and target coverage exists. The found limits of FLASH dose delivery have to be translated to HPTC machine data. Additionally, the impact of increasing sizes of beam current fluctuations is coupled to the PBSDR98 and the -10% dose rate constraint. For all field sizes and pencil beam spacings an identical allowed random beam current error was found. The allowed beam current fluctuation is within easier reach than the allowed error in pencil beam position. The PBSDR consists of several shortcomings so in the next uncertainty analysis the change in beam current should be coupled to the FER_{eff}. Proton FLASH radiotherapy is viable under the influence of machine-intrinsic uncertainties. In the future, it should be studied for which field sizes.

Before FLASH proton therapy is translated to the clinic, the uncertainty analysis should also be performed on a generated FLASH treatment plan. Several concept ideas are formulated to analyse the impact of pencil beam placement errors on FLASH dose delivery with an adjoint analysis. First, the accuracy of the adjoint analysis concerning CT transformations and HU perturbations has to be studied. In the end, various sizes of pencil beam displacements should be related to the ICRU metrics and constraints for dose delivery. Ultimately, beam current fluctuations should be added to the 3D uncertainty analysis as well.

This research forms the first base of knowledge for the relation between machine intrinsic uncertainties and proton FLASH radiotherapy. The developed analytical dose model can be used for QA purposes. The results offer insights into the boundaries of FLASH dose delivery. The challenge is to adhere to both the FLASH thresholds and the ICRU constraints.

References

[1] Isabel Abril et al. “Lateral spread of dose distribution by therapeutic proton beams in liquid water”. In: *Nuclear Instruments and Methods in Physics Research Section B: Beam Interactions with Materials and Atoms* 352 (2015), pp. 176–180.

[2] Gabriel Adrian et al. “The FLASH effect depends on oxygen concentration”. In: *The British journal of radiology* 92.1106 (2020), p. 20190702.

[3] Claudia Allemani et al. “Global surveillance of cancer survival 1995–2009: analysis of individual data for 25 676 887 patients from 279 population-based registries in 67 countries (CONCORD-2)”. In: *The Lancet* 385.9972 (2015), pp. 977–1010.

[4] Claudia Allemani et al. “Global surveillance of trends in cancer survival 2000–14 (CONCORD-3): analysis of individual records for 37 513 025 patients diagnosed with one of 18 cancers from 322 population-based registries in 71 countries”. In: *The Lancet* 391.10125 (2018), pp. 1023–1075.

[5] MJ Berger et al. “Stopping-power and range tables for electrons, protons, and helium ions, NIST Standard Reference Database 124”. In: *National Institute of Standards and Technology (NIST), Physical Measurement Laboratory (PML)* (2017).

[6] Thomas Bortfeld. “An analytical approximation of the Bragg curve for therapeutic proton beams”. In: *Medical physics* 24.12 (1997), pp. 2024–2033.

[7] Jean Bourhis et al. “Clinical translation of FLASH radiotherapy: Why and how?” In: *Radiotherapy and oncology* 139 (2019), pp. 11–17.

[8] Jean Bourhis et al. “Treatment of a first patient with FLASH-radiotherapy”. In: *Radiotherapy and oncology* 139 (2019), pp. 18–22.

[9] Sebastiaan Breedveld et al. “iCycle: Integrated, multicriterial beam angle, and profile optimization for generation of coplanar and noncoplanar IMRT plans”. In: *Medical physics* 39.2 (2012), pp. 951–963.

[10] Tiberiu Burlacu, Danny Lathouwers, and Zoltán Perkó. “A deterministic adjoint-based semi-analytical algorithm for fast response change computations in proton therapy”. In: *Journal of Computational and Theoretical Transport* 52.1 (2023), pp. 1–41.

[11] Benjamin Clasie et al. “Golden beam data for proton pencil-beam scanning”. In: *Physics in Medicine & Biology* 57.5 (2012), p. 1147.

[12] Emily C Daugherty et al. “FLASH radiotherapy for the treatment of symptomatic bone metastases (FAST-01): protocol for the first prospective feasibility study”. In: *JMIR Research Protocols* 12.1 (2023), e41812.

[13] Simon Deycmar et al. “The relative biological effectiveness of proton irradiation in dependence of DNA damage repair”. In: *The British journal of radiology* 93.1107 (2020), p. 20190494.

[14] M. Diepeveen. *Biological computational model for FLASH proton therapy*. [Master thesis, Technical University of Delft], 2022.

[15] Edward R Epp, Herbert Weiss, and Ann Santomasso. “The oxygen effect in bacterial cells irradiated with high-intensity pulsed electrons”. In: *Radiation research* 34.2 (1968), pp. 320–325.

[16] Vincent Favaudon et al. “Ultrahigh dose-rate FLASH irradiation increases the differential response between normal and tumor tissue in mice”. In: *Science translational medicine* 6.245 (2014), 245ra93–245ra93.

[17] Michael M Folkerts et al. “A framework for defining FLASH dose rate for pencil beam scanning”. In: *Medical physics* 47.12 (2020), pp. 6396–6404.

[18] Benedick A Fraass. “The development of conformal radiation therapy”. In: *Medical physics* 22.11 (1995), pp. 1911–1921.

[19] David Robert Grimes. “Estimation of the oxygen enhancement ratio for charged particle radiation”. In: *Physics in Medicine & Biology* 65.15 (2020), 15NT01.

[20] David Robert Grimes and Mike Partridge. “A mechanistic investigation of the oxygen fixation hypothesis and oxygen enhancement ratio”. In: *Biomedical physics & engineering express* 1.4 (2015), p. 045209.

[21] David Robert Grimes et al. “A method for estimating the oxygen consumption rate in multicellular tumour spheroids”. In: *Journal of The Royal Society Interface* 11.92 (2014), p. 20131124.

[22] Intensity Modulated Radiation Therapy Collaborative Working Group et al. “Intensity-modulated radiotherapy: current status and issues of interest”. In: *International Journal of Radiation Oncology* Biology* Physics* 51.4 (2001), pp. 880–914.

[23] Steven Habraken et al. “Trade-off in healthy tissue sparing of FLASH and fractionation in stereotactic proton therapy of lung lesions with transmission beams”. In: *Radiotherapy and Oncology* 175 (2022), pp. 231–237.

[24] John H Hubbell and Stephen M Seltzer. *Tables of X-ray mass attenuation coefficients and mass energy-absorption coefficients 1 keV to 20 MeV for elements Z= 1 to 92 and 48 additional substances of dosimetric interest*. Tech. rep. National Inst. of Standards and Technology-PL, Gaithersburg, MD (US), 1995.

[25] Jonathan R Hughes and Jason L Parsons. “FLASH radiotherapy: current knowledge and future insights using proton-beam therapy”. In: *International journal of molecular sciences* 21.18 (2020), p. 6492.

[26] ICRU. “Recording and Reporting Photon Beam Therapy”. In: *ICRU report* 50 (1993).

[27] ICRU. “recording and reporting photon beam therapy (supplement to ICRU report 50)”. In: *ICRU report* 62 (1999).

[28] Simon Jolly et al. “Technical challenges for FLASH proton therapy”. In: *Physica Medica* 78 (2020), pp. 71–82.

[29] Eunsin Lee et al. “Ultrahigh dose rate pencil beam scanning proton dosimetry using ion chambers and a calorimeter in support of first in-human FLASH clinical trial”. In: *Medical Physics* 49.9 (2022), pp. 6171–6182.

[30] Daniel A Low et al. “A technique for the quantitative evaluation of dose distributions”. In: *Medical physics* 25.5 (1998), pp. 656–661.

[31] Patricia van Marlen et al. “Ultra-high dose rate transmission beam proton therapy for conventionally fractionated head and neck cancer: treatment planning and dose rate distributions”. In: *Cancers* 13.8 (2021), p. 1859.

[32] Anthony E Mascia et al. “Proton FLASH radiotherapy for the treatment of symptomatic bone metastases: The FAST-01 nonrandomized trial”. In: *JAMA oncology* 9.1 (2023), pp. 62–69.

[33] Radhe Mohan and David Grosshans. “Proton therapy–present and future”. In: *Advanced drug delivery reviews* 109 (2017), pp. 26–44.

[34] Pierre Montay-Gruel et al. “Long-term neurocognitive benefits of FLASH radiotherapy driven by reduced reactive oxygen species”. In: *Proceedings of the National Academy of Sciences* 116.22 (2019), pp. 10943–10951.

[35] Frank WJ Olver. *NIST handbook of mathematical functions hardback and CD-ROM*. Cambridge university press, 2010.

[36] Karl Otto. “Volumetric modulated arc therapy: IMRT in a single gantry arc”. In: *Medical physics* 35.1 (2008), pp. 310–317.

[37] Harald Paganetti. *Proton therapy physics*. CRC press, 2018.

[38] Eros Pedroni et al. “The 200-MeV proton therapy project at the Paul Scherrer Institute: Conceptual design and practical realization”. In: *Medical physics* 22.1 (1995), pp. 37–53.

[39] Kristoffer Petersson et al. “A quantitative analysis of the role of oxygen tension in FLASH radiation therapy”. In: *International Journal of Radiation Oncology* Biology* Physics* 107.3 (2020), pp. 539–547.

[40] R.J. Santo. *Novel optimization strategies for clinical FLASH proton therapy*. [Master thesis, Technical University of Lisbon], 2021.

[41] Rodrigo José Santo et al. “Pencil-beam delivery pattern optimization increases dose rate for stereotactic FLASH proton therapy”. In: *International Journal of Radiation Oncology* Biology* Physics* 115.3 (2023), pp. 759–767.

[42] Emil Schüler et al. “Ultra-high dose rate electron beams and the FLASH effect: From preclinical evidence to a new radiotherapy paradigm”. In: *Medical physics* 49.3 (2022), pp. 2082–2095.

[43] Emil Schüler et al. “Very high-energy electron (VHEE) beams in radiation therapy; Treatment plan comparison between VHEE, VMAT, and PPBS”. In: *Medical physics* 44.6 (2017), pp. 2544–2555.

[44] Pascal Storchi and Evert Woudstra. “Calculation of the absorbed dose distribution due to irregularly shaped photon beams using pencil beam kernels derived from basic beam data”. In: *Physics in Medicine & Biology* 41.4 (1996), p. 637.

[45] Steven Van De Water et al. “Towards FLASH proton therapy: the impact of treatment planning and machine characteristics on achievable dose rates”. In: *Acta oncologica* 58.10 (2019), pp. 1463–1469.

[46] Marie-Catherine Vozenin et al. “The advantage of FLASH radiotherapy confirmed in mini-pig and cat-cancer patients”. In: *Clinical Cancer Research* 25.1 (2019), pp. 35–42.

[47] S.M.A. Wendt. *Uncertainty analysis in proton therapy using adjoint theory*. [Master thesis, Technical University of Delft], 2020. URL: `https://repository.tudelft.nl/islandora/object/uuid%3A8a8c1489-bef8-4ca4-8b4e-a449a0aa4a9c?collection=education`.

[48] Christopher P Wild et al. “Cancer Prevention Europe”. In: *Molecular oncology* 13.3 (2019), pp. 528–534.

[49] Susu Yan et al. “Global democratisation of proton radiotherapy”. In: *The Lancet Oncology* 24.6 (2023), e245–e254.

[50] Guoliang Zhang, Wenchao Gao, and Hao Peng. “Design of static and dynamic ridge filters for FLASH–IMPT: A simulation study”. In: *Medical Physics* 49.8 (2022), pp. 5387–5399.

A

Backward Euler Derivation

The ODE is as follows:

$$\frac{dC(t)}{dt} = -\kappa\dot{D} + \frac{1}{\tau}\Big(C_{ss} - C(t)\Big) \tag{A.1}$$

In backward Euler form this becomes:

$$\frac{C(t+dt) - C(t)}{dt} = \frac{-\kappa\dot{D}C(t+dt)}{C(t)} + \frac{\Big(C_{ss} - C(t+dt)\Big)}{\tau} \tag{A.2}$$

Note that the depletion term has a factor C(t+dt)/C(t), which keeps the function differentiable and above zero. Rewriting the equation further to a C(t+dt) and C(t) part yields:

$$\frac{C(t+dt)}{dt} + \frac{C(t+dt)}{\tau} + \frac{\kappa\dot{D}C(t+dt)}{C(t)} = \frac{C_{ss}}{\tau} + \frac{C(t)}{dt} \tag{A.3}$$

$$C(t+dt)\Big(\frac{1}{dt} + \frac{1}{\tau} + \frac{\kappa\dot{D}}{C(t)}\Big) = \frac{C_{ss}}{\tau} + \frac{C(t)}{dt} \tag{A.4}$$

This can be written to the final form relating C(t+dt) with the previous solution to C(t):

$$C(t+dt) = \frac{\Big(\frac{C_{ss}}{\tau} + \frac{C(t)}{dt}\Big)}{\Big(\frac{1}{dt} + \frac{1}{\tau} + \frac{\kappa\dot{D}}{C(t)}\Big)} \tag{A.5}$$

B

Dimension Analysis of the ODE

The ODE is as follows:

$$\frac{dC(t)}{dt} = -\kappa\dot{D} + \frac{1}{\tau}\Big(C_{ss} - C(t)\Big) \tag{B.1}$$

Set $C'(t) = C(t)/C_{ss}$, $t' = t/\tau$ and substitute in ODE:

$$\frac{dC'}{dt'} = \frac{-\kappa\dot{D}\tau}{C_{ss}} + \frac{C_{ss}\tau}{\tau C_{ss}} - \frac{C'C_{ss}\tau}{\tau C_{ss}} \tag{B.2}$$

This is dimensionless:

$$\frac{[mol * m^{-3} * s]}{[s * mol * m^{-3}]} = \frac{[mol * m^{-3} * Gy^{-1} * Gy * s^{-1} * s]}{[mol * m^{-3}]} + \frac{[mol * m^{-3} * s]}{[s * mol * m^{-3}]} - \frac{[mol^2 * m^{-6} * s]}{[mol^2 * m^{-6} * s]} \tag{B.3}$$

From Eq. B.2 it is evident that the ODE depends only on one characteristic term.

$$\frac{dC'}{dt'} = \frac{-\kappa\dot{D}\tau}{C_{ss}} + 1 - 1 \tag{B.4}$$

The following dimensionless system can be set up:

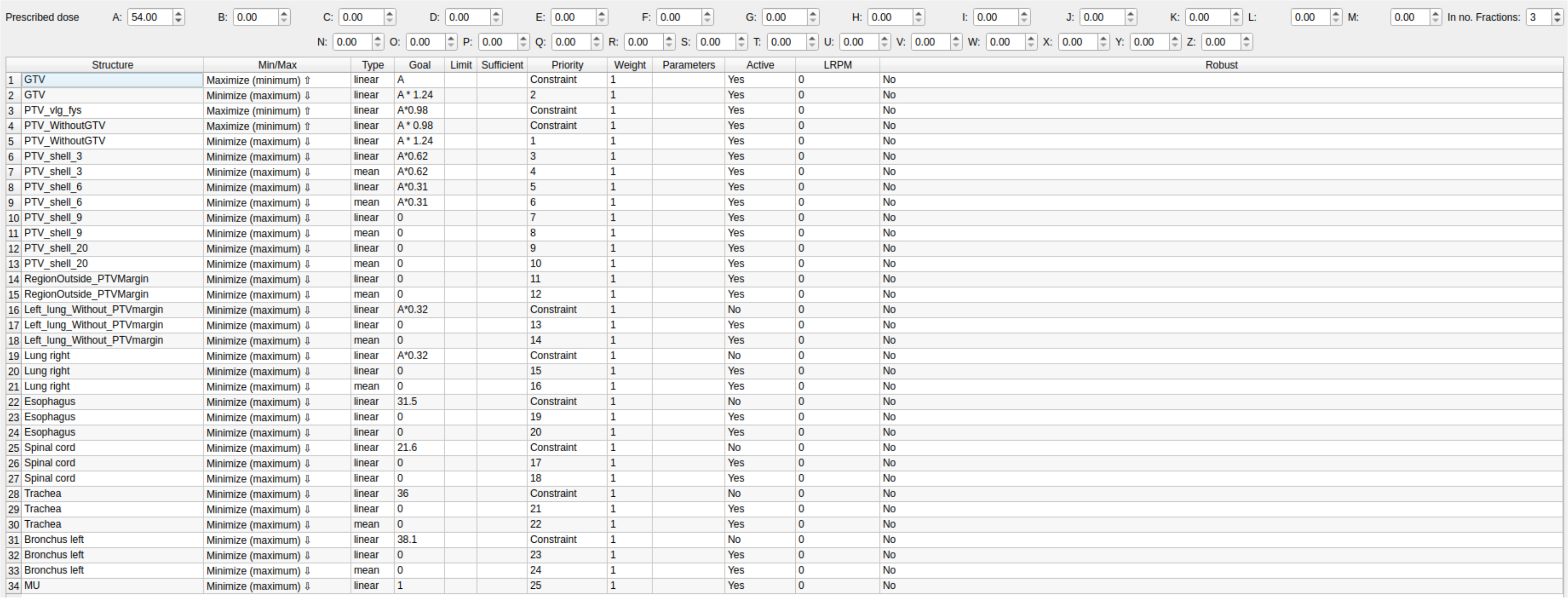
$$C' = C'(\frac{-\kappa\dot{D}\tau}{C_{ss}}) \tag{B.5}$$

The system depends on the varying dose rate over time. The dose rate is defined by the finite difference in the dose. The system parameters κ, τ and C_{ss} are all linked together and can be substituted by one global system parameter: λ = $\frac{\kappa\tau}{C_{ss}}$. This will result in the following system:

$$C' = C'(-\lambda\dot{D}) \tag{B.6}$$

C

Patient Specific iCycle Wish-list



Prescribed dose A: 54.00 B: 0.00 C: 0.00 D: 0.00 E: 0.00 F: 0.00 G: 0.00 H: 0.00 I: 0.00 J: 0.00 K: 0.00 L: 0.00 M: 0.00 In no. Fractions: 3

N: 0.00 O: 0.00 P: 0.00 Q: 0.00 R: 0.00 S: 0.00 T: 0.00 U: 0.00 V: 0.00 W: 0.00 X: 0.00 Y: 0.00 Z: 0.00

	Structure	Min/Max	Type	Goal	Limit	Sufficient	Priority	Weight	Parameters	Active	LRPM	Robust
1	GTV	Maximize (minimum) ⇧	linear	A			Constraint	1		Yes	0	No
2	GTV	Minimize (maximum) ⇩	linear	A * 1.24			2	1		Yes	0	No
3	PTV_vlg_fys	Maximize (minimum) ⇧	linear	A*0.98			Constraint	1		Yes	0	No
4	PTV_WithoutGTV	Maximize (minimum) ⇧	linear	A * 0.98			Constraint	1		Yes	0	No
5	PTV_WithoutGTV	Minimize (maximum) ⇩	linear	A * 1.24			1	1		Yes	0	No
6	PTV_shell_3	Minimize (maximum) ⇩	linear	A*0.62			3	1		Yes	0	No
7	PTV_shell_3	Minimize (maximum) ⇩	mean	A*0.62			4	1		Yes	0	No
8	PTV_shell_6	Minimize (maximum) ⇩	linear	A*0.31			5	1		Yes	0	No
9	PTV_shell_6	Minimize (maximum) ⇩	mean	A*0.31			6	1		Yes	0	No
10	PTV_shell_9	Minimize (maximum) ⇩	linear	0			7	1		Yes	0	No
11	PTV_shell_9	Minimize (maximum) ⇩	mean	0			8	1		Yes	0	No
12	PTV_shell_20	Minimize (maximum) ⇩	linear	0			9	1		Yes	0	No
13	PTV_shell_20	Minimize (maximum) ⇩	mean	0			10	1		Yes	0	No
14	RegionOutside_PTVMargin	Minimize (maximum) ⇩	linear	0			11	1		Yes	0	No
15	RegionOutside_PTVMargin	Minimize (maximum) ⇩	mean	0			12	1		Yes	0	No
16	Left_lung_Without_PTVmargin	Minimize (maximum) ⇩	linear	A*0.32			Constraint	1		No	0	No
17	Left_lung_Without_PTVmargin	Minimize (maximum) ⇩	linear	0			13	1		Yes	0	No
18	Left_lung_Without_PTVmargin	Minimize (maximum) ⇩	mean	0			14	1		Yes	0	No
19	Lung right	Minimize (maximum) ⇩	linear	A*0.32			Constraint	1		No	0	No
20	Lung right	Minimize (maximum) ⇩	linear	0			15	1		Yes	0	No
21	Lung right	Minimize (maximum) ⇩	mean	0			16	1		Yes	0	No
22	Esophagus	Minimize (maximum) ⇩	linear	31.5			Constraint	1		No	0	No
23	Esophagus	Minimize (maximum) ⇩	linear	0			19	1		Yes	0	No
24	Esophagus	Minimize (maximum) ⇩	linear	0			20	1		Yes	0	No
25	Spinal cord	Minimize (maximum) ⇩	linear	21.6			Constraint	1		No	0	No
26	Spinal cord	Minimize (maximum) ⇩	linear	0			17	1		Yes	0	No
27	Spinal cord	Minimize (maximum) ⇩	linear	0			18	1		Yes	0	No
28	Trachea	Minimize (maximum) ⇩	linear	36			Constraint	1		No	0	No
29	Trachea	Minimize (maximum) ⇩	linear	0			21	1		Yes	0	No
30	Trachea	Minimize (maximum) ⇩	mean	0			22	1		Yes	0	No
31	Bronchus left	Minimize (maximum) ⇩	linear	38.1			Constraint	1		No	0	No
32	Bronchus left	Minimize (maximum) ⇩	linear	0			23	1		Yes	0	No
33	Bronchus left	Minimize (maximum) ⇩	mean	0			24	1		Yes	0	No
34	MU	Minimize (maximum) ⇩	linear	1			25	1		Yes	0	No

Figure C.1: The iCycle transmission FLASH proton therapy specific wish-list for a lung patient with a relatively small CTV.

D

Literature Study

BM51010: BME Literature Research

Fast and accurate dose distribution calculations: Applications & Possibilities

by

Noor Odin Kruithof

Student Number: 4817168

Supervisors:	dr. ir. D. Lathouwers (TUD), dr. S. Habraken (LUMC), K. Spruijt (HPTC)
Project Duration:	December, 2022 - June, 2023
Faculty:	Faculty of Applied Sciences, Delft
Group:	Medical Physics and Technology

TU Delft

Abstract

For perfect patient treatment, accurate dose calculations are required. With full-scale MC simulation, the most precise dose matrix is computed. Examples of full-scale MC algorithms are FLUKA and GEANT4. Unfortunately, full MC is computationally very expensive and hence is not feasible in the clinic. In this work, first, the development of clinical dose calculations is reviewed. Throughout the years, the available computational resources grew, allowing for better dose calculations. Eleven quite recently developed dose engines are described thoroughly. Accuracy is coupled to the gamma pass rate and dose difference metrics. In these studies, full-scale MC simulated dose matrices and clinical treatment planning system dose matrices were used as benchmarks. The speed of the models is characterised by the total calculation time for a clinical treatment plan. No clear correlation is found between accuracy and computation speed. By considering the underlying structure of these engines, they were grouped into four classes: deep learning, GPU programming, phase-space and track-repeating. These classes display various advantages and disadvantages. Deep learning algorithms often require complex input types and must be trained properly. GPU programming is complex and evaluation of CPU against GPU performance is needed. Phase-space models consist of some energy approximations, drastically decreasing accuracy in low-mass-density regions. Track-repeating algorithms do not work well for in-homogeneous volumes. The discussed algorithms are suited for a big variety of clinical applications and create possibilities for further improvement of patient care. Four applications are described with possible algorithms aiding in the development. These applications were beam-angle optimisation, robust treatment planning, model-based plan comparison and online adaptive radiotherapy. Improvement of robust treatment planning and model-based plan comparison is within easy reach. Beam-angle optimisation and online adaptive radiotherapy are more complex procedures and require more effort. In general, more research has to be conducted before any of the models can be implemented clinically.

Contents

1

Introduction

Radiotherapy is an important modality in cancer treatment together with systemic therapy and surgery. These modalities are often combined for optimal patient care. The aim of radiotherapy is to gain regional control of the disease and to limit side effects with irradiation of the cancerous tissue with photons (e.g. X-rays) or charged particles (e.g. protons). The energy deposited with the irradiation will damage both tumorous cells as well as normal tissue (healthy) cells. The main goal of radiotherapy for each patient is to maximise tumour control while minimising possible side effects due to normal tissue damage. Therefore, the calculation of energy deposition (dose) in patients demands high accuracy. The most accurate way to compute the dose is by fully tracking the energy release for millions of charged particles or photons along its path. This is known as a full-scale Monte Carlo (MC) simulation. The output of MC simulations is a probabilistic description that maps charged particle or photon fluence to energy deposition per voxel. Yet, full-scale MC simulations are not generally used in standard clinical practice as the computation is too expensive. [3, 5] There are accelerated MC algorithms used in the clinic, but the accuracy is still not as good as full-scale MC-modelled energy deposition. For that reason, a vast amount of other and faster ways to calculate the energy deposition with full-scale MC accuracy are continuously being developed. [11]

1.1. Background

Historically speaking, several types of dose computation algorithms can be distinguished. In the starting phase of clinical conventional radiotherapy, doses were calculated correction-based on a homogeneous patient. The idea was to simply map the energy released in a perfect homogeneous water phantom to in-homogeneous patients. Many rough approximations were included in this procedure, which made it inaccurate compared to current standard clinical practice. Especially, in patients irradiated at very in-homogeneous treatment sites (e.g. lungs), this procedure performs poorly. This fueled the urge to make the switch to model-based dose calculation algorithms, where the energy deposition per photon is modelled along its path. These model-based algorithms consist of the pencil beam convolution (PBC), the analytical an-isotropic algorithm (AAA) and the collapsed cone convolution (CCC). These model-based algorithms are pretty accurate and are still being used clinically today in various treatment planning systems (TPS). Many constraints developed for treatment planning with photons and charged particles are based on these algorithms. Switching to more accurate dose computation algorithms will inevitably affect these constraints. Currently, deterministic-analytical dose calculations based on the linear Boltzmann transport equations (LBTE) are becoming popular as they display good dose computation accuracy. [3] In Fig. 1.1 the historical overview of these types of dose calculation algorithms is given with their relative accuracy.

Each of the model-based approaches tries to establish a correct relation between nuclear interactions and dose deposition. For photon irradiation, it is possible to formulate the dose deposition with the help of a response function, known as a dose spread kernel. In Eq. 1.1 a simple description for dose deposition with the inclusion of such a kernel is given. The dose is given by the convolution between the TERMA (T) and the dose spread kernel (h) for a given volume (V) and energy spectrum (E). The

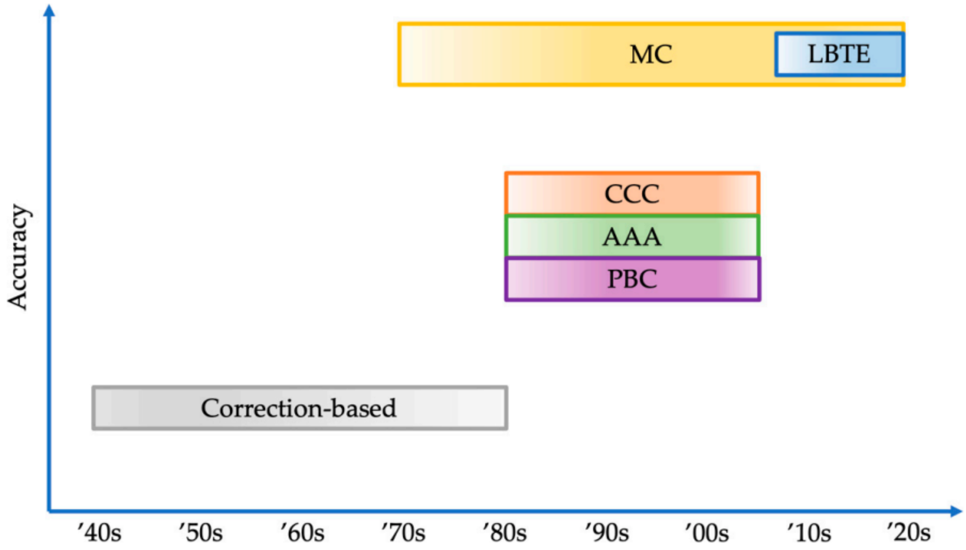


Figure 1.1: An overview of the various types of dose rate algorithms with their relative accuracy over time. This figure is adapted from [3].

TERMA accounts for the energy release at the primary absorption point of the incoming photon. The dose spread function accounts for energy deposition due to other interactions, e.g. secondary particles or multi-scattering. The TERMA can be calculated with a fluence model of the treatment machine and an attenuation coefficient in the patient. For photons, the fluence model is based on the settings of the linear accelerator. CT Hounsfield units (HU) are mapped to the attenuation coefficient with predefined conversion tables. With this attenuation factor, the absorption of photons, and thus deposition of energy, in the patient is characterised. The dose spread kernel is also predefined and is a probabilistic expression of energy deposition because of nuclear interactions. This kernel is often determined with MC simulation to increase the accuracy of the model-based algorithms. [3]

$$D(r) = \int_E \int_V T(s,E)h(r-s,E)dsdE \tag{1.1}$$

Model-based algorithms can also be used to calculate energy deposition for charged particles (e.g. protons), but are clinically not used frequently. The reason behind this is that these types of calculations are even more sensitive to inhomogeneities compared to photon dose calculations and that particle semi-MC computations can be done quickly enough to be implemented into the clinical routine. [3] The dose calculation based on charged particles is based on the stopping power of particles, instead of the attenuation coefficient for photons. This stopping power is determined for a patient by mapping CT HU values to stopping powers with conversion tables. The stopping power equals energy loss per unit path length rather than a loss of photons. In combination with an input fluence model of the number of charged particles, the total dose can be calculated.

One of the approximations for the dose spread function for photons is the pencil beam kernel used in PBC. For this kernel, the energy released over the complete path of a photon is mapped. Lateral energy deposition components are determined by lateral integration at each point in depth. In essence, this kernel represents an average of the energy deposition of infinite photons over depth. The kernel is often created with the help of MC simulation. When the kernel is computed, the dose deposition can be calculated relatively quickly. Unfortunately, the accuracy is low compared to full-scale MC point spread kernel-based dose calculations. Lateral heterogeneities are not perfectly taken into account when computing the energy deposition at each depth position. [3] This heterogeneity problem is partly tackled with AAA. Here the pencil beams are rewritten in both lateral and depth components. The lateral energy spreading from the primary interaction point is written in spherical coordinates and subdivided into small angular $(d\theta)$ bins. It has been shown that this is a better functioning implementation for modelling the dose deposition. [12, 17]

Another well-known dose computation algorithm for photons is the CCC algorithm. At the main interaction point of a photon, the surrounding volume is divided into equal cones. The spreading energy deposition is mapped on the central line of these cones. Scaling of the dose deposition is done according to the heterogeneity inside the cones. The collapsed cones nicely resemble a dose point spread function instead of a pencil beam kernel and outperform both PBC and AAA in dose calculation accuracy. [3, 12]

Still, the three types of approximations of the dose spread kernel mentioned above display a lack of accuracy compared to full-scale MC simulations, such as GEANT4 and FLUKA. Newly developed dose algorithms are often benchmarked against these engines. An emerging field is the LBTE dose computation. With the help of numerical and analytical solutions to the transport equations, a complete phase-space of particles is formed. This phase-space is applicable to both photons and charged particles like protons. From the phase-space, the dose deposition under certain conditions in a volume can be computed. These Boltzmann-based dose solvers show similar results to full-scale MC dose determination but are computationally less expensive. [2]

In the past accuracy was often sacrificed, for instance with the three convolutional approaches, for total computation time. Currently, the goal is to shorten computation time while maintaining MC dose quantification accuracy. The everlasting development of faster computer chips helps to achieve this goal. [11]

1.2. Purpose

This literature study will focus on fast and accurate dose computation algorithms. However, instead of emphasizing the methodology behind these algorithms, this study will line out the possibilities and applications. The algorithms will be drawn out of their mathematical framework and be looked at from a more clinical perspective.

1.3. Overview

First, several methods to calculate the dose fast and accurately will be explained with results from recently published articles in chapter 2. Thereafter, in chapter 3, the classes of algorithms will be discussed together with created clinical possibilities and applications of the algorithms. In the last chapter, a conclusion is drawn.

1.4. Literature Search

The included articles in this literature study are found on Google Scholar. Articles about both photon and particle radiotherapy were taken into account. The starting search consisted of ”clinical application” together with ”algorithm” and ”fast dose”. This search returned a total of 173 papers. After scanning these, 29 of them were accessible and potentially useful. In the end, 19 of these were included in this literature study. Several references from these papers were also included, yielding a total of 25 papers.

2

Fast & Accurate dose calculation algorithms

Many efforts have been made to decrease the computation time of dose calculation algorithms concerning full-scale MC simulation. Within this active field of research, the developed algorithms can be divided into several sub-classes. From the literature search, one can distinguish algorithms based on deep learning, GPU programming, the phase-space and track-repeating. To quantify the performance of algorithms on 3D dose distributions, gamma pass rates are most frequently used. The algorithms are often benchmarked against full-scale MC simulated or treatment planning system (TPS) generated dose-deposition matrices. For 3D dose-deposition matrices, gamma analysis is the golden standard for comparison. However, for clinical implementation and validation of an algorithm, other metrics are also used. These consist of absolute dose differences, dose-volume-histogram comparison and dose distribution statistics.

The gamma index was originally introduced by Low et al. (1998) [10] The gamma index for a given simulated voxel position (r_s) compared with a reference position (r_c) is given by Eq. 2.1

$$\gamma(r_s) = \text{min}\{\Gamma(r_s, r_c)\}\forall\{r_c\} \tag{2.1}$$

In this equation, the $\Gamma(r_s, r_c)$ for a voxel is given by Eq. 2.2. The $\Gamma(r_s, r_c)$ comprises two parts. The first term relates to the distance-to-agreement and the second is the dose difference. In this equation, the r is the distance between the reference and the simulated voxel. The δ is the dose difference. Both terms are normalised with a user-selected margin in distance and dose difference (respectively Δd_M and ΔD_M). After the full gamma analysis of two dose-deposition matrices, every voxel has a value for $\gamma(r_s)$. If $\gamma(r_s)$ of a voxel is smaller than one, the two compared dose-deposition matrices lie within the selected margins. If $\gamma(r_s)$ is bigger than one, the reference and simulated matrices can not be matched within the margins. The gamma pass rate is defined as the fraction of the simulated dose matrix with a $\gamma(r_s)$ below one. A graphical representation of the gamma index is shown in Fig. 2.1. There are two types of gamma pass rates used, the local and global gamma pass rates. With a local gamma analysis, two different dose-deposition matrices are directly compared with each other and analysed according to the set margins. A global analysis varies slightly. Here, the dose-deposition matrices are compared to a single dose value, e.g. the prescribed dose. Additionally, a lower dose threshold is often used for the gamma analysis. With the threshold, low-dose areas, which might disturb the gamma analysis are excluded. In literature, both local and global gamma pass rates are used and are reported with distance margins of 1-mm, 2-mm and 3-mm together with dose margins of 1%, 2% and 3%. [1, 7, 13, 16, 21, 22, 23, 25] To compare the performance of different fast dose algorithms, ideally gamma pass rates with equal margins and equal reference doses are needed.

$$\Gamma(r_s, r_c) = \sqrt{\frac{r^2(r_s, r_c)}{\Delta d_M^2} + \frac{\delta^2(r_s, r_c)}{\Delta D_M^2}} \tag{2.2}$$

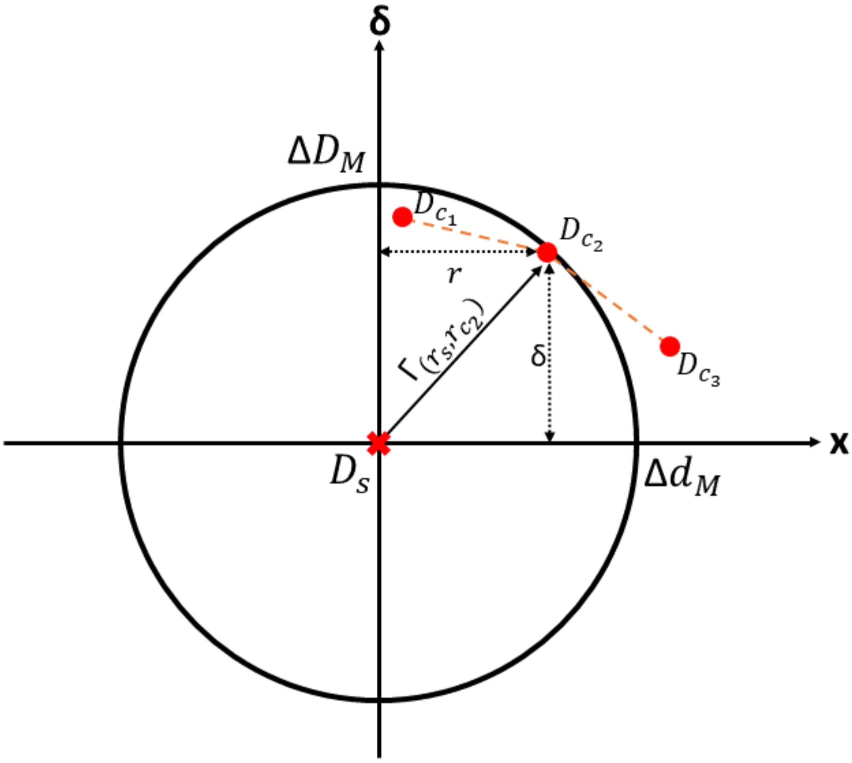


Figure 2.1: 2D representation of the gamma index between a simulated (D_s) and three reference dose points (D_{c_i}). Note that D_{c_1} is located within the margins in dose and distance $(\Delta D_M, \Delta d_M)$, D_{c_2} on the margin boundary and D_{c_3} outside the margin space. This implies that the $\Gamma < 1$ for point D_{c_1}, $\Gamma = 1$ for D_{c_2} and $\Gamma > 1$ for D_{c_3}. The minimal Γ value equals the value for γ. In this example, the gamma index is passed as at least one dose point is located within the margin space. This figure is adapted from [10].

In this chapter, the identified sub-classes will be explained together with examples from recently published articles. In chapter 3 the clinical relevance and possibilities of the algorithms will be discussed thoroughly.

2.1. Deep Learning

Deep Learning (DL) is a subdomain of artificial intelligence and machine learning which focuses on mimicking human logic with self-learning algorithms. These algorithms comprise complex networks which represent the human neural network architecture. The comparison with biological neurons is made as the convolutional networks should be capable of performing similar tasks. With the help of convolutional neural networks, the input data is rerouted to the desired output. These self-learning algorithms have to be trained such that high generalisability of the model can be achieved. In radiotherapy, this is of importance as the variability among patients is large. With the help of an activation and loss function, the most optimal weights of each individual 'neuron' in the network are found. After training the model, it should be able to solve a problem correctly and return the desired output with high accuracy. [18]

In the last years, DL progressively found its way into various aspects of radiotherapy. DL is also being used to compute fast and accurate dose distribution. Several approaches can be taken to simulate the dose with the help of DL. The DL model can either focus on converting physical quantities to the dose (e.g. fluence maps or CTs to dose), on the transformation of a dose from a water phantom to complex heterogeneous patients, or on mapping the dose from a coarse CT grid to a high-resolution grid without losing accuracy. The latter shows promising results for dose calculations in the penumbra region of proton beams. [19]

Fan et al. (2020) created a DL-based algorithm that can map an input photon fluence map to a dose distribution. [4] Within this algorithm, 2D fluence maps are first converted to their 3D equivalent with a ray traversal algorithm. Thereafter, the 3D fluence map together with an input CT is mapped to a dose matrix. The training of the network is done with approximately 2000 samples of individual beam doses calculated with the clinical Pinnacle TPS. The results acquired with this DL algorithm are benchmarked against dose distributions calculated with the CCC algorithm available in the Pinnacle TPS. Fan et al. (2020) report calculation times of several seconds. [4] For a total of 47 patients, the relative deviation between the dose calculated with the TPS and the DL-based dose engine is $0.17\% \pm 2.28\%$. High dose deviations are found in regions with relatively high density gradients, e.g. bone-to-air. The main drawback of this presented algorithm is the fact that it is benchmarked against a CCC algorithm instead of a full-scale MC simulation. On a positive note, Fan et al. (2020) point out that the algorithm is multi-functional since it depends on only a 2D fluence map and CT scan. The proposed algorithm can therefore calculate photon doses as well as particle doses. [4]

Guerreiro et al. (2021) set up a DL network for model-based comparison between photon and proton therapy for two types of abdominal cancers. [8] The DL network requires ten types (10 channels in the DL network) of patient-specific inputs to predict a 3D dose distribution. Among these, important inputs are the planning-CT, the internal target volume (ITV), several OAR and vertebra contours. The predicted photon and proton doses are compared with respectively clinical VMAT and PBS doses optimised within RayStation. Both plans were created by one experienced treatment planner. After full training of the network, a 3D dose distribution is simulated in approximately 0.5s. Guerreiro et al. (2021) report relative dose differences for the ITV and several OARs. [8] For the ITV, the average relative deviation in the $D_{98\%}$ is $-0.4\% \pm 0.9\%$ for proton PBS and $-0.8\% \pm 0.5\%$ for photon VMAT. For OARs, the difference in average dose was a bit higher and equals several percentage points for both PBS and VMAT. The drawbacks of this DL-based algorithm are that predicted doses are benchmarked against not 100% accurate TPS doses and that it requires patient-specific input structures. Forming these inputs can be time-consuming. Contrarily, the algorithm does return reasonable deviations and therefore shows potential for speeding up the model-based comparison between photon and proton treatment plans.

Zhang et al. (2022) constructed a convolutional neural network that is capable of simulating 3D photon dose distributions from the TERMA and electron density (EC) distributions. [22] The TERMA is calculated with a simple ray-tracing algorithm, which considers the linear attenuation coefficient and the EC distribution. [24] The TERMA calculation takes approximately 50-ms. The simulated dose-deposition matrices are compared with full-scale MC simulations. The gamma pass rates for both a 7 fields and 9 fields IMRT plan are calculated. The gamma pass rates for the 7 fields plan are 97.06% ($\gamma_{1\%/2mm}$), 99.10% ($\gamma_{2\%/2mm}$) and 99.52% ($\gamma_{3\%/2mm}$). The gamma pass rates for the 9 fields plan are 98.16% ($\gamma_{1\%/2mm}$), 99.61% ($\gamma_{2\%/2mm}$) and 99.89% ($\gamma_{3\%/2mm}$). The reported calculation time is in the sub-second region for full IMRT plans. Zhang et al. (2022) also prove that the described DL algorithm shows great generalisability. [22] This was achieved by using mixup in the training phase of the neural network. With this procedure, new TERMA and dose data distributions are formed by taking a linear combination between two randomly selected training samples.

This DL-based photon dose engine is further improved by Zhu et al. (2023) [25] Input TERMA distributions were generalised by including a pre-computed MC kernel. The dose engine described by Zhang et al. (2022) [22], functioned well for 6 MV spectrum photon beams, but lost accuracy at higher energies due to scattering. With the described generalisation method by Zhu et al. (2023) of the TERMA, accurate dose distributions can be computed for photon beams up to 24 MV. [25] Average gamma pass rates of 94.08% ($\gamma_{2\%/2mm}$) and 98.46% ($\gamma_{3\%/3mm}$) were found for 24 MV IMRT field tests. The reported computation time is again in the sub-second region.

Pastor-Serrano and Perkó (2022) constructed DoTa, a convolutional DL-based particle dose engine. [13] The network is capable of mapping an input 3D CT to an energy-dependent 3D dose distribution.

The input 3D CT is sub-sampled to small-volume compartments in the direction of an input beam. For each compartment, the energy transfer of particles is predicted. The final 3D dose is a dynamic combination of all compartment doses. DoTa dose distributions are compared with 2-field 9 full-scale MC treatment plans (3x Lung plans, 3x H&N plans and 3x Prostate plans). The $\gamma_{1\%/1mm}$ pass rates range from 94.71% to 96.93%, the $\gamma_{2\%/2mm}$ pass rates between 99.39% and 99.82% and the $\gamma_{3\%/3mm}$ pass rates lie between 99.81% and 99.99%. DoTa outperforms all the evaluated pencil beam algorithms in their own work. For a 2-field plan consisting of many individual pencil beams, the CPU runtime is around 10 minutes. When running on a GPU, the full dose distribution is generated in less than 15 seconds. A single pencil-beam dose can be computed in 5 milliseconds on average.

Apart from mapping a physical input to a dose matrix output, Zhang et al. (2022) build a DL network to achieve the opposite. [23] The described network is capable of mapping an input 3D dose distribution to a particle beam weight matrix. This inverse relationship can be of importance in the re-optimisation of treatment plans. The performance of the network was tested by applying pencil beam weight perturbations on a known clinical IMPT plan and analytically forming the resulting weight map. From the perturbed samples, the spot weight matrix is predicted with the network. This spot matrix is then converted back to a dose map with the dose-deposition equation given in Eq. 2.3

$$D_i = \sum_j K_{ij} w_j^2 \tag{2.3}$$

Here, D_i corresponds to the dose deposition in voxel i by all spots j. The dose-deposition matrix K represents the dose contribution per unit spot weight and w is the weight spot matrix. As K is fixed, the resulting dose distribution can directly be compared with the sampled perturbed distribution. For fluctuations up to 10% of the weight matrix, the global gamma pass rates were 97.6% ($\gamma_{1\%/1mm}$) and 99.9% ($\gamma_{3\%/3mm}$) for the vast majority of the compared samples. Spot matrices could be computed in approximately 1 second.

It is worthwhile to mention a paper related to DL-related dose engines by Peng et al. (2022) [14] The performance of a DL network is related to the given training data. Intuitively, low-quality training data inevitably results in a badly functioning DL network. Peng et al. (2022) proved that the performance of a DL model strongly correlates with the homogeneity of training data. [14] In general, the homogeneity of training data should be based on the clinical purpose of the DL model. [18]

2.2. GPU Programming

In the past twenty years, GPUs became more important in the field of medical physics. A GPU is in essence a chip consisting of up to thousands of small processing cores, whereas a CPU consists of just several cores. Because a GPU has so many cores, parallelisation of computation tasks is possible. This yields tremendously lower execution times compared to CPU times, with similar resources. [9] Currently, growing energy consumption outweighs further improvement of CPUs. Developers focus on GPU development as the potential is higher. In Fig. 2.2 the number of articles related to GPUs in two leading medical physics research journals is depicted. Since the beginning of this century, the number of articles exponentially increased.

A GPU is suited for a big variety of radiotherapy-related tasks. [9] The calculation time of model-based dose algorithms, as well as MC-based algorithms, can be further reduced by executing them on GPU. Optimisation of treatment plans is also highly efficient on GPUs due to their parallelisation ability. Additionally, a GPU is also suited for voxel-wise comparison of treatment plans, for instance in a gamma analysis. [9] Recent published other dose algorithms are most often already GPU compatible. [4, 13, 21, 22, 23] Below articles are described which are closely related to GPU execution.

Giordanengo et al. (2019) parallelised a clinically used ion beam kernel-based dose engine. [7] The kernel is formed by constructing look-up tables with MC simulations for the energy deposition along the path of various particle energies. The energy deposition is based on the phase-space information of an ion beam. The energy deposition per voxel can be computed with a linear combination of the quantities

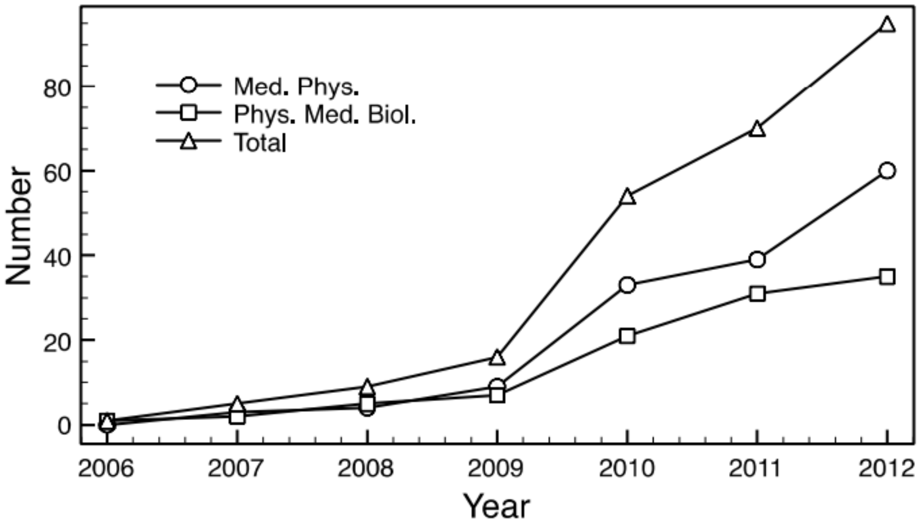


Figure 2.2: The number of articles related to GPUs per year. The two articles are Medical Physics and Physics in Medicine & Biology. This figure is adapted from [9].

in the look-up tables (superposition). [15] The parallelisation of this algorithm allowed for execution on a GPU and faster dose computations. Giordanengo et al. (2019) report dose calculation times of 0.3 seconds and dose comparison times of 0.4 seconds. [7] The difference between the CPU and GPU calculated doses is negligible. These fast computations allow for online monitoring of the dose deposition in a patient after each given pencil beam. The kernel-based algorithm is fast but lacks accuracy. When benchmarked against full-scale MC simulation, the 3%,3mm gamma pass rate equals 95.5% in the PTV for a clinical neuro case. [15]

Schiavi et al. (2017) set up a simplified proton dose version of full-scale MC simulations which is executable on a GPU. [16] The most important energy deposition influencing terms were fully taken into account. Approximations were made in the energy fluctuations and multiple coulomb scattering of protons. The algorithm has recently been commissioned in two proton therapy centres (CCB in Krakow and EMORY in Atlanta). [6] In this commissioning study they benchmarked the semi-analytic MC algorithm against dose QA measurements in a heterogeneous phantom with a 2D dose detector (MatriXX). Additionally, they also compared the QA data with a TPS (RayStation) proton dose engine. For all measurements, the local $\gamma_{2\%/2mm}$ was above 99% for the semi-MC engine, whereas the RayStation dose engine only yielded a 93.2% pass rate (8.4% σ). The total computation time for a clinical plan is approximately 2 minutes and 30 seconds, significantly longer than previously described dose algorithms.

2.3. Phase-Space

Algorithms based on phase-space information provide dose calculation accuracy similar to full-scale MC simulation while reducing computation time. Such phase-space models can be built for both photons and charged particles. A phase-space model is a complete set of functions describing the behaviour of photons or particles passing through a medium. This complete set consists of information about the direction, energy, position and number of photons or particles at a given point in space. From the complete phase-space model, the energy deposition can be found. This energy deposition is location specific and depends on other factors like energy and fluence. [1] The accuracy of the computed dose depends on both the complexity of the phase-space model. The highest accuracy is achieved with phase-space models generated with full-scale MC simulations. The discussed article by Giordanengo et al. (2019) included such a phase-space constructing approach [7]. This again computationally very expensive and clinically not viable. Another quicker option for photons is to fill a phase-space based on the simple geometrical properties of the linear accelerator. For particles, the Boltzmann transport equation can be simplified and solved analytically to fill the phase-space. Below are two phase-space-related articles described for fast and accurate dose computations.

Bedford et al. (2022) created a simple geometry-based phase-space model for photon radiotherapy. [1] The phase-space is formed by assuming two Gaussian source inputs, one at the treatment head and one at the flattening filter (virtual source). Furthermore, it only includes an energy-dependent fluence. The phase-space model depends on the positioning of the two Gaussian sources. The phase-space model was integrated into a clinical semi-MC dose calculation algorithm and for two IMRT patients (prostate and lung) the calculated dose was compared with a TPS (RayStation) dose matrix. The reported global gamma pass rates (2%/2mm) are 79.2% and 66.2%. With the phase-space approach, a dose is computed in three minutes. From the gamma pass rates, one can deduce that the two compared dose-deposition matrices are not similar. A possible explanation is the simplicity of the model. Gaussian source inputs work reasonably well but also contribute to the loss of accuracy.

Burlacu et al. (2023) created a semi-analytic model based on the linear Boltzmann transport equation. [2] The transport equation is approximated by a combination of two proton flux ordinary differential equations. With this approximation, a full proton phase-space is established. From there, the energy deposition in a volume can be calculated. The performance of the semi-analytic engine has been compared to FLUKA full-scale MC simulations. Dose differences or gamma pass rates are not reported. However, visually the energy deposition of both methods is very similar. The computation time is quicker than full-scale MC and lies around 0.1 seconds per pencil beam. From this, one can deduce that a typical treatment plan consisting of thousands of pencil beams will take several minutes. The main advantage of this semi-analytic model is the fact that a direct adjoint system can be set up. In other words, energy deposition changes due to density shifts can be calculated based on an initial solution. A function response can be formed from one of the flux terms and the parameters describing the system. For small Hounsfield unit perturbations (± 40 HU), the maximal percentage error between the adjoint and re-calculated solution is 0.0036%.

2.4. Track-Repeating

Another way to create accurate and fast proton dose-deposition matrices is with so-called track-repeating algorithms. With this approach, one first uses MC simulations to create the energy deposition of proton tracks in a homogeneous water phantom. From there, based on system parameters, the energy deposition is re-scaled to a given volume of interest. [20] In my literature search, I came across one track-repeating dose algorithm.

Yepes et al. (2008) formed a proton track database with MC-simulation (GEANT4) and used a re-scaling function based on material and mass density. A good agreement between full-scale MC and the described dose engine was achieved by fitting the re-scaling function. [20] The fitting was done with the minimisation of a chi-squared function between the full-scale MC dose and engine dose. In a follow-up study, Yepes et al. (2016) validated the fast dose engine on clinical patient data. [21] For 23 patients with varying treatment sites, the performance of the algorithm was benchmarked against full-scale MC simulation (GEANT4). Inside the target, the local gamma pass rates are 71.66% ($\gamma_{1\%/1mm}$), 99.78% ($\gamma_{2\%/2mm}$) and 100% ($\gamma_{3\%/3mm}$) on average. Noticeable is the difference between the 1%/1mm and 2%/2mm pass rates. This indicates that the accuracy is in the order of magnitude of 1 % and 1 mm. Intuitively, the engine performs best in areas with a similar mass density to water. The computation time of a full patient dose takes over a minute.

3

Discussion

In the previous chapter, the computation speed and accuracy were described for all the dose computation algorithms found in my literature search. A direct comparison of the computation speed and accuracy between described algorithms is not possible. The total computation time correlates with the available hardware and the gamma pass rate or absolute dose difference depends on the benchmarked dose-deposition matrices. Furthermore, the performance of an algorithm also highly depends on the treatment plan complexity and on the structure of the algorithm itself. All these factors vary between the described algorithms. In the previous chapter, the described algorithms were also subdivided into several classes. The stated computation times together with the accuracy per class provide an overall image of the performance of various dose computation options.

In this chapter, the identified classes of algorithms are generalised in section 3.1. Thereafter, the created clinical possibilities and applications of the algorithms are lined out in section 3.2. In Table 3.1 the identified classes are listed with an estimation of their computation speed and accuracy.

Table 3.1: An overview of the identified types of dose algorithms together with an overall speed and accuracy estimation. (*) Only two papers describe the gamma pass rate and margins vary; the accuracy is a rough estimation in this case.

	Order of Magnitude: Calculation Times (s)	Computation Accuracy	Number of Studies
Deep Learning	[0.1, 1.0]	$\gamma_{2\%/2mm}$ $\approx 99.0\%$	6
GPU Programming	1.0	$\gamma_{2\%/2mm}$ $\approx 95.0\%$	3
Phase-Space	100.0	$\gamma_{2\%/2mm}$ $\approx 80.0\%$	3*
Track-Repeating	100.0	$\gamma_{2\%/2mm}$ $\approx 99.0\%$	1

3.1. Class Descriptions

3.1.1. Deep Learning

From the literature study, it is clear that DL-based algorithms can compute a dose the quickest with exquisite accuracy. But, DL-based algorithms only function under perfect conditions. The networks have to be trained properly and should be given the right error-free input. The pitfalls of DL-based algorithms are mentioned below.

First of all, the type of input varies for every DL-based algorithm. In other words, the networks only will function well with the correct input type. In my literature search, the DL-based algorithms all required various types of inputs. These consisted of fluence maps, CT scans, OAR contours, TERMA distributions, electron density distributions and dose distributions. Therefore, having all the right inputs for a DL-based algorithm can be problematic. In a clinical setting, TERMA distributions and electron density distributions are not always available. Ideally, a DL-based algorithm only requires one type of input besides a CT scan. For three algorithms, this holds. [4, 13, 23] The other described algorithms required two [22, 25] or ten [8] types of inputs. Especially for the latter, the clinical importance is questionable. With an increasing number of input channels, the generalisability becomes worse. In other words, using a DL algorithm at various institutes for a broad spectrum of patients becomes more difficult with increasing input complexity.

Besides input complexity, the training phase of DL networks is a source of class-specific drawbacks as well. The training of a network can take hours up to weeks and requires a lot of training input data. [8, 22, 23, 24] As mentioned before, the homogeneity of the training data should reflect the application of the algorithm. With a high level of homogeneity, the network is problem-specific and will be able to solve a problem very accurately. Again, the generalisability of the network can be problematic in such cases. Ideally, the input training data should be 'error-free'. When training a network with a faulty training database, a bias will be built into the network. For dose calculations, accuracy is negatively affected by this bias. To account for this, bigger margins in treatment planning are necessary. Therefore, acquiring an 'error-free' training database is of great importance. However, this can be challenging as it requires many MC calculations or a large patient group.

In general, a DL-based algorithm is very fast (sub-second regime) and accurately tackles the problems faced. However, some caution is needed in clinical applications of the algorithms as they can require complex inputs and can be trained incorrectly or too specific. Before implementing a DL-based algorithm in a clinical routine it should be validated with care. [18]

3.1.2. GPU Programming

As explained before, developers mainly focus on improving GPUs instead of CPUs as more computational power can be achieved with GPUs. Making clinical computations GPU-compatible is another way to acquire fast and accurate results. Accuracy is not influenced by the GPU directly and is fully inherent to the structure of the algorithm itself. When an algorithm is functioning poorly, running it on a GPU will not help to improve accuracy. In general, time is an important factor in the clinic. Therefore, some previously described algorithms are GPU-compatible. [4, 13, 21, 22, 23] When utilising the GPU correctly, the computation time can be reduced by several orders of magnitude with GPU programming. [1, 13] This gain makes GPUs suited for real-time clinical applications. Future real-time clinical procedures will most likely in some way involve GPUs.

Writing an efficient parallelised algorithm can be challenging. Therefore, it is important to compare the execution of an algorithm on both a CPU and a GPU throughout the development of an algorithm. With constant optimisation of an algorithm, computation times can be reduced further. The required computation times are problem-specific, making the parallelisation of an algorithm a dynamic and not always necessary procedure.

In conclusion, GPU algorithms outperform their CPU counterparts regarding computation times when correctly implemented. The accuracy is not affected by parallel programming and fully depends on the complexity of the algorithm itself. The required accuracy and speed are problem-specific and thus vary quite a lot between the described algorithms. GPUs will most likely be exploited for real-time clinical calculations as computation speed is the most important factor here.

3.1.3. Phase-Space

A full description behaviour of photons or charged particles inside an area of interest is constructed with a phase-space model. Again, the accuracy and speed of the algorithm strongly depend on the problem one is facing. Therefore, the speed and accuracy metrics vary for the described phase-space algorithms. The calculation of full-scale phase-space is a quite costly process making it slower than DL-based and some GPU algorithms.

The deposited energy per unit weight, i.e. dose, along the path of either a photon or particle fully depends on all physical interactions taking place. With a full-scale MC simulation, these interactions are perfectly described. In phase-space models, probabilistic kernels are introduced in order to speed up the full simulation process. These kernels will inevitably decrease the accuracy compared to full-scale MC simulation. Especially, when dealing with large-density gradients (e.g. air to bone) or low-density regions (e.g. air in lungs), the accuracy of dose calculations is affected relatively severely. Therefore, it is advised to carefully interpret phase-space-based dose calculations for these types of patients. When the calculated deposited dose is heavily distorted or exceeds an organ tolerance dose level, recompute the dose with another, more suited, model.

3.1.4. Track-Repeating

With a track-repeating algorithm, the dose deposition inside a patient is coupled to full-scale MC energy deposition calculations in a water phantom. By correctly re-scaling particle tracks from water to in-homogeneous volumes, the dose inside a patient can be found. The computation times (minutes) and gamma pass rates ($\gamma_{1\%/1mm} : 71.66$) of the single article are not perfect. Additionally, constructing the required energy deposition database of particles in water takes minutes to hours, depending on available resources. [20, 21]

Track-repeating algorithms function perfectly for volumes consisting of equal mass density as water. However, patients are in-homogeneous and particle tracks almost always consist of multiple density gradients. Therefore, these types of algorithms display a lack of accuracy in dose calculation for patients. For treatment plans consisting of multiple beams, the accuracy is affected even more as relative track in-homogeneity increases. In my opinion, track-repeating algorithms are not suited for patients treated with multiple particle beams. A slight increase in accuracy is achieved compared to some commercial analytical dose algorithms but track-repeating algorithms will most likely not become standard practice in the future. There are more promising alternatives available.

3.2. Clinical Applications & Possibilities

The reported metrics in table 3.1 regarding speed and accuracy do not differ that much between classes. Selecting an algorithm is problem and application specific. In this section, several clinical applications are sketched together with possible algorithms aiding in these applications.

3.2.1. Beam-Angle Optimisation

Almost all treatment plans are formed automatically and inversely planned (apart from some brachytherapy treatments). In other words, planning is based on a set of constraints and objectives. With automated planning, the most optimal treatment plans (Pareto front) are found that meet the set requirements. For each patient group, there are ICRU guidelines which help to form the constraints and objectives. However, in most automated treatment planning systems the angle at which particle or photon beams enter the patient is fixed and is required as input for the system. Experienced treatment planners are needed to find beam angles which result in the most optimal treatment plans. Constantly optimising treatment plans with different angles is very time costly, making automatic beam-angle optimisation a hot topic in radiotherapy treatment planning. Some algorithms described in this literature study can potentially help with beam-angle optimisation.

As many calculations are needed for an optimal beam-angle search, dose calculation times in the (sub)second regime are a prerequisite. Additionally, the beam angle should be related to the dose calculation procedure. These conditions make four of the DL-based algorithms suited for beam-angle optimisation. The models of Fan et al. (2020) [4], Zhang et al. (2022) [22], Zhu et al. (2023) [25] and Pastor-Serrano and Perkó [13]. The first three require a fluence or TERMA matrix input, which can be formed with ray-tracing algorithms from input beam directions. The latter works directly with an input beam direction. The parallelised algorithm by Gioranengo et al. (2019) [7] might also be suited for beam-angle optimisation as it depends on individually simulated particle tracks in a volume. With the mapping of particle tracks to full particle beams, a beam-angle search can be constructed. All other algorithms are either not related to beam-angle directions or are not able to calculate dose-deposition matrices quickly enough.

For a single beam plan, the angle optimisation is straightforward. Iteratively calculate dose deposition in a patient for all possible sub-sampled beam angles and then determine plan quality according to clinical constraints violated and objectives met. The problem becomes more complex with multiple beam plans. The number of possible beam setups depends exponentially on the number of possible beams. The total beam-angle optimisation time could be reduced by implementing constraints on possible beam setups. For instance, particle beams should at least be separated by 90 degrees. Alternatively, the beam-angle search could start on a coarsely sampled angle grid and from there work towards the optimal solution in multiple constrained searches.

3.2.2. Robust Treatment Planning

Another crucial aspect of clinical treatment planning is robustness. With robust planning, target coverage and dose level constraints are evaluated for several scenario dose distributions. For photon treatment planning, beam placement errors are introduced as scenarios. The margin is added to account for both systematic and random errors present in radiotherapy. In robust planning for particles, range uncertainties are added as well. When the clinical constraints are not violated for all the scenario doses, the treatment can safely be given to a patient. The error is often modelled as a combination of possible 3D shifts and is equal to several millimetres. In this case, robustness will increase with an increasing modelled error size. The range uncertainty is several percentages of the calculated particle range and depends on robustness settings as well.

When considering all possible equidistant 3D shifts inside a cube, 27 scenarios are possible (8 vertices, 6 faces, 12 edges and 1 nominal scenario). Including a range uncertainty will yield 81 scenarios (particle undershoot, particle overshoot and the nominal case). With a maximum of 81 dose calculations, a computation time in the second regime is needed again. Furthermore, the structure of the algorithm has to allow setup and range shifts for particles. A setup error can be simulated by shifting input beam directions. Range errors can be mimicked with input beam energy variations or CT density shifts. Currently, robust evaluation is a tedious process taking several minutes. This makes a fast and accurate dose engine suited for robust evaluations in the clinic.

The stated conditions are satisfied by the same algorithms as described in the beam-angle optimisation application. By manually adjusting beam input positions, shifted scenario doses can be simulated by creating the right engine inputs. Range uncertainties can be added by adjusting the input energy spectrum of these models. Besides these algorithms, the phase-space dose engine created by Burlacu et al. (2023) [2] has the potential to speed up robust planning as well. A full treatment plan takes several minutes, but with the adjoint analysis, range uncertainties and beam shifts can be modelled directly.

Robust treatment planning mainly focuses on whole beam shift errors (i.e. the added error is equal for all pencil beams). However, adding errors of individual pencil beams might make robust evaluation more realistic. The algorithm by Giordanengo et al. (2019) [7] and Burlacu et al. (2023) [2] are based on individual particle tracks and thus have the architecture available to simulate errors in individual pencil beam placement. With the addition of these errors, the robust evaluation becomes computationally more expensive as more scenario doses have to be computed. Whether or not the increase in computation time is worth the effort, should be investigated in the future.

3.2.3. Model-Based Plan Comparison

Within the Dutch healthcare system, patients are eligible for proton therapy based on a plan comparison (apart from a few standard indications) between a photon and a proton treatment plan. The patient is treated with the most beneficial plan, quantified by predicted treatment outcome metrics. The plan comparison requires input from multiple medical professionals from multiple institutes. In addition, dose-deposition matrices on which the treatment plans are based, are computed within different clinical treatment planning systems. In order to accelerate the plan comparison procedure and make the results more comparable, a single engine capable of computing both photon and proton dose-deposition matrices with good accuracy is needed. For ideal patient care, the accuracy of photon and proton dose distributions should be equal, making the plan comparison unbiased.

In theory, only one photon and one proton treatment plan have to be formed for the comparison. This allows for dose computation and plan optimisation times of minutes and results in the fact that accuracy is the most important factor. The discussed DL-based algorithm by Guerreiro et al. (2021) [8] is constructed for treatment plan comparison and thus can be implemented clinically to accelerate plan comparison. However, as it requires ten types of input types, the clinical implementation might be challenging. Besides this algorithm, the model by Fan et al. (2020) [4] mentions that it is also suited for both photon and proton dose calculations. All other papers do not explicitly mention that photon and proton calculations are possible. Nevertheless, some of the algorithms might still be suited with minor adjustments to the structure. In essence, track-repeating algorithms can be made suitable for both photons and protons as well as DL-based algorithms based on TERMA distributions. For phase-space models, it is more complicated, as the geometrical treatment setup is often taken into account. This treatment setup varies for photon and proton treatments.

Considering the accuracy of the described algorithms, DL-based algorithms are the most promising class to accelerate plan comparison. The accuracy of an individual algorithm for both photons and protons has been studied in the work of Guerreiro et al. (2021) [8]. Both photon and proton dose-deposition matrices show similar accuracy compared to a benchmarked clinical treatment system. It is reasonable to assume that the other mentioned algorithms also will display similar accuracy for photon and proton dose calculations. Whether this truly holds, has to be studied in the future before clinical implementation.

3.2.4. Online Adaptive Radiotherapy

Adaptive radiotherapy is required to account for any change in dose delivery to a patient due to a combination of anatomical changes, treatment machinery output changes and patient setup changes during the course of a full treatment. There are two types of adaptive radiotherapy: offline- and online adaptive radiotherapy.

With offline adaptive radiotherapy, inter-fractional dose delivery changes are taken into account. To allow for normal tissue recovery after irradiation, the time between the two fractions is at least 8 hours. Fast and accurate dose engines are not strictly necessary to account for these inter-fractional changes. Current clinically implemented patient QA functions properly.

In the field of online adaptive radiotherapy, where dose delivery is adjusted according to intra-fractional changes, some of the described algorithms may help. Changes in patient setup during the delivery of a fraction are already mostly corrected with patient immobilisation and patient motion tracking. The output of the treatment machine is also monitored during treatment and the delivery is adjusted accordingly. In vivo, anatomical variations are most difficult to correct for and require fast and accurate dose quantification during treatment. Anatomical changes can be tracked with a cone-beam CT during the fraction delivery. The cone-beam CT has limited resolution but scans relatively fast. With this intra-fractional patient CT data, the given fraction can be adjusted to guarantee good dose delivery.

Dose calculations projected on the cone-beam CT scan have to lie in the (sub)second regime to adjust for anatomical changes as quickly as possible. As mentioned before, the resolution of cone-beam CTs is poor compared to full CT scans. This implies, that treatment delivery changes have to be robust and account for changes with care. Computation time is the limiting factor for dose calculations during treatment. Four of the algorithms have the potential to be useful in the development of online adaptive radiotherapy for anatomical changes. The engine by Pastor-Serrano and Perkó (2022) [13] might help as it can calculate the dose deposition from a 3D CT input in around 15 seconds. The GPU accelerated algorithm by Giordanengo et al. (2019) [7] is suited as well. The computation time is less than a second but the accuracy is not great due to kernel approximations. With the engine of Burlacu et al. (2023) [2]

anatomical changes can be related to the adjoint analysis. For small anatomical changes, the treatment delivery can be adjusted very accurately. Lastly, with the engine from Zhang et al. (2022) [23] pencil beam weights can be mapped to a dose distribution. When anatomical changes occur, pencil beam weights can be manipulated such that the updated dose on the cone-beam CT equals the dose on the planning CT. Unfortunately, these four algorithms are mainly developed for particle dose calculations. Efforts have to be made to allow for photon dose calculations. Alternatively, another model has to be found or built from scratch.

The mentioned four algorithms show great potential. However, which of the engines can account for anatomical variations during treatment delivery the best, is not clear. To answer this question, more research is needed to compare and optimise various online adaptive radiotherapy procedures. Complete online adaptation is not readily available yet. Scientific advances in all facets of patient QA will shape online adaptive radiotherapy further.

4

Conclusion

The eleven described fast and accurate dose algorithms are able to compute a dose-deposition matrix from under a second to several minutes and result in comparable gamma pass rates and dose differences. Consequently, a clear correlation between computation time and accuracy is not noticeable. The recognised classes of algorithms all show their advantages and disadvantages. Depending on the problem at hand, one should carefully select and validate a model. The algorithms are suited for a variety of applications within the field of photon and charged particle radiotherapy, depending on time and accuracy constraints. From the four described applications and possibilities, clinical robust treatment planning and model-based plan comparison can directly be developed further with the allocated algorithms. The other two, beam-angle optimisation and online adaptive radiotherapy require more work before clinical translation is possible. Some of the described algorithms are not mentioned in the application and possibilities, nonetheless, they might still solve other, more specific, clinical problems. For further improvement of patient treatment, future research should focus on clinical testing and implementation of fast and accurate dose engines for the mentioned applications and possibilities.

References

[1] James L Bedford et al. “A phase space model of a Versa HD linear accelerator for application to Monte Carlo dose calculation in a real-time adaptive workflow”. In: *Journal of Applied Clinical Medical Physics* 23.9 (2022), e13663.

[2] Tiberiu Burlacu, Danny Lathouwers, and Zoltán Perkó. “A deterministic adjoint-based semi-analytical algorithm for fast response change computations in proton therapy”. In: *Journal of Computational and Theoretical Transport* (2023), pp. 1–41.

[3] Fortuna De Martino et al. “Dose calculation algorithms for external radiation therapy: An overview for practitioners”. In: *Applied Sciences* 11.15 (2021), p. 6806.

[4] Jiawei Fan et al. “Data-driven dose calculation algorithm based on deep U-Net”. In: *Physics in Medicine & Biology* 65.24 (2020), p. 245035.

[5] Irina Fotina et al. “Advanced kernel methods vs. Monte Carlo-based dose calculation for high energy photon beams”. In: *Radiotherapy and oncology* 93.3 (2009), pp. 645–653.

[6] Jan Gajewski et al. “Commissioning of GPU–accelerated Monte Carlo Code FRED for clinical applications in proton therapy”. In: *Frontiers in Physics* 8 (2021), p. 567300.

[7] S Giordanengo et al. “RIDOS: A new system for online computation of the delivered dose distributions in scanning ion beam therapy”. In: *Physica Medica* 60 (2019), pp. 139–149.

[8] F Guerreiro et al. “Deep learning prediction of proton and photon dose distributions for paediatric abdominal tumours”. In: *Radiotherapy and Oncology* 156 (2021), pp. 36–42.

[9] Xun Jia, Peter Ziegenhein, and Steve B Jiang. “GPU-based high-performance computing for radiation therapy”. In: *Physics in Medicine & Biology* 59.4 (2014), R151.

[10] Daniel A Low et al. “A technique for the quantitative evaluation of dose distributions”. In: *Medical physics* 25.5 (1998), pp. 656–661.

[11] S Muraro, G Battistoni, and AC Kraan. “Challenges in Monte Carlo simulations as clinical and research tool in particle therapy: a review”. In: *Frontiers in Physics* 8 (2020), p. 567800.

[12] Jarkko J Ojala et al. “Performance of dose calculation algorithms from three generations in lung SBRT: comparison with full Monte Carlo-based dose distributions”. In: *Journal of applied clinical medical physics* 15.2 (2014), pp. 4–18.

[13] Oscar Pastor-Serrano and Zoltán Perkó. “Millisecond speed deep learning based proton dose calculation with Monte Carlo accuracy”. In: *Physics in Medicine & Biology* 67.10 (2022), p. 105006.

[14] Yiru Peng et al. “Accuracy Improvement Method Based on Characteristic Database Classification for IMRT Dose Prediction in Cervical Cancer: Scientifically Training Data Selection”. In: *Frontiers in Oncology* 12 (2022).

[15] G Russo et al. “A novel algorithm for the calculation of physical and biological irradiation quantities in scanned ion beam therapy: the beamlet superposition approach”. In: *Physics in Medicine & Biology* 61.1 (2015), p. 183.

[16] Angelo Schiavi et al. “Fred: a GPU-accelerated fast-Monte Carlo code for rapid treatment plan recalculation in ion beam therapy”. In: *Physics in Medicine & Biology* 62.18 (2017), p. 7482.

[17] L Tillikainen et al. “A 3D pencil-beam-based superposition algorithm for photon dose calculation in heterogeneous media”. In: *Physics in Medicine & Biology* 53.14 (2008), p. 3821.

[18] Liesbeth Vandewinckele et al. “Overview of artificial intelligence-based applications in radiotherapy: Recommendations for implementation and quality assurance”. In: *Radiotherapy and Oncology* 153 (2020), pp. 55–66.

[19] Yewei Wang et al. “A generalization performance study on the boosting radiotherapy dose calculation engine based on super-resolution”. In: *Zeitschrift für Medizinische Physik* (2023).

[20] Pablo Yepes et al. “Monte Carlo fast dose calculator for proton radiotherapy: application to a voxelized geometry representing a patient with prostate cancer”. In: *Physics in Medicine & Biology* 54.1 (2008), N21.

[21] Pablo P Yepes et al. “Validation of a track repeating algorithm for intensity modulated proton therapy: clinical cases study”. In: *Physics in Medicine & Biology* 61.7 (2016), p. 2633.

[22] Bailin Zhang et al. “Convolution neural network toward Monte Carlo photon dose calculation in radiation therapy”. In: *Medical Physics* 49.2 (2022), pp. 1248–1261.

[23] Guoliang Zhang et al. “SWFT-Net: a deep learning framework for efficient fine-tuning spot weights towards adaptive proton therapy”. In: *Physics in Medicine & Biology* 67.24 (2022), p. 245010.

[24] Jinhan Zhu, Xiaowei Liu, and Lixin Chen. “A preliminary study of a photon dose calculation algorithm using a convolutional neural network”. In: *Physics in Medicine & Biology* 65.20 (2020), 20NT02.

[25] Jinhan Zhu et al. “Feasibility of the photon spectrum generalisation model for rapid Monte Carlo dose calculation with a deep learning-based framework”. In: *Radiation Physics and Chemistry* 202 (2023), p. 110587.