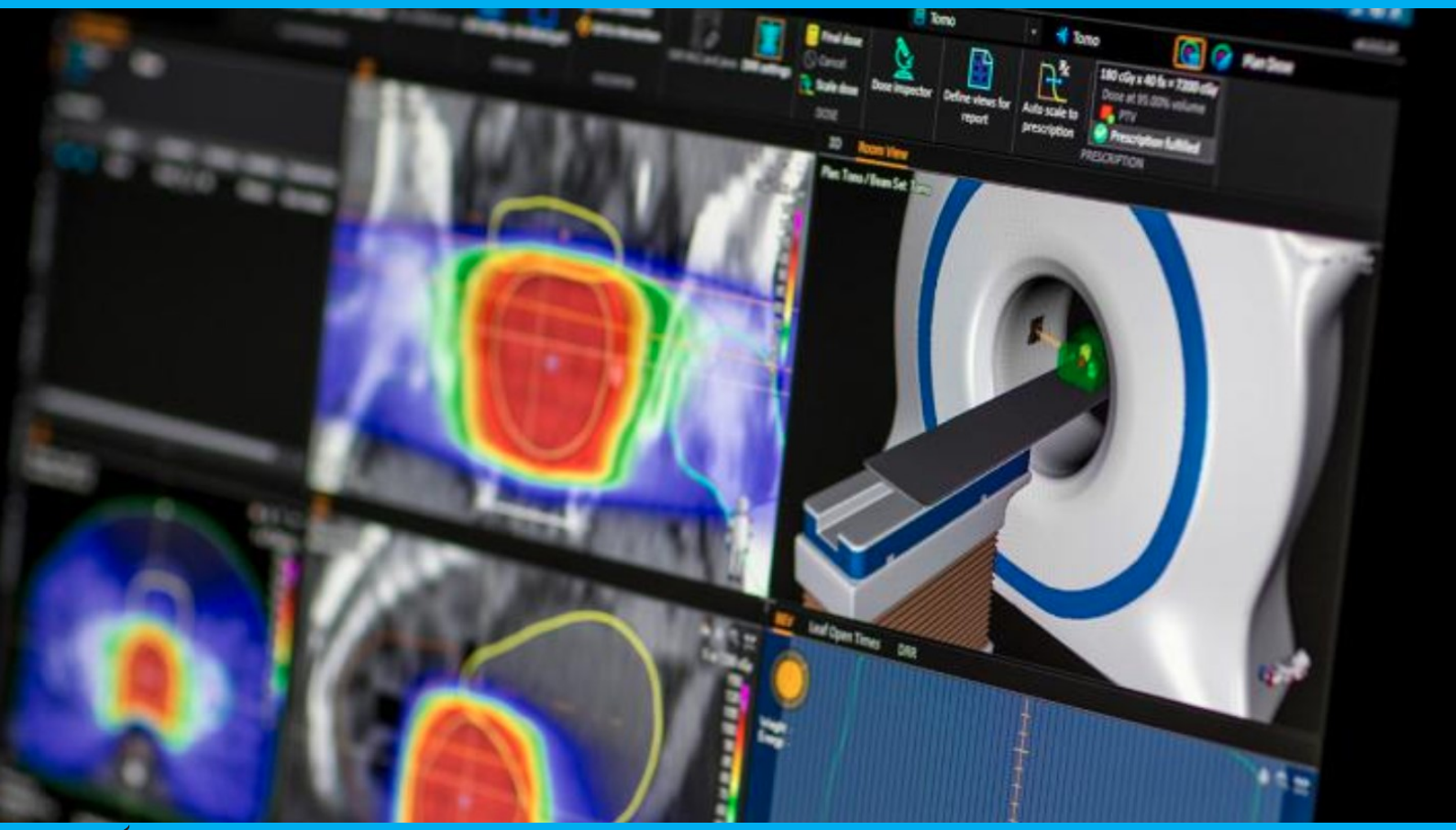
Towards clinically feasible iterative probabilistic treatment planning of IMPT

with polynomial chaos expansion

S. Mastrogianni

to obtain the degree of Master of Science in Biomedical Engineering, to be defended publicly on Thursday July 8, 2021 at 03:00 PM.



TU Delft

Towards clinically feasible iterative probabilistic treatment planning of IMPT

with polynomial chaos expansion

by

S. Mastrogianni

to obtain the degree of Master of Science
at the Delft University of Technology,
to be defended publicly on Thursday July 8, 2021 at 15:00.

Student number:	5058953	
Project duration:	November, 2020 – July, 2021	
Thesis committee:	Prof. dr. M.S. Hoogeman,	TU Delft & Erasmus MC
	Dr. ir. D. Lathouwers,	TU Delft
	Dr. S. J. M. Habraken,	Erasmus MC
	J. Rojo-Santiago, M.Sc,	Erasmus MC

This thesis is confidential and cannot be made public until July 8, 2021.

An electronic version of this thesis is available at `http://repository.tudelft.nl/`.

TU Delft

Acknowledgments

First of all, I would like to thank my direct supervisors. Mischa, thank you, for giving me the opportunity to join the project, and also for all your valuable comments. Steven, thank you for your guidance and supervision. In our weekly discussions, you always had new interesting ideas, and the patience to explain all the concepts. Jesus, thank you for helping me from the first till the last day. You were the person who introduced me to all the concepts, and you were always in the next room at Erasmus to solve my questions. Danny thank you for your time to be on the thesis defense committee.

S. Mastrogianni
Delft, July 2021

Abstract

Radiotherapy is among the most popular modalities used for cancer treatment. Proton therapy is a promising kind of radiotherapy, which uses protons characteristic maximum dose deposition to a specific tissue depth, for precise tumor irradiation. However, in comparison to conventional photon radiotherapy, proton therapy is sensitive to more treatment uncertainties, like errors in proton range and geometrical errors. To account for these uncertainties robust optimization and robustness evaluation have been developed, to obtain and guarantee the robustness of the treatment plans against potential error. Robust optimization uses a fixed number of scenarios, for which the plan is optimized using the worst-case scenario. However, a proper weighting of the sampled scenarios, with the corresponding probabilities is missing using this approach. Probabilistic treatment planning can resolve this limitation, but it requires significant time resources.
To evaluate the quality of a treatment plan, one must quantify the effect of errors in dose distributions. However, dose distributions calculations are time-consuming, therefore in this work, polynomial chaos expansion (PCE) is used, as a dose meta-model [21], for fast and advanced dose analysis. This model uses a series of expansion in terms of polynomials, to evaluate the dose distribution when different errors occur.

The first aim of this work was to improve the PCE construction speed and accuracy. To build the model a fixed number of dose scenarios from a dose engine are required. Currently, PCE dose meta-model is constructed using dose distributions with 1% Monte Carlo (MC) noise level. The trade-off between the noise level (↑ faster model) and the model's accuracy, was investigated. The accuracy for models built with larger noise levels (2 & 3%) was compared, to conclude that the default value is the most efficient choice. Additionally, PCE built using the dose differences (between a scenario with no errors and a shifted scenario), was used to improve models' accuracy for complex anatomies. However, probably due to the larger impact of MC noise when dose differences (smaller dose input value) are considered, the PCE accuracy was not high as expected.

The second aim of this work is to evaluate robustness and trade-offs made in treatment planning in a clinically robust neuro-oncological patient, and a clinically complex patient case. As mentioned, robust optimization does not consider the occurrence probability of uncertainties, therefore the treatment plans might be over-conservative. Using PCE robustness evaluations, we compare treatment plans with different robustness settings. From the results, we concluded that a further reduction of the settings was possible for the clinically robust patient.

Third, for five robust skull base patients, we investigated the trade-off between homogeneity of target dose, robustness, and dose to healthy tissue. PCE robustness evaluations were used for comparison of inhomogeneous (120% maximum target dose) and homogeneous (107% maximum target dose) treatment plans, for different robustness settings. From these planning approaches, no intrinsic difference in the degree of robustness was observed. However, the inhomogenous treatment plan resulted in more healthy tissue sparing overall, at the expense of homogeneity of the target dose.

The final aim of this project is to use PCE robustness evaluations towards a clinically feasible iterative probabilistic treatment planning. We attempted to find a linear relationship between the optimal robustness settings which meet a probabilistic goal, and a scaling factor, using iterations of PCE robustness evaluations. For our method, the overall scaling factor α was used is based on the assumptions that the actual margin recipe is approximately linear, and that the real underlying robustness recipe is a scaled version of the photon therapy margin recipe (van Herk's formula [30]). The conclusion was that for one iteration an overall linear scale factor allows for a substantial gain. However, linear modeling and one iteration do not suffice to find the overall robustness settings optimum.

Contents

1 Introduction

Radiotherapy is a commonly used treatment for various cancer types. During therapy, high doses of ionizing radiation are used to kill cancer cells or shrink the tumors. More than 60% of cancer patients receive radiation in the form of definitive, adjutant, or palliative treatment [5]. Radiotherapy is often used, in a combination with other treatment modalities, like surgery or chemotherapy. Surgery is usually used locally to remove tumor volumes. Radiation therapy is a loco-regional treatment method that effectively reduces the risk of local recurrence. Finally, chemotherapy is an adjuvant systemic therapy used to treat metastatic disease or non-solid tumors.
Modern external beam radiotherapy relies on an intensity modulation (IMRT) technique that aims to deliver high-dose gradients to cancerous tissues while respecting healthy tissue tolerance. The ratio of the probabilities for tumor eradication and normal tissue complication is called therapeutic ratio, and it depends on tumor stage, type and localization. However, when using conventional photon therapy is not possible to avoid exit dose from the target, as photons release energy along their path.

In contrast to photons, proton therapy has been introduced as a novel and promising alternative approach to external radiotherapy, with therapeutic potential. The main advantage of proton therapy is the characteristic maximum dose deposition at a certain depth in tissue. Therefore the beam precisely irradiates the tumors, fully sparing tissue behind the target, while it also reduces integral dose. Proton therapy appears to be superior in the reduction of side effects.
Despite the dosimetric advantage and the therapeutic potentials, proton therapy requires special considerations compared to the processes used for photon treatment planning [5]. The finite range of the protons makes proton therapy more sensitive compare to photons in both treatment planning and delivery uncertainties. For instance, during proton therapy planning, the tissue densities along the particle path have to be accurately determined to obtain the desired dose distribution. Figure 1.1 depicts a comparison of an expected outcome dose distribution and the impact of uncertainties introduced during proton treatment. Uncertainties that can influence the treatment plan are associated with imaging artifacts, tissue heterogeneities, volumes of interest delineation in 3D imaging, patient immobilization and setup, inter-and intrafractional patient and organ motion, physiological changes, and treatment delivery [8].

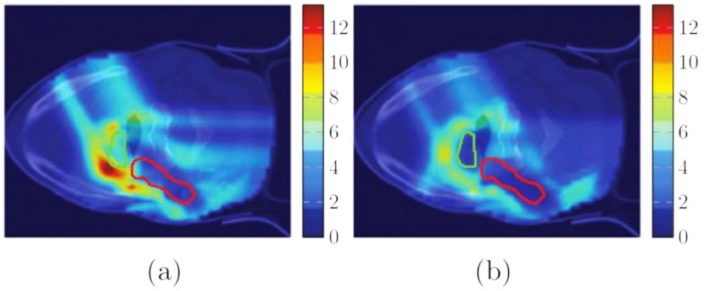


Figure 1.1: Proton therapy: Desired dose distribution (a), dose distribution when errors occur (b)[21]

Unfortunately, it is not possible to eliminate all uncertainties in treatment planning and delivery. In

photon therapy, margins are used as an extension of the Clinical Target Volume (CTV) to create a Planning Treatment Volume (PTV), which provides adequate target coverage under the presence of uncertainties. The assumption that when CTV moves within the PTV limits, it receives the prescribed dose is not valid for proton therapy due to the finite proton range. When intensity-modulated proton therapy (IMPT) used as a delivery technique, range and geometrical uncertainties can cause misalignment of dose contributions of different pencil beams.

To overcome this limitation a method called robust optimization has been developed [48]. Conventional robust optimization aims to account for uncertainties related to range, and geometrical uncertainties, e.g. patient misalignment with respect to the planning CT (setup). There are different approaches aiming at practical and clinically feasible robust optimization solutions. The ones implemented in a commercial planning system are probabilistic approach, minimization of the expected dose value, or optimizing according to the worst-case scenario (Appendix). For scenario-based robust optimization dose distribution is calculated explicitly for each scenario and minimized through the use of planning objectives [11]. Even when the plan is optimized according to robust optimization objectives, there is still a necessity for clinical robustness evaluation. Scenario-based robustness evaluations have been already developed, and give information of how the dose distribution changes, along with additional error scenarios, to guarantee the robustness of the plan, as robust optimization cannot always be constrained. Currently, the clinical robustness evaluation is in a preliminary stage.
Previous work by Perko et al.[21] has shown that it is possible to use Polynomial Chaos Expansion (PCE) as a computationally efficient, patient- and treatment plan-specific, scenario-dose calculation. The model has successfully been used for robust evaluation, as it is capable of simulating 100,000 complete fractionated treatments, within minutes, reducing the statistical error to below 0.3% [84].
Generally, robust optimization uses a few discrete scenarios for optimization. Probabilistic optimization is proposed to improve the optimization process. This optimization method does not use a number of predefined scenarios, but the uncertainties distributions, therefore is a promising method, since it can handle many more scenarios with their occurrence probability taken into account as well. However, the computational time resources for this approach are still really high [83].

1.1. Project's Goal

Previous work [83] has shown that it is possible to take advantage of PCE's fast sampling and investigate the potentials of a clinically feasible probabilistic treatment plan. This fast dose calculation model allows for multiple scenario evaluation within minutes. Therefore the model can work as a bridge towards probabilistic treatment planning.

The purpose of this work is to assess the potential for clinically feasible probabilistic treatment planning, using PCE-based robustness evaluation in a loop with a clinical treatment planning system. The method proposed aims to use PCE to quantify the effect of errors on dosimetric parameters and to use this information to determine the minimum treatment system plan robustness settings, for which sufficient target coverage is achieved. The goal is to create the optimal treatment plan using probabilistic objectives, for individual patients. In addition, with this study, we also aim to evaluate robustness and trade-offs made in treatment planning in a straightforward (clinically robust neuro-oncological target) and a clinically complex patient case. Furthermore, the trade-off between homogeneity of target dose, robustness, and dose to healthy tissue is investigated for a five-patient cohort.

This thesis is organized as follows. In Chapter 2, a background on proton therapy is given, and the relevant uncertainties for neurological patients are discussed. Chapter 3 analyze the used treatment planning systems and the methods of including robustness in the process. Additionally, the potentials and the limitations of probabilistic planning against conventional robust planning are presented. Following Chapter 4, the Polynomial Chaos Expansion model for dose simulation is explained. The methods applied in this thesis are presented in Chapter 5. The Result section begins with the optimization of PCE accuracy and construction time. Probabilistic approaches are computationally demanding. Therefore, even further improvement in the model's speed can facilitate probabilistic plans. Next, the effect of margin selection in the treatment plan's robustness is investigated, for both a robust and a non-robust patient case. Other ways of achieving robustness are tested, by studying the trade-off between dose homogeneity and robustness. The final part of the results is dedicated to implementation and evaluation of a feasible iterative approach to probabilistic

treatment planning. Finally, in the last chapter, a discussion of the results and the conclusions of this work are provided.

2

Proton Therapy

As mentioned in the introduction, proton therapy can potentially reduce radiotherapy side effects due to the less irradiation of the healthy tissue. Even if the treatment can be sensitive to anatomy changes and range uncertainties than photon therapy, the mentioned benefits make the treatment particularly interesting for a **specific patient population**, as it can reduce both early and late radiation side effects. Depending on the patient's age, the late radiation effects should be considered, as they are usually permanent. In the Netherlands, the patients eligible for proton therapy are selected through the national indication protocol for proton therapy [74]. The patients are divided into 'standard indications' (directly eligible for proton treatment) and so-called 'model-based indications' (must have a substantial reduction in dose to healthy tissues to be eligible for proton therapy). Skull base and brain tumors are frequently eligible for proton therapy due to the reduction of the mean brain dose. In this work, we focus on skull-base tumors.

2.1. Proton Energy deposition

The first steps toward clinical use beams of accelerated protons for radiotherapy were in 1946, when Robert Wilson proposed the use of high-energy proton beams from the cyclotron to treat deep tumors. His article [72], described the qualitative properties of proton beams, the propagation techniques of the beam, and the treatment benefits of using charged particles for treatment. The first human proton therapy was conducted in 1954 at Lawrence Berkeley Laboratory, to control metastatic breast cancer [2]. Later in 1962, at Harvard Cyclotron Laboratory, proton therapy was introduced to cure eye tumors and larger malignancies [3]. Harvard's research program continued for more than 40 years, with physicists and clinicians further improving the safety and effectiveness of the therapy. However, the development and use of protons were limited, due to cost and technical difficulties compared to conventional photon therapy.

As previously discussed, the major advantage of proton therapy relies on the proton dose deposition within the patient. When protons travel through the tissues, they lose energy mainly as a result of Coulomb interactions with electrons and nuclei [6]. Due to protons' positive charge, momentum is transferred to the electrons of the medium. This results in the excitation or ionization of the atoms in the medium. At the end of their range maximum interaction with electrons occurs, causing maximum energy release [6]. When the proton is completely stopped, the deposition of the energy is maximized, a location usually known as Bragg Peak. A priori, this dosimetric property makes proton therapy beneficial for cancer treatment. In contrast, high-energy photons do not reach the maximum dose on the skin surface, but build-up to a maximum dose deposition a few cm under the skin, and after gradually decreases till it exits the body (dose fall off). Figure 2.1 shows the difference between proton and photon therapeutic beam. The slope of photons dose fall-off is decreasing with increasing photon energy.

The range of protons depends on their initial energy. Since the Bragg peak of a quasi-monoenergetic (pristine) pencil beam are not wide enough to achieve tumor volume coverage, several pristine Bragg curves are combined to compose a Spread-Out-Bragg-Peak (SOBP). The width of an SOBP depends on the tumor's volume. Figure 2.1 highlights the dosimetric advantage of the SOBP in comparison to the conventional photon therapy, as the dose is minimized after the distal SOBP edge. Therefore, healthy tissue located deeper than the tumor has a dose deposition close to zero. Even in shallower depths, healthy tissue income dose is reduced compared to the photon beam. Scattering foils, brass apertures, and energy modulation techniques can be utilized to form an SOBP [7]. This form of proton therapy is called passively scattered proton therapy

(PSPT), and it is highly studied and used.

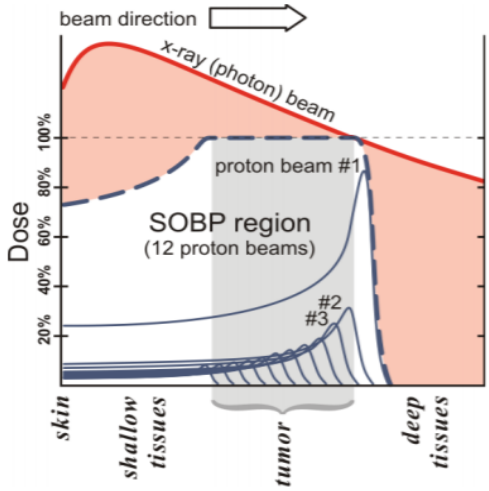


Figure 2.1: Comparison of SOBP (consisting by 12 beams) as a sum of individual Bragg peaks in different depth, and X-ray dose depth diagram. Pink area represent the additional dose that photons deposit to healthy tissue.

In this work, intensity-modulated proton therapy (IMPT) is used, which is analogous to IMRT. IMPT uses pencil-beam scanning (PBS) technology for radiotherapy delivery, and it is a more complex mode than PSPT. In contrast to passive scattering delivery methods, IMPT **does not** require custom-made materials (like compensators). This therapy uses different pencil beams, a few millimeters wide, which scan the tumor in layers of spots. This method is capable of treating with high precision. However, IMPT is highly sensitive to range uncertainties. High precision is a critical characteristic for proton therapy, as the dose deposition is significantly high for the tumor area, making any misalignment of the target-beam system, high risk for the surrounding healthy tissue.
The following sections of this chapter provide an overview of the uncertainties that occur for neurological tumors, along with the method used in each treatment plan to account for them. A more detailed analysis of uncertainties is given in the Appendix.

2.2. Uncertainties in Proton Therapy

There are several sources of uncertainties that can compromise radiotherapy treatment accuracy. They can be divided into patient-related uncertainties that apply to all treatment modalities (like target motion) and technology-based uncertainties that vary between different modalities (like proton range uncertainty). In general, all uncertainties may vary between treatment centers, because even the patient-related uncertainties depend on the immobilization devices, patient setup, and plan adaptation protocols.

The delivery dose is highly sensitive to target positioning relatively to the beams, uncertainty in the patient density data, or uncertainty regarding the location of the cancer cells as the tumor's boundaries at a microscopic level is hard to define. The errors in radiotherapy are decomposed as systematic, which occur during treatment preparation, and they remain the same for each fraction during the treatment, and random errors during the treatment execution, which have random nature so they change value for each fraction.
Figure 2.2 depicts the coordinate system that is going to be used throughout the paper concerning the patients' position. Dorsoventral (y) direction indicates the back of the head to the nose direction, the Left-Right (x) refers from ear to ear axis, while Craniocaudal (z) corresponds to the top and the bottom of the head direction.

2.2.1. Geometrical Uncertainties

The first category of discussed errors that affect both photon and proton treatment is inter-fraction geometrical uncertainties, called geometrical uncertainties which split into a systematic and a random component for a given fractionated treatment. These errors can occur due to variations in patient setup and anatomy, due to registration, or due to isocenter misalignment of the patient during irradiation against the position during planning CT. The differences in position between the treatment plan and the final irradiation position are called shifts and they can occur in different combinations of craniocaudal, lateral, and dorsoventral

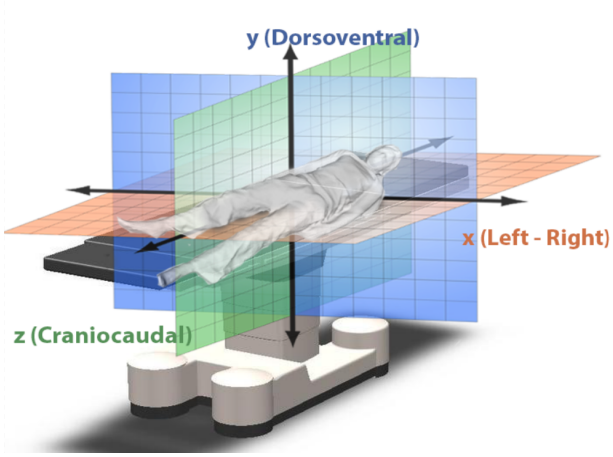


Figure 2.2: Couch coordinate system.

directions.
The systematic component of the error derives from the fact that even after stabilization techniques or the use of markers for beam alignment, the positioning of the patient during treatment might differ from the position of planning CT. This causes a constant error over multiple fractions at each axis which has a standard deviation Σ expressed with millimeter (mm) as a unit. The random error component is related to misjudgments occurring during one of the treatment fractions and is similarly expressed with a standard deviation σ in (mm). Figure 2.3 is a schematic representation of systematic and random setup errors in which the nominal position (scenario free of errors) is in black, the added systematic error corresponds to the difference between the blue and the black point in patient positioning, while the additional random errors shift the patient to the corresponding red dots.

The final total positioning error relative to the reference position, for one fraction, is a sum of systematic and random sigma. As a result, many studies select to consider them as a combined error for each direction.

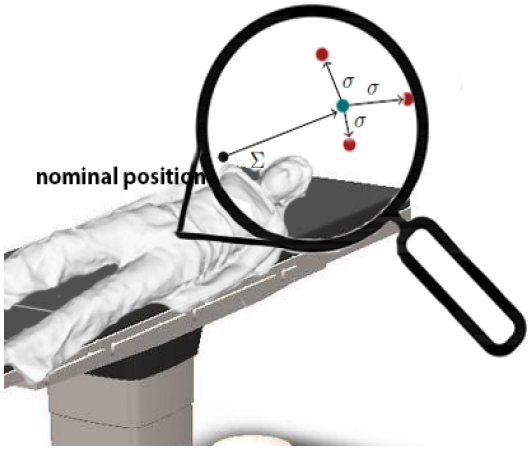


Figure 2.3: Setup error three shifts.

2.2.2. Range errors

The second category of errors that affects IMPT, is the range errors. As mentioned above these errors are "unique" to charged particle therapy. Range errors can be diveded into relative and absolute errors. Relative range errors arise from the conversion of CT Hounsfield units (HU) in proton stopping power. For instance, the calibration curves between CT and stopping proton power, sometimes different HU from different tissue densities leads to the same stopping power values [85]. Uncertainties in range, that are not related to CT - stopping power conversion, are called absolute range errors. These errors arise from CT image artifacts, anatomical, and contour changes, weight, or gain loss [12]. In this work, only relative range errors are con-

sidered.
All setup and range errors in this study are considered to be **Gaussian distributions** with a standard deviation depending on the patient site. The type of distribution is selected following the central limit theorem, which states that combining many mutually uncorrelated factors lead to normal Gaussian distribution. Combining all the errors mentioned, an error scenario is formed, that consists of three setup and the range error.

2.3. Skull based tumors

According to the United States National Brain Tumor Society, there are more than 120 types of brain tumors that can be diagnosed (by using a combination of medical imaging, histological, pathological, and molecular essays). Likely, the probability to develop these types of tumors in a person's lifespan is small (1%), however, brain tumors account for 90% of all primary central nervous system tumors [70]. Except for the known dosimetric advantage of proton therapy, the clinical benefit depends on the tumor's location and the patient's age. As mentioned, in the Netherlands patients eligible for proton therapy can be [74]:

- skull base tumors
- brain tumors, for which both brain and spinal cord must be irradiated
- brain tumors with a favorable prognosis without the possibility of high-precision radiation therapy

According to literature [55], patients with frontal-parietal tumor or malignancies involving large parts of the brain already showed clinical interesting results. Figure 2.4 shows that a significant volume of healthy tissue is spared with the use of IMPT compared to conventional photon therapy.

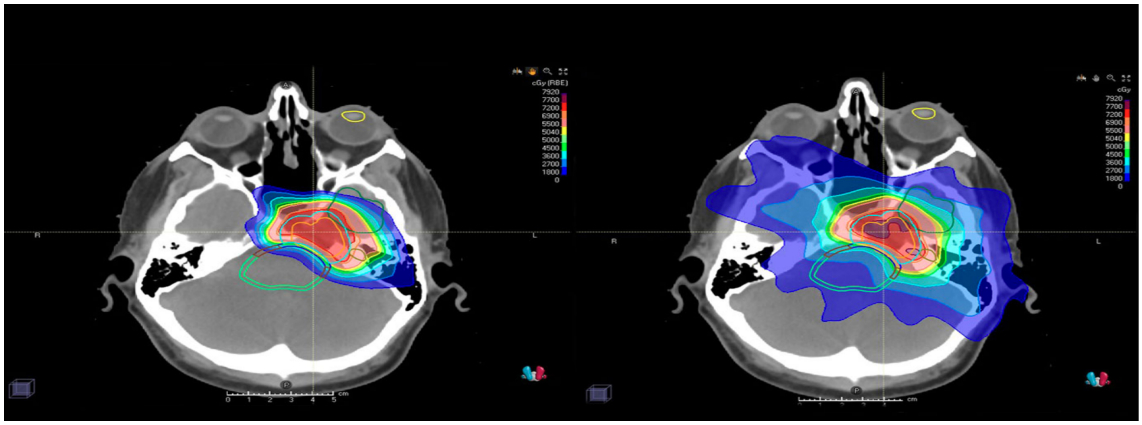


Figure 2.4: Comparison of IMPT (left) and traditional photon therapy (right) [58]

Skull base tumor, which will be this work, are located in close proximity to sensitive structures, called organs at risk (OAR). In Figure 2.5, a patient volume delineation is illustrated, in which CTV along with some OARs are shown.

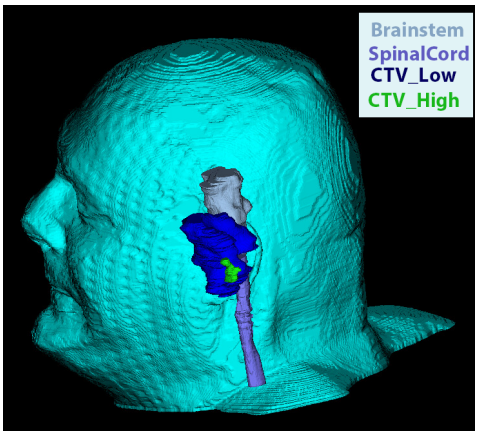


Figure 2.5: Organs-at-risk and CTV for a skull based chordoma patient (using Erasmus RTStudio)

The clinical dose constraints for OARs must be strictly followed, as serious toxicities can occur. The radiosensitive structures that are often considered for skull base tumors are listed as followed:

- *Brainstem*: toxicity has been reported, and it can affect the cognitive functions
- *Spinal cord*: toxicity is not commonly reported after brain irradiation, however, it has to be considered for later risks [56]
- *Optical chiasm* / *Optical nerves*: patients often lose vision of one-half or a quadrant of the visual field from injuring a part of the optic nerves or chiasm during radiation therapy [76]
- *Pituitary gland*: overdosing the pituitary gland can lead to reduced or completely stop hormone production
- *Hippocampus*: function is related with memory processing. Therefore, damage to this structure can lead to complications in neuro-cognitive function

3

Treatment Planning

In this Chapter, different treatment parameters are presented, which are used to assess and evaluate the adequacy of a given treatment plan. Secondly, the step followed to obtain a treatment are listed. Additionally, dose engines used for this study are briefly analyzed. Finally, the robust and the probabilistic treatment planning used in proton therapy are explained.

3.1. Dosimetric Parameters

-D$_x$
This parameter indicates the maximum dose that is received by at least x% of the volume. Often two values of x% are used for treatment [16]:

- $D_{98\%}$ is called near-minimum dose, and corresponds to the dose received by 98% the PTV volume
- $D_{2\%}$ is called near-maximum dose, and corresponds to the dose received by 2% of the PTV volume.

and

-D$_{pres}$
The total dose that the target should receive during the treatment.

-V$_y$
A complementary measure that gives the fraction of the volume that receives at least a specified dose, and is determined from the following expression:

$$V_y = \frac{1}{N_{voxels}} \sum_{i=1}^{N_{voxels}} \delta(D^i \geqslant yD_{pres}) \tag{3.1}$$

where N_{voxels} is the number of voxels included in the examined volume fraction, D^i is the voxel dose, which is connected with a multiplication of the fraction y and the prescribed dose with a Kronecker Delta (δ). This relationship indicates the percentage (y) of the prescribed dose that a volume (V) will receive [85]. For instance V_{95} for the CTV gives the volume that receives 95% of the prescribed dose.

3.2. Treatment Planning

In IMPT a treatment plan prescribes the number, intensity, the lateral (xy) position and the angle of the pencil beams, that will be used for treatment, depending on the prescribed dose to the CTV and OARs for a specific patient [85].
Firstly, a CT scan of the patient is used for treatment planning (planningCT or pCT), from which a clinician will delineate the organs at risk and the target volume. The structure delineation is important for determining the clinical objectives and constraints. However, tumor delineation, in particular, is very difficult, because most of the structure is not visible on a CT. In delineation, the radiation oncologist combines clinical investigation of the patient with different imaging modalities. Therefore the expertise and the experience

of the clinician can affect the treatment plan quality. The goal of the treatment plan goal is to achieve a sufficient dose to the CTV while minimizing the dose deposition for the healthy tissue.
A treatment plan can be made using different Treatment Planning Systems (TPS). The delineated CT and the clinical goals are used as input to the TPS, which optimizes the plan's parameters to achieve the objectives. The output treatment plan is examined by the treatment planner. In this work, two treatment planning systems will be used.
Finally, the optimized beam angles and weights are used for treatment. However, beam-angle optimization, using TPS is often time-consuming, so often a beam-angle configuration is given, manually chosen by the treatment planner or a medical physicist (e.g. brain tumors), or by using standard beam arrangement with minor variation (e.g. head and neck tumors). Figure 3.1 illustrates a proton therapy gantry - couch system, for which the angles are optimized.

Figure 3.1: Delivery system: Gantry beam and couch 3D illustration.[57]

3.2.1. Uncertainties in conventional treatment plan

The goal of the treatment plan is to ensure that the tumor's coverage probability is higher as possible while all OARs receive the minimum possible dose. In this research, the aims is the tumor volume to receive at least 95% of prescribed dose. To ensure dose deposition in the desired area the plan has to be robust against uncertainties.

Photon therapy
For photon beams, the impact of (rigid) shifts on the dose distributions can be approximated by a static shift of the dose cloud, relative to the patient. Static dose cloud approximation assumes dose invariance under small shifts. For homogeneous treatment sites, this is considered an accurate approximation. The use of static cloud dose approximation implicitly assumes the use of geometric margins which allow a margin generalization for the creating of a robust treatment plan. A margin recipe as a linear combination of standard deviation of systematic Σ and random σ uncertainty, for photon therapy is defined by van Herk's formula [30]:

$$M = 2.5\Sigma + 0.7\sigma \tag{3.2}$$

Figure A.7 provides an example of a Gross Target Volume (GTV), surrounded from the CTV margin to account for microscopic tumor expansions, and finally expanding to a margin to ensure robustness against treatment uncertainties.

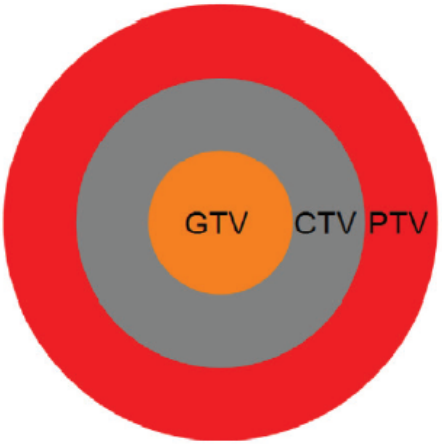


Figure 3.2: Schematic representation of GTV, CTV and PTV [17]

Proton therapy
For IMPT, the static dose cloud approximation has been proven insufficient to account for treatment uncertainties [69]. Therefore, for this treatment, an alternative to margin-based approaches is proposed to account for both setup and range uncertainties. The importance of range errors is displayed in Figure 2.1, in which when the SOBP shifts relative to the tumor, a tumor undershooting or healthy tissue overshooting might occur. In contrast to other treatment sites with anatomical variations i.e., lungs or prostate cancers, organ motion will not be considered in the following section as an additional uncertainty. This thesis will discuss robustness against uncertainties in IMPT for neuro and skull chordoma patients, where the effect of anatomical uncertainties is negligible.
A minimax robust optimization method is introduced in treatment planning to make a treatment plan robust against patient error scenarios.

3.3. Treatment Planning System

In this section, the optimization methods used in the two TPSs are briefly analyzed. The two treatment planning systems used are iCycle and RayStation. RayStation treatment planning system is currently used in clinics for proton therapy, while iCycle, which is a fully automated treatment planning system, is expected to be used soon for proton therapy patient selection.

3.3.1. RayStation

RayStation is a clinical TPS developed by RaySearch Laboratories, which enables the design and optimization of proton treatment plans for actively scanned pencil beams. It is equipped with both pencil beam algorithm dose engine and Monte Carlo (MC) dose calculations. Even if MC dose calculations are more time-consuming, they can be useful in complex anatomies like nasal cavities or changes in densities where pencil beam accuracy is lower due to scattering effects (Appendix [81]).
From the available dose engines, MC is used in this research. The TPS optimizes the treatment plan giving different weights to constraints and objectives, which increase respectively with the priority of them, in the optimization process.

3.3.2. Erasmus-iCycle

The second TPS is a fully automated treatment planning system developed in Erasmus Medical Center, in which a multi-criteria optimization based on patient individual wish-list is also used. The wish-lists contains predefined constraints and objectives. The constraints must be met during the optimization, while the objectives get prioritized. For example, a sensitive OAR that is in high proximity to the tumor will have higher priority. The wish-list depends on the treatment area, and not on the patient.
After the wish-list is functional and complete, the treatment plan can be made. For beam selection using iCycle, a proton beam resampling is used [71, 85]. Figure 3.3 is a schematic representation of iCycle multicriterial beam optimization. First, a number of pencil beams are randomly selected, from a very fine grid.

These beams are included in an optimization iteration. Next, an optimization is performed, from which the beams with the lowest contributions are removes and replaced with random new ones. This is repeated for each optimization iteration. The optimization repeats till the treatment plan meets all the constraints, and the objectives are optimized to their fullest extent [12], creating a Pareto optimal plan.

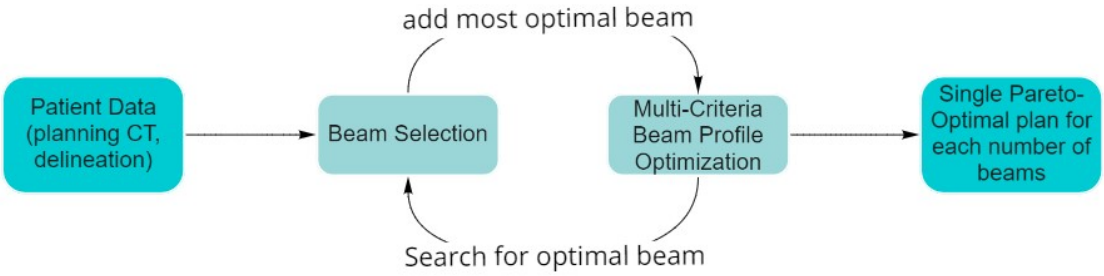


Figure 3.3: Erasmus-iCycle optimization: Beam directions are added to the plan,followed by a multicriterial optimization of beam profiles, resulting in a Pareto-optimal plan [82]

Erasmus-iCycle use the Astroid proton dose engine, which was developed at Massachusetts General Hospital - Harvard Medical School (based on pencil beam algorithms) [12].

3.4. Robust Treatment Planning

In Chapter 2, the need, for an alternative method to account for uncertainties present in proton therapy, was expressed. Robust treatment planning aims to ensure adequate CTV dose under the presence of any treatment uncertainty. Robust planning can be divided into robustness optimization and robustness evaluation (Appendix [A.5]). For robust optimization, both RayStation and iCycle use the minimax worst-case optimization.
Minimax optimization uses a number of planning scenarios, which are a combination of setup and range errors. For conventional treatment planning (non-robust), the optimization in the TPS is performed only for a planning scenario with no errors, called error-free of nominal scenario. For brain tumors, the nominal scenario is given by zero setup error, and the range error will be a systematic undershooting of 1.2% [84]. Minimax optimization uses additional planning scenarios, for which the dose is evaluated and the worst-case value for each objective is optimized [12]. **RayStation** uses 21 planning scenarios are included during robust treatment planning [84]. The planning scenarios will be defined from the robustness setting. The robustness settings are divided as follows [85]:

- **Setup Robustness**: defines a shift in both the positive and negative direction (for all three directions), given in (mm)
- **Range Robustness**: sets relative range shifts in both positive and negative direction, given in percentage.

Additionally, the nominal scenario is considered in the optimization. All the scenarios combinations are included in the optimization. Thus in total 21 planning scenarios are included, which correspond to:

*(6 directions [SR positive & negative] + 1 nominal [zero setup error]) * (2 range errors [RR positive & negative] + 0% range error)*

More scenarios can be used for robust optimization in RayStation, however, the time requirements, for treatment plan construction, increase exponentially.

-iCycle: allows a selection of which objectives and constraints to plan robustly and which not, the remaining objectives are optimized for the nominal scenario [12]. By default, iCycle uses 19 scenarios for optimization. More specifically:

*1 nominal scenario + (6 directions [SR positive & negative]) * (2 range errors [RR positive & negative] + 0% range error)*

Scenario-based robustness evaluation , is often used, in a combination with robust optimization to:

- assess the sensitivity of a given treatment plan to errors
- guarantee that the final result will be sufficiently robust, as robust optimization cannot always be constraints
- as an independent check, often also involving scenarios not used in the optimization

In this work PCE is used for robust evaluation, as it provides fast and accurate dose distributions, with minimum time requirements.

3.5. Probabilistic Treatment Planning

Robust treatment planning ensures a treatment plan which achieves adequate target coverage for all the predefined scenarios. However, the worst-case scenario, which is chosen to be optimized, does not necessarily correspond to the scenario in which the patient is going to be treated. In fact, a proper weighting of the scenarios sampled with corresponding probabilities is lacking in clinical robust optimization and robustness evaluation. That leads the treatment plans to be often over-conservative. If lower setup robustness can be used with the same target coverage, then less dose deposition at the healthy tissue might be possible.
Probabilistic treatment planning resolves this issue. The challenge in making probabilistic treatment plans feasible is the computation time and to proper define probabilistic metrics. In probabilistic planning, each uncertainty parameter is assumed to be continuously distributed according to a probability density function. As mentioned in Chapter 2, the actual uncertainty data, using the central limit theorem, justify the use of Gaussian distributions. Using a continuous function to describe the scenarios, one can overcome the restriction of the robust treatment plan, in which a specific number of scenarios are taken into account. Additionally, different statistical measures, like *median, percentile, expected value, conditionally expected value*, can be used for optimization. To create a probabilistic plan, a statistical measure that minimizes a stochastic variable is needed [83]. Examples of stochastic variables are the dose or the minimum dose. However, there are not protocols yet to determine which statistical measure or stochastic variable is the most suitable for probabilistic treatment planning.
In this work, Near-minimum dose ($D_{98\%,CTV}$) is selected as a stochastic variable, and the percentile which corresponds to 98% population coverage is selected as statistical measurement. Using a probabilistic treatment plan a probabilistic statement about the probability to achieve the stochastic quantity. For instance: using specific setup robustness the tumor coverage meets the clinical constraints with a 98% probability.
However, since there is not an analytical expression for the dose deposition matrix ($D(\vec{\xi})$) as a function of the uncertainty, an approximation of the statistical measure is necessary. As discussed in the appendix, Monte Carlo simulations can solve the problem. This method is brute force, but computationally efficient MC scenario dose calculation for up to 1,000 scenarios. However, compared to the polynomial chaos expansion dose meta-model (Chapter 4), the MC calculations are rather slow. Even when PCE was used, J.H. Salverda et al. [83], reported 3 weeks for a probabilistic plan, using iCycle (8 processors & 60 GB memory reserved).
Based on these results, a need for computationally efficient tools is highlighted to obtain a probabilistic treatment plan. Therefore, an optimization of PCE construction time is attempted in Chapter 6 of this work.

4

PCE

As mentioned in the previous chapter, fast dose calculations are necessary for clinically feasible robustness evaluations. This chapter covers the basic theory behind Polynomial Chaos Expansion (PCE) mathematical model. A PCE dose distribution meta-model can be used for fast and advanced dose analysis. This section covers the needed theory about the model.

4.1. Polynomial Chaos Expansion

The basic principle of Polynomial Chaos Expansion (PCE) is to express the dependency of the stochastic response R($\vec{\xi}$) of a deterministic system to stochastic noise $\vec{\xi}$, as a series expansion of basis vectors Ψ_n multiplied by the corresponding deterministic expansion coefficient r_n. The relationship is given in Equation 4.1.

$$R(\vec{\xi}) = \sum_{n=0}^{\infty} r_n \Psi_n(\vec{\xi}) \tag{4.1}$$

In this research, the output of the PCE model will be the dose distribution (R_i, for voxel i) , as a function of the uncertainty vector. The uncertainty vector is given from Equation 4.2.

$$\vec{\xi} = (\xi_1, \xi_2, ..., \xi_N) \in \mathbb{R}^N \tag{4.2}$$

where N is the number of input parameters. The random variables ξ_i are assumed to be independent Gaussian distributions, with mean μ and standard deviation σ. Therefore, the probability density function of each ξ is given from Equation 4.3.

$$p(\xi) = \frac{1}{\sigma\sqrt{2\pi}} e^{\frac{1}{2}\frac{\xi-\mu}{\sigma}^2} \tag{4.3}$$

Since the Gaussian distributions of ξ are independent, the joint probability density is given as a multiplication of the individual probability density functions:

$$p_\xi(\xi) = \prod_{i=1}^{N} p_{\xi i}(\xi_i) \tag{4.4}$$

In order to solve the Equation 4.1, the basis vector Ψ_n and the expansion coefficients r_n have to be determined.

4.2. Basis Vectors

The PCE basis vectors are polynomials from the subspace of polynomials, in all possible combinations of the random variables ξ. Depending on the probability distribution function ($p_{\xi i}(\xi_i)$) of each random input uncertainty, different polynomials are selected. For the polynomial selection, Wiener-Askey polynomial chaos scheme [61] is used. The scheme suggests the polynomial type for each random uncertainty, which reduces systems dimensionality and leads to exponential convergence of the error. If one used a different polynomial than the one suggested from the scheme, higher polynomial order should be used. The continuous part of the scheme is given in Table 4.1.

Random Variable	Wiener–Askey chaos	Support
Gaussian	Hermite	(-∞,∞)
gamma	Laguerre	[0,∞)
betta	Jacobi	[a,b]
uniform	Legendre	[a,b]

Table 4.1: Continuous types of Wiener–Askey polynomial chaos and their underlying variables

Following the scheme, for the input uncertainties (ξ), which are described by Gaussian distributions, Hermite polynomials should be selected. The probabilists' Hermite polynomials are given by:

$$He_n(\xi) = (-1)^n e^{\frac{\xi^2}{2}} \frac{d^n}{d\xi^n} e^{\frac{-\xi^2}{2}} \tag{4.5}$$

The multiplication of the Hermite polynomials $He_n(\xi)$ with Gaussian distributions results in orthogonal polynomials. Therefore, using Hermite polynomials for polynomial chaos construction, the basis vectors will also be orthogonal:

$$< \Psi_n, \Psi_l >= \int_{D(\theta)} \Psi_n(\vec{\xi}) \Psi_l(\vec{\xi}) p_{\vec{\xi}}(\vec{\xi}) d\vec{\xi} = h_n^2 \delta_{n,l} \tag{4.6}$$

where $D(\theta)$ the domain of random variables and ${h_n}^2$ is the norm of the basis vector Ψ, and kronecker delta ($\delta_{n.l}$)[9, 12].
The polynomial chaos basis vectors Ψ_n are constructed by tensorisation of the univariate polynomials $\psi_{i,\gamma_{n,i}}$ correspond to the different random variables(ξ_i).

$$\Psi_n(\vec{\xi}) = \prod_{i=1}^{N} \psi_{i,\gamma_{n,i}}(\xi_i) \tag{4.7}$$

where i represents the uncertainty input ξ_i, and $\gamma_{n,i}$ is used to define the polynomial orders, which will be included to the expansion [83]. For a set of $o-th$ order basis vectors are given by:

$$\Gamma_o = \{\prod_{i=1}^{N} \psi_{i,\gamma_{n,i}} : \sum_{i=1}^{N} \gamma_{n,i} = o\} \tag{4.8}$$

And for a full O-th order polynomial chaos vectors are given by:

$$\Gamma_O == \{\prod_{i=1}^{N} \Psi_{i,\gamma_{n,i}}(\xi_i) : \sum_{i=1}^{N} \gamma_{n,i} \leqslant O\} = \{\Psi_n(\vec{\xi}) : \sum_{i=1}^{N} \gamma_{n,i} \leqslant O\} \tag{4.9}$$

Then, the series expansion of Equation 4.1 is limited to P+1 vectors in the expansion. Which is expressed by:

$$P + 1 = \frac{(N+O)!}{N!O!} \tag{4.10}$$

where O corresponds to the combined order of all multi-dimensional polynomials, and N is the number of input uncertainties.

4.3. Expansions coefficients

After determining the basis vectors for the PCE, from the Equation 4.1 only the expansion coefficients r_n are unknown. Using the orthogonality of Equation 4.6, the coefficient expansion is given by:

$$\begin{aligned} r_n &= \frac{< R(\vec{\xi}), \Psi_n(\vec{\xi}) >}{< \Psi_n(\vec{\xi}), \Psi_n(\vec{\xi}) >} \\ &= \frac{\int_{D(\theta)} R(\vec{\xi}) \Psi_n(\vec{\xi}) p_{\vec{\xi}} d\vec{\xi}}{< \Psi_n(\vec{\xi}), \Psi_n(\vec{\xi}) >} \\ &= \frac{1}{h_n^2} \int_{D(\theta)} R(\vec{\xi}) \Psi_n(\vec{\xi}) p_{\vec{\xi}} d\vec{\xi} \\ &= \frac{1}{h_n^2} \int_{D(\theta)} R(\vec{\xi}) \prod_{i=1}^{N} \psi_{i,\gamma_{n,i}}(\xi_i) p_{\vec{\xi}} d\vec{\xi} \end{aligned} \tag{4.11}$$

The final integral of the Equation 4.11 contains the unknown response $R(\vec{\xi})$, as an input. Therefore, a non-analytical solution is necessary to approximate the integral. To this end, there are different approaches to solve this equation and accurately determine the PCE coefficients. In this work, both progression and regression approaches are used:
– For the iCycle's dose engine, which uses a pencil beam dose algorithm, a projection approach is used to determine these coefficients. Each PCE coefficient is determined using cubature to approximate the multi-dimensional integral
– For RayStation's dose engine, which uses an MC dose engine, a regression approach is used to determine these coefficients. The PCE coefficients are estimated by minimizing the mean square error of the difference between the TPS and the PCE dose distribution.

4.3.1. Projection Approach

-Quadratures
Quadrature methods are used for numerical approximation of one dimensional, aiming to attain a level of precision with the fewest possible function evaluations:

$$I^{(1)}f = \int_a^b f(\xi)p_\xi(\xi)d\xi \approx Q_{lev}^{-1}f = \sum_{i=1}^{n_{lev}} f(\xi_{lev}^{(i)})w_{lev}^{(i)} \tag{4.12}$$

where ξ_{lev} are the predefined quadrature points and w_{lev} the selected weights.
The quadrature methods rely on evaluating the integrand $f(\xi)$ on a finite predefined integration points ξ_i, called quadratures, and taking into account the weighted sum of the outputs (w_{lev}). In Equation 4.12 the index lev indicates the accuracy level of the method, with a higher number of levels (lev) means that more points (n_{lev}) are used for the approximation. The number of evaluations increases when using more quadrature points. The points and weights used in the Equation 4.12, must aim to minimization of integration error. For this selection, the probability density function ($p_\xi(\xi)$) is combined with the uncertainty variable (ξ) and one of the quadrature rules.
There are also, several quadrature rules for numerical integration with different properties. Most important are the accuracy and nestedness which refers to the number of reoccurring points in each level [85]. When moving to a higher quadrature level, if nestedness is high, some quadrature points have already been calculated in a lower level, so they are reused. When all points from the previous level are reused we have full nestedness, while when for each quadrature level all points are different from the previous levels, zero nestedness is observed. In Figure 5.5, in Clenshaw-Curtis rule full nestedness is observed for an [-3,3] interval, while for Gauss-Hermite limited nestedness is noted, with only one point to reoccur between the levels. That means that for Clenshaw-Curtis fewer function evaluations are necessary for the higher levels, however, Gauss-Hermite is more accurate for the studied interval.

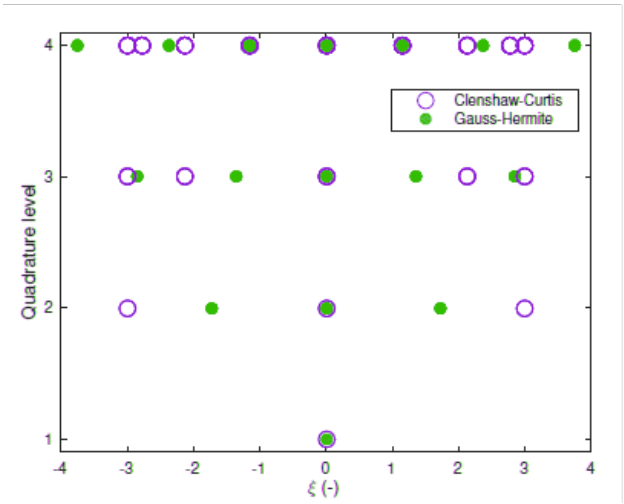


Figure 4.1: Different levels of nestedness plotted together. The Gauss-Hermite rule only re-uses the origin, while the fully nested Clenshaw-Curtis each point reoccurs [85]

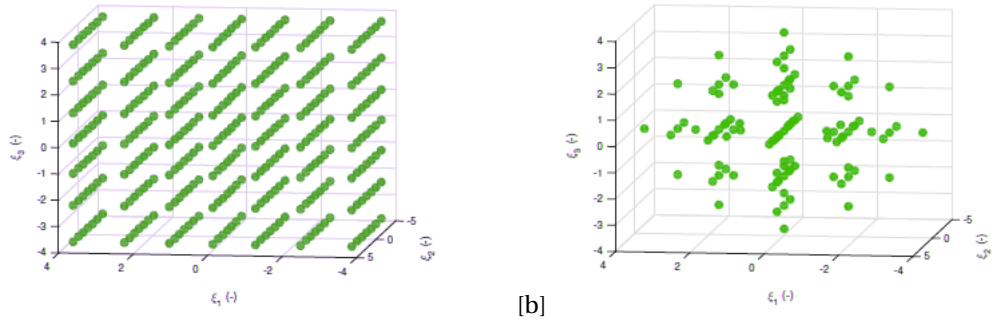
When Hermite polynomials are used as basis vectors, Gauss-Hermite rule's high accuracy is desired. The

Gaussian-Hermite rule is capable to accurately integrate polynomials up to order $4lev$ - 3 and contain n_lev = $2lev$ - 1 quadrature points, for known level. As mentioned only the origin point is reused in all the higher levels, leading to low nestedness.
However, quadrature rules will only be an accurate approximation to a single dimension integral. For Equation 4.11 a multidimensional integral has to be approximated, which is the tensorization of the different quadrature, called cubatures.

-Cubatures
Cubatures are an extension of quadrature to handle multi-dimensional integrals, and derive from quadrature tensorization given by the Equation 4.13:

$$Q_{lev}^{(N_{dim})} f = (Q_{lev_1}^{(1)} \otimes Q_{lev_2}^{(1)} \otimes ... \otimes Q_{(1)}^{(N)}) f \tag{4.13}$$

where lev in this formula is a multi-index indicating the different quadrature levels which each dimension uses. The cubatures points are derived by taking all the possible quadrature point combinations for the different dimensions. Figure 4.2a displays an example of a three-dimensional cube which is constructed when the same quadrature levels are used in all directions.



[a] [b]

Figure 4.2: Cubature points for full grid (a) and Smolyak Sparse grid (b)[85]

However, it is not necessary for all the quadrature levels lev_i to have the same direction. For that reason, a grid notation is used to describe the different dimensions that each quadrature use. For instance, a grid (3,3,3,3) indicates that a level of 3 is used in each direction, while is the grid can be replaced by (4,3,3,3), making the accuracy of the first dimension higher [85]. The cubatures needed to approximate an N-dimensional integral are given by:

$$\begin{aligned} I^{(N)} f &= \int_{D(\theta)} f(\xi) p_\xi(\xi) d\xi \\ &\approx Q_{lev}^N f \\ &= (Q_{lev_1}^{(1)} \otimes Q_{lev_2}^{(1)} \otimes ... \otimes Q_{lev_N}^{(1)}) f \\ &= \sum_i^n f(\xi^{(i)}) w^i \end{aligned} \tag{4.14}$$

Using the cubature methods the expansion coefficients can be approximated, however based on Equation 4.14 the grid leads to many function evaluations and more specifically $\prod_{i=1}^{N_{dim}} n_{lev_i}$ evaluations. To reduce the number of cubature points, Smolyak Sparse Grids are used as shown in Figure 4.2b.

Smolyak Sparse Grids
These grids are based upon the sparsity-of-effects principle, which proposes that lower-order terms are more likely to be dominant in the majority of the models [62]. This means that not all cubature points are equally important, allowing for a higher-order reduction. From the different categories of sparse grids, Smolyak sparse grids are selected for this work, when pencil beam dose engine is used. With $Q_0^(1) f = 0$ the formula is given by:

$$\Delta_{lev}^{(1)} f = Q_{lev}^{(1)} f - Q_{lev-1}^{(1)} f \tag{4.15}$$

Sparse cubature formula for a multi-index level $\vec{lev}$ for cubature tensorization is given by:

$$\begin{aligned}\Delta_{\vec{lev}}^{N_{dim}} f &= \sum l_1 = 1^{lev_1} ... \sum l_N = 1^{lev_N} (\Delta_{lev_1}^{(1)} \otimes ... \otimes \Delta_{lev_n}^{(1)}) f \\ &= \sum \vec{lev} \leqslant GO + N_{dim} - 1 (\Delta_{lev_1}^{(1)} \otimes ... \otimes \Delta_{lev_n}^{(1)}) f\end{aligned} \tag{4.16}$$

where GO indicates the grid order. Increasing the grid order means that quadratures with higher levels should be included, increasing simultaneously the number of function evaluations needed. Table 4.2 shows the number of points in a sparse grid based on Gauss-Hermite rules, indexed by level, for dimension and GO up to 6.

	GO=lev$_{max}$						
N$_{dim}$	**0**	**1**	**2**	**3**	**4**	**5**	**6**
1	1	3	7	15	31	63	127
2	1	5	22	75	224	613	1570
3	1	7	37	161	608	2070	6507
4	1	9	57	289	1268	4994	18076
5	1	11	81	471	2341	10367	41957
6	1	13	109	713	3953	19397	86522

Table 4.2: Number of points is a sparse grid based on Gauss-Hermite rules [64]

-Extended sparse grids
The use of sparse grid allows dealing with the so-called curse of dimensionality, reducing the number of the examined function evaluations and making the calculations feasible within a reasonable time. To improve the accuracy in one direction, we can extend the grid for this single direction (extra level [EL]). Thus, a higher quadrature level will be used for the extended direction, while the multi-dimensional grid remains the same. The number of the additional calculations added by the grid are only 2·lev$_{extra}$·N$_{dim}$.

Hyperbolic trimming
When the dimentionality of the polynomial basis vectors is bigger than the grid order, the expansion coefficients are not accurately approximated [10]. To overcome this limitation, a hyperbolic trimming factor (q) is used to determine the basis vectors which satisfy the quasi norm in Equation 4.17, that will be included.

$$||\gamma_n||_q = (\sum_{i=1}^{N} \gamma_{n,i}^q)^{\frac{1}{q}} \leqslant GO \tag{4.17}$$

The q value range from 0 to 1. For q=1 all the polynomial chaos basis sets will be used.

4.3.2. Regression Approach

For RayStation TPS a MC dose engine is used. When MC noise is present in the dose engine, projection approach for expansion coefficient calculation, is replaced from regression method as it is less sensitive to noise.
For computational purpose the expansion in Equation 4.1 is truncated to a corresponding set $S = \{s_1, ..., s_{|S|}\}$ of multi-indices, given by:

$$R(\xi) = \sum_{n=1}^{|S|} r_n \Psi_{s_n}(\xi) \tag{4.18}$$

For the regression methods, linear least squares are used to determine the coefficients r_n. When the number of regressors $|S|$ is smaller than the number k of observations, r_n is determined by solving the following [65]:

$$r_n = arg_r min \sum_{i=1}^{k} (r^{(i)} - R(\xi^{(i)})) = arg_r min \sum_{i=1}^{k} (r^{(i)} - \sum_{n=1}^{|S|} r_n \Psi_{s_n}(\xi^{(i)})) \tag{4.19}$$

The solution reads:

$$r_n = (S^T S)^{-1} S^T \begin{bmatrix} r_1 \\ r_2 \\ \vdots \\ r_k \end{bmatrix} \tag{4.20}$$

where $S_{i,n} = \Psi_{sn}(\xi^{(i)}), i = 1,...,k, n = 1,...,|S|$. This approach aims to minimize the sum of the squares of the errors generated from the differences in the PCE dose value and the value anticipated from the TPS MC dose engine. This method almost averages out the MC noise from the expansion coefficients. However, linear regression is susceptible to overfitting.

5

Methods

This chapter describes the methods used in this work. First, the construction of the PCE dose model is presented. After, the methods used for validation of the results, obtained from the thesis, are analyzed. Furthermore, the proposed method to determine patient individual robustness settings is explained. Subsequently, the potential of inhomogeneous treatment planning is discussed. Finally, a flowchart for the implementation of a feasible iterative approach towards probabilistic treatment planning is presented.

5.1. PCE construction

PCE provides a computationally efficient patient and plan-specific model of dose, which can be used to quantify the effects of uncertainties on the delivered dose in IMPT. To build the PCE dose model, one must select a dose engine, and define the PCE input parameter settings. In this research, an in-house Matlab toolbox was used for the construction of the PCE model, developed at TU Delft [21]. Thus, this toolbox was coupled for both RayStation and iCycle dose engines used in this study. Figure 5.1 is a flowchart of the PCE construction. Depending on the dose engine selection, different approaches are selected for the approximation of the expansion coefficients.

5.1.1. iCycle

When iCycle dose engine is selected, the PCE construction can be described by the following steps:

- *PCE input parameters*: standard deviation of the uncertainties taken into account, polynomial order (PO), grid order (GO) and extra level (EL) of basis vectors, value of hyperbolic trimming (q), dose cut-off :
  Standard deviations of the uncertainties: The standard deviations of the uncertainties discussed in Chapter 2, are input as a 4 element array, in which systematic (Σ) and random (σ) setup standard deviation, are quadratically summed for each direction. The expression 5.1, consists of the total setup standard deviation for each axis, along with the relative range uncertainty:

$$\left[\sqrt{\Sigma_x^2+\sigma_x^2} \quad \sqrt{\Sigma_y^2+\sigma_y^2} \quad \sqrt{\Sigma_z^2+\sigma_z^2} \quad \rho_{rel}\right] \tag{5.1}$$

  PO: higher polynomial order allows more complex distribution modeling, but also requires higher grid order for accurate calculation of expansion coefficients. Therefore, there is a trade-off between accuracy and calculation time. From the literature review (Appendix), PO equal to five is suggested as the most efficient basis vector polynomial order for skull-base tumor patients.
  GO: the level of quadrature points that are included in the grid is $PO-1=4$
  EL: one extra level is used, for higher accuracy along a single dimension axis.
  q: hyperbolic trimming factor equal to 0.861 is selected (determined from a previous study for the patient cohort used in this research [12])

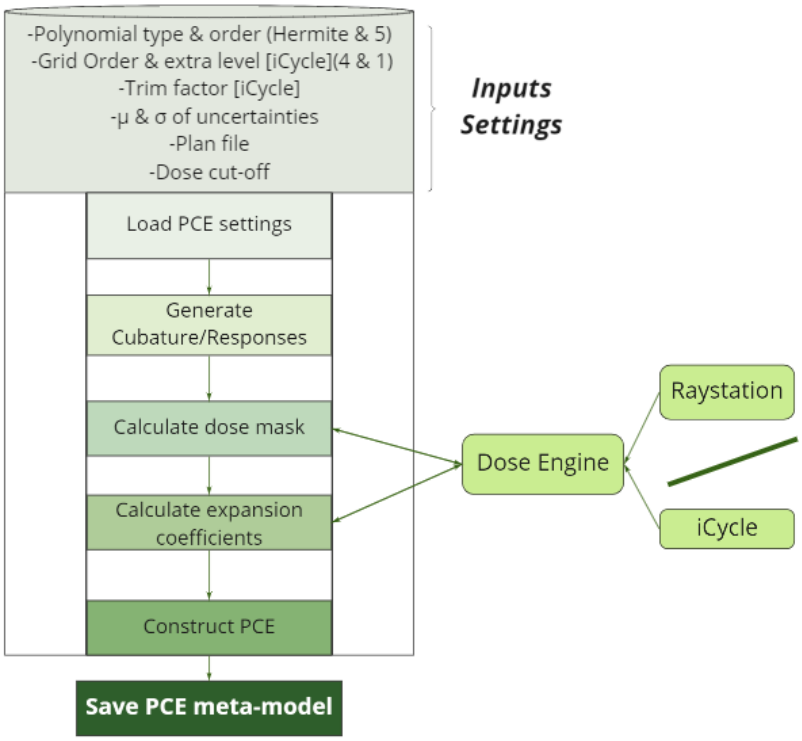


Figure 5.1: PCE construction flowchart, For Raystation a regression approach is followed for approximation of the expansion coefficients. For iCycle projection methods are used

- *Construction of cubatures*:
  Cubatures are generated for the numerical approximation of the integral to obtain the expansion coefficients.
- *Construction of basis vectors*:
  Basis vectors are constructed from all multidimensional polynomials up to a maximum degree PO, and a basis vector hyperbolic trim is performed
- *Determine the voxels included in the PCE model (dose mask)*
  An additional cut-off dose is being set to reduce the number of the considered voxels and make the model computationally efficient. In a previous work [18], Voort et al showed that 95% of the voxels from the planning CT receive almost zero dose. As a result, dose cut-off value is set to exclude voxels that receive dose below 0.01 Gy. This limit is set not only for the dose-related to the nominal scenario, but to all scenarios included in the optimization (19 for iCycle, 21 for RayStation). The dose is calculated for the voxels in every scenario, using the coupled TPS (iCycle or RayStation). This results in a logical array, called dose mask, which indicates if a voxel should be included or not in the PCE model.
- *Calculate expansion coefficient*:
  The cubature points are used to construct the error scenarios, which will be used as an input in the dose engine, to calculate the dose responses for all the voxels included in the dose mask. The PCE coefficients are calculated using cubature rules applied to these dose responses.

5.1.2. RayStation

As discussed in Chapter 4, when an MC dose engine is used, the projection approach for calculating the expansion coefficient is replaced by a regression method, due to the presence of MC noise. With the regression method different scenario sampling approaches could be used for the PCE construction, however, currently, the construction method described previously is used. The only difference is that instead of cubature rules, interpolation is used for approximating the multi-dimensional integral of Equation 4.9.
Finally, the PCE model is saved and it is able to calculate dose distributions for any patient error scenario. The model works as a function in which the uncertainty scenario is imported, in a format identical to Equation 4.11.

5.2. Validation of PCE

After the construction of the PCE model, they are validated with different methods in comparison with the TPS dose engine. The validation tools used for this research are discussed in this section.

5.2.1. Dose Volume Histograms

Dose volume histogram (DVH) is a cumulative histogram that shows the fraction of the volume which receives a certain dose. DVH's are frequently used in clinical practice as they can be constructed for different structures, determining the relationship between radiation dose to the tissue volume. Almost always the clinical goals, both for targets and OARs are expressed in terms of DVH parameters. DVHs from different PCE dose models, for the same error scenario, can compare to the actual TPS dose distributions, and therefore give information of PCE model accuracy. The DVH can be constructed for dose comparisons in both nominal scenario and error scenarios.

5.2.2. DVH dependencies

An additional method to evaluate PCE models' accuracy for the CTV, is to study treatment dose parameters ($D_{2\%}$, $D_{50\%}$, $D_{98\%}$), as a function of error shifts. More specifically the dosimetric parameters from different PCE models are compared to the actual TPS dose engine calculation, for a number of shifted scenarios. Using this method the DVH dosimetric dependencies can be studied immediately under the shift scenarios.

5.2.3. Absolute Dose Differences (2-norm Relative Error)

A voxel-dose error can be determined by calculating the differences between the dose distributions under shifted scenarios and the corresponding dose from the TPS dose engine. However, this method is sensitive when large dose gradients are considered. Instead of using the differences between the PCE dose distribution and the TPS dose, a 2-norm relative voxel dose evaluation can be used. The 2-norm relative voxel dose (%) is given from Equation 5.2:

$$RN(\%) = 100\% \cdot \sqrt{\frac{\sum_{i=1}^{i=N}(D_{PCE} - D_{TPS})^2}{\sum_{i=1}^{i=N} D_{TPS}^2}} \tag{5.2}$$

where N is the number of scenarios, for which the PCE's voxel dose is calculated and validated against the TPS dose.
The relative norms will be plotted against different shifts, to determine the accuracy of each PCE model.

5.3. Optimization of PCE construction

The first aim of this work is to validate PCE performance, regarding not only the geometrical and range uncertainties assumed, but also the impact of the MC noise level from the clinical dose engine. Different PCE dose models are generated in order to study the MC noise impact. For this work, thirty shifted scenarios are considered, including shifts along the main axes, where PCE is optimized during construction, and also additional off-axis shifts [84]. Finally, a proposed PCE construction method, based on voxel dose differences, is implemented.

Iterative use of PCE in treatment planning requires considerable time resources. Therefore this research attempts to speed up the existing PCE model construction, to facilitate methods towards clinically feasible probabilistic treatment plans. The goal is to investigate the impact of stochastic Monte Carlo (MC) noise in PCE construction for proton voxel dose calculation. The objective is to reduce the computational cost and speed up the construction of PCE, by using a larger MC noise level (2 and 3%), while maintaining model's accuracy. Previous study results [84], suggested that there is space for further investigation around the default value of MC noise (1%), that is currently used in research. Depending on the number of uncorrelated treatment uncertainties taking into account, and the tumor's size (more voxels receiving dose), the time for PCE construction varies, and a further reduction will be beneficial for clinical practice. To evaluate the performance of PCE built with different MC noise levels, a comparison of DVH and DVH dependencies for these models against the TPS dose engine will be made.

5.3.1. MC noise impact

For patients treated at HollandPTC for meningioma, grade-I glioma, grade II-III oligodendroglioma with 1p/19q codeletion, and grade-II astrocytoma with IDH-mutation between September 2018 and September 2019, **one** patient treated to 54 GyRBE in 30 fractions using MC 1% dose engine were randomly chosen for validation, from the patient cohort used by Rojo-Santiago et al [84].
The PCE performance under different MC (1%, 2%, 3%) noise level, was determined by DVHs comparisons [84]. We are considering the nominal scenario (free of geometrical errors) and range error of 1.2%, to account for the systematic error of SECT calibration. Ideally, PCE constructed with MC noise 0.1% could describe the exact TPS dose. However, this is not feasible, due to the required model's construction time. The clinical TPS (RayStation) is equipped with MC dose engine, therefore it is selected. Figure 5.2 shows the main steps followed for model's performance validation.

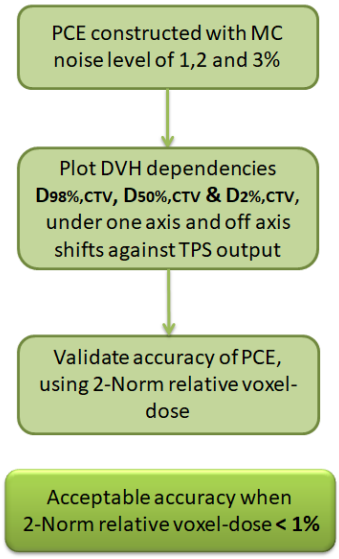


Figure 5.2: Flowchart of accuracy validation of PCE models built with different MC noise level

The relative voxel dose differences will be given as a function of single-axis shift (Left-right, dorsoventral, craniocaudal), and also for combined shift (X=Y, Y=Z, Z=X) directions. For the selected patient, the DVH dependencies $D_{2\%,CTV}$, $D_{50\%,CTV}$ and $D_{98\%,CTV}$ are calculated for 30 shift scenarios.

5.3.2. PCE construction using dose differences

PCE is currently mainly used for anatomies with a steep DVH gradient. In this section, an approach to account for more complex anatomies is implemented. The goal was to construct an accurate PCE model, using the dose differences between the nominal and the shifted doses. The method is expected to improve the model's ability to reconstruct the areas that a larger dose gradient is observed, in which the dose differences are larger. To achieve this the assumption of the static dose cloud approximation is used, and we interpolate using the nominal TPS scenario, for all the geometrical directions and different scenarios. To do this, a calculation of the difference between the dose from PCE MC 1%, and the interpolated dose from the nominal TPS scenario for each of the scenarios, is made. The PCE is then constructing based on those differences.

To evaluate the model's performance, a DVH comparison is made with PCE MC 1%, for the nominal and different geometrical shifted scenarios.

5.4. Optimal Robustness settings

The second aim of this work is to use PCE evaluations to quantify the effect of errors on dosimetric parameters, and to use this information to determine the optimal robustness settings for each patient.

5.4.1. Patient set

The patients from two different tumor groups were selected. More specifically, a complex skull-base chordoma, and a straightforward skull-base meningioma patient. Chordoma tumors are characterized as complex clinical cases, due to high prescribed dose to the CTV, which is well above any tolerance of a healthy brain.

-Skull base chordoma

RayStation TPS was used to study a skull base chordoma patient. This treatment site is considered challenging due to the high prescription dose to the CTV close to hard constraint OARs. The objectives and the constraints for this patient are given in Table 5.1.

Structure	Criterion
CTV_{high}	$D_{98\%,y} \geqslant 0.95 \cdot 74Gy = 70.3Gy$
CTV_{high}	$D_{2\%,y} \leqslant 1.07 \cdot 79.2Gy$
CTV_{low}	$D_{98\%,y} \geqslant 0.95 \cdot 59.2Gy = 56.24Gy$
CTV_{low}	$D_{2\%,y} \leqslant 1.07 \cdot 59.2Gy = 63.34Gy$
	Objectives
Brainstem Surface	$D_{max} \leqslant 60Gy$
Brainstem Core	$D_{max} \leqslant 54Gy$
Spinal Cord Surface	$D_{max} \leqslant 60Gy$
Spinal Cord Core	$D_{max} \leqslant 50Gy$
Optical Nerves	$D_{2\%} \leqslant 60Gy$
Chiasm	$D_{2\%} \leqslant 60Gy$
Hippocampi,combined	$D_{40\%} \leqslant 60Gy$
Pituitary gland	$D_{mean} \leqslant 30Gy$

Table 5.1: Plan Constraints & Objectives for Skull-base Chordoma patient

Skull-base meningioma

Similarly, RayStation TPS was used to validate the results against the non-robust chordoma case. The objectives and the constraints for this patient are given in Table 5.2.

Structure	Criterion
CTV	$D_{98\%,y} \geqslant 0.95 \cdot 50.4Gy = 47.88Gy$
CTV	$D_{2\%,y} \leqslant 1.07 \cdot 53.93Gy$
	Objectives
Chiasm	$D_{2\%} \leqslant 55Gy$
Optical Nerves	$D_{2\%} \leqslant 55Gy$
Brainstem	$D_{max} \leqslant 55Gy$
Hippocampi,combined	$D_{40\%} \leqslant 7.3Gy$

Table 5.2: Plan Constraints & Objectives for Skull-base meningioma patient

5.4.2. PCE-based robustness evaluation

For each patient in this study, treatment plans, with different Setup Robustness, were created. The plan file along with the corresponding TPS, was used to build a PCE model for each case, using the clinical uncertainties as an input scenario. Table 5.3 shows the values of the clinical uncertainties, as reported in the literature for brain tumors [84].

The systematic geometrical error is derived from errors in CT isocenter, MR registration, and residual error,

Systematic Geometrical Mean (1 SD)	[0 0 0] (0.74,0.74,0.74)
Random Geometrical Mean (1 SD)	[0 0 0] (0.83,0.83,0.83)
Range (1SD)	1.2 (1)

Table 5.3: Clinical Uncertainties for skull base tumors (1SD, refers to standard deviation)

while the random component is a result of gantry isocenter, couch position, and online matching errors. As mentioned above the range error is a systematic undershooting error to account of SECT calibration.
The effects of setup and range errors on CTV dose distribution for each plan are observed using PCE evaluation for different input standard deviations in the model. More specifically, uncertainty scenarios are used, in which the standard deviation of random setup error of the x-axis, was changing with a step of 0.5 from 0.5 to 3. The format of the error scenarios used for evaluation is given in Table 5.4:

Systematic Setup	(0.74,0.74,0.74)
Random Setup	$(\sigma_x, 0, 0)$
Range	1

Table 5.4: Error scenarios used for skull base tumors ($\sigma_x : 0.5 : 0.5 : 3$)

-α

We define as α (alpha), an overall scaling factor, based on the assumptions that (i) the actual margin recipe is approximately linear and (ii) that the real underlying robustness recipe is a scaled version of van Herk's formula 3.2. We used PCE to estimate the scaling factor. To study a plan's robustness, α is given by the multiplies of the clinical setup error vector. The clinical uncertainties, which consists of systematic error $(\Sigma_x, \Sigma_y, \Sigma_z)$ = (0.74,0.74,0.74) and random error $(\sigma_x, \sigma_y, \sigma_z)$ = (0.83,0.83,0.83), correspond to α equal to one. The previously introduced alpha multiples are going to be used for PCE evaluation of different treatment plans. Table 5.5 show the different values of alpha that are used for evaluation.

α	Systematic Component $(\Sigma_x, \Sigma_y, \Sigma_z)$	Random Component $(\sigma_x, \sigma_y, \sigma_z)$
0.5	(0.37,0.37,0.37)	(0.42,0.42,0.42)
1	(0.74,0.74,0.74)	(0.83,0.83,0.83)
1.25	(0.93,0.93,0.93)	(1.04,1.04,1.04)
1.5	(1.11,1.11,1.11)	(1.25,1.25,1.25)
1.75	(1.3,1.3,1.3)	(1.45,1.45,1.45)
2	(1.48,1.48,1.48)	(1.66,1.66,1.66)

Table 5.5: Values corresponding to α

Using PCE dose model, for each patient 100,000 complete fractionated treatments were simulated. Depending on the objective (of each section), for every treatment plan. Different uncertainty scenarios are used, for PCE evaluations, each time:

- systematic setup error
- random setup error
- systematic range error
- alpha (multiplies of clinical uncertainties)

The resulting 100,000 dose distributions, are evaluated to obtain population dose distributions of $D_{98\%,CTV}$ and $D_{2\%,CTV}$, along with dosimetric dependencies related with the OARs ($D_{0.03cc}$). More specifically, the mean dose, 2 and 98 percentile of $D_{98\%,CTV}$, $D_{2\%,CTV}$ or $D_{0.03cc}$, will be plotted as a function of the uncertainties scenarios. The method is summarized in Figure 5.3

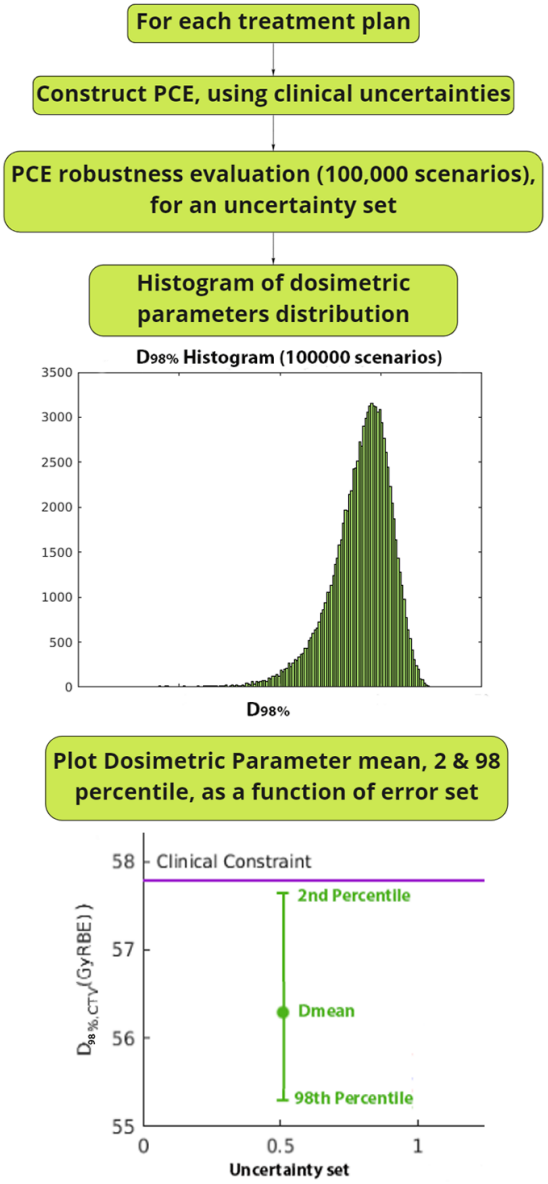


Figure 5.3: Flowchart used for robustness comparison of plans with different robustness settings

5.4.3. Rescaled Dose

to compare plan quality for the complex skull-based chordoma patient, a dose rescaling is recommended. Almost none of the plans meets the CTV criteria when evaluating under random setup uncertainties. Rescaling the plans as such 98% of $D_{98\%,CTV}$ ($D_{98\%,min}$) meets the clinical constrain, using different error scenarios, allows to distinguish which plan is superior to others, and also till which error scenario this is true. The Scaling Factor (SF) is calculated from:

$$SF = \frac{0.95 D_{pres}}{D_{98\%,min}} \tag{5.3}$$

When plan has SF less than one is considered robust. However, after evaluating the different treatment plans (clinical plan for different SR & reconstructed plans from scratch), using PCE, we did not manage to achieve a robust plan due to the trade-off between target coverage and OARs. Therefore this research changed focus to only skull-base meningioma patients.

5.5. Inhomogeneous planning

Five patients, from the previous patient cohort of skull-base meningioma, were used to investigate the trade-off between homogeneity of target dose and robustness. Erasmus-iCycle was used for this part of the research, as it enables rapid prototyping and testing of stereotactic planning.

5.5.1. iCycle Wish-list

From a previous work [12], the required wish-list was available for this patient cohort. Table 5.6 reports the used values.

Priority	**Structure**	**Priority**	**Goal**
Constraint	CTV	Maximum	0.9850.4 Gy
Constraint	Optical Nerve$_{Left}$	Minimum	0.9855 Gy
Constraint	Optical Nerve$_{Right}$	Minimum	0.9855 Gy
Constraint	Optical Chiasm	Minimum	0.9855 Gy
1	CTV	Minimum	1.0750.4 Gy
2	CTV$_{ring_{2-5mm}}$	Minimum	0.850.4 Gy
2	CTV$_{ring_{5-7mm}}$	Minimum	0.6550.4 Gy
2	Hippocampi	Minimum	7 Gy
3	CTV$_{ring_{5-10mm}}$	Minimum	0.550.4 Gy
4	Hippocampi$_{left}$	Minimum	1 Gy
4	Hippocampi$_{right}$	Minimum	1 Gy
5	Brainstem	Minimum	55Gy
5	Cerebellum	Minimum	35Gy
6	Cochlea$_{right}$	Minimum	45Gy
6	Cochlea$_{left}$	Minimum	45Gy
7	Retina$_{right}$	Minimum	20Gy
7	Retina$_{left}$	Minimum	20Gy
8	Lens$_{right}$	Minimum	10Gy
8	Lens$_{left}$	Minimum	10Gy
9	Lacrimal Gland$_{right}$	Minimum	26Gy
9	Lacrimal Gland$_{left}$	Minimum	26Gy

Table 5.6: Wish-list Constraints & Objectives for Skull based meningioma patients

For each patient different beam angles were selected.

5.5.2. Maximum Target dose

Previous study results from Petit et al. [59] on lung and liver patients, showed a robustness gain, when inhomogeneous dose delivery is used in the treatment plan, when evaluating for different range errors. Similarly, we considered five skull-base meningioma patients and focused on the range uncertainties. When inhomogeneous dose distribution is an objective the maximum allowed dose to the CTV increases to 120% of the prescribed dose (50.4 Gy

The method is described by the following consecutive steps:

- We obtain treatment plans for both homogeneous and inhomogeneous dose delivery
- Construct the DVH of the different treatment plans
- Calculate the mean dose to the OARs
- PCE dose model constructed for all the treatment plans
- PCE evaluations using clinical setup uncertainties and range error from 0 to 4%, as showed in Table 5.7.

Systematic Setup	(0.74,0.74,0.74)
Random Setup	(0.83,0.83,0.83)
Range	ρ_{rel}

Table 5.7: Clinical Setup Uncertainties & range error $\rho_{rel}(0-4\%)$

- obtain population dose distributions of $D_{98\%,CTV}$ and $D_{2\%,CTV}$, like the previous section

5.6. Optimal Margin

Chapter 3 introduces the advantages of probabilistic treatment planning. In the last section of this work, a step towards probabilistic treatment planning will be attempted. The proposed method aims to determine the minimum optimal margin for which a treatment plan is robust against the stochastic variable of minimum dose.
In photon therapy the optimal margin is selected using van Herk's formula, which linearly connects the margin with the standard deviations of the clinical setup uncertainties. Based on this, using PCE evaluations to determine the 98% population coverage of the $D_{98\%}$ ($D_{98\%,min}$), we aim to examine if we can determine an optimal margin (SR_{opt}) which meet the objective, using only the multiples of the clinical uncertainties (α) and the plan's margin, shown in Figure 5.4. If we use the Figure 5.4 as a reference, the goal is having only the results from iterative evaluations of the red plot (right), to determine the optimal margin value (left).

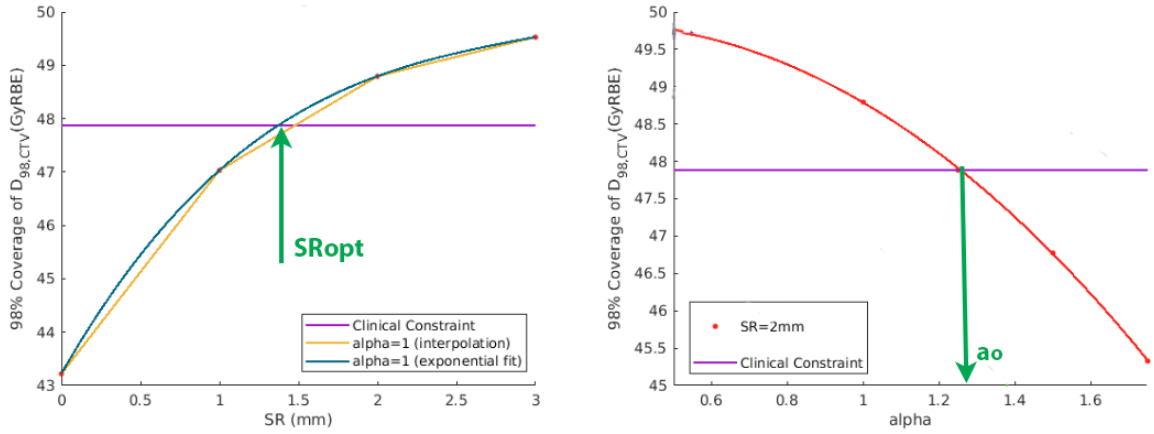


Figure 5.4: Example of the plots used in this method, along with the variables of interest

In the first plot of Figure 5.4, four plans with different setup robustness are optimized for each patient, using iCycle TPS. Next, PCE models are constructed and the plan is evaluated for the CTV. Finally, the $D_{98\%,min}$ is plotted against the setup robustness, and the margin for which the plan meets the clinical constrain is determined (SR_{opt}). For each patient, the average time required for this diagram is estimated at 20 hours when the treatment plans run in order, and a minimum of 7 hours, if the computational requirements are available (plans run in parallel).
The second plot in Figure 5.4, uses only one treatment plan. The PCE model is constructed, followed by plan's evaluation. However, in this plot PCE is used to evaluate for additional multiplies scenarios of clinical uncertainties. Finally, the $D_{98\%,min}$ is plotted against the multiplies alpha. To validate the observations the

procedure is repeated for plans with different SR. Figure 5.5, illustrates a scheme with the description of the followed strategy.

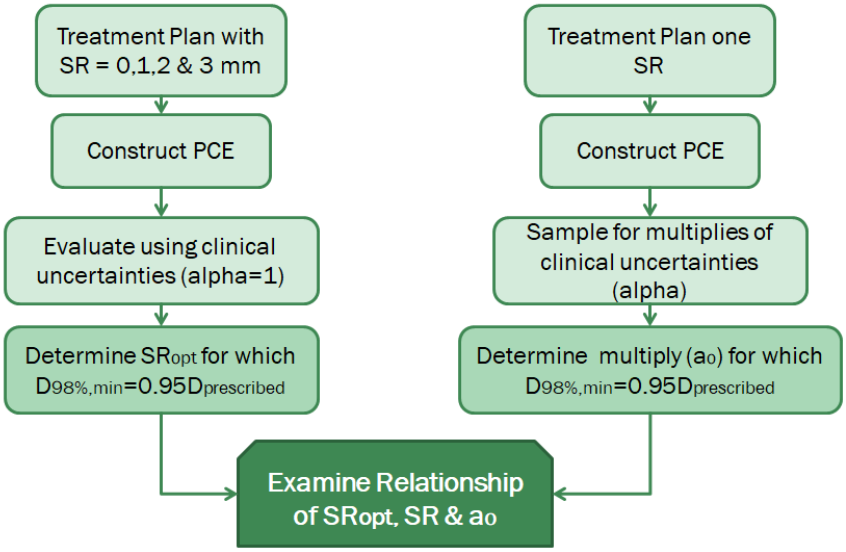


Figure 5.5: Determine the optimal margin flowchart

Only integer SR is used, as the current version of iCycle does not support a non-integer SR. Because of that the behavior of the margin which meets the clinical constraints in the graphs cannot be evaluated.

6

Results

This chapter starts with the results of PCE construction time optimization, using different levels of Monte Carlo (MC) noise. After the construction of PCE meta-model for complex anatomies is attempted. Subsequently, the effect of SR in plan's robustness at patient individual level is examined, using PCE coupled with RayStation TPS. Thereafter, for five skull base meningioma patients the effect of dose inhomogeneity in the plan's robustness is presented. The final section of this chapter shows the first results towards clinically feasible probabilistic treatment planning, using PCE-based robustness evaluation in a loop with a clinical treatment planning system.

6.1. Validation of PCE construction

This section starts with the comparison of PCE models, constructed with different MC noise levels, against the dose distribution from the RayStation clinical TPS. The impact of the MC noise is illustrated through DVHs and 2nd relative norm voxel dose evaluation. Additionally, the results from the PCE model constructed with dose differences, for a neurological patient, are presented.

6.1.1. MC noise impact in DVH

Figure 6.1 illustrates the DVHs for CTV and the relative OAR for the nominal scenario and one off-axis shifted scenario for MC 1,2 and 3% against the TPS dose. The plots are focus on higher doses between 45 and 60 Gy(RBE), to emphasize the CTV dose distribution tail.

The resulting diagrams indicate that larger deviation occur when MC noise level increases, which is more noticeable in the tails of the CTV dose distributions, and also at the OAR's dose distributions. Therefore, the dosimetric parameters $D_{98\%,CTV}$ and $D_{2\%,CTV}$ are expected to be more sensitive to noise. For the shifted scenarios, OAR's tend to be less robust than CTV for all levels of noise. However, PCE with a noise level of 1% appears almost unaffected by the shifts. Moreover, for the nominal case, all three models appear capable to describe the TPS doses. An example of a shifted scenario (total shift of 2.9 mm) also indicates that CTV dose distribution remains accurate for every model. The DVH comparison showed that PCE MC1% outperforms the models constructed with higher MC noise levels.

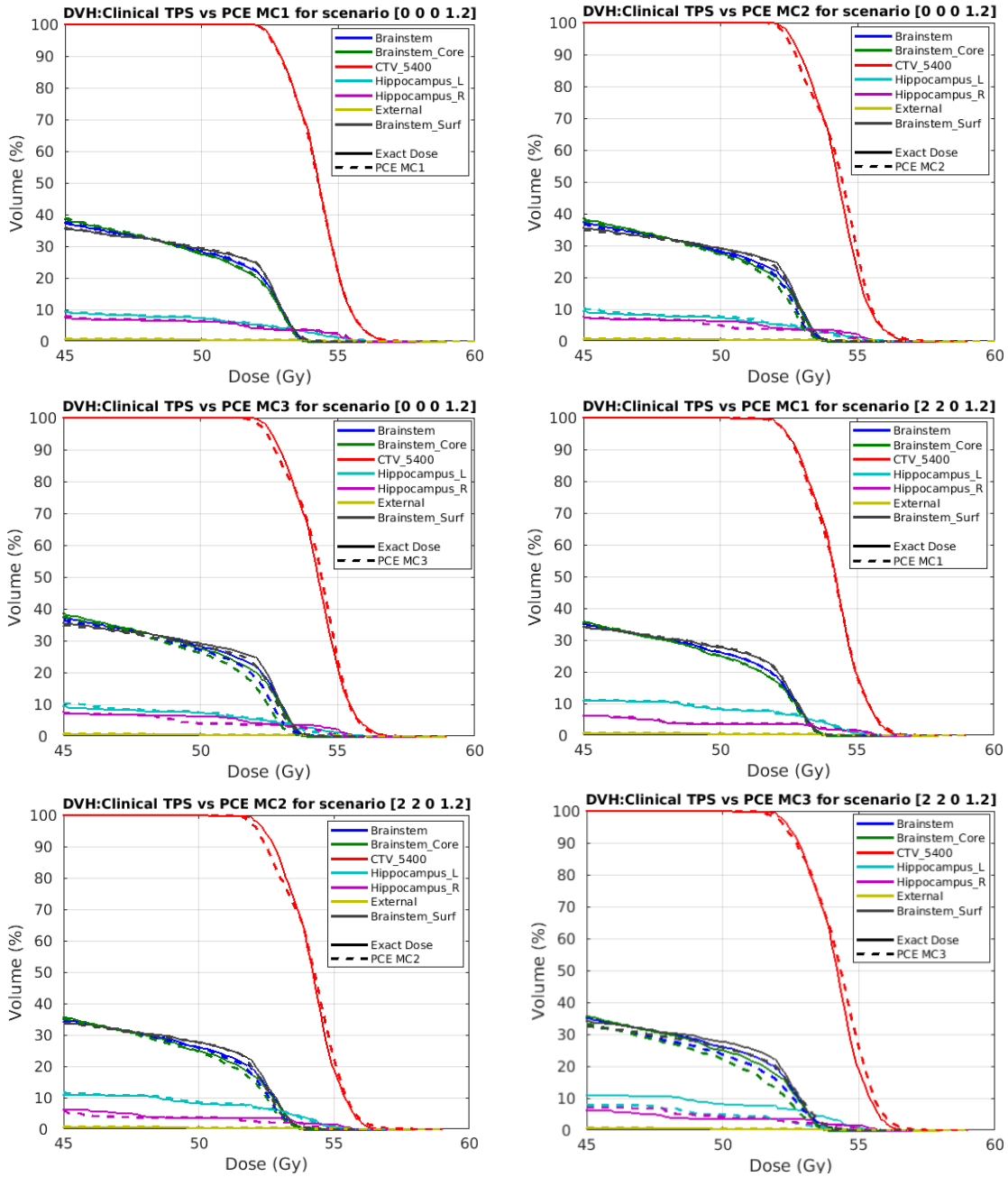


Figure 6.1: DVH diagrams for TPS dose engine and PCE dose model constructed with different MC noise levels (1%, 2% and 3%) for nominal case (top) and an off axis shifted scenario (bottom) with range uncertainty of 1.2%

None of the models shows a large deviation from the exact CTV dose. Therefore further analysis on the dosimetric parameters is suggested. To validate the noise influence in the PCE model, the accuracy of the voxel-dose for the dosimetric parameters will be evaluated for different shifted scenarios along the main patient axes, and additional off-axis scenarios, for which the PCE is not optimized during construction.

6.1.2. DVH Dependencies

The first thing to compare is the nominal dose distribution obtained by the exact dose engine and by the PCE, for geometrical shifts sampled in both the main directions and off-axis directions. As mentioned in Figure 6.1, the tails of the DVH distribution, appear to be more sensitive to MC noise. To investigate the effect of the noise in the CTV dosimetric parameters, the behavior of $D_{2\%}$, $D_{50\%,CTV}$ and $D_{98\%,CTV}$ under geometrical errors are studied.

One axis shifts:
For the selected patient, the DVH dependencies $D_{2\%,CTV}$, $D_{50\%,CTV}$ and $D_{98\%,CTV}$ are calculated for 30 shift scenarios among x, y, and z-directions. The shifts range from -4.2 mm to 4.2 mm at craniocaudal, lateral, and dorsoventral direction each time, with an additional range uncertainty of 1.2%. Figure 6.2 displays each model's performance against one axis shift.

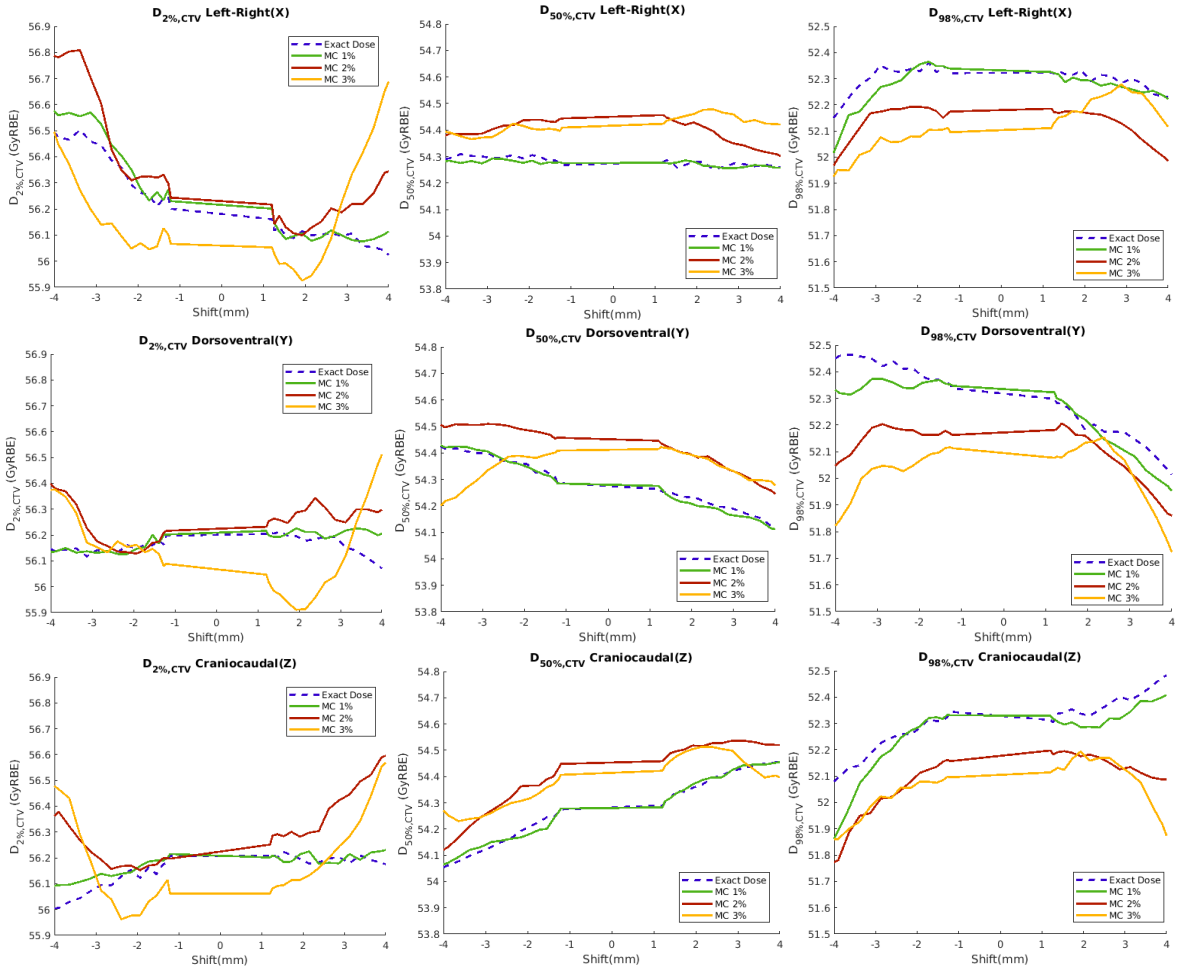


Figure 6.2: $D_{2\%,CTV}$, $D_{50\%,CTV}$ and $D_{98\%,CTV}$ dose dependencies calculated from the TPS dose engine and the PCE model. From top to bottom, the DVH dependencies are shown with PCE models constructed with MC noise levels of 1%, 2% and 3% respectively

From Figure 6.2, PCE MC1% outperforms in all cases the other models. For higher MC noises an overestimation of $D_{50\%,CTV}$ and $D_{2\%,CTV}$ (with exception of MC3%) and underestimation of $D_{98\%,CTV}$ values is visible in all directions. Also, the plots highlight that by increasing the MC noise, the dosimetric parameters fluctuation is increased. Additionally, the plots for $D_{50\%,CTV}$ and $D_{98\%,CTV}$ showed that PCE built with MC noise 2%, has more similar behavior with the model constructed with 3% noise, rather than with PCE MC 1%, for these dosimetric parameters. However, the plot of $D_{2\%,CTV}$, indicates similar performance for PCE constructed with 1 and 2 % level noise. Finally, as expected, $D_{50\%,CTV}$ appears more robust against the shifts, as is not determined from the dose distribution tails.

Using the 2-norm relative voxel dose evaluations when taking into consideration all voxels which received dose above 1 Gy(RBE), the absolute dose differences are plotted in Figure6.3(up). The observation is that PCE constructed with MC 1%, provides dose distributions that are accurate below a 1% voxel-dose error within 3 mm scenario shift range for shifts along the main axis, and is slightly increases for shifts till 4mm. With accuracy below 3.9% in voxel-dose error, PCE MC2% is following, while MC3% indicates a weak performance even in minimum shifts.
From the same Figure 6.3 (bottom), when only CTV's voxels are considered, both PCE MC1% and PCE MC2% maintain an relative error below 1%.

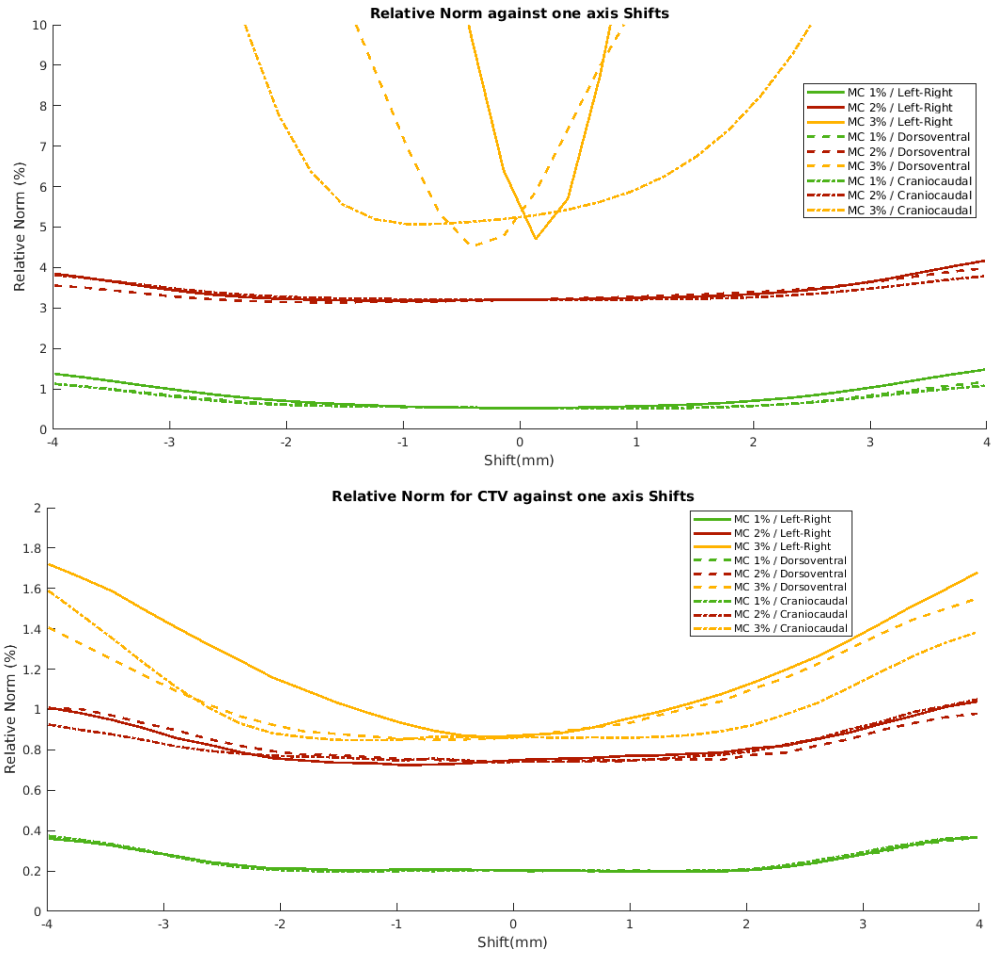


Figure 6.3: 2-Norm relative voxel-dose against one axis shifts for all dose voxels receiving dose larger than 1 Gy (up) and 2-Norm for CTV voxels (bottom)

Off-axis shifts:
For a complete evaluation of model's robustness under different shifts, the evaluations were repeated for combined shift directions. The combined total shifts range up to ±5.5mm. More specifically an equal shift for two-axis was used, for all axis combinations. The DVH dependencies are plotted in Figure 6.4.

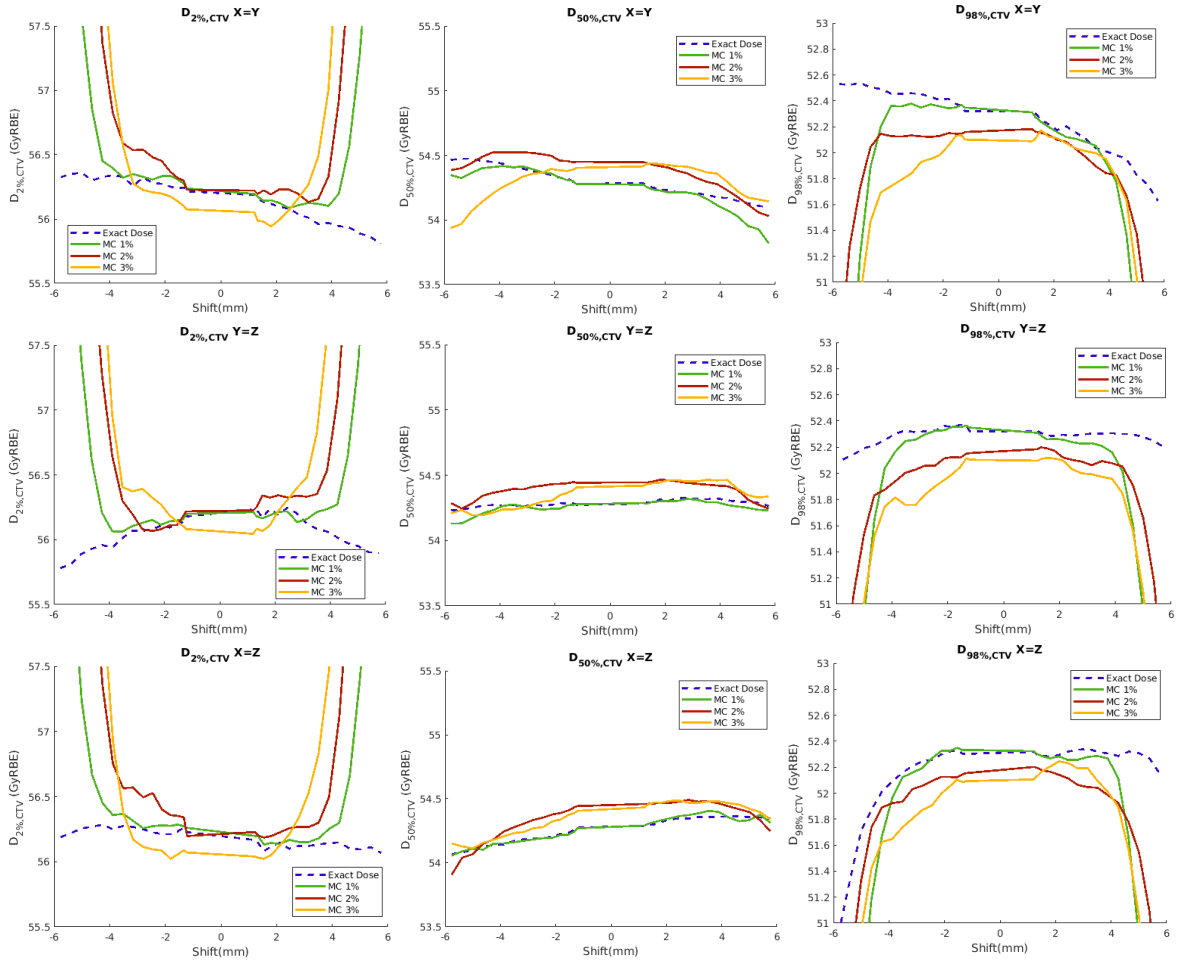


Figure 6.4: $D_{2\%,CTV}$, $D_{50\%,CTV}$ and $D_{98\%,CTV}$ dose dependencies calculated from the TPS dose engine and the PCE model. From top to bottom, the DVH dependencies are shown with PCE models constructed with MC noise levels of 1%, 2% and 3% respectively for combined shift directions (X=Y,Y=Z,X=Z)

From Figure 6.4, the PCE dose distributions seem to be more unstable than for patient scenarios that contain one direction shift. However, the overall distribution behavior of the models appears similar. Due to the magnitude of the total uncertainties vector, a total shift larger than 4mm is reached, causing a steep increase of the dose parameters in the tales of the plots. However, errors larger than 4mm are not expected in clinical practice for neurological tumors (probability $< 10^{-4}$) [84].

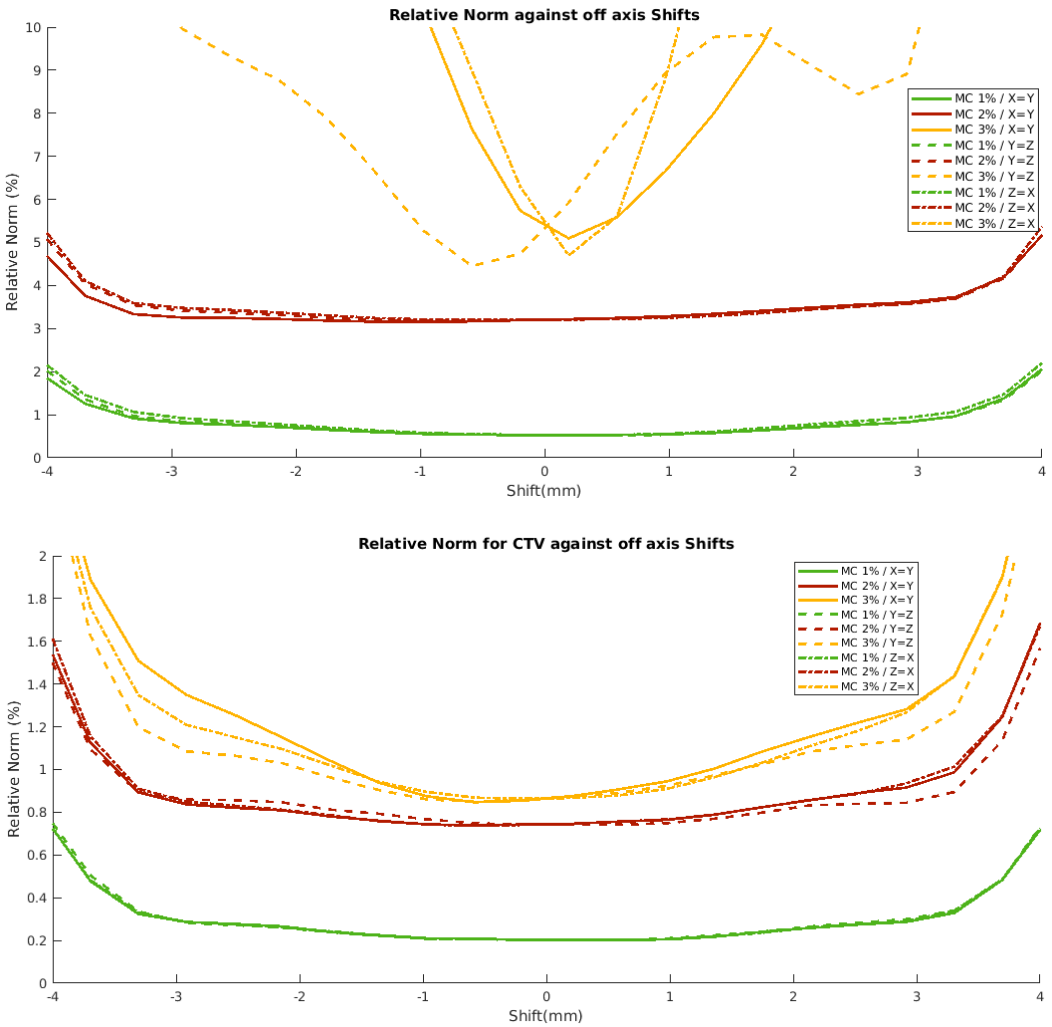


Figure 6.5: 2-Norm relative voxel dose against off axis shifts, for all dose voxels receiving dose larger than 1 Gy (up) and 2-Norm for CTV voxels (bottom)

Figure 6.5 suggests that PCE is less accurate when the geometrical error is in a combination of directions. Simultaneously, the decrease in model accuracy appears earlier around ±3mm, while after an absolute shift of 4mm all PCE models lose their ability to accuretly simulate TPS dose. When all the voxels, which receive dose larger than 1 (GyRBE), are considered only PCE MC 1% is below the limit error of 1%.

6.1.3. Dose Differences PCE

This section shows how the PCE model performs when dose differences have been used for its construction. In this case, the construction of PCE was slightly more time-consuming, due to the additional interpolation of the nominal dose distribution using the dose invariant assumption, which needed 0.5 hours more for the PCE construction using the same patient data compared to the previous PCE configuration. Dose volume histograms were used to the dose distribution for the CTV and the important structures. Figure 6.6 shows the DVHs of the nominal dose distributions of both the exact dose engine and the dose differences PCE. From the figure, we observe that the dose distributions overlap for every structure. However, more the dose distributions under shifts are also examined to make safer conclusions for the model's potentials.

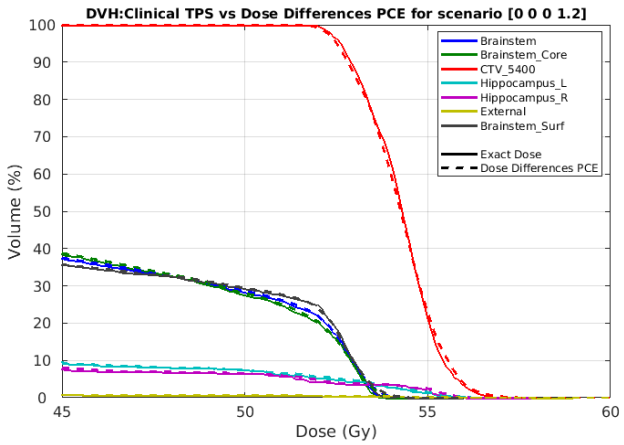


Figure 6.6: DVH diagrams for PCE dose model constructed with dose differences(dashed) and TPS Dose Engine(solid)

In order to evaluate the model's performance, a DVH comparison is made with PCE MC 1% for different geometrical shift scenarios. The results from Figure 6.7 (up), indicate that the meta-model build with PCE dose differences performs worse than PCE MC 1% in CTV accuracy, for small one-direction shifts. In the case of combined geometrical error, with a total shift of 2.5mm, the difference between the two models' accuracy for CTV dose distribution, becomes wider.

From Figure 6.7, PCE dose differences model seems to be less precise than expected. The model simulates poorly the dose distributions for the OARs, especially for larger shifts. While the PCE MC 1% continues to be aligned completely with the output line of the TPS dose engine. Finally, four scenarios of positive and negative shifts have been displayed, to highlight that the model performs differently for the same absolute error. For instance, positive shifts affect more model's performance for the left hippocampus, while negative shifts have a larger effect on right hippocampus accuracy. This asymmetry is also noticeable in the Figure 6.2 of the previous section, and it is probably related to the location of the tumor volume (e.g. positive shifts might be closer to the beams).

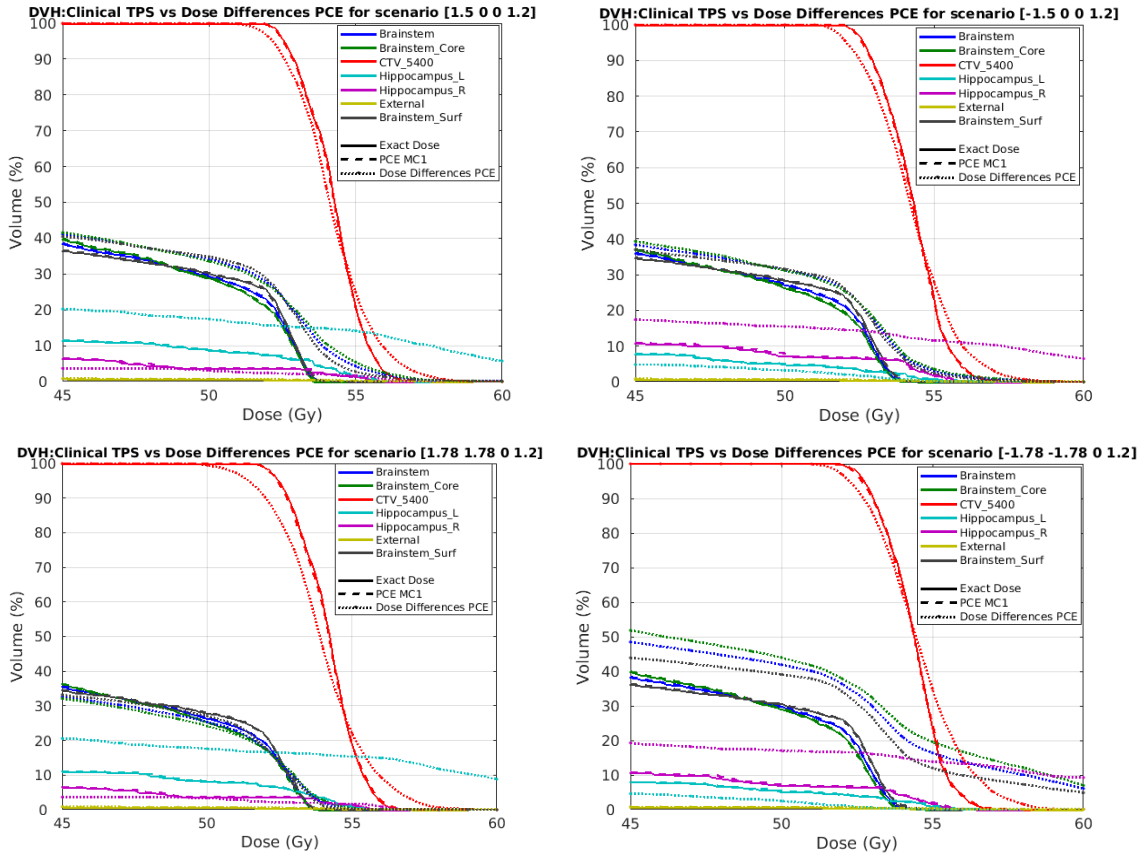


Figure 6.7: DVH diagrams for 4 shifted scenarios of PCE constricted with dose differences compared with PCE MC 1% and TPS dose engine

6.2. Reduction of Setup Robustness

In this section, the results of different PCE evaluations for treatment plans with different SR are presented. First, the evaluation results for a robust skull-based meningioma patient are shown, as discussed in the methods section. After, the dosimetric parameter graphs for the non-robust skull-base chordoma patient are presented, along with plans' validation using the scaling factor.

6.2.1. Robust Skull-base meningioma patient

Plans with three different margins (SR = 1, 2, and 3mm), using RayStation dose engine, were used for PCE evaluation under geometrical uncertainties. The dosimetric dependencies used in this section are the CTV near-minimum ($D_{98\%}$) and near-maximum ($D_{2\%}$) dose. From Figure 6.8 we observe the behavior of each plan, using PCE evaluations. The graphs, display the average value of each dosimetric parameter with a dot, while the 2% and the 98% percentiles are the top and the bottom value of each error bar, respectively. The first two graphs (up) correspond to PCE evaluations under different random setup x standard deviations. The other two plots illustrate the treatment plans behavior under alpha uncertainty scenarios. Figure 6.8, indicates that the clinically suggested SR=3mm showed similar robustness with the treatment plans made with SR=2mm. More specifically, replacing the margin from 3 to 2 mm for both cases (σ_x, alpha), the desired tumor coverage is lost at approximately the same uncertainty scenario. The SR=1mm, was used to indicate that the margin cannot be infinitely small, as the robustness is lost for lower uncertainty values.

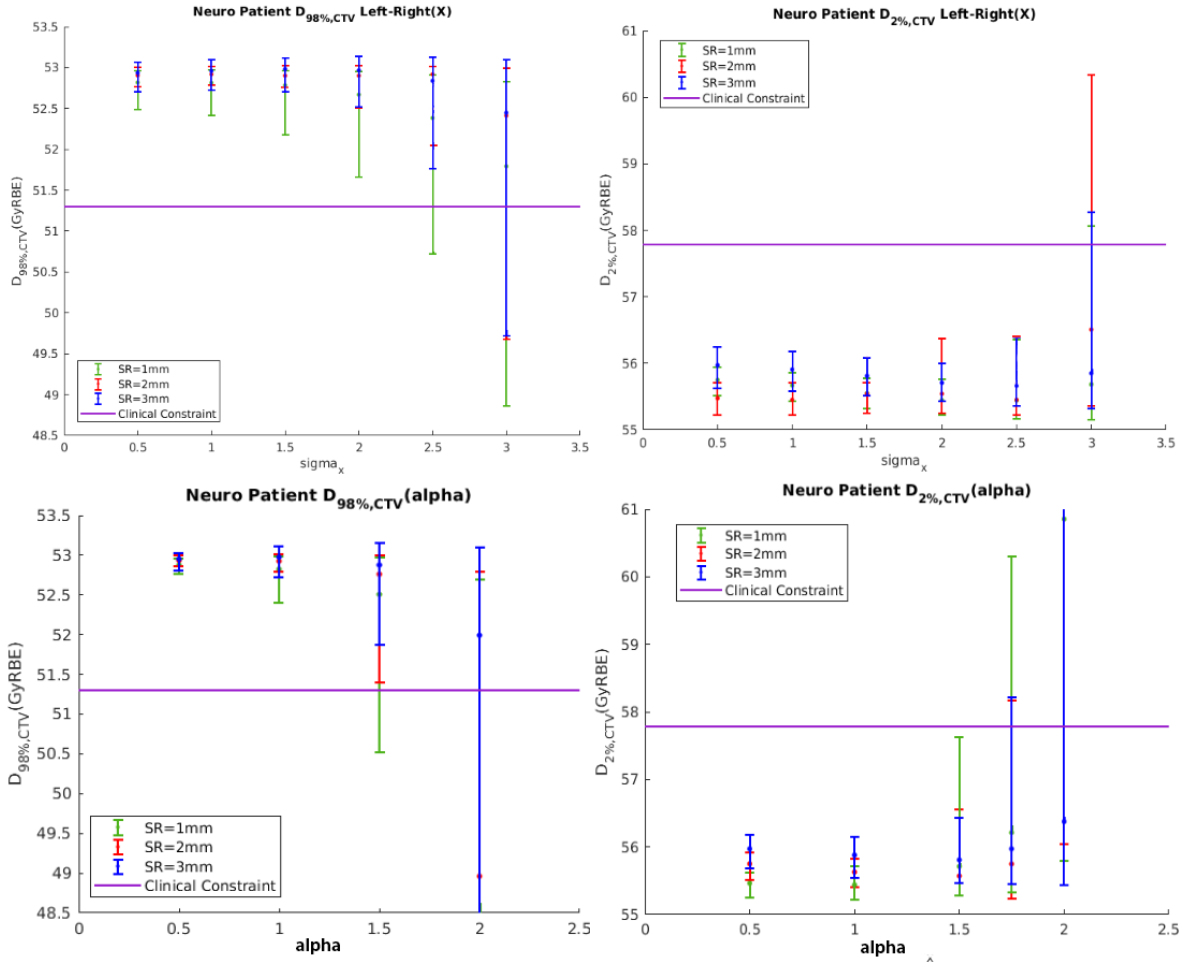


Figure 6.8: $D_{98\%,CTV}$ and $D_{2\%,CTV}$ dose dependencies as a function of random sigma in x direction(up), and as a function of alpha (down)

6.2.2. Non-Robust Skull-base Chordoma Patient

For the skull-base chordoma patient, two optimization plans were made. One treatment plan used the initial clinical plan's inputs, and the second was constructed from scratch. For each of these two plans, we re-plan with different SR (1,2 & 3mm). Thus, a total of six treatment plans were used, to study the influence of SR in the treatment plan's robustness.

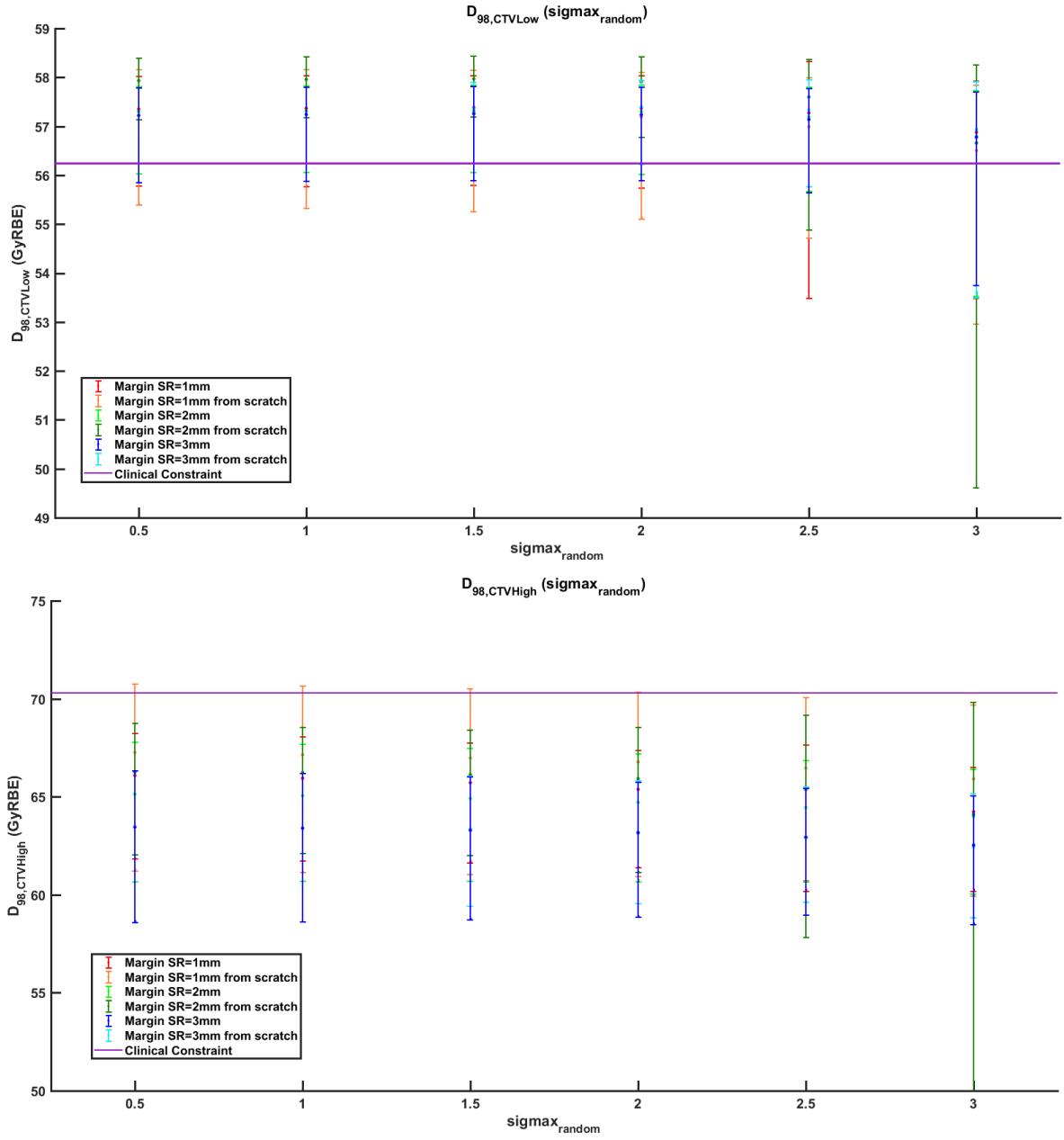


Figure 6.9: $D_{98\%,CTV}$ as a function of random setup (x axis) error, for non robust chordoma patient

From Figure 6.9 comparable robustness is observed between 2 and 3mm. However, none of the plans met the clinical constraint. To validate the differences between the treatment plans, the scaling factor (introduced in Chapter 5), is plotted against the standard deviation of the random setup error (x). Figure 6.10 illustrates, the required SF, for which the 98% coverage of $D_{98\%}$ of the treatment plan will meet the clinical constraint, as a function of σ_{random_x}.

From the graph 6.10, it is clear that none of the treatment plans are robust, with an exception of the treatment plan with SR=2mm when constructed from scratch. However, even this plan develops a steep increasing trend compared to the others, for which no explanation was found. For that reason, it can be considered misleading or an outlier. This is further discussed in the Discussion section.

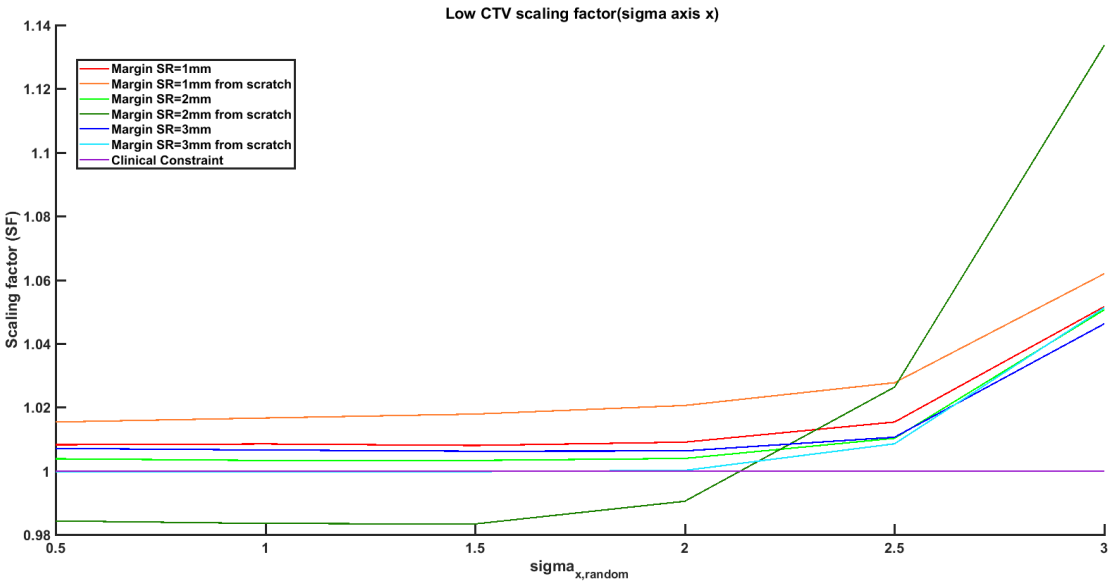
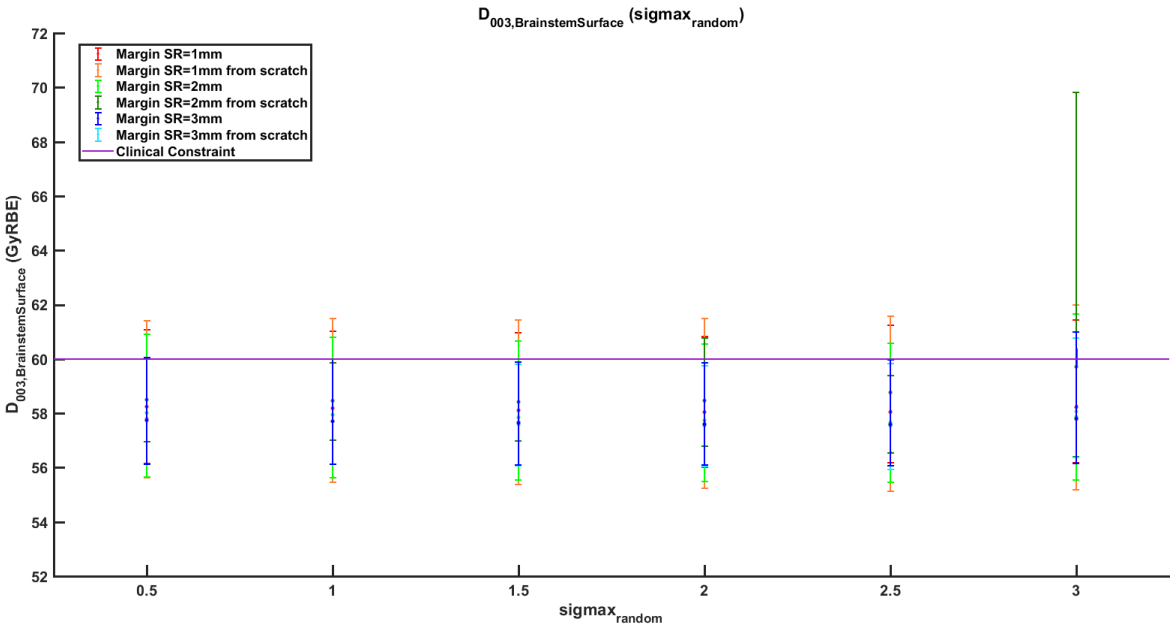


Figure 6.10: Scaling Factor for low CTV, as a function of random sigma x

The non-robust treatment plans resulted from the trade-off between the dose given to CTVs and to OARs. Following the same evaluation, the D0.03c for the Brainstem and the Spinal cord (Surface and core area), as a function of random sigma on x-direction is given in Figures 6.11 and 6.12.



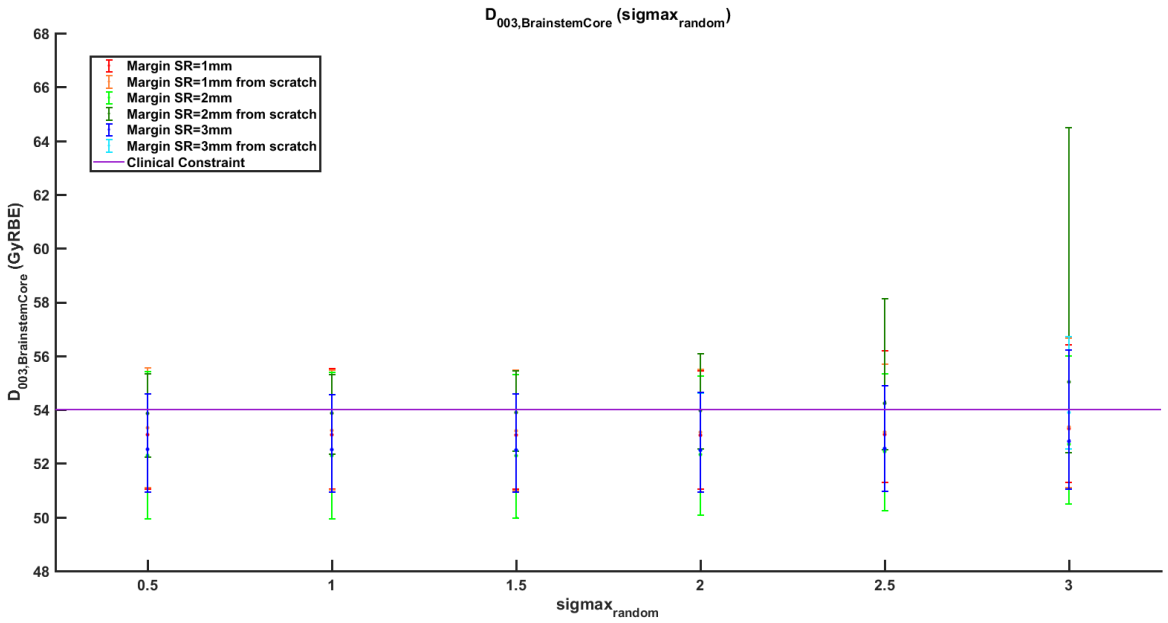
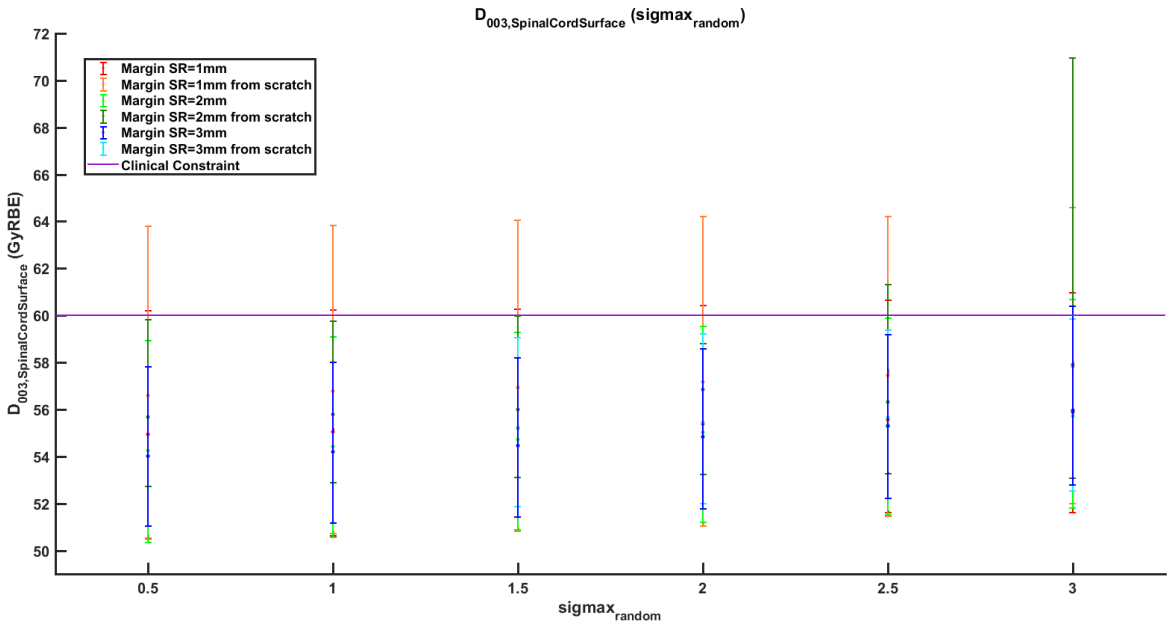


Figure 6.11: D_{003cc} for Brainstem Surface (top), Brainstem Core (bottom), as a function of sigma x



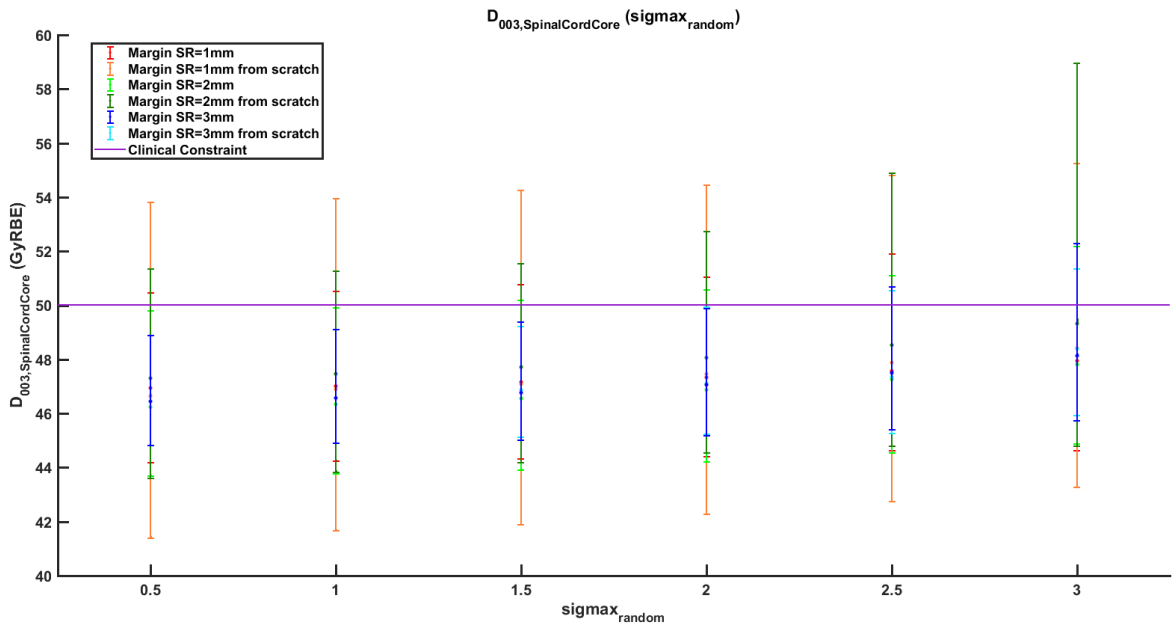


Figure 6.12: $D_{0.03cc}$ for Spinal Cord Surface (top), Spinal Cord Core (bottom), as a function of sigma x.

Figures 6.11 and 6.12, indicate that the treatment plans with SR=1mm, is unable to meet the OARs objectives. Treatment plans with SR=3mm, appears superior in terms of OARs objectives, for both OARs.

The conclusions drawn from these observations showed unclear trade-off information between the OARs and the CTV, along with the tumors complex geometry eventually resulted in the decision to switch tumor sites and to focus the research on robust patients.

6.3. Inhomogeneity vs Robustness

Using Erasmus-iCycle as a dose engine and the five-patient cohort, described in Methods Chapter, robustness dependency of dose inhomogeneity delivery was investigated.

6.3.1. Inhomogeneous Plan

This part of the study is focusing only on the range uncertainty. A homogeneous (maximum dose constraint 107% D_{pres}) and an inhomogeneous (maximum dose constraint 120% D_{pres}) treatment plans were constructed. Figure 6.13 shows the dose volume histograms of the nominal dose distributions of both the homogeneous and the inhomogeneous dose treatment plan, for the first patient. The graph illustrates not only the CTV but also the most important OARs. Additionally, the areas around the CTV, called CTV rings, are displayed in order to study the effect of the dose delivery to the surrounding healthy tissue.

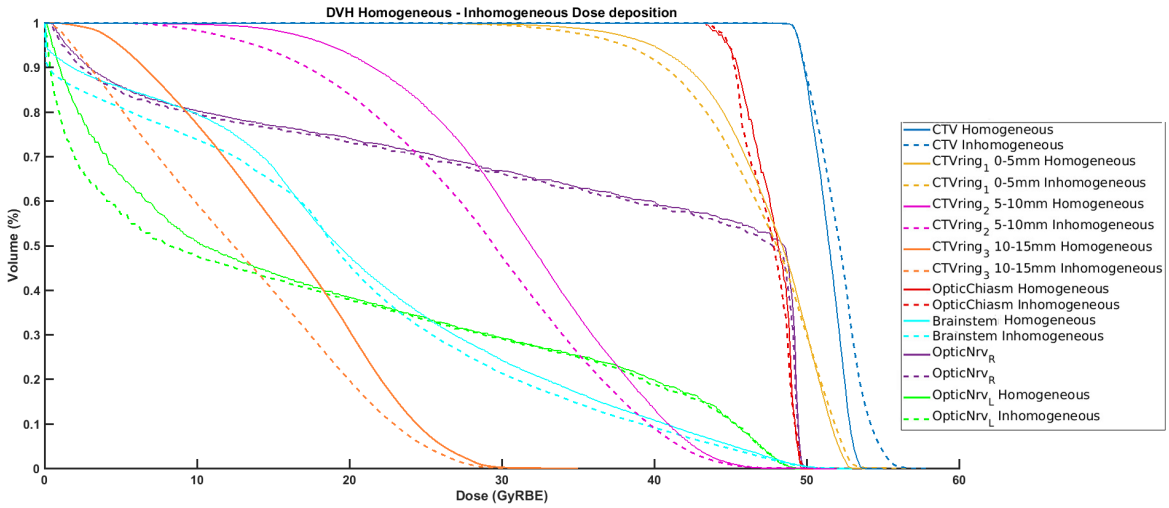


Figure 6.13: Homogeneous and Inhomogeneous beam nominal scenario DVH for CTV, CTV rings and OARs

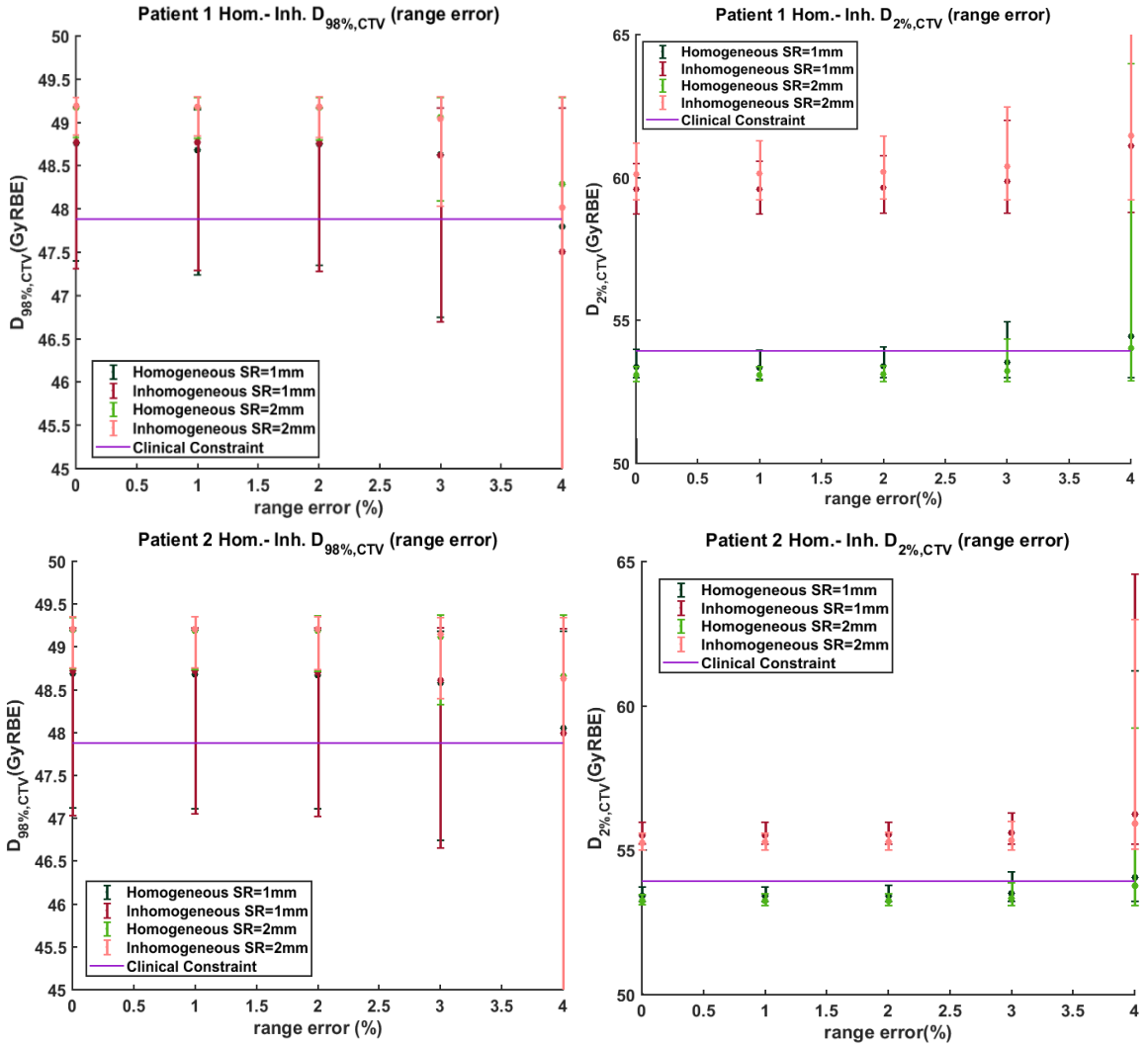
The mean values for the CTV rings and the OARs, for all five patients, are given in Table 6.1. The mean dose for these structures was reduced from 0 to almost 3 Gy, depending on the patient and the structure. A dosimetric gain in the mean dose value for the OARs is observed when an inhomogeneous dose is selected.

Mean dose (Gy)										
Structure	Patient 1		Patient 2		Patient 3		Patient 4		Patient 5	
	Hom.	Inh	Hom	Inh	Hom	Inh	Hom	Inh	Hom	Inh
$CTV_{ring0-5mm}$	46.80	45.37	47.40	46.91	45.56	44.49	46.33	45.51	47.21	46.64
$CTV_{ring5-10mm}$	30.64	28.92	31.54	28.87	26.04	23.54	29.23	26.83	31.69	29.93
$CTV_{ring10-15mm}$	15.44	12.75	15.82	12.70	10.87	8.88	14.14	12.00	16.24	15.01
Brainstem	6.5	5.48	20.88	19.49	16.08	14.54	13.55	12.78	13.81	12.78
Optical Chiasm	42.47	40.42	47.65	47.44	27.48	26.57	44.16	42.04	45.51	43.29
Opt. $Nerve_{Right}$	10.21	9.88	34.92	34.43	3.32	2.55	17.16	15.96	37.76	34.86
Opt. $Nerve_{Left}$	34.88	33.65	18.17	17.11	0.73	0.56	7.38	7.48	19.37	18.34

Table 6.1: Mean Doses for CTV rings and OARs. For both Table's plan SR=2mm

From the dosimetric dependencies of Figure 6.13, in agreement with previous research ([59]), when conventional maximum dose constraint (107% $D_{prescribed}$) is used to ensure robustness, the excess dose is depositing around the normal healthy tissue. However, when a higher maximum tumor dose is used, this excess dose can be steered to the tumor itself, achieving better normal tissue sparing.

For each of the patient, the $D_{98\%,CTV}$ and $D_{2\%,CTV}$ from each PCE evaluation, are plotted against range errors (used for the evaluations) in Figure 6.13.



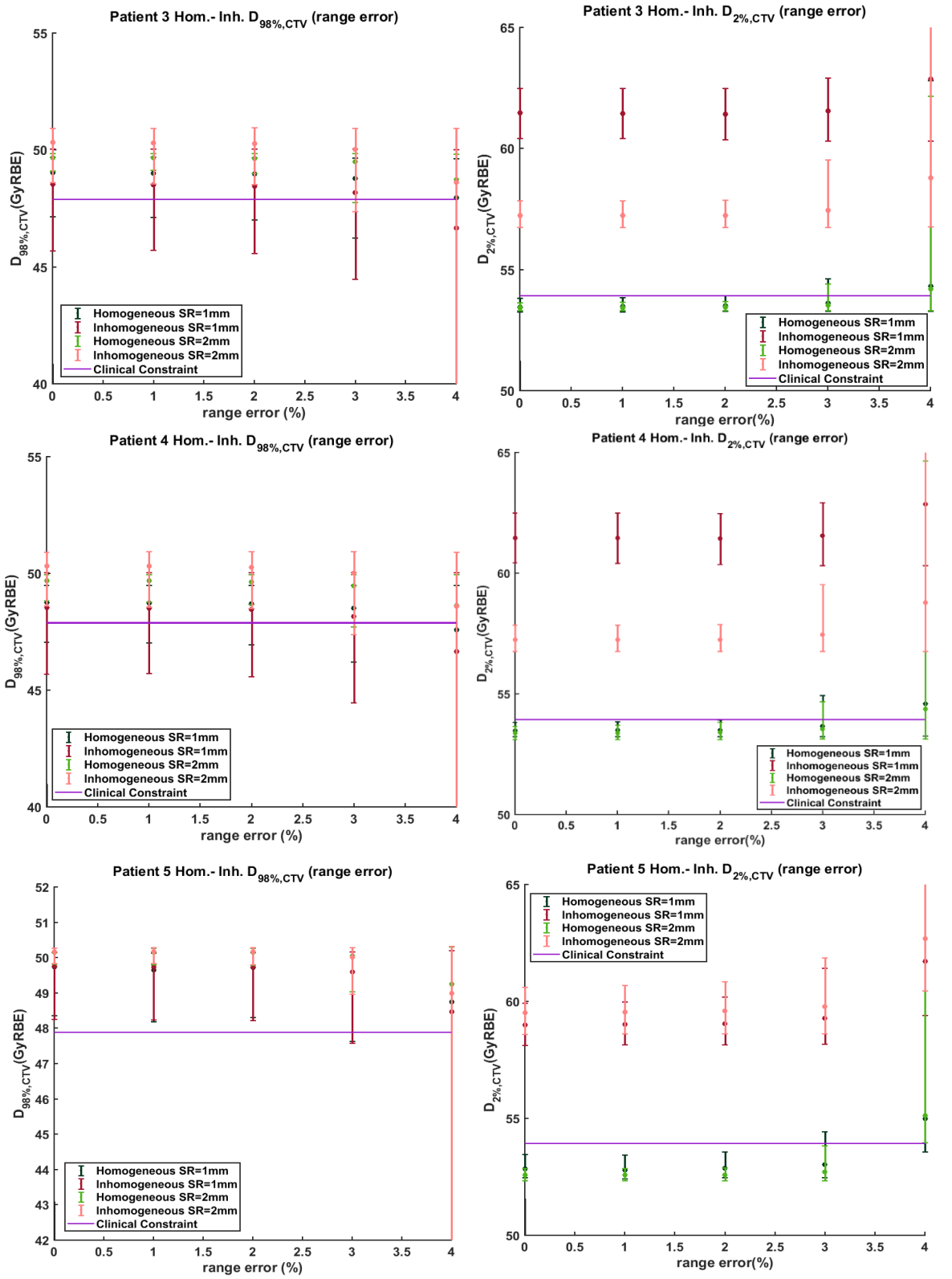


Figure 6.14: Homogeneous and Inhomogeneous beam $D_{98\%,CTV}$, $D_{2\%,CTV}$ dose dependency as a function of random range error

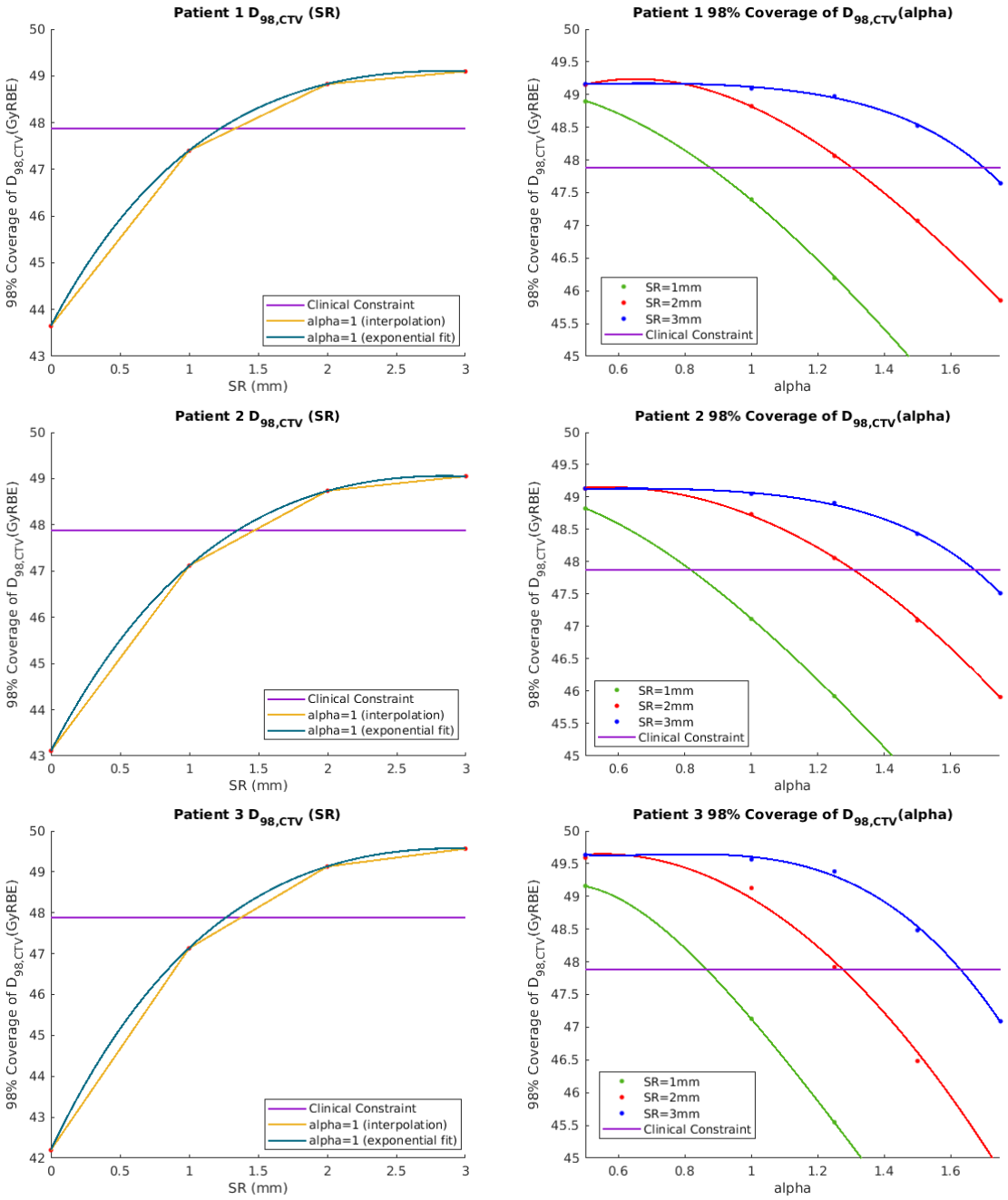
Depending on the patient different robustness behavior is observed. However, when the geometrical margin is equal to 1mm, no treatment plan is robust in any of the patients.
It is worth to mention, that the robustness achieved is almost identical in the same plan of Figure 6.14. That is not in agreement with the previous work from Petit et al. [59], who reported a robustness gain when an

inhomogeneous treatment plan is used. In contrast, from Table 6.1, we confirmed that robustness leads to more normal tissue dose, however, the overall plan's robustness remained the same.
For the first, second, and fifth patients both treatment plans have similar robustness when SR=2mm is used. The third and fourth patient the inhomogeneous plan also lose the robustness for approximately the same range error values as the homogeneous dose plans, however, the average $D_{98\%,CTV}$ mean values are higher. For the larger margin (2mm) in every patient case, the robustness is lost at range error between 3 and 3.5%.

6.4. Towards Probabilistic Planning

The goal of this section, PCE robustness evaluations in a loop with a clinical treatment planning system to assess the potential for clinically feasible probabilistic treatment planning. The method proposed in Chapter 5, includes the use of probabilistic measurement 98th percentile of $D_{98\%}$, to investigate if there is a relationship between the margin size and the clinical constraints vector.

98% Population coverage of $D_{98\%}$ ($D_{98\%,min}$), in a relationship with the α parameter and as a function of SR, for all patients are given in Figure 6.15.



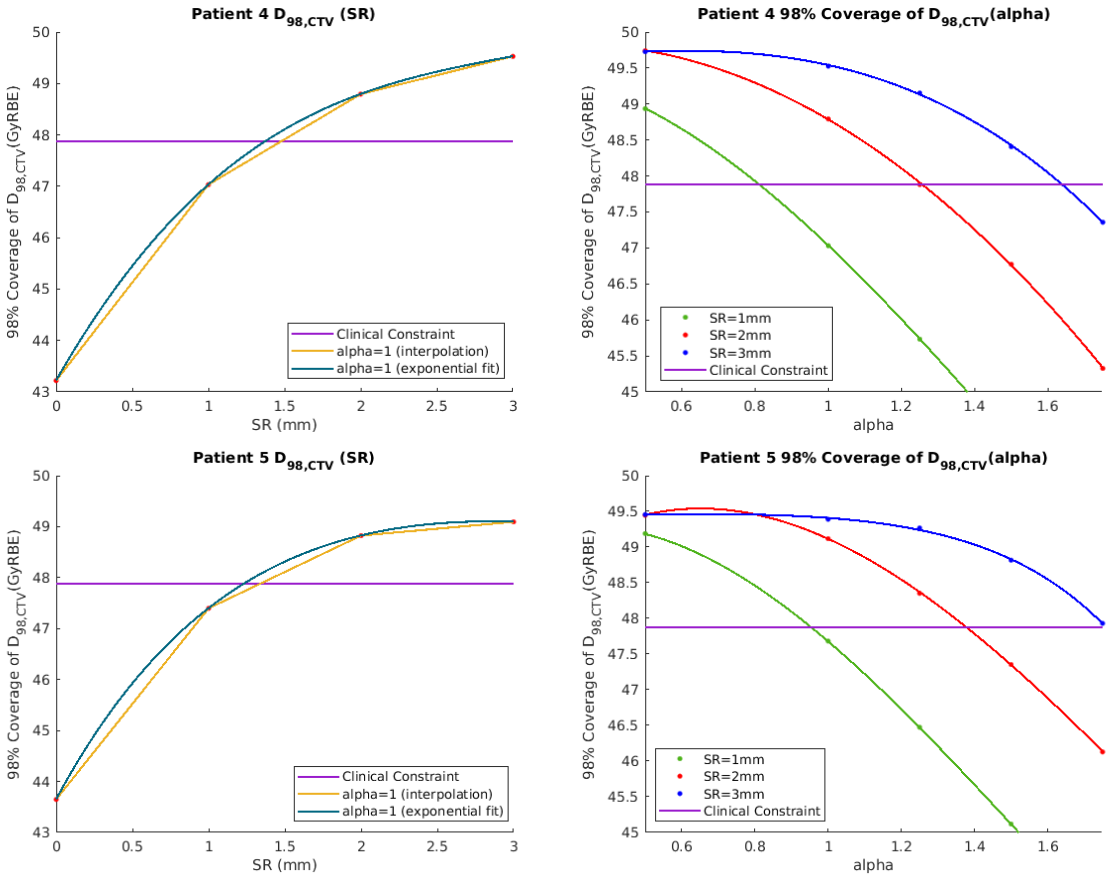


Figure 6.15: 98% population coverage of $D_{98\%,CTV}$ as a function of SR (left), and as function of alpha (Right)

Each PCE evaluation results to one $D_{98\%,min}$ point. The plots are fitted using third order polynomials given by 6.1 and 6.2.

$$D_{98\%,min} = x_1 SR^3 + x_2 SR^2 + x_3 SR + x_4 \tag{6.1}$$

$$D_{98\%,min} = x_1 \alpha^3 + x_2 \alpha^2 + x_3 \alpha + x_4 \tag{6.2}$$

The indicators of each polynomial are given in the Tables 6.2 and 6.3.

Parameters of polynomial function $D_{98\%,min}$(SR)				
Patient	x_1	x_2	x_3	x_4
1	0.195	-1.747	5.303	43.65
2	0.180	-1.732	5.556	43.11
3	0.229	-2.154	6.857	42.2
4	0.170	-1.536	5.178	43.22
5	0.193	-1.792	5.724	43.05

Table 6.2: Parameters for all 5 patients, for $D_{98\%,min}$(SR)

Parameters of polynomial function $D_{98\%,min}(\alpha)$												
SR(mm)	1				2				3			
Patient	x_1	x_2	x_3	x_4	x_1	x_2	x_3	x_4	x_1	x_2	x_3	x_4
1	0.665	-4.041	1.865	48.9	0.881	-5.49	6.024	47.4	-1.78	4.324	-3.49	50.05
2	0.648	-3.586	0.83	49.23	0.428	-3.78	4.127	47.96	-1.78	4.27	-3.44	50.01
3	2.194	-9.05	5.681	48.31	2.416	-11.59	12.2	46.09	-2.31	4.864	-3.31	50.36
4	0.785	-4.006	0.825	49.43	-0.055	-1.97	1.126	49.68	-0.71	0.268	0.46	49.52
5	1.247	-5.85	3.58	48.7	0.886	-5.521	6.06	47.69	-1.79	4.35	-3.52	50.35

Table 6.3: Parameters for all 5 patients, for $D_{98\%,min}(\alpha)$

For each patient range of optimal margin is provided, from $D_{98\%,min}$(SR) plot (Left plots in Figure 6.15). The average value of the suggested optimal margin is 1.35 Additionally, the value α_o is determined from $D_{98\%,min}$(alpha) plot (Right plots in Figure 6.15). The ratio of between the α_o and the SR, along with the suggested optimal margin SR_{opt} are given at the Table 6.4.

Patient	SR_{opt}(mm)	SR_1/a_o(mm)	SR_2/a_o(mm)	SR_3/a_o(mm)
1	1.26-1.35	1.19	1.54	1.78
2	1.35-1.47	1.23	1.53	1.80
3	1.27-1.38	1.16	1.57	1.84
4	1.37-1.48	1.23	1.58	1.84
4	1.33-1.42	1.20	1.56	1.83

Table 6.4: Optimal Margin SR_o & Ratio of SR and α_o

From the values reported in the Table 6.4 there no linear relationship is observed between the SR_{opt}, α_o and SR. However, the plots indicate that iCycle is a robust optimizer (smooth curves). The results are discussed further in the following section.

7

Discussion & Conclusion

7.1. Discussion

The initial ambition for the project was the implementation and evaluation of a feasible iterative approach towards probabilistic treatment planning. During this research, multiple additional topics were addressed, starting with optimization of construction speed of PCE dose simulation model. Additionally, evaluation robustness and trade-offs made in treatment planning, for neurological patients, were investigated. This chapter discusses the results from each part of this research, then presents recommended directions for future work, finishing with the main conclusions.

7.1.1. PCE optimization

-Monte Carlo Noise level

In the first section of the results, optimization of PCE construction time was attempted. By increasing the MC noise level, the constructing time is reducing significantly. However, since there is a trade-off between accuracy and noise level, this work validated the model's performance under different noise levels. Therefore PCE was built with MC noise levels of 1, 2, and 3%. The accuracy of the different models was evaluated, using several shifted scenarios in both one-axis shifts and off-axis shifts. For all the models, their performance was worse when tested against off-axis shifts. That was an expected observation since PCE is mainly optimized along the main axis during its construction. Additionally, we noticed that PCE dose distributions are accurate within a range of ±3mm. However, this is not expected to negatively influence neurological tumor patients, as the expected clinical treatment errors are smaller.

From the DVH comparisons, with the exact TPS dose, dose distributions from PCE models constructed with higher noise levels almost overlapped with the exact dose values. Therefore, further results validation was conducted.

When only CTV voxels are considered, the results from the 2-norm relative voxel-dose evaluations, indicated that both PCE built with one and two percent had acceptable accuracy. However, since the CTV is not the only constraint of the treatment plan, to evaluate the models' performance for the OARs objectives, we also consider CT voxels that receive a dose superior to 1 Gy. In the second case, when off-axis shifts are considered only PCE build with MC 1% is precise enough to remain below the 1% voxel-dose error limit.

According to the results, an increase in MC noise level during the PCE construction is not suggested as the model's accuracy will be reduced noticeably. When the worst-case scenario is considered (off-axis shifts & CT voxels with dose above 1 (Gy)), the PCE constructed with MC 1%, is four times more accurate than the second-best model. Thus, we confirmed that the default noise value (1%) is the best fit for this treatment area.

-PCE dose differences

The goal of a PCE model designed to model changes relative to a static dose-cloud approximation, was to accurately simulate the dose in complex anatomies, where the dose gradient in DVHs is not a steep line. In those areas, the dose difference would be higher, therefore the PCE would be expected to perform better.

The results of this section were not encouraging. The poor performance of the model is derived from:

- the assumption of static cloud dose approximation for the nominal scenario, while the noise sampled in the calculating shifted dose (increases the impact of noise)

- mainly due to MC noise effect, when dose differences are considered. Since the dose difference values are way smaller than the previous doses themselves, the effect of the noise is expected to be higher.

The MC noise effect is expected to remain, even if we use a different resampling approach. Additionally, the selected data concerns a neurological tumor, which is considered an "easy" treatment site. In case of the presence of larger uncertainties and higher dose gradients, like the case of head and neck, PCE built with dose differences expected to perform even worse.
Therefore, the method based on dose differences is concluded to not be as accurate as expected.

7.1.2. Setup Robustness in Treatment planning

In this section, for a clinically robust neuro-oncological patient, further SR reduction was suggested after PCE robustness evaluation. The evaluations from the meningioma patient suggested that the currently used margin was over-conservative and it could be replaced by a margin of 2mm without robustness loss. Additionally, it was observed that SR=1mm is not robust in any plan. It is also important to avoid really small margins, as there is a probability of tumor underdosing. Since, the Tumor Control Probability (TCP) is a very steep dose curve, choosing a very small margin can eventually lead towards small parts of the tumor getting underdosed, which can drastically decrease the TPC.

When the complex case of skull-based chordoma patient was considered, the results from the evaluations did not lead to clear conclusions. First, almost none of treatment plans were not robust, with one exception of the treatment plan with SR=2mm, which was constructed from scratch. This plan achieved target's coverage when the standard deviation of the error scenarios was below 2.25. However, the overall plan's performance is not fully understandable, as after losing robustness it follows a much sharper increase compare to all other five treatment plans (Figure 6.10). Additionally , for this patient in the final clinical treatment plan higher dose was allowed to the OAR objectives. Therefore, one can question how *strict* are the clinical objectives for the OARs, for complex patients. From the skull-base chordoma patient, we also noticed that the treatment plan with SR=3mm, has similar robustness to the ones having SR=2mm.

7.1.3. Inhomogeneity vs Robustness

This research investigated an additional factor that can influence treatment plans' robustness, dose delivery inhomogeneity.
The finding of this section was particularly interesting. When an inhomogeneous dose is used the average dose to OARs, along with the healthy tissue surrounding the CTV, appeared to be systematically lower for all the patients, with a maximum reduction of 2.8 Gy. The increased dose to the CTV leads to a dose distribution focus more in this structure, minimizing the dose to the surrounding tissue. Nonetheless, the robustness behavior for $D_{98\%,min}$, was the same for all the patients. For the studied patient cohort, similar robustness was noticed, as the $D_{98\%,min}$ is comparable in both homogeneous and inhomogeneous treatment plan.

7.1.4. Setup Robustness Selection

In the final part of the study the potential of using probabilistic variables, for the determination of the minimum robust margin, was examined. Influenced by van Herk's formula, the objective of this section was a linear relationship between setup robustness and clinical uncertainties multiplies. The aim was to determine minimum setup robustness, for which 98% coverage of $D_{98\%,CTV}$ is achieved, for the clinical uncertainties.
A relationship between the examined quantities (SR, SR_{opt}, α_o) did not appear to be linear. Therefore, it was not possible to use PCE robustness evaluation to determine the optimal SR, using the α quantity. For the SR=1mm, the α_o was smaller than one, suggesting a margin increment, while for both SR=2mm and SR=3mm the α_o was above one, which indicates that the currently used margins are over-conservative. The optimal margin from the plot of the minimum $D_{98\%}$ and the SR, suggested a margin around 1.3 average for all the patients. Even if the pointing direction of the method was correct (SR=1mm suggested margin increment, SR=2,3mm decrements), there is not an obvious connection between the results. Additionally, it was difficult to confirm the observations for the non-integer margins, as these feature is not implemented yet to the dose engine (Erasmus-iCycle). Although, the curves in the plots appear to be smooth, suggesting that a prediction can be made for the intermediate values, and their behavior is not expected to be much different. Finally, iCycle appears to give consistent optimization as all five patients' similar behavior was observed.

7.1.5. Future work

From the results of this work, we observe the potential of using inhomogeneous plan for this treatment site, especially in patients with hard OARs constraints, as we observe the dose reduction for these areas. It would have been interesting to compare the results, using PCE evaluations for more treatment sites, and different error scenarios. More specifically, the challenging non-robust chordoma patient discussed in this research. Additionally, in our results similar robustness was observed between the plans with the same margin. Petit et al. [59], concluded a robustness gain we inhomogeneous treatment plan is used. The main difference between the works is the treatment site. However, a further comparison of the methods might provide further explanation.

As for the probabilistic approach for determining the minimum margin, the most important question could be how the quantity α could be defined to allow the creation of a recipe, able to use the probabilistic goals for prediction of the optimal margin for an individual patient. In our case, multiplies of the clinical uncertainties were used, and $D_{98\%,min}$ was the probabilistic variable. However, there is still not an acceptable protocol to indicate which probabilistic goal, should be used for a plan to be considered robust (e.g. 98% population coverage of $D_{98\%,CTV}$ or 95% coverage of $D_{98\%,CTV}$). Therefore, guidelines like these would provide ground truth for future researchers.

7.2. Conclusion

This work has confirmed that the currently used noise setting for polynomial chaos expansion (PCE) metamodel provides the optimal accuracy and time trade-off. PCE evaluation has been used to determine the optimal robustness setting for skull-based meningioma and chordoma patients at individual level, using Raystation for Intensity Modulated Proton Therapy. The results from the skull-based meningioma reveal space for margin reduction. While for the complex case of skull-based chordoma, where trade-offs between CTV and OAR dose are unavoidable, better metrics could be helpful and provide more grip on the treatment planning.

Changing the treatment planning system to iCycle and using PCE evaluations, no intrinsic difference was observed in the degree of robustness of homogeneous and inhomogeneous planning approach. However, inhomogenous plan allowed for more OAR sparing overall, at the expense of homogeneity of target dose.

Finally, the potential for a probabilistic treatment plan which ensures that at least 98% of the treated population receives a near-minimum dose D98% of at least 95% of the prescribed dose, using a loop of PCE-based robustness evaluation was examined. We concluded that the approach with an overall linear scale factor allows for a substantial gain in one iteration. Still, for the relatively large difference between the actual and optimal robustness, linear modeling and one iteration do not suffice to find the overall optimum. The results, also indicate that the currently used margin is conservative enough for this patient category.

A

Literature Study

This review provides an overview of the different state of the art methodologies used for robust treatment planning in Intensity-Modulated Proton Therapy (IMPT), focusing on *"Proton therapy uncertainties & development of robust optimization"*, proposed in the literature. First, delivery uncertainties in radiotherapy are discussed, with attention given to neurological and head and neck patients (HNC). Next, robust optimization accounting for anatomical uncertainties in the head and neck treatment area is reviewed. Besides, computationally efficient proton dose calculation approaches are presented. Finally, an introduction to polynomial chaos expansion as a tool for robust evaluation is made.

A.1. Literature research

To obtain this information a literature search was carried out using search queries on PubMed and Scopus, which resulted in 60 unique articles for uncertainties in head and neck or neurological tumors in IMPT. From those, 37 studies were focused on robust optimization, probabilistic approach, proton uncertainties, or PCE for HNC and considered eligible after screening the titles and abstracts. Full-text screening resulted in 24 studies that all inclusion criteria and, hence, were eligible for the comparison analysis. Figure A.1) illustrates the process for the selection of relevant literature.

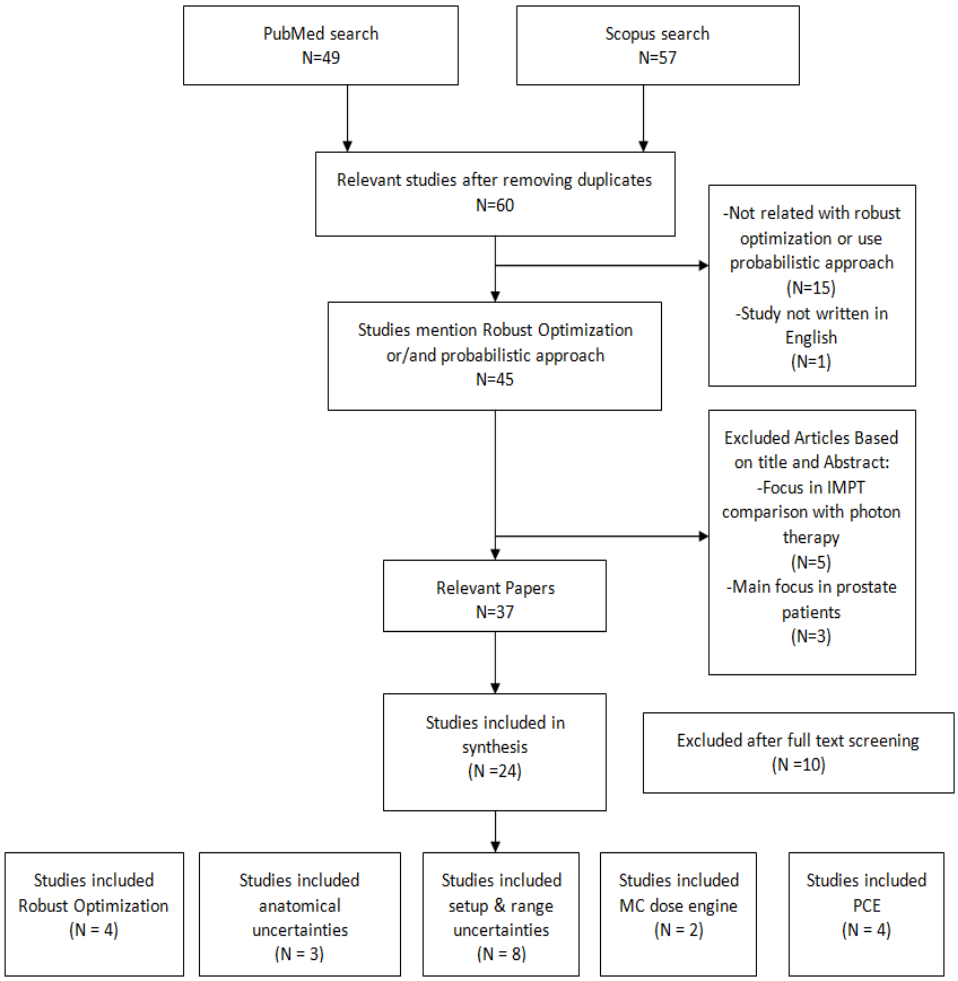


Figure A.1: Flowchart of Methods used in this review

A.2. Radiation therapy for cancer treatment

Radiotherapy aims to deliver doses of ionizing radiation to the tumor site, sparing tissue or minimizing dose as much as possible to the surrounding healthy tissue. Uncertainties and errors are unavoidable, and the treatment planning and delivery need to be robust against these. Over the past 25 years, the use of volumetric modulated arc therapy (VMAT) has allowed considerable improvement in treatment conformality and reduction of high doses to neighboring critical structures. Consequently, this has drastically reduced the incidence of major forms of toxicity, like xerostomia [66]. However, the improvements in physical delivery of photon therapy cause a trade-off of alternative toxic effects such as fatigue, nausea, hair loss, oral mucositis that further improvement in the therapeutic window requires alternative methods of radiation delivery.

In this context, proton therapy has emerged as a novel means to reduce toxicity and potentially further improve tumor control. IMPT has the potential to deliver a significantly lower dose to Organs at Risk (OARs), while achieving similar or even better tumor coverage [27, 29, 36]. Additionally, Blanchard and Frank [31] made a review study comparison of IMPT survival to Intensity-Modulated Radiotherapy (IMRT) and concluded that there could be an advantage for skull-based or sinonasal malignancies. However, due to protons nature and the characteristic energy deposition at a specific depth (Bragg peak), the treatment plan can severely be compromised by uncertainties in patient setup and proton range, along with anatomical uncertainties, that have a larger impact on the actual dose distribution, as shown in Figure A.2. Therefore, it is necessary to achieve a conform dose distribution in IMPT, the plan to be robust against any of these uncertainties [21].

Under the presence of range uncertainties as shown in Figure A.2 conventional margins used for photon therapy do not work, as if the range of the beam moves deeper or shallower than the tumor an undershooting or healthy tissue overshooting might occur. Robust optimization methods, replace the conventional margins and create a treatment plan robust against uncertainty effects.

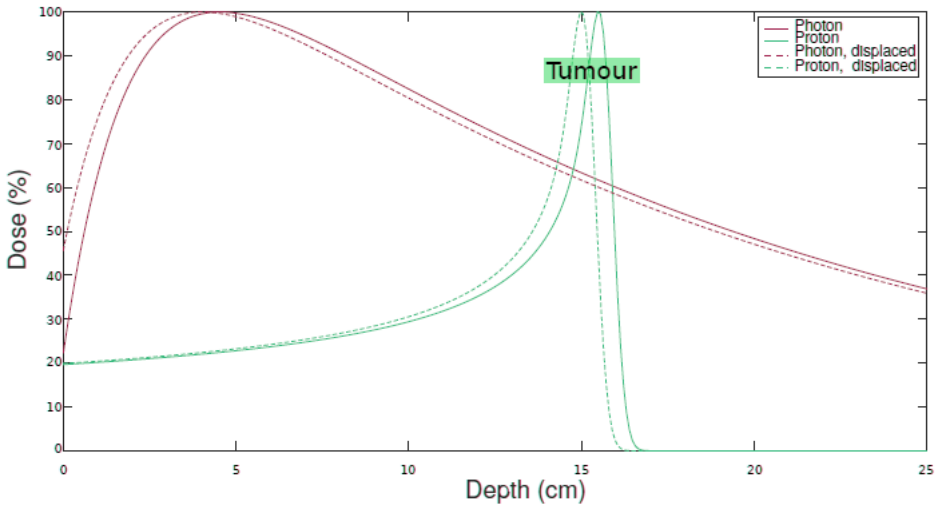


Figure A.2: Difference in the delivered dose when a shift occurs [85]

In recent years many studies have proposed different robust optimization approaches, with non of them being superior to the others in every aspect. Given the variability observed between studies, a comprehensive review is needed to infer the main limitations and improvements, mainly in uncertainties for which optimization accounts, and dose computational methods. Conventional robust optimization in proton therapy accounts for range and setup errors. Since the presence of some uncertainties might be stronger or weaker for each treatment site (breathing motion affects lung tumors), different site-specific additional uncertainties can be considered to improve the optimization. To highlight the importance of accounting for additional uncertainties, we compare site-specific optimization approaches for HNC patients.

A.3. Uncertainties Proton Therapy

As discussed, both proton and photon treatment plans can be compromised by uncertainties. Figure A.3 provides a summary of the different uncertainty types that concern the modern literature, while it shows the influence of each uncertainty source in each treatment type.

Factor	Proton Therapy	Photon therapy
CT coversion & stoppping power accuracy	-Sensitive, affects range, distal target coverage or distal normal tissue sparing	-No sensitive
Normal structure motion orthogonal to beam	-Affects range, dose distribution distal to structure	-Minimal effect
Target motion normal to beam	-Affect margin, may affect dose distribution distal to target	-Affect Margin
Complex inhomogeneities	-Perturb dose distributions, degrade distal edge	-Not strong effect
Anatomy changes over course of RT	-Affect dose distribution	-Minimal effect
Plan Evaluation	-Impact of uncertainties significant, PTV concept not valid, validity of initial nominal plan questionable	-PTV concept valid, dose distributions relatively invariant to uncertainties

Figure A.3: Comparison of different factors influence in IMPT and IMRT [67]

Usually, models which include uncertainty such as range and set up errors are mainly considered for robust IMPT optimization [48]. Anatomical uncertainties and organ motion can also be taken into account. However, depending on the tumor's location, different uncertainties are taken into consideration. For instance, lung motion is considered in many studies, while for HNC, organ motion is not considered in any reviewed

case. In contrast, anatomical changes during the fractionated treatment have been often reported in the literature [32–35], mainly from research on adaptive proton therapy.
The type of uncertainties can be also categorized into intra and inter-fraction errors [36]. Inter-fraction uncertainties are considered the displacement of the patient or the OARs between each faction delivery. Intra-fraction refers to changes during the fraction delivery, like swallowing during the treatment. In the case of head and neck patients, this kind of motion during the treatment has not been investigated from any of the reviewed papers.

A.3.1. Geometrical Uncertainties

As displayed in the previous section, geometrical uncertainties occur in both proton and photon therapy like patient motion, delineation, imaging inaccuracies, gantry isocenter, and setup uncertainties. Several proton treatment studies [68, 72], select to focus mainly on range uncertainty due to proton nature. This report addresses the setup, range, and anatomical uncertainties, that are the main focus in HNC area.

Set-up Uncertainties
Before each treatment fraction, the patient is positioned on the treatment couch, using skin tattoos, immobilization devices, and laser beams of the machine, to align the internal target volume to the isocenter of beam application [88]. However, even when the lasers are aligned perfectly with the skin tattoos, the internal anatomy does not necessarily remain the same. Therefore during proton therapy delivery, image guidance is always used for the setup process. Most of the proton therapy systems are equipped with orthogonal 2D X-ray imaging. The method uses low energy photons (keV) source, orthogonal to the patient, and creates an image right before the treatment.

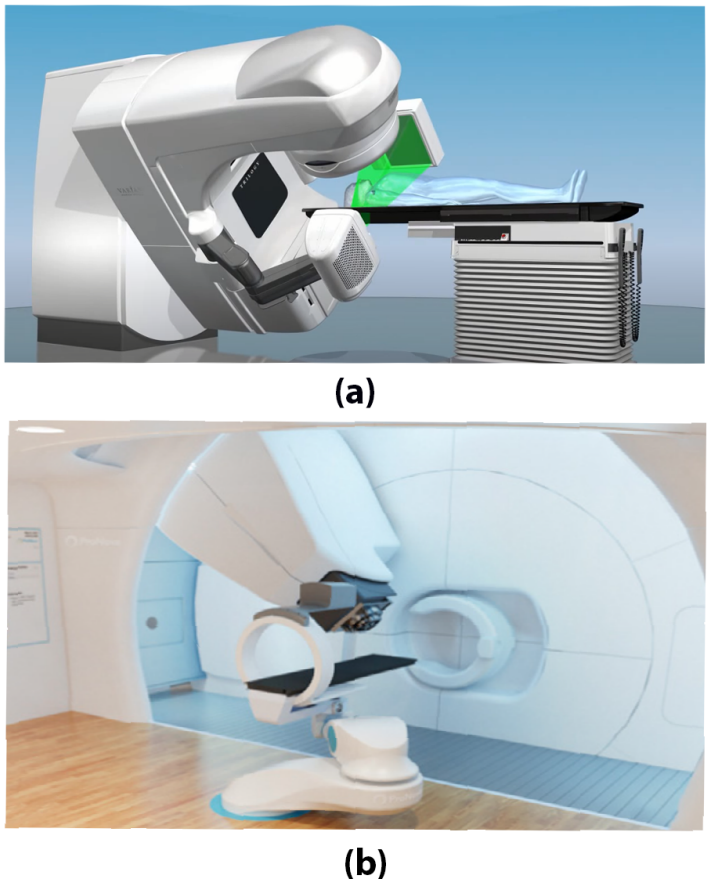


Figure A.4: (a) Linac - Volumetric Imaging image guided radiation therapy, keV portal (green beam), (b) Proton therapy - Volumetric Imaging CBCT

Recently, volumetric imaging for proton therapy has clinically implemented, using either in-room X-ray computed tomography (CT) or cone-beam CT (CBCT) imaging devices [89–91]. Linear Accelerator (Linac) CBCT (photon therapy) and a proton pencil beam imaging system, for proton therapy, are illustrated in Fig-

ure A.4.
Prior to the treatment accurate patient **alignment** is achieved by registration of the in-room control image and the initial planning CT. CT images provide sufficient contrast for the bony anatomy, but the generated images do no give any soft-tissue contrast information. To surrogate for the soft tissue in planar imaging, metal fiducial markers are often surgically implanted.
Inter-fraction setup uncertainties correspond to geometric shifts that can occur during treatment due to misalignment of the patient's isocenter with the isocenter taking into account during the planning CT, which can lead to CTV underdose or OAR overdosing. To measure the setup errors, the planning CT image and the orientation of the beams are used, to digitally reconstruct the expected output image called digitally reconstructed radiograph (DRR). This reference image is compared with the image produced during treatment from the portal imaging device. Applying a daily correction remote CT, Paul Scherrer Institute estimated the position error for head and neck to be below 2.4mm [37]. To characterize the errors, information about their distribution is required. The setup error can be distinguished to systematic and random [12]. Both systematic and random errors can appear in either of the three patient dimensions (x,y,z) as illustrated in Figure A.5.

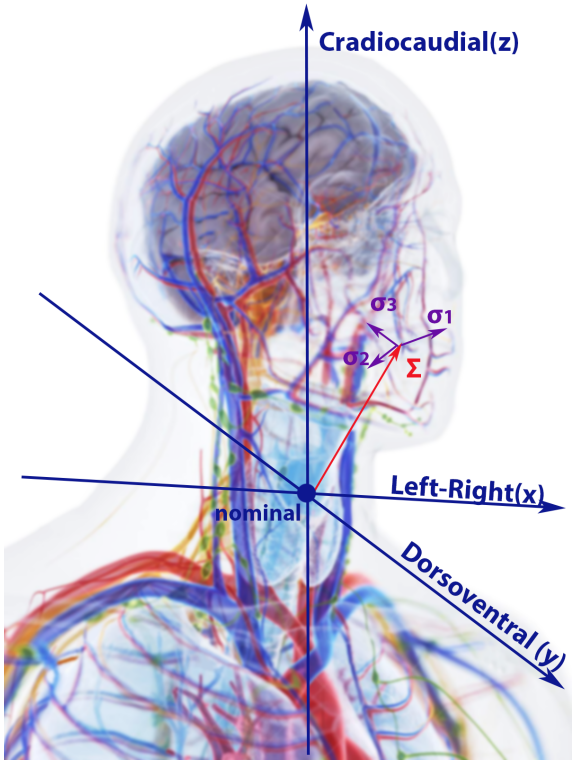


Figure A.5: Systematic (red) and random (purple) setup uncertainties in x, y, z coordinate system

-Systematic: by definition, systematic error refers to an error that remains the same over a measurement. Similarly, in the case of radiotherapy, systematic setup errors are defined as the position misalignment that remains the same over each fraction that the patient receives. For all reviewed studies, the setup errors assumed to be represented by a Gaussian distribution following the central limit theorem which states that combining many distributions asymptotically leads to normal Gaussian distribution. The standard deviation of the systematic setup error is expressed as Σ measured in millimeters (mm), which is the resultant systematic error origin from errors in each patient direction (Σ_x,Σ_y,Σ_z). In Figure A.5, total systematic error is displayed with a red arrow.
-Random: random setup error refers to the displacement that can occur within one fraction and differs from fraction to fraction. This difference in position can again appear in either of the axis and has a standard deviation of σ, which components are displayed in purple in Figure A.5.

Systematic setup errors can cause a dose distribution shift during the treatment, while random error can cause a blur in dose distribution of each fraction [12]. The studies that consider in their analysis both systematic and random components, quadratically sum the standard deviations in each axis, to achieve a combined setup error in each direction.

A limitation is that the systematic error is known for each patient only after the treatment (as it is determined from the average of all errors in a patient). However, it is possible to evaluate how accurate the current procedures are.
Figure A.6 illustrates a simple schematic representation of a possible set of patients, for whom the setup errors registered in 2 dimensions. For each day, the error is registered for each fraction. The steps followed to validate the current procedure are:

- historically collection of data
- determine the mean error for every patient
- calculate the standard deviation of how much the random errors vary around the mean
- after repeating for all the patients, the overall standard deviation for the random errors is given from $\sigma_x = \sqrt{\sum_{i=1} \frac{\sigma_{i,x}}{N}}$ and $\sigma_y = \sqrt{\sum_{i=1} \frac{\sigma_{i,y}}{N}}$, where N is the number of the fractions and i, the number of patients
- among the patients, the systematic error is calculated
- the distribution of the systematic errors is calculated
- calculate average random and systematic error which are characterized by standard deviations σ and Σ, respectively

The expected mean systematic setup error should be close to zero.

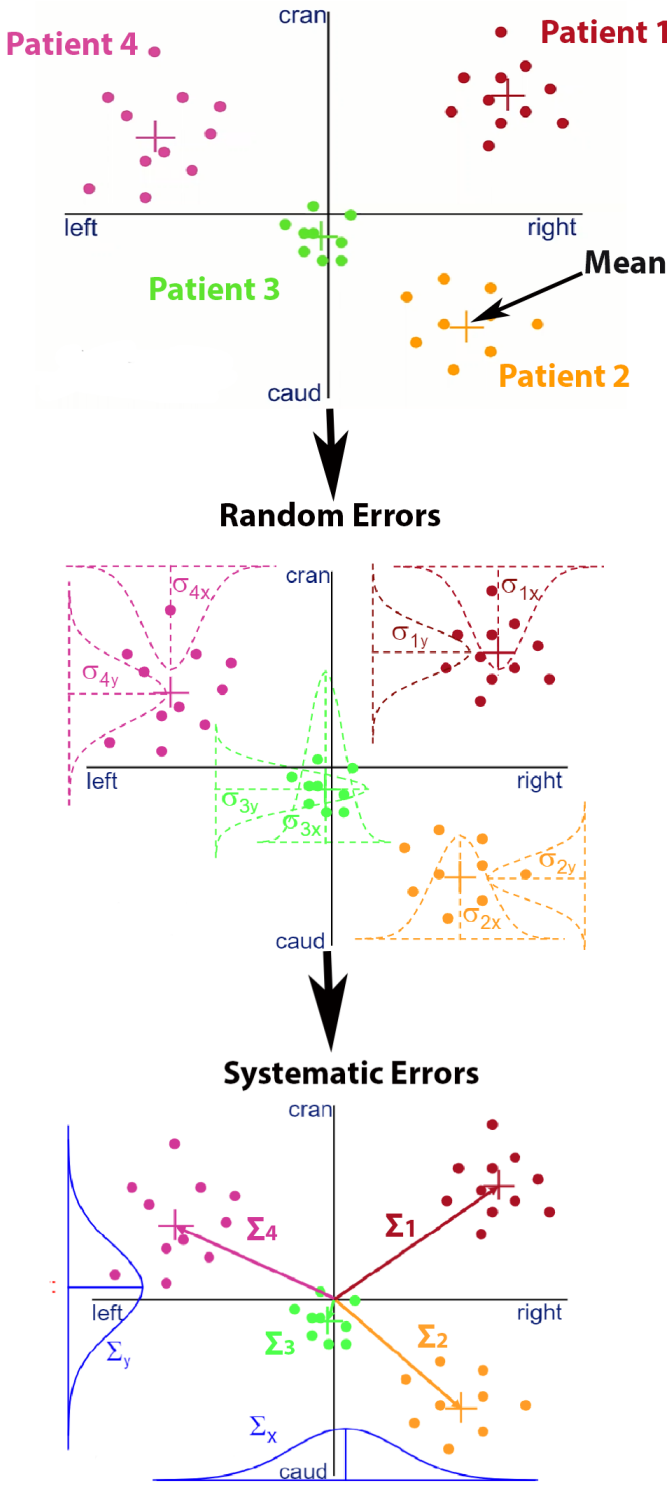


Figure A.6: 2D set-up errors relative to planning for a patient population, different colors correspond to different patients, each fraction (•) [21]

A.3.2. Range Uncertainties

Range uncertainties in proton therapy can be substantial (mm), while is an additional degree of freedom to treatment planning. The errors mainly arise from the conversion of CT Hounsfield units to proton-stopping power (finite proton range), but they can also be a result of CT image artifacts, anatomical and contour changes, weight, or gain loss. We can distinguish the range error to relative (measured in %) and absolute (measured in mm).

The most important range errors, which are considered in robust optimization error scenarios, is the relative error. This error arises from the lack of correspondence between Hounsfield units and tissue materials [20, 25]. Due to the lack of an analytical solution for the conversion to stopping power, Monte Carlo simulations or measurements are used. As a result, a systematic error is created. Other contributions can be eliminated by proper quality assurance and by using a plan adaptation protocol while the stopping-power prediction error is fundamental.

Over the past decades, this error had remained constant, with its values to be determined around 1.5% for head and neck patients. So far a single X-ray CT was used to determine the Stopping Power Ratio (SPR). Last year, a review of the status, benefits, and potentials of dual-energy CT for SPR concluded that for brain tumors the relative range error should be a systematic underdosing $1.2 \pm 1\%$ [38, 84].

Finally, for head and neck metallic dental implants, which contain high-density material, limit CT's potential causing artifacts and increase range uncertainties (3–10%) [22].

The uncertainties to proton beam range when HU conversion is based on stoichiometric calibration is below 0.5mm according to Paganetti [46].

The second category, called absolute range error is given as an error in range that is not related to the CT image. None of the examined studies consider the absolute range error in the optimization approaches.

A.3.3. Anatomical Uncertainties

Significant variations in anatomy during the treatment of head and neck patients depend on the tumor site (like nasal filling) or can be caused by changes in tumor volume during the treatment period. The main focus of literature studies is rarely the non-rigid uncertainties, as it is difficult to predict and model. Besides, tumor shrinkage, during fractionated treatment, has been usually reported as an additional anatomical uncertainty for some treatment sites [17–20], from research on adaptive proton therapy. To tackle this, some groups are using anatomical variation models that incorporate morphological changes of the planning CT, comparing them subsequently with classical-robust optimization approaches [13, 14, 23]. The results highlighted the benefit of anatomical Robust Optimization(aRO).

Uncertainty HNC	Set-up	Range	Anatomical
Cause	CT isocenter, registration	covenversion of CT HU to stopping power, misalignment of dose contributions from different beam directions[47], density heterogeneities	nasal filing, tumor shrinkage
Type	Systematic & Random[18]	Absolute & Relative	Rigid (not studied) & not Rigid

Table A.1: Main uncertainties in head and neck area

A.4. Accounting for uncertainties in Radiotherapy

The treatment plan's goal is to ensure that the tumor's coverage probability is higher as possible while all OARs receive the minimum possible dose. Several parameters have to be determined in the plan, including beam position, the number of beams, and the settings for dose delivery. To ensure dose deposition in the desired area the plan has to be robust against uncertainties.

Photon therapy

For photon beams, the dose distribution can be described by static dose cloud approximation. Static dose cloud approximation assumes dose invariance under small shifts. For homogeneous treatment sites, this is considered an accurate approximation. The use of static cloud dose approximation implicitly assumes the use of geometric margins which allow a margin generalization for the creating of a robust treatment plan. A margin recipe as a linear combination of standard deviation of systematic Σ and random σ uncertainty, for photon therapy is defined by van Herk's formula [30]:

$$M = 2.5\Sigma + 0.7\sigma \tag{A.1}$$

Figure A.7 provides an example of a Gross Target Volume (GTV), surrounded from the Clinical Target Volume (CTV) margin to account for microscopic tumor expansions, and finally expanding to a margin to ensure robustness against treatment uncertainties.

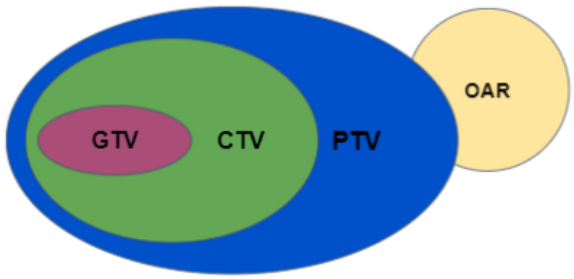


Figure A.7: Schematic representation of GTV, CTV and the treatment margin (PTV)

Proton therapy

In IMPT, due to the range uncertainty nature, the static dose cloud approximation cannot be used to account for these types of error. Therefore, for this treatment, an alternative to margin-based approaches is proposed to account for both setup and range uncertainties. To decrease the plan's sensitivity against errors a technique called robust planning was introduced. The method is analyzed in the following section.

A.5. Robust Planning

Robust planning in the literature is divided into robustness optimization and robustness evaluation. Robust optimization studies, consist of the description of the process to obtain a plan which is robust against the errors, while robustness evaluation, even if it is combined with robust optimization:

- assesses the sensitivity of a given treatment plan to errors
- guarantee that the final result will be sufficiently robust, as robust optimization cannot always be constraints
- is used as an independent check, often also involving scenarios not used in the optimization

For IMPT given a large number of degrees of freedom associated with conformal treatment techniques, treatment planning has become a complex optimization problem and advanced mathematical methods have been developed to solve this. In the studies [48] dose distribution $d = Dx$ is given as a linear function of the incident fluence x, with D defined as dose-influence matrix whose elements D_{ij} store the dose contribution of beamlet j to voxel i for unit fluence. In this review, we focus on the uncertainty in the dose influence matrix D for neurological and head and neck patients. Uncertainty in D means that for the same treatment plan we might end up with different dose distribution.
Different robust planning methods have been developed to ensure adequate dose to the CTV, in which treatment uncertainties are incorporated in the optimization procedure of the clinical treatment planning system. Traditionally, to make IMRT plans less sensitive in uncertainties and superior in terms of OARs, Planning treatment volume(PTV) and planning risk volumes(PRV) were used. Liu et al [26] made a comparison of these methods for IMPT against robust optimization, for head and neck treatment site. PTV approach considers only set up error and sets a margin around the target volume, while robust optimization takes into consideration both set up and range errors. From this study, when HNC cases were considered, OAR sparing in robust optimization was superior compared to PTV-based optimization with also better target dose robustness and homogeneity.
When margins are used the plan is optimized expanding a single nominal scenario (without errors), while in robust optimization dose distributions are optimized for n scenarios, for which different treatment uncertainties are addressed.
The most commonly used method from the literature for robust optimization is via scenario-based optimization or the probabilistic approach. In this approaches the treatment plan is optimized either using all the scenarios at once or by using a combination of them assigning probabilities[73].

A.5.1. Approaches to robust optimization

The most common robust optimization approaches that are implemented and used in commercial and in research treatment planning systems (TPS), along with their mathematical representation are displayed in Table A.1. An extensive study around the robust optimization methods has been conducted by Unkelbach et all [48].

Optimization Method	Worst Case Minimax Optimization [78]	Voxel-wise worst case dose distribution [75]	Stochastic Optimization [47]
Planning System	RayStation & iCycle	Eclipse (Varian)	Pinnacle(Philips)
Description	Determines the pencil beam intensities such that the dose will be as good as possible for the worst error scenario included in the optimization [50]. (RayStation is also equipped with Monte Carlo dose engine)	Evaluates the objective function by using the minimum dose for the target voxel and the maximum for normal tissue [49]	Assigns probabilities to the error scenarios and optimizes based on those
Objective function		$f = \sum_n \frac{\alpha_n}{V_n} \sum_{(i \in V_n)} (d_i - d_i^{pres})^2$	
Mathematical Representation	$minimize_x[max_s f(d_s(x))]$	$minimize_x f(d_s(x))$	$minimize_x \sum_s p_s f(d_s(x))$

Table A.2: Summary of Optimization Methods (s:scenarios, d:dose distribution, f: objective function, x:pencil beam intensities, p:probabilities to error scenarios, D^{presc}:prescribed dose, V: number of voxels, n:index of interest, a: weight of objectives)

Other approaches, like probabilistic minimax optimization, are proposed in the literature [77].

A.6. Limitations of Robust optimization

After its introduction in the clinical treatment planning system (TPS) in 2014, robust optimization has been increasingly used in clinical practice [73]. However, there are still some limitations to overcome to improve robust planning. The areas of improvement are related with:

-**Lack of evaluation protocol**: commonly accepted methods and metrics to quantify the degree of robustness of a given plan [73]. As mentioned in the following sections there is not a common rule of how the acceptable limits for a plan to consider robust. From Table A.2 other studies consider acceptable a CTV dose equal to $0.95D_{prescribed}$ [23], while this plan would have been non-robust for the study of Yang et al, where the same limit is set to $0.98D_{prescribed}$.

-**Additional uncertainties beyond setup and range**: Section A.5 introduces briefly proton therapy uncertainties, and studies that develop site-specific uncertainty models approaches, used for anatomical robust optimization for HNC, are compared.

-**Need for fast TPS tools**: In section A.6, fast dose simulation tools are discussed, which facilitate both robust optimization and robust evaluation.

A.6.1. Scenario based evaluation

The TPS creates robust clinical treatment plans, however, there is still a necessity for clinical robustness evaluation, which gives the information of how the dose distribution changes compare to nominal scenarios. Currently, the clinical robustness evaluation is in a preliminary stage. In 2019, Korevar et al [11] suggested a robust evaluation workflow able to integrate into the clinics. In their work, a scenario-based evaluation method was followed in which dose metrics based on acceptance criteria of error scenarios are used to quantify the uncertainties in Dose Volume Histograms (DVHs).

However, there is no generally accepted protocol or approach on how to assess and report the robustness of

proton plans.

A.6.2. Anatomical Robust Optimization

Anatomical uncertainties are difficult to parametrize and to be included in scenario robust optimization [21]. In Anderson Cancer Center in Houston, about 40% of head and neck patients who undergo IMPT require adaptive planning after re-evaluation of the original IMPT plans on verification CTs during treatment due to anatomical changes, and about 10% of head and neck patients require more than one plan adaptation [14]. Head and neck, as a treatment site that can benefit from the integration of anatomical uncertainties in the robust optimization citeWater,Cub,Yang, reducing considerably the need for online plan adaptation. In Table A.2, the studies that introduced anatomical uncertainties in the scenario-based robust optimization, are displayed

Paper	van de Water(2018) [13]	Cubillos-Mesías(2018) [23]	Yang (2020) [14]
Patient	20 HNC (HNSCC)	10 HNC	10 HNC
Constraints & Objectives for CTV	D_{min}=0.99$D_{prescribed}$, D_{max}=1.06 $D_{prescribed}$	$D_{98\%,CTV} \geqslant 0.95 D_{prescribed}$ $D_{98\%,CTV} \geqslant 1.07 D_{prescribed}$	$D_{98\%,CTV} \geqslant 0.98 D_{prescribed}$ $D_{98\%,CTV} \geqslant 1.05 D_{prescribed}$
Uncertainties modeled	nasal filling	random anatomical changes	inter-fractional uncertainties
aRO	Planning CT & (2,3 or 4) artificial CTs	Planning CT & first two weekly CTs (63 scenarios)	treatment plan which meets dose criteria of primary and first adaptive CT
Comparison with	SFUD/PTV & online plan adaptation	cRO (21 different scenarios), PTV-plan based	three selective robust plan using planning CT, first and second adaptive CT respectively
Conclusions	Improvement compare to SFUD in sparing OARs, worse than online plan adaptation	Dose perturbation effect of anatomical changes is at least the same order of magnitude as setup and range uncertainties	Significantly more robust dose compare to selective robust plans
Remarks	Online plan adaptation outperforms the other models, but faster dose calculation is necessary for dose calculation	for 20% of patients avoid replanning due to use of aRO	limitation is that this optimization could currently only help to improve the plan robustness for the first adaptive plan

Table A.3: Studies which consider anatomical uncertainties for head and neck patients

From these studies, it can be concluded that anatomical uncertainties should be accounted for in the robust optimization for treatment sites susceptible to patient anatomical variations. The three papers used different approaches to account for the anatomical errors. Van de Water et al [13], did not intrinsically take anatomical uncertainties into account, but instead, they used synthetic CTs to account for nasal filling in HNC patients. The second study from Cubilllos-Mesias et al [23], used the planning and the first two weekly CTs to account for anatomical variation, which restricts their method to account for possible future large anatomical variations. Finally, the optimization method developed by Yang et al [14], in which the treatment plan is constructed by using the dose criteria for the planning and at least the first adaptive CT, can be useful only after the first adaptive plan.

A.7. Computationally efficient dose simulation

One of the main limitations in robust optimization planning is that, when more scenarios are included in the optimization, the required implementation time will rapidly increase. Additionally, strategies that account for anatomical variation in the robust optimization inherently increases the computational time required to perform a treatment plan [21]. In fact, fast dose calculations are required to account for treatment uncertainties in treatment planning. In this section, approaches to reduce computational time in dose calculation are explained. As mentioned in Table A.2, clinical machines use either analytical algorithms or stochastic sim-

ulations to calculate the proton dose. Three papers that use different approaches for fast optimization are discussed.

Pencil Beam:

Pencil beam algorithms are analytical solutions that rely on kernels to model proton range in density-scaled water equivalent material [81]. The performance of the algorithm depends on the treatment site considered. In the presence of complex geometries, the accuracy of analytical dose calculation techniques is comprised from heterogeneous tissue interface, like air cavities (lung tumors) or bony structures (skull-base tumors) with soft tissue, because large density differences between the materials increase scattering effects [81]. However, pencil beam simulation is still comparable with more time-consuming techniques [81](etc. Monte Carlo), making it valuable for clinical applications.

Monte Carlo:

Monte Carlo (MC) methods are considered the most accurate dose calculation methods for radiotherapy. Generally, MC belongs to the category of statistical models based on random sampling, which can be implemented for every problem category, and their accuracy increases along with the number of iterations for model evaluation. The tracks of a sufficiently large number of protons are simulated using the random number generated probability distribution governing the individual physical processes [92]. MC used for dose calculation is limited mainly by the delineation of the patient's anatomy from the planning CT. Using MC instead of pencil beam algorithms can reduce the relative range error in half when we consider highly heterogeneous anatomies [52]. However, in contrast with conventional photon therapy, the dose has to be recalculated for each error scenario in proton therapy treatment planning[24]. As MC has a random nature, the resulting dose distributions are limited by statistical uncertainty, which prevents making reliable clinical decisions. This noise is reduced when using a huge number of simulated protons, but it leads to significantly greater computation times. Therefore, there is a trade-off between the computation time and the noise level in MC dose maps [53], which could lead to limitations to be implemented in the clinical treatment planning, especially in the case that the clinical decision focus on the DVH tails, when the noise impact is more significant. Despite the disadvantage, in case that enough samples are available, this method is used as reliable ground truth for comparison with other methods and it is currently used in research studies [15]. When enough computational resources are available, it is feasible to be used in the clinics (RayStation TPS is equipped with MC dose engine).

Recently, the accessibility in graphics processing units (GPUs) has lead to an effort of computationally accelerating dose calculation in radiotherapy. Since 2012 fast GPU-based Monte Carlo code has been implemented, with a significant decrease in required time while keeping high the statistical accuracy. MCsquare (2016) and gPMC (2012) provide more feasible MC solutions that can integrate into the clinical workflow. Sterpin et all [52], report promising time for dose calculation, for head a neck patient, using an MC dose engine (MCsquare).
Currently, there is an effort of gradual replacement of analytical dose calculation techniques with a sampling-based method and more specifically with Monte Carlo (MC) simulations [15, 39–42, 44, 45, 53]. However, this transition has to be done carefully, as the absolute calibration between MC and pencil beam may differ, while MC noise has an impact on relevant dosimetric parameters. Since 2018, TPS RayStation (version 6.0.0.24) is equipped with a MC dose engine, which is proven superior in handling complex conditions [93] in comparison with pencil beam.

Improvements in Pencil Beam and Monte Carlo dose calculation speed are beneficial for robust optimization used in TPS [52, 80]. Allowing for more scenarios to be included in robust optimization methods. However, based on the reported dose calculation times in Table A.3 from the GPU accelerated models, when robust evaluation is an objective, these methods are capable of evaluating dose scenarios in the order of hundreds.
In the following section, PCE meta-model of dose distribution is introduced, which is capable of reducing the dose distribution time to 0.01s per scenario maintaining high accuracy. This characteristic, makes the model a great tool for robust evaluation, as it is capable to to calculate 100.000 fractionated treatments in a reasonable time(depending on the complexity of the tumor), reducing the statistical uncertainty. Figure A.8 is an example of a dosimetric dependency histogram, in the case of 1000 and 100000 scenarios evaluations.

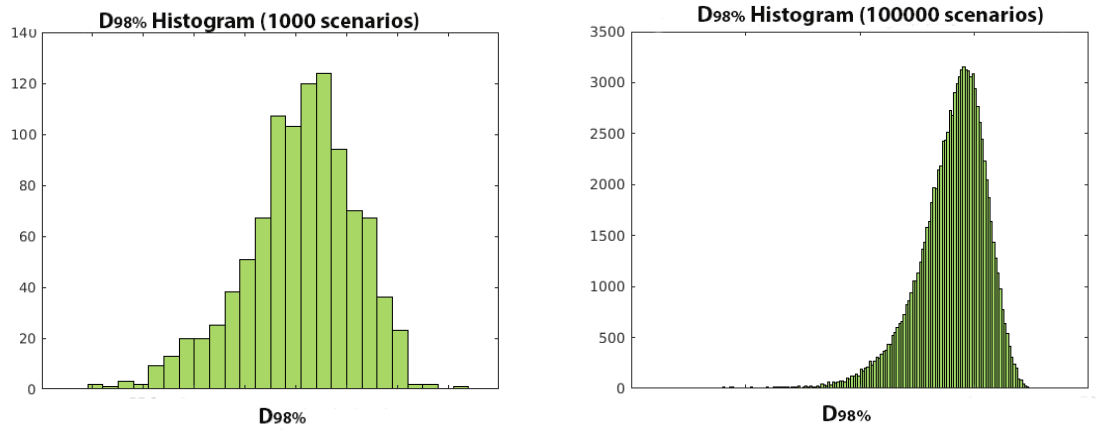


Figure A.8: $D_98\%$ Histogram, for 1000 scenarios (left) and 100000 scenarios (right) evaluation

A.8. Polynomial Chaos Expansion

From previous work [12, 21, 85], it has been proven that Polynomial Chaos Expansion (PCE) can be used as a meta-model of the dose distribution, for fast proton dose calculations. PCE belongs to the category of spectral models in which the function of interest, in our case the dose (D), can be constructed as a function of its input variables(ξ), which corresponds to the considered uncertainties. The choice of the number of input uncertainties is depending on the treatment site. This reconstruction of the desired outcome is obtained by a series expansion of basis vectors Ψ_n multiplied by the corresponding deterministic expansion coefficient r_n. The uncertainties that are currently used in research for the PCE models are four-dimensional and consist of the total setup error in each of the axis, along with the range error.

$$D(\xi) = \sum_{n=0}^{\infty} r_n \Psi_n(\xi)$$

Based on the Wiener-Askey scheme [61], the selected polynomial from the literature is the normalized Hermite. This is based on the fact that the input uncertainties are distributed as Gaussian functions [12, 21, 85]. The model is able to account for uncertainties with a distribution that can be approximated by a well-defined mathematical function, with limited classes [21]. Therefore a limitation of the model is that some binned distributions cannot be approximated by an analytical distribution.
Knowing the polynomial used, the only unknown factor to calculate the output D is the expansion coefficients. Using the basis vector orthogonality of PCE, r_n is given by:

$$r_n = \frac{< D_n(\xi), \Psi_n(\xi) >}{< \Psi_n(\xi), \Psi_n(\xi) >} = \frac{\int\int ... \int D_n(\xi)\Psi_n(\xi) p_\xi(\xi)\, d\xi_1 .... d\xi_n}{< \Psi_n(\xi), \Psi_n(\xi) >}$$

Where $p_\xi(\xi)$ is the probability density function of the variable ξ. As this integral cannot be solved analytically, as it includes the output value of the model which is not priory known, an approximation is necessary. In order to build the model, known dose distributions in limited scenarios are required. Therefore, a TPS is selected for pre-calculated dose distributions in 217 scenarios [21]. Depending on the TPS, a different approach is used for the calculation of the expansion coefficients.

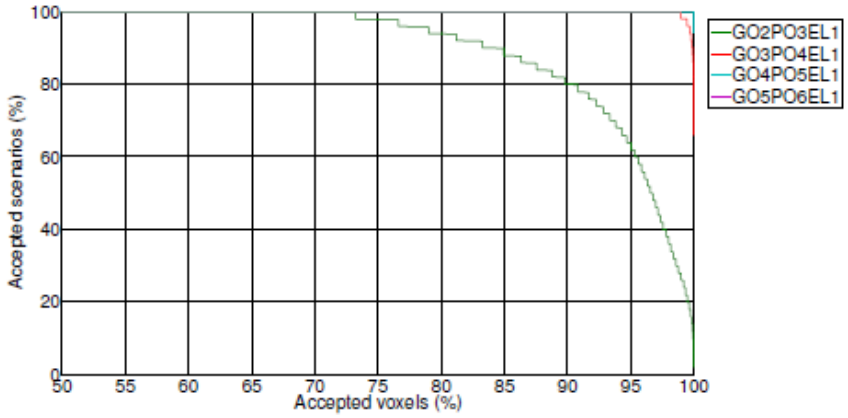
In order to build the model, known dose distributions in limited scenarios are required, therefore the model cannot replace a TPS which also contains information like patient characteristics, delineation, dose algorithm. However, PCE is a very useful tool for robust evaluation as shown in Figure A.8.

Paper	Perko, PC model (2016) [21]	Yao (2019) [80]	Souris [52](2016)
Dose simulation	PCE	fast robust optimizer using GPU or CPU	many core MC (MCsquare)
Errors	Setup & Range (Systematic & Random)	Setup & Range	Setup & Range
Uncertainties	Gaussian (μ=0,σ)	Gaussian (μ=0,σ)	Gaussian (μ=0,σ)
Comparison engine	Erasmus iCycle	Varian Eclipse(version 13.3)	GEANT4
Robust optimization	Erasmus iCycle ASTROID pencil beam dose engine, minmax approach	pencil beam dose engine (in house), gradient method to minimize objective function	NA
Patient	6 HNC (3 unilateral and 3 bilateral cases)	one HNC, one lung & prostate case	one HNC
Model's accuracy	90% correct dose prediction in 95% of scenarios (681 test scenarios, Γ evaluation 1mm/0.1 Gy value) & 99.94% (3mm/3% criteria in Γ evaluation)	fast-proton-dose calculation algorithm	GPU accelerated MC
Equipment	2 core 64bit Intel Xeon X5675, 24GB memory	GPU:Nvidia Titan V CPU:E5 2686V3	Simultaneously use of 240 Xeon Phi threads & 32 host CPU threads
Average time requirements (HNC	dose engine for 217 scenarios and to build PCE requires 80 min	25 times faster than Varian Eclipse	13h for 300 scenarios MC
Average time required to dose calculation (HNC)	less than 0.01s (PCE)	16s (CPU), 3s (GPU)	153s per dose scenario

Table A.4: Studies on fast dose calculations focus on HNC

A.8.1. PCE Parameters

From the previous studies [12, 18], a trade-off between accuracy and computational time of PCE model have been reported. The most important parameters to modify the model´s accuracy and speed are related to the number of polynomials used in PCE and to the number of sampled error scenarios where the dose distribution has to be previously calculated[12]. In Figure A.9, the results from Te Haar et al work [12] are displayed. According to the report, when using gamma evaluation for a skull-based patient case, with dose criterion at 0.1 Gy and distance-to-agreement criterion at 1 mm, with error scenarios from the 99% confidence ellipsoid, the number of voxels that passed the evaluation for increasing polynomial order appear in. From these images, a large trade-off between time and number of polynomials(PO) included in the PCE, is observed. Based on that, 5th order polynomials are usually selected. However, other treatment areas, like the prostate, may require the additional time that higher-order polynomials require, as less number of voxels is acceptable when 5th PO is used. During our thesis work, we also conducted a study to speed up the construction



PO	accepted voxels
3	3%
4	76%
5	98%
6	100%

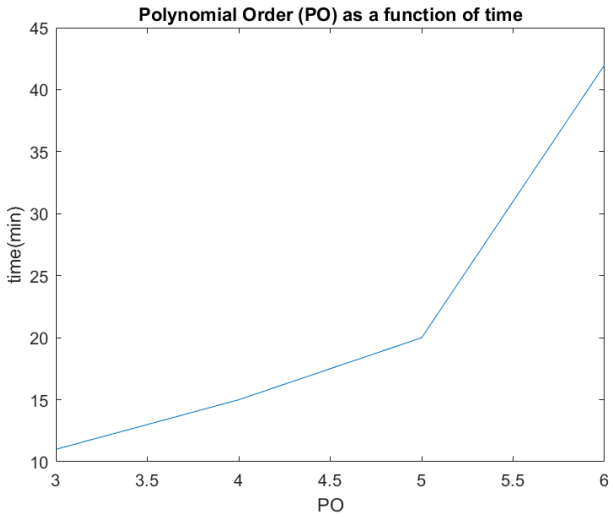


Figure A.9: Percentage of accepted voxels as a function of accepted scenarios for different polynomial orders (above), Polynomial order as a function of time (below) [12]

time of the model. From this, we estimated the optimal Monte Carlo noise level that can be used to improve the model's construction time.

A.9. Discussion

The current literature review provides an overview of the proton delivery uncertainties and approaches that have been developed and implemented to deal with these. Proton therapy is sensitive to all type of uncertainties and there is not feasible to eliminate all of them before treatment. As mention in the first section, the majority of the research around robust optimization includes setup and range errors, which are usually integrated into the models as Gaussian distributions. Clinically setup data confirm that indeed, Gaussian distributions are a good approximation for these uncertainties. This is considered as a real-world example of the central limit theorem, where if sufficiently many uncorrelated factors contribute to a single outcome, the results will always be a normal distribution. Depending on the treatment site additional uncertainties should be included in the optimization, to reduce the need for an adaptive treatment plan.

This review, compared studies that focus on the importance of anatomical uncertainties for HNC patients. The developed anatomical robust optimizations(aRO) proved that accounting for these types of uncertainties, was beneficial (less need for adaptation) to at least 20 percent of the patients(in every study). The location of the tumor was crucial to determine if the patient will be benefited from the use of aRO. In general, it is advisable to avoid irradiation through a path that is highly inhomogeneous or known anatomical changes are present (etc. nasal cavity). The studies showed that when the changes are localized and can be well modeled, anatomical robust optimization can help to reduce the need for online adaptation and in some cases, they can even account for setup and range errors. In every case, additional CTs to the planning CT was necessary for the optimization. As a specific patient population can benefit from this method, it would have been an interesting research topic in which are the criteria that this optimization is beneficial. An important general remark from these studies was the necessity of models that allows generating dose evaluations fast and accurately.

Currently, advantages in computational resources and the development of GPUs, have significantly improved computational time for both analytical algorithms, like pencil beam, and MC dose engines.
Even if MC is more time-consuming, the computation time does not scale necessarily with complexity, which means higher uncertainties and more error scenarios can be considered in the simulation. However, even the times reported for robust evaluation of head and neck cancer patient, (around 13 hours for 300 required scenarios) makes it only useful for a final evaluation. That suggests that there is an improvement needed in the statistical criteria used to make the MCsquare faster or even larger investment in computational equipment.

On contrary, the PCE meta-model appears to be computationally superior to its competitors for robust evaluation. Considering the examined treatment area the model presents high accuracy in a fraction of time. The total time needed for building the model is on average 80min while the time required to calculate a dose scenario is 0.01s against the 153s of the MCsquare, and 100 faster than the dose engine. Someone can argue that PCE requires a specific number of input scenarios for its construction (around 200), which depending on the dose engine, create additional time requirements. However, as mentioned in the literature, the number of the scenarios that MC models required to work is around 300. When the scenarios needed for PCE are available, the model's construction is not a time-consuming procedure. After building the PCE model, the user has a patient- and treatment- plan-specific modeling of the dose under the presence of treatment uncertainties, which further reduces statistical errors and allows to calculate far more than 300 error scenarios in a couple of minutes. The main advantage of PCE is the good and accurate statistics on clinically relevant dosimetric parameters. When 300 scenarios are considered the statistical noise is on the order of 5-10%. PCE evaluating 100000 error scenario reduces the error to below 0.3%. Additionally, the model allows accounting for fractionation effects. However, due to the model's nature, PCE complexion increases when uncertainties difficult to parametrize (like anatomical errors) are considered. Additionally, some treatment sites like prostate case [21], required higher-order polynomials for sufficient accuracy increasing significantly computational time of PCE.

Finally, MC is indeed considered the most accurate dose calculation method for both proton and photon therapy. However, it is worth highlight that MC dose engines are not the only accurate method for dose simulation used in commercial TPSs. For instance, the Eclipse treatment planning system (Varian) has developed and implemented computationally efficient deterministic linear Boltzmann transport equation solvers for photon therapy, with comparable accuracy to MC dose calculation (they account for the effect of electron transport), but faster [93].

This literature study highlights some important improvements that have been made in the last years, to resolve proton therapy's robust planning limitations. There are still limitations to overcome, like reach a consensus in the scientific community on which methods and metrics are the most appropriate for robust optimization and evaluation [73]. However, this review showed that with the right tool combination significant time reduction can be achieved. As mentioned, the pencil beam algorithm reports similar dose distribution with MC for most of the treatment sites. If a fast pencil beam algorithm is used to generate the required doses needed for PCE construction, a remarkable time improvement of robust evaluation is observed. In the case that more complex uncertainties have to be taken into consideration MC engines could be useful.

Future of robust planning
This work highlights fast changes and improvements of robust treatment planning for proton therapy in the last years. Most of the studies are conducted over the last ten years, indicating how fast radiotherapy is improving. Based on the overall observations from the different papers, we estimate that in the near future probabilistic optimization approaches, that are currently restrained from computational resources, will be dominant. Additionally, the optimization is expected to account for additional uncertainties, like anatomical variations or internal models through advanced modeling (Principal component analysis, mechanical modeling). Furthermore, considering the increased interest in online adaptive replanning, the role of robustness might change. As it can become more focused on a number of fractions and not the whole treatment, therefore the errors taken into account will differ. For instance, systematic setup error values will not consider the average of all fractions but will change after a set of fractions.

A.10. Conclusion

This review provides an overview of the status of the robust treatment plan in IMPT. The research highlights that there is not a global solution to account for uncertainties in IMPT, depending on the treatment site and even in the same region different approach should be followed. Models for anatomical robust optimization for the HCN area were discussed. Additionally, efficient dose computation methods for robust optimization and evaluation were presented. Finally, we conclude that in the last years, there is a noticeable improvement in robust treatment planning.

Bibliography

[1] WILSON RR. Radiological use of fast protons. Radiology. 1946 Nov;47(5):487-91. doi: 10.1148/47.5.487. PMID: 20274616.

[2] Tian X, Liu K, Hou Y, Cheng J, Zhang J. The evolution of proton beam therapy: Current and future status. Mol Clin Oncol. 2018;8(1):15-21. doi:10.3892/mco.2017.1499

[3] Wouters BG, Skarsgard LD, Gerweck LE, Carabe-Fernandez A, Wong M, Durand RE, Nielson D, Bussiere MR, Wagner M, Biggs P, Paganetti H, Suit HD. Radiobiological intercomparison of the 160 MeV and 230 MeV proton therapy beams at the Harvard Cyclotron Laboratory and at Massachusetts General Hospital. Radiat Res. 2015 Feb;183(2):174-87. doi: 10.1667/RR13795.1. Epub 2015 Jan 14. Erratum in: Radiat Res. 2015 Apr;183(4):e51. PMID: 25587741.

[4] Fukuda K, Okumura T, Abei M, Fukumitsu N, Ishige K, Mizumoto M, Hasegawa N, Numajiri H, Ohnishi K, Ishikawa H, Tsuboi K, Sakurai H, Hyodo I. Long-term outcomes of proton beam therapy in patients with previously untreated hepatocellular carcinoma. Cancer Sci. 2017 Mar;108(3):497-503. doi: 10.1111/cas.13145. PMID: 28012214; PMCID: PMC5378259.

[5] Jagodinsky, Justin C., Paul M. Harari, and Zachary S. Morris. "The promise of combining radiation therapy with immunotherapy." International Journal of Radiation Oncology* Biology* Physics 108.1 (2020): 6-16.

[6] Bortfeld, Thomas, Harald Paganetti, and H. Kooy. "MO-A-T-6B-01: Proton Beam Radiotherapy—The State of the Art." Medical Physics 32.6Part13 (2005): 2048-2049.

[7] Moreno AC, Frank SJ, Garden AS, Rosenthal DI, Fuller CD, Gunn GB, Reddy JP, Morrison WH, Williamson TD, Holliday EB, Phan J, Blanchard P. Intensity modulated proton therapy (IMPT) - The future of IMRT for head and neck cancer. Oral Oncol. 2019 Jan;88:66-74. doi: 10.1016/j.oraloncology.2018.11.015. Epub 2018 Nov 21. PMID: 30616799; PMCID: PMC6615027.

[8] Paganetti, Harald, ed. Proton therapy physics. CRC press, 2018.

[9] Crestaux, Thierry, Olivier Le Maitre, and Jean-Marc Martinez. "Polynomial chaos expansion for sensitivity analysis." Reliability Engineering System Safety 94.7 (2009): 1161-1172.

[10] Wens Kong,"Margin recipes for stereotactic body radiation therapy for prostate cancer treatments", TU Delft, 2019

[11] Korevaar EW, Habraken SJM, Scandurra D, Kierkels RGJ, Unipan M, Eenink MGC, Steenbakkers RJHM, Peeters SG, Zindler JD, Hoogeman M, Langendijk JA. Practical robustness evaluation in radiotherapy - A photon and proton-proof alternative to PTV-based plan evaluation. Radiother Oncol. 2019 Dec;141:267-274. doi: 10.1016/j.radonc.2019.08.005. Epub 2019 Sep 3. PMID: 31492443.

[12] C.E. ter Haar:Robustness Recipes for Proton Therapy. Polynomial Chaos Expansion as a tool to construct robustness recipes for proton therapy, Delft University of Technology (2018)

[13] van de Water S, Albertini F, Weber D C, Heijmen B J M, Hoogemann M S and Lomax A J 2018 Anatomical robust optimization to account for nasal cavity filling variation during intensity-modulated proton therapy: a comparison with conventional and adaptive planning strategies Phys. Med. Biol. 63 025020

[14] Yang Z et al 2020 Multiple-CT optimization: an adaptive optimization method to account for anatomical changes in intensity-modulated proton therapy for head and neck cancers Radiother. Oncol. 142 124–32

[15] C.M.H. Hartman:Sensitivity and uncertainty analysis in proton therapy plan using Polynomial Chaos Methods, Rotterdam (2014)

[16] Mrozowska, Marzena, and Paweł Kukolowicz. "Relationships between various indices of doses distribution homogeneity." Reports of Practical Oncology and Radiotherapy 20.4 (2015): 278-283.

[17] Jones, Bleddyn. "Proton radiobiology and its clinical implications." ecancermedicalscience 11 (2017).

[18] Sebastian Van Der Voort, Steven Van De Water, Zoltán Perkó, Ben Heijmen, Danny Lathouwers, and Mischa Hoogeman. Robustness Recipes for Minimax Robust Optimization in Intensity Modulated Proton Therapy for Oropharyngeal Cancer Patients. International Journal of Radiation Oncology Biology Physics, 95(1):163–170, 2016. ISSN 1879355X. doi: 10.1016/j.ijrobp.2016.02.035.

[19] S.R. van der Voort, Danny Lathouwers, Dr. Ir. S. van de Water, Zoltán Perkó, and Mischa S. Hoogeman: Application of polynomial chaos in proton therapy, Rotterdam (2015)

[20] Malyapa R, Lowe M, Bolsi A, Lomax AJ, Weber DC, Albertini F. Evaluation of Robustness to Setup and Range Uncertainties for Head and Neck Patients Treated With Pencil Beam Scanning Proton Therapy. Int J Radiat Oncol Biol Phys. 2016 May 1;95(1):154-162. doi: 10.1016/j.ijrobp.2016.02.016. Epub 2016 Feb 11. PMID: 27084638.

[21] Perkó, Zoltán, et al. "Fast and accurate sensitivity analysis of IMPT treatment plans using Polynomial Chaos Expansion." Physics in Medicine Biology 61.12 (2016): 4646.

[22] Beddok A, Vela A, Calugaru V, Tessonnier T, Kubes J, Dutheil P, Gérard A, Vidal M, Goudjil F, Florescu C, Kammerer E, Bénézery K, Hérault J, Bourhis J, Thariat J; Gortec et de trois centres français. Protonthérapie des carcinomes épidermoïdes des voies aérodigestives supérieures : de la physique à la clinique [Proton therapy for head and neck squamous cell carcinomas: From physics to clinic]. Cancer Radiother. 2019 Sep;23(5):439-448. French. doi: 10.1016/j.canrad.2019.05.015. Epub 2019 Jul 26. PMID: 31358445.

[23] Cubillos-Mesías M, Troost EGC, Lohaus F, Agolli L, Rehm M, Richter C, Stützer K. Quantification of plan robustness against different uncertainty sources for classical and anatomical robust optimized treatment plans in head and neck cancer proton therapy. Br J Radiol. 2020 Mar;93(1107):20190573. doi: 10.1259/bjr.20190573. Epub 2019 Nov 28. PMID: 31778315; PMCID: PMC7066968.

[24] McGowan S E, Albertini F, Thomas S J and Lomax A J 2015 Defining robustness protocols: a method to include and evaluate robustness in clinical plans Phys. Med. Biol. 60 2671–84

[25] Cao W, Lim GJ, Lee A, et al. Uncertainty incorporated beam angle optimization for IMPT treatment planning. Med Phys. 2012;39(8):5248-5256. doi:10.1118/1.4737870

[26] Liu W, Frank SJ, Li X, Li Y, Zhu RX, Mohan R. PTV-based IMPT optimization incorporating planning risk volumes vs robust optimization. Med Phys. 2013 Feb;40(2):021709. doi: 10.1118/1.4774363. PMID: 23387732; PMCID: PMC3562272.

[27] van de Water TA, Bijl HP, Schilstra C, Pijls-Johannesma M, Langendijk JA. The potential benefit of radiotherapy with protons in head and neck cancer with respect to normal tissue sparing: a systematic review of literature. Oncologist. 2011;16:366–377.

[28] Lomax AJ, Goitein M, Adams J. Intensity modulation in radiotherapy: photons versus protons in the paranasal sinus. Radiother Oncol. 2003;66:11–18.

[29] Stenekar M, A, Schneider U. Intensity Modulated photon and proton therapy for the treatment of head and neck tumors. Radiother Oncol. 2006;80:263–267

[30] van Herk M. Errors and margins in radiotherapy. Semin Radiat Oncol. 2004 Jan;14(1):52-64. doi: 10.1053/j.semradonc.2003.10.003. PMID: 14752733

[31] P. Blanchard, S.J. Frank, Place de la protonthérapie en cancérologie ORL, Cancer/Radiothérapie, Volume 21, Issues 6–7, 2017, Pages 515-520, ISSN 1278-3218

[32] Wu Y, Chi P, Chen Q, Krauss DJ, Yan D, Martinez A. Adaptive replanning strategies accounting for shrinkage in head and neck IMRT. Int J Radiat Oncol Biol Phys 2009;75:924–32.

[33] Elstrøm UV, Wysocka BA, Muren LP, Petersen JB, Grau C. Daily kV cone-beam CT and deformable image registration as a method for studying dosimetric consequences of anatomic changes in adaptive IMRT of head and neck cancer. Acta Oncol 2010;49:1101–8.

[34] Hansen EK, Bucci MK, Quivey J, Weinberg V, Xia P. Repeat CT imaging and replanning during the course of IMRT for head-and-neck cancer. Int J Radiat Oncol Biol Phys 2006;64:355–62.

[35] Barker Jr JL, Garden AS, Ang KK, O'Daniel JC, Wang H, Court LE, et al. Quantification of volumetric and geometric changes occurring during fractionated radiotherapy for head-and-neck cancer using an integrated CT/linear accelerator system. Int J Radiat Oncol Biol Phys 2004;59:960–70.

[36] Lomax AJ. Intensity modulated proton therapy and its sensitivity to treatment uncertainties 2: the potential effects of inter-fraction and inter-field motions. Phys Med Biol. 2008 Feb 21;53(4):1043-56. doi: 10.1088/0031-9155/53/4/015. Epub 2008 Jan 29. PMID: 18263957.

[37] Bolsi A, Lomax AJ, Pedroni E, Goitein G, Hug E. Experiences at the Paul Scherrer Institute with a remote patient positioning procedure for high-throughput proton radiation therapy. Int J Radiat Oncol Biol Phys. 2008 Aug 1;71(5):1581-90. doi: 10.1016/j.ijrobp.2008.02.079. PMID: 18640501.

[38] Wohlfahrt P, Richter C. Status and innovations in pre-treatment CT imaging for proton therapy. Br J Radiol. 2020;93(1107):20190590. doi:10.1259/bjr.20190590

[39] Petti PL Evaluation of a pencil-beam dose calculation technique for charged particle radiotherapy. Int J Radiat Oncol Biol Phys. 1996 Jul 15; 35(5):1049-57

[40] Schaffner B, Pedroni E, Lomax A Dose calculation models for proton treatment planning using a dynamic beam delivery system: an attempt to include density heterogeneity effects in the analytical dose calculation. Phys Med Biol. 1999 Jan; 44(1):27-41

[41] Szymanowski H, Oelfke U Two-dimensional pencil beam scaling: an improved proton dose algorithm for heterogeneous media.Phys Med Biol. 2002 Sep 21; 47(18):3313-30.

[42] Tourovsky A, Lomax AJ, Schneider U, Pedroni E Monte Carlo dose calculations for spot scanned proton therapy.Phys Med Biol. 2005 Mar 7; 50(5):971-81.

[43] Soukup M, Alber M Influence of dose engine accuracy on the optimum dose distribution in intensity-modulated proton therapy treatment plans.Phys Med Biol. 2007 Feb 7; 52(3):725-40.

[44] Titt U, Sahoo N, Ding X, Zheng Y, Newhauser WD, Zhu XR, Polf JC, Gillin MT, Mohan R Assessment of the accuracy of an MCNPX-based Monte Carlo simulation model for predicting three-dimensional absorbed dose distributions.Phys Med Biol. 2008 Aug 21; 53(16):4455-70.

[45] Paganetti H, Jiang H, Parodi K, Slopsema R, Engelsman M Clinical implementation of full Monte Carlo dose calculation in proton beam therapy.Phys Med Biol. 2008 Sep 7; 53(17):4825-53

[46] Paganetti. Range uncertainties in proton therapy and the role of Monte Carlo simulations. Phys Med Biol. 2012; 57(11):R99-R117

[47] Unkelbach J, Bortfeld T, Martin BC, Soukup M. Reducing the sensitivity of IMPT treatment plans to setup errors and range uncertainties via probabilistic treatment planning. Med Phys. 2009 Jan;36(1):149-63. doi: 10.1118/1.3021139. PMID: 19235384; PMCID: PMC2673668.

[48] Unkelbach J, Alber M, Bangert M, Bokrantz R, Chan TCY, Deasy JO, Fredriksson A, Gorissen BL, van Herk M, Liu W, Mahmoudzadeh H, Nohadani O, Siebers JV, Witte M, Xu H. Robust radiotherapy planning.Phys Med Biol. 2018 Nov 12; 63(22):22TR02

[49] Zaghian, Maryam, et al. "Comparison of linear and nonlinear programming approaches for "worst case dose" and "minmax" robust optimization of intensity-modulated proton therapy dose distributions." Journal of applied clinical medical physics 18.2 (2017): 15-25.

[50] Paganetti, H., Beltran, C., Both, S., Dong, L., Flanz, J., Furutani, K., ... Lomax, T. (2021). Roadmap: proton therapy physics and biology. Physics in Medicine Biology, 66(5), 05RM01.

[51] Fredriksson A, Forsgren A, Hårdemark B. Minimax optimization for handling range and setup uncertainties in proton therapy. Med Phys. 2011 Mar;38(3):1672-84. doi: 10.1118/1.3556559. PMID: 21520880.

[52] Souris K, Lee JA, Sterpin E. Fast multipurpose Monte Carlo simulation for proton therapy using multi- and many-core CPU architectures. Med Phys. 2016 Apr;43(4):1700. doi: 10.1118/1.4943377. PMID: 27036568.

[53] Cohilis, Marie, et al. "A noise correction of the -index method for Monte Carlo dose distribution comparison." Medical physics 47.2 (2020): 681-692.

[54] Eekers DB, In 't Ven L, Roelofs E, Postma A, Alapetite C, Burnet NG, Calugaru V, Compter I, Coremans IEM, Høyer M, Lambrecht M, Nyström PW, Méndez Romero A, Paulsen F, Perpar A, de Ruysscher D, Renard L, Timmermann B, Vitek P, Weber DC, van der Weide HL, Whitfield GA, Wiggenraad R, Troost EGC; "European Particle Therapy Network" of ESTRO. The EPTN consensus-based atlas for CT- and MR-based contouring in neuro-oncology. Radiother Oncol. 2018 Jul;128(1):37-43. doi: 10.1016/j.radonc.2017.12.013. Epub 2018 Mar 13. PMID: 29548560.

[55] Weber DC, Lim PS, Tran S, et al. Proton therapy for brain tumours in the area of evidence-based medicine. Br J Radiol. 2020;93(1107):20190237. doi:10.1259/bjr.20190237

[56] Naziri, Jason, and Steven J. DiBiase. "Toxicity of Cranial and Spinal Cord Irradiation." Brain and Spinal Tumors-Primary and Secondary. IntechOpen, 2019.

[57] Feain I, Shieh CC, White P, O'Brien R, Fisher S, Counter W, Lazarakis P, Stewart D, Downes S, Jackson M, Baxi S, Whelan B, Makhija K, Huang CY, Barton M, Keall P. Functional imaging equivalence and proof of concept for image-guided adaptive radiotherapy with fixed gantry and rotating couch. Adv Radiat Oncol. 2016 Nov 8;1(4):365-372. doi: 10.1016/j.adro.2016.10.004. PMID: 28740907; PMCID: PMC5514241.

[58] Marc Shapiro, Proton Therapy Delivers Targeted Treatment to Brain and Spinal Cord Tumors, John Hopkins Medicine, Spring/Summer 2020

[59] Petit, Steven, Joao Seco, and Hanne Kooy. "Increasing maximum tumor dose to manage range uncertainties in IMPT treatment planning." Physics in Medicine Biology 58.20 (2013): 7329.

[60] Albertini F, Hug E B and Lomax A J 2011 Is it necessary to plan with safety margins for actively scanned proton therapy? Phys. Med. Biol

[61] D. Xiu and G. E. Karniadakis, "The Wiener–Askey polynomial chaos for stochastic differential equations," SIAM Journal on Scientific Computing, vol. 24, no. 2, pp. 619–644, 2002.

[62] C. F. JeffWu andMichaelHamada. Experiments: Planning, analysis, and parameter design optimization. JohnWiley Sons, Inc., page 112, 2000.

[63] Blatman, Géraud, and Bruno Sudret. "Adaptive sparse polynomial chaos expansion based on least angle regression." Journal of computational Physics 230.6 (2011): 2345-2367.

[64] Volker Barthelmann, Erich Novak, Klaus Ritter, High Dimensional Polynomial Interpolation on Sparse Grids, Advances in Computational Mathematics, Volume 12, Number 4, 2000, pages 273-288.

[65] Torre, Emiliano, et al. "Data-driven polynomial chaos expansion for machine learning regression." Journal of Computational Physics 388 (2019): 601-623.

[66] Cilla, Savino et al. "Volumetric modulated arc therapy (VMAT) and simultaneous integrated boost in head-and-neck cancer: is there a place for critical swallowing structures dose sparing?." The British journal of radiology vol. 89,1059 (2016): 20150764. doi:10.1259/bjr.20150764

[67] Jatinder R Palta PhD,Understanding the Uncertainties in Proton Therapy, Department of Radiation Oncology Virginia Commonwealth University National Radiation Oncology Program(10P4H) Veterans Health Administration Richmond, VA

[68] España S, Paganetti H. The impact of uncertainties in the CT conversion algorithm when predicting proton beam ranges in patients from dose and PET-activity distributions. Phys Med Biol 2010;55:7557–71.

[69] Unkelbach J, Bortfeld T, Martin BC, Soukup M. Reducing the sensitivity of IMPT treatment plans to setup errors and range uncertainties via probabilistic treatment planning. Med Phys. 2009;36(1):149-163. doi:10.1118/1.3021139

[70] Reilly KM. Brain tumor susceptibility: the role of genetic factors and uses of mouse models to unravel risk. Brain Pathol. 2009;19(1):121-131. doi:10.1111/j.1750-3639.2008.00236.x

[71] S. van de Water, A. Kraan, S. Breedveld, W. Schillemans, D. Teguh, H. Kooy, T. Madden, B. Heijmen, and M. Hoogeman, "Improved efficiency of multi-criteria IMPT treatment planning using iterative resampling of randomly placed pencil beams," Physics in medicine and biology, vol. 58, no. 19, pp. 6969–6983, 2013

[72] Schaffner B, Pedroni E. The precision of proton range calculations in proton radiotherapy treatment planning: experimental verification of the relationship between CT-HU and proton stopping power. Phys Med Biol 1998;43:1579–92.

[73] Hernandez V, Hansen CR, Widesott L, Bäck A, Canters R, Fusella M, Götstedt J, Jurado-Bruggeman D, Mukumoto N, Kaplan LP, Koniarová I, Piotrowski T, Placidi L, Vaniqui A, Jornet N. What is plan quality in radiotherapy? The importance of evaluating dose metrics, complexity, and robustness of treatment plans. Radiother Oncol. 2020 Dec;153:26-33. doi: 10.1016/j.radonc.2020.09.038. Epub 2020 Sep 25. PMID: 32987045.

[74] Ligtenberg, Ms Dr G., and Ms S. Kleijnen. "Report Indications for proton therapy (part 1):• Intraocular tumours."

[75] Fredriksson A, Forsgren A and Hardemark B 2011 Minimax optimization for handling range and setup uncertainties in proton therapy Med. Phys. 38 1672–84

[76] Mayo, Charles, et al. "Radiation dose–volume effects of optic nerves and chiasm." International Journal of Radiation Oncology* Biology* Physics 76.3 (2010): S28-S35.

[77] Fredriksson A 2012 A characterization of robust radiation therapy treatment planning methods-from expected value to worst case optimization Med. Phys. 39 5169–81

[78] Pflugfelder D, Wilkens J J and Oelfke U 2008 Worst case optimization: a method to account for uncertainties in the optimization of intensity modulated proton therapy Phys. Med. Biol. 53 1689–700

[79] Tilly D, Ahnesjö A. Fast dose algorithm for generation of dose coverage probability for robustness analysis of fractionated radiotherapy. Phys Med Biol 2015;60:5439–54.

[80] Xu Y, Chen J, Cao R, Liu H, Xu XG, Pei X. A fast robust optimizer for intensity modulated proton therapy using GPU. J Appl Clin Med Phys. 2020;21(3):123-133. doi:10.1002/acm2.12835

[81] Paganetti, Harald, et al. "Comparison of pencil-beam and Monte Carlo calculated dose distributions for proton therapy of skull-base and para-spinal tumors." World Congress on Medical Physics and Biomedical Engineering 2006. Springer, Berlin, Heidelberg, 2007.

[82] Komninakidis,"A planning study using iCycle for lung cancer patients", TU Delft, 2017.

[83] J.H. Salverda, "Probabilistic Treatment Planning with Polynomial Chaos Expansion for Proton Therapy",TU Delft, 2019.

[84] Jesús Rojo-Santiago, Steven J. M. Habraken, Danny Lathouwers, Alejandra Méndez Romero, MD, Zoltán Perkó, and Mischa S. Hoogeman: Accurate assessment of a Dutch practical robustness evaluation protocol in clinical PT with pencil beam scanning for neurological tumors, Rotterdam (2021)

[85] S.R. van der Voort, Danny Lathouwers, Dr. Ir. S. van de Water, Zoltán Perkó, and Mischa S. Hoogeman: Application of polynomial chaos in proton therapy, Rotterdam (2015)

[86] International Commission on Radiation Units and Measurements. ICRU Report 62. Prescribing, recording, and reporting photon beam therapy (Supplement to ICRU Report 50). Bethesda, MD: ICRU, (1999)

[87] Wohlfahrt, P., Möhler, C., Richter, C., Greilich, S. (2018). Evaluation of stopping-power prediction by dual-and single-energy computed tomography in an anthropomorphic ground-truth phantom. International Journal of Radiation Oncology* Biology* Physics, 100(1), 244-253.

[88] Kurz, C., Süss, P., Arnsmeyer, C., Haehnle, J., Teichert, K., Landry, G., ... Richter, C. (2018). Dose-guided patient positioning in proton radiotherapy using multicriteria-optimization. Zeitschrift Für Medizinische Physik.

[89] Stuetzer, K., et al. “SU-F-J-203: Retrospective Assessment of Delivered Proton Dose in Prostate Cancer Patients Based On Daily In-Room CT Imaging.” Medical Physics 43.6Part12 (2016): 3455-3455.

[90] Rit, Simon, et al. "Filtered-backprojection reconstruction for a cone-beam computed tomography scanner with independent source and detector rotations." Medical physics 43.5 (2016): 2344-2352.

[91] Maeda, Yoshikazu, et al. "Positioning accuracy and daily dose assessment for prostate cancer treatment using in-room CT image guidance at a proton therapy facility." Medical physics 45.5 (2018): 1832-1843.

[92] Kan MW, Yu PK, Leung LH. A review on the use of grid-based Boltzmann equation solvers for dose calculation in external photon beam treatment planning. Biomed Res Int. 2013;2013:692874. doi:10.1155/2013/692874

[93] Widesott, Lamberto, et al. "Improvements in pencil beam scanning proton therapy dose calculation accuracy in brain tumor cases with a commercial Monte Carlo algorithm." Physics in Medicine Biology 63.14 (2018): 145016.